A BIM-based ontological seismic multi-objective evaluation and optimisation design for buildings

Shang Gao

Cardiff School of Engineering

Cardiff University

Thesis submitted for the degree of Doctor of Philosophy

January 2025

ACKNOWLEDGEMENT

I would like to extend thanks to my supervisors, Professor Haijiang Li and Professor Monjur Mourshed for their tremendous academic support, scholarly guidance and keen observations, which have enabled me to complete this research work objectively and comprehensively. I would like to give my special thanks to Professor Haijiang who introduced me to the research world of BIM and Artificial Intelligence. With his consistent inspiration, valuable advice and enthusiasm provided from the first day I began my PhD career. I would also like to express my gratitude to Professor Monjur for guiding the research with his vast knowledge.

A big thanks to my colleagues of Cardiff BIM team, and faculty and staff at Cardiff school of Engineering. Without our interesting discussion, enjoyable time and moral support, the research work would not have been as enjoyable.

Finally, but by no means least, I would thank the unconditional love and support from my parents. Also, my cat comforted me a lot during my struggling time. I dedicate this thesis to them.

ABSTRACT

The proposal of Performance-Based Seismic Design (PBSD) theory improves the efficiency of simultaneously designing and evaluating structures in earthquake engineering. Leveraging digital tools to enhance the quality and efficiency of engineering application is an important proposition for the information reform in the field of seismic design. Based on PBSD theory and with the help of Building Information Modeling (BIM), semantic web, Artificial Intelligence (AI) and other technologies, this thesis realises the automated evaluation and optimization design for individual buildings to analyse their seismic performance. Additionally, it predicts the seismic damage of groups of building in a specific location. The research will provide effective guidance for the overall and detail-oriented regional seismic precaution.

In 2001, the Applied Technology Council (ATC) received the initial contract from the Federal Emergency Management Agency (FEMA) to create advanced PBSD for both newly constructed and pre-existing structures. The main outcome of this project is a collection of volumes, supporting documents, and digital resources known as the FEMA P-58 Seismic Performance Assessment of Building, Methodology and Implementation. This thesis utilises BIM technology to seamlessly integrate and convey detailed technical information at the component level, following the guidelines set by above documents in its first section. Then, Ontology is utilised to articulate the evaluation content and reasoning, while also organising, storing, associating and interacting with the many and disparate data sources for evaluation in a cohesive manner. This enables the automated evaluation of seismic performance for individual buildings. Therefore, the seismic optimisation design, guided by the “Return on Investment” (ROI) criterion, aims to achieve an equilibrium between the initial building expense and the anticipated earthquake damage. The multi-objective genetic algorithm, known as NSGA-II, is employed to carry out the optimisation iterations at the building’s component level. The second section focuses on multi-scale regional seismic precaution and establishes a seismic response prediction model using Artificial Neural Network (ANN). This model not only expedites the rapid acquisition of seismic performance distribution for building groups, but also provides a framework for more comprehensive seismic design and evaluation of individual buildings with significant damage.

Ultimately, this thesis demonstrates the enhancement of seismic performance assessment quality for building and the optimisation degree of seismic design through the application of practical cases. Furthermore, the operational efficiency of both has been improved. Moreover, this thesis not only guarantees the precision of seismic response prediction, but also expands the model's applicability by facilitating the adoption of PBSD from individual buildings to regional groups.

Keywords: PBSD, Ontology, BIM, ANN

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

AI	**Artificial Intelligence**
API	**Application Programming Interface**
ANN	**Artificial Neural Network**
ATC	**Applied Technology Council**
BIM	**Building Information Modelling**
BSPEF	**Building Seismic Performance Evaluation Framework**
DXF	**Data Exchange Format**
EDP	**Earthquake Demand Parameter**
EBPN	**Error Back Propagation Network**
FA	**Function Approximation**
FEMA	**Federal Emergency Management Agency**
FMP	**Feedforward Multi-layer Perception**
HCI	**Human-Computer Interaction**
IDR	**Inter-story Drift Ratio**
IDA	**Incremental Dynamic Analysis**
IFC	**Industry Foundation Classes**
MOGA	**Multi-Objective Genetic Algorithm**
NSGA	**Non-dominated Sorting Genetic Algorithm**
PBSD	**Performance-Based Seismic Design**
PEER	**Pacific Earthquake Engineering Research**
PFA	**Peak Floor Acceleration**
PGA	**Peak Ground Acceleration**
PPP	**Public Private Partnership**
PR	**Pattern Recognition**

RC Reinforced Concrete

RDF Resource Description Framework

CHAPTER 1 INTRODUCTION

1.1 BACKGROUND

Earthquakes are widely recognised for causing significant economic losses, extensive casualties, and disruptions to buildings and infrastructure systems. These impacts occur annually, affecting both developing and developed regions worldwide (Schmitz et al., 2021). Besides, the global occurrence of earthquakes surpasses one million annually, resulting in an average rate of nearly two earthquakes per minute (Geiß et al., 2015). With the in-depth research and practice of calculation theory and experiment, certain progress has been made in the field of seismic precaution.

The author of this thesis participated in a research project, cooperated by Shanghai Urban Construction Design and Research Institute and Sichuan Provincial Government, which aimed to study how to predict earthquake damage for regional building groups and then to improve seismic performance of the individual building of building groups that have suffered serious damage. This research project originated from a very serious earthquake occurred in 2008 in Wenchuan, a small city of Sichuan province. The earthquake caused a total of 69,227 deaths, 17,923 missing, 374,643 injuries, 19.9303 million people lost their homes, and the total population affected reached 46.256 million (GSMMA, 2008). After learning about the tragic situation of the Wenchuan earthquake, the research region of this thesis is chosen as China to reduce the recurrence of such disasters to some extent. Additionally, the building types provided by Sichuan Provincial Government for study cases are all Reinforced Concrete (RC) frame structures. Hence, the thesis only discusses RC frame structures and does not involve other building types.

At present, the seismic design of buildings meets the basic seismic precaution objectives of “three levels” standard through the “two-stage” method as outlined in the current Chinese code *Code for seismic design of buildings GB50011-2010* (CSI, 2016). The term “three levels” standard refers to a classification system based on the structure’s ability to withstand earthquake. Level 1 indicates that the building will remain in its original condition during a mild earthquake. Level 2 means that the building may be repaired after a moderate earthquake. Level 3 signifies that the building will not collapse during a large earthquake.

The “two-stage” method refers to a design process that involves two distinct stages. In the first design stage, the building must meet the requirements of level 1 and level 2. The seismic action should be designed according to the ground motion parameters of frequent earthquakes for structure analysis and seismic internal force calculation. It is important to take into account the analysis coefficients and load combination coefficients when calculating the cross-section, reinforcement, and controlling the structural elastic displacement. Structural measures are taken to ensure the ductility of the structure to meet the second level of deformation capacity, thus achieving "building will remain in its original condition during a mild earthquake" and "building may be repaired after a moderate earthquake". According to GB50011-2010, the building categories are divided based on the importance of the building's use function and are specifically divided into four earthquake-resistant fortification categories: Class A, Class B, Class C and Class D (CSI,2016). Class A buildings should belong to major construction projects and buildings that may cause serious secondary disasters during earthquakes. Class B buildings should belong to buildings whose use functions cannot be interrupted or need to be restored as soon as possible during earthquake. Class C buildings are general buildings other than Class A, B and D buildings. Class D buildings are those whose earthquake damage will not affect Class A, B, or C buildings, and whose social impact and economic losses are minor. They are generally single-story warehouses with low-value stored goods and few human activities. In the second design stage, the plastic deformation calculation of the weak layer should be carried out for the Class A/particularly irregular building structure with weak earthquake resistance or high seismic requirements during the earthquake. Efforts should also be made to improve the load-bearing capacity and deformation capacity of the weak layer. The utilisation of this conventional seismic design approach may undoubtedly guarantee the structural integrity in the event of a significant earthquake, so effectively preventing any loss of life. Nevertheless, it possesses specific constraints. On the one hand, it fails to acknowledge the significant economic losses resulting from structural damage and functional disruption of building during small and moderate earthquake. On the other hand, the precaution target is limited in its ability to produce a customised seismic design plan to meet unique project requirements (Li et al., 2021). Furthermore, the fundamental issue in earthquake engineering revolves around the twin challenge of assessment and design. Conventional design methods have not yet offered a realistic and viable way of assessing the real seismic performance of buildings, hence hindering the capacity to fulfil the growing demands of stakeholders for customisation (Baris, Atakan, Turgay & Cemil, 2023).

The implementation and utilisation of PBSD theory has effectively addressed the constrains of conventional seismic design (Xiong & Huang, 2019). The theory provides a more detailed explanation and categorisation of seismic and performance levels. It establishes a connection between structure response to earthquake and the desired performance standard, allowing for the achievement of specific goals such as the ability to be repaired after a moderate earthquake. In 2001, the Applied Technology Council (ATC) received the initial contract from the Federal Emergency Management Agency (FEMA) to create advanced PBSD for both newly constructed and pre-existing structures. The main outcome of this project is a collection of volumes, supporting documents, and digital resources known as the FEMA P-58 Seismic Performance Assessment of Building, Methodology and Implementation. FEMA P-58 is an implementation of PBSD theory. On this premise, it suggests using observable data like maintenance cost, maintenance time and other significant data as performance indicators (Xiong & Huang, 2019). Using this collection of performance assessment methodologies, designers can conduct seismic design of buildings for a specific seismic level and ensure they fulfill predetermined performance standards. Additionally, they have the capability to assess the precise seismic performance of a specific structure in the face of a particular earthquake threat. The outcomes are comprehensive and intuitive. Simultaneously, PBSD suggests using the “Return on Investment” (ROI) criterion to achieve a harmonious equilibrium between the initial construction expense and the anticipated losses from earthquakes in the future through its structure design (initial construction expense and the anticipated losses are named as “two major costs”). This approach introduces novel concepts and techniques to optimise the seismic performance of structures (Xiong & Huang, 2019). Besides, because of the unpredictable nature of earthquakes and the extensive computational resources required for building seismic assessment, predicting earthquake damage will serve as a crucial method for rapidly obtaining the distribution of seismic performance among buildings in a given region. It will offer decision support for doing thorough evaluations of each building performance and designing optimisation strategies. Hence, finding effective ways to integrate the aforementioned processes is a crucial notion for the digitalisation, automation, intelligent advancement of the conventional seismic design sector.

1.2 MOTIVATION

With the acceleration of urbanisation, PBSD-based building seismic research and application is seen as a new trend with a lot of potential. On the basis of traditional seismic design model, it fully considers the requirements of economy, safety and other aspects of building performance, which can be fortified before and after earthquakes. Nevertheless, the present study on construction seismic performance assessment, seismic design optimisation based on the ROI criterion, and seismic damage prediction still encounter issues such as inadequate evaluation efficiency, subpar optimisation outcomes, and significant constraints in prediction. Hence, it is vital to continue conducting methodical investigation and making enhancements.

The emergence of numerous digital technologies has facilitated the provision of solutions for the aforementioned concerns. In the civil engineering research field, the extensive use of BIM technology enables the integration and sharing of information throughout the whole project's lifespan by portraying the physical structure through a three-dimensional central mode. In the last century, there has been a significant transformation in building designs and construction methods. Buildings have evolved to become more advanced, incorporating intricate and interconnected system (Kubba, 2017). Industry Foundation Classes (IFC) is proposed to establish a set of data exchange standard of BIM software in the construction engineering industry, which can realise the information transfer of BIM data among various professional software, greatly improving design efficiency and quality (Koo & Shin, 2018). Semantic Web technology facilitates the creation of a cohesive semantic environment where diverse data from several sources may be organised, stored, associated, and interlinked (Zangeneh & McCabe, 2020). Ontology, a crucial technology, is extensively employed to direct the development of the semantic framework system (Stadnicki, Pietron & Burek, 2020). In the context of developing a seismic performance evaluation application, it is crucial to efficiently organize and establish relationships between building information of component level and a substantial volume of seismic damage data. Existing research is deficient in terms of utilising information technology to enhance system efficiency and quality.

The engineering application of Artificial Intelligence (AI) also provides novel strategies for optimising seismic design and predicting earthquake damage. Optimisation design, which is based on the ROI criterion, should primarily consider the competing objectives of minimising the initial building expense and minimising the expected seismic losses in the future. Multi-objective optimisation algorithm is well suited for this feature as it considers both the entire and partial structural features simultaneously. This allows for the establishment of a strict functional relationship between the cost and design plan, leading to an improvement in the level of optimisation. Artificial Neural Network (ANN) makes it possible to predict earthquake damage under any combination of multiple structural parameters and seismic parameters (Huang & Huang, 2020). To some extent, it solves the limitation that traditional predictions cannot consider related factors such as amplitude modulation coefficients, then to expand the application range of predictive models.

Hence, this thesis is grounded in the notion of performance-based design, integrating digital techniques like BIM, semantic web, and AI, to investigate research from two perspective of individual buildings and regional building groups. It possesses specific practical significance and can also enhance the implementation of performance evaluation in engineering. At the same time, it also holds the importance of novelty. Essentially, the thesis is motivated by two main factors:

Firstly, using BIM and semantic web technology, the information needed for evaluating the seismic performance of a building according to FEMA P-58 guidelines is organised and linked together in a unified way. This allows for the creation of an evaluation framework and the automation for the evaluation process. Furthermore, when coupled with the AI optimisation algorithm, a multitude of enhanced design schemes are generated, hence broadening the range of options and improving the efficiency of optimisation.

Secondly, when considering the regional building groups, it is important to take into account the building structural feature, seismic wave characteristics, and amplitude modulation factors in a thorough manner. The earthquake damage prediction model is constructed using ANN to accurately forecast the seismic performance distribution of the group buildings under more comprehensive earthquake scenarios.

1.3 PROBLEM STATEMENT

Despite the significant possibility of building earthquake research based on PBSD, global studies regarding earthquake performance assessment and prediction of earthquake-induced influence in structure predominantly depend on manual processes, leading to low levels of informatisation, diminished assessment efficiency and predictive outcomes (Nia, Moradi & Yang, 2023; Gunes, 2022; Ruiz-García & Olvera, 2021). The primary causes are expressed as follows:

1. Initially, there is an absence of a comprehensive technological framework characterised by a significant level of informatisation integration. The intricacy of the assessment process renders engineers very subjective in the used of PBSD, with the efficacy dependent on the user's comprehension and skill (Fathi-Fazl, Fazileh & Cai, 2022).
2. The implementation of PBSD requires fundamental structure information, seismic response outcomes, earthquake information and so on. Therefore, it is essential to efficiently obtain, analyse and use heterogeneous information from several sources using advanced technology. However, the information and knowledge on the evaluation of structure seismic performance are typically scattered and maintained in disorganized formats, for example, unstructured text (Zhong, 2019).
3. For structural engineers, tools that contributed to inform design-decisions based on component-level damage assessment of design solutions are still lacking.
4. As for buildings' seismic performance assessment, the timely acquisition and transmission of information greatly affects the evaluation efficiency and quality. The classification of performance clusters is still performed manually even with the help of BIM technologies (Liu, 2022). Therefore, it is crucial to establish a standardized semantic structure to achieve compatibility and facilitate querying of diverse knowledge sources, enabling efficient digital management for seismic assessment process.
5. Current research focuses on seismic optimisation design, treating it as a single-objective optimisation issue based on "return on investment". Participants are limited to accepting the results passively and cannot make compromises depending on their preferences. Moreover, the current calculation models are still reliant on the comprehensive evaluation of the structure, therefore failing to account for the impact of adjusting component sizes in the design plan on the optimisation outcome (Dong, Garcia & Pilakoutas, 2024). Additional investigation is still required.
6. The existing studies regrading seismic damage prediction do not consider the issue of seismic wave amplitude modulation, resulting in poor prediction result when subject to a sequence

of seismic waves (Xiong, Li & Lu, 2020). Expanding the application range of the prediction model is challenging and necessitates thorough research.

Based on the identified research gaps, the research problem statement could be summarized as follow: **the efficiency and accuracy of conventional PBSD practice is relatively low because there is a lack of efficient computer aided tools for managing fragmented information and knowledge related to building seismic performance assessment, optimisation design, and earthquake damage prediction. This makes it difficult to assess design schemes using quantitative terms and consider multiple criteria holistically.**

1.4 RESEARCH HYPOTHESIS AND QUESTIONS

Based on the problem statement, a study hypothesis may be formulated as follows:

With leverage of several digital technologies to provide domain information/knowledge, and to qualify the design schemes with quantitative terms, structure seismic performance assessment, optimisation design and seismic damage prediction based on PBSD theory can be optimised for both individual buildings and regional building groups.

Given the stated hypothesis, research question that arises of how to verify it. Below are five research questions, with the answers to each questions presented independently in distinct chapters of the thesis:

Question 1 (RQ1): What are the important areas of knowledge that a structural engineer should take into account for PBSD optimisation? (Chapter 2)

Question 2 (RQ2): How to establish a Building Seismic Performance Evaluation Framework (BSPEF) for the automation evaluation process with the combination of ontology and BIM? (Chapter 4)

Question 3 (RQ3): How to optimise multi-objective seismic design based on ROI criterion with leverage of multi-objective optimisation algorithm? (Chapter 5)

Question 4 (RQ4): How to optimise seismic damage prediction of multi-scale regional building clusters based on ANN? (Chapter 6)

Question 5 (RQ5): How to validate the developed systems? (Chapter 7)

1.5 RESEARCH OBJECTIVES

This thesis intends to use digital technologies, including BIM, semantic web and AI to enhance seismic performance design for both individual buildings and regional building groups, and take typical RC frame structures as case studies. To accomplish the overarching purposes and address the research inquiries outlined in Section 1.4, the research objectives have been established as below:

Objective 1: Identify domain knowledge, methodology and current practice of PBSD.

Objective 2: Explore information technologies to provide the essential groundwork for the integration of information technology and engineering. Then use them to create a knowledge model that enable a standardized semantic format, allowing for interchange and querying of diverse information and knowledge from several sources.

Objective 3: A BSPEF would be proposed to realise the automation of building evaluation process. This framework utilises BIM and ontology to express the evaluation process and logic through organising the key concepts in performance evaluation and the relationship between the concepts.

Objective 4: Establish a Multi-objective Seismic Optimisation Design Method (MSODM) for RC frame according to the BSPEF. The method can enable automatic calculation based on the ROI criterion to work out a sequence of better scheme set that can be selected by designers.

Objective 5: Establish an Earthquake Damage Prediction Method (EDPM) of multi-scale regional RC frame based on ANN and BSPEF. It can realise the prediction of regional building groups seismic performance distribution under more comprehensive seismic conditions.

Objective 6: Validate the application effect of BSPEF, MSODM and EDPM. Therefore, the improved efficiency and quality of seismic design, evaluation and prediction by using these methods can be verified.

1.6 SUMMARY OF RESEARCH METHODOLOGY

As for the underpinning research methodology, the research employed exploratory study. Prototype system development is employed in conjunction with case study evaluation. A simplified framework of the research methodology adopted has been displayed in Figure 1.1.

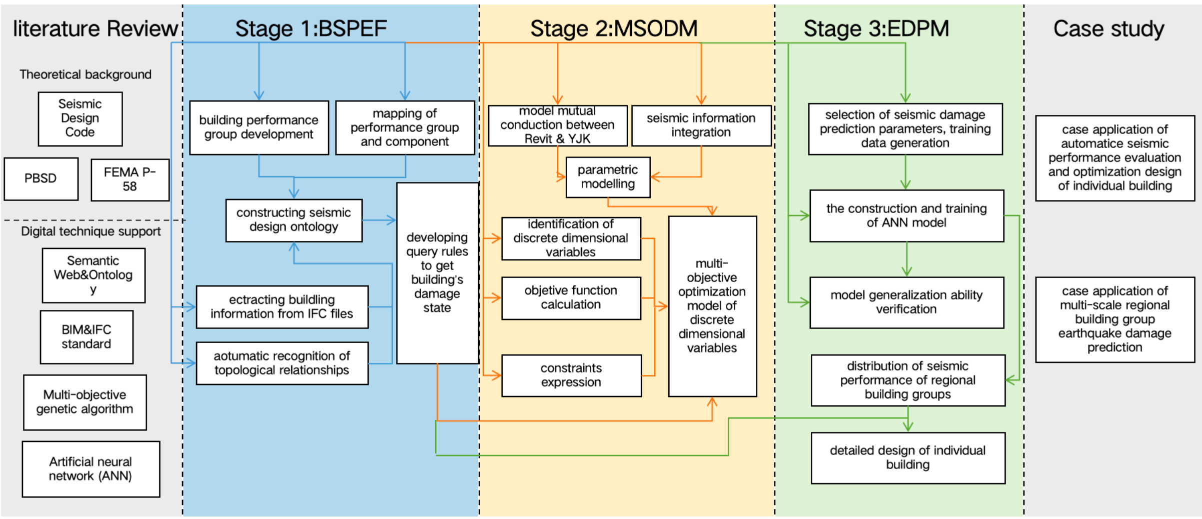


FIGURE 0.1 SIMPLIFIED RESEARCH METHODOLOGY FRAMEWORK

The main research content and methodology are concluded as follows:

Literature Review

Comprehend the theoretical foundation of PBSD and sorting out the latest generation of FEMA P-58 method to clarify the assessment process and data requirements, which has yielded the fundamental insight to the research domain. Then explain the research significance, present status, and challenges of seismic performance assessment, seismic optimisation design grounded in the "return on investment" principle, and earthquake damage prediction, while delineating the research gaps and difficulties this thesis seeks to address.

Subsequently, various advanced technologies are examined to address the identified research gaps. Understand the principles and application techniques of technologies such as BIM, semantic web, ontology, ANN and multi-objective algorithm, and provide essential groundwork for the integration of information technology and engineering.

BSPEF Development

At stage one, this thesis presents a seismic performance assessment framework based on BIM and ontology, in accordance with the FEMA P-58 performance assessment approach. The framework systematically arranges the essential concepts of performance evaluation and their interrelations to articulate the assessment process and logic, while concurrently storing the diverse, multi-source information necessary for the assessment, encompassing fundamental building data at component level that are extracted from IFC files, structural analysis results, cost and maintenance estimates, and additional earthquake damage information. Simultaneously, ontology alignment and rule reasoning functionalities are used to facilitate entity mapping and information interaction across ontologies, hence automating the assessment process. For example, building component are divided into different performance clusters and automatic mapping of performance cluster and component can be realised. Additionally, query rules are also developed for obtaining the structure damage state automatically.

MSODM Development

Based on BSPEF developed at stage one, this thesis focusses on RC frame structures and presents a BIM-based multi-objective seismic optimisation design method. The core of this method lies in a multi-objective optimisation model using discrete size variables. Utilising an appropriate multi-objective optimisation algorithm, the structural seismic optimisation design challenge, grounded in the ROI criterion, is reformulated to achieve equilibrium between the two opposing sub-objectives of initial construction expenditure and anticipated earthquake losses. The optimisation outcomes consist of a collection of superior solution sets from which designers might choose. The two major expenses are directly aligned with the individual design schemes, and automated calculations are facilitated by BIM and ontology technology, significantly enhancing iteration efficiency while achieving seismic optimisation design at component level.

EDPM development

Initially, with the help of ANN technology, a structural response model is developed to enhance the prediction of a certain building type subjected to various seismic waves, incorporating the amplitude modulation factor, thereby achieving a comprehensive prediction of seismic performance distribution for regional building groups under more extensive earthquake scenarios. The selection of model input parameters thoroughly accounts for structural characteristics, seismic wave characteristics, and amplitude modulation factors, with the output parameter defined as the Maximum Inter-story Drift Ratio (MIDR), which signifies the overall damage level of the RC frame structure. The thesis simultaneously examines the influence of various network configurations and training algorithms on the model predictive results, while also validating the model's generalisation capability. The method also reflects the idea of multi-scale regional seismic fortification. The overall distribution of earthquake damage within the building groups can be swiftly determined from the earthquake response prediction results. Conversely, from a local perspective, specific buildings exhibiting significant earthquake damage can be identified, enabling more comprehensive performance assessment and optimisation design based on BSPEF and MSODM outlined at stage one and stage two.

Case Study

First, a real RC frame building is selected as the study case provided by Sichuan Provincial Government, and the automated performance assessment system developed at stage one is used to compute the maintenance loss across different earthquake magnitudes to validate the efficacy of this technique. On this basis, the multi-objective optimisation model with discrete size variables developed at stage two is applied to perform seismic optimisation design for this case. The ultimate optimisation outcome is expressed as a collection of alternative design schemes, illustrating the trade-off between the two major costs.

Then, 30 RC frame building cases that meet the structural parameter requirements at stage three are selected, and 60 seismic wave series that include the seismic parameters as comprehensively as possible are captured to generate a sample data set. Simultaneously, by evaluating various network configurations and training algorithms, a seismic response prediction model is developed to assess earthquake impacts within a specified range, and the model predictive efficacy and generalisation capability are validated. Finally, taking the combination of a certain earthquake level and a certain seismic wave record as an example, the seismic performance distribution of regional building is demonstrated, reflecting the concept of multi-scale regional seismic fortification.

1.7 THESIS OUTLINE

The thesis has eight chapters. Each chapter's substance is introduced in just a few sentences as below.

The introduction of the research background, research motives, research hypothesis, research objectives, and overview in Chapter 1 establishes the fundamentals of the thesis.

Chapter 2 introduces a comprehensive literature review pertinent to the research issue. An exploratory study is carried out to evaluate the present status of research objectives, BIM technology, Semantic Web technology, multi-objective optimisation algorithms, and ANN. These technologies' rapid development facilitates the digitization of the Architecture, Engineering, and Construction (AEC) industry.

Research technique is expressed in Chapter 3 and includes research philosophy, research design, research strategy, methodological options, time horizons, and data collecting.

Chapter 4 studies the approach for assessing the performance of building during seismic events using BIM and ontology. It introduces a new approach called BSPEF. The framework organises and expresses the content and logic of performance evaluation in FEMA P-58 by building ontology knowledge base. The basic building information is acquired through the analysis of the IFC files and the analysis model is promptly generated through the model conversion between the BIM software and the structural software to obtain the seismic response. The aforementioned information, with other necessary information required for evaluation, are stored in the ontology and subsequently utilised to achieve automated seismic performance evaluation through interactive realisation.

Chapter 5 carries out the research on the multi-objective seismic optimisation design based on Chapter 4. It proposed a design method for RC frames that incorporates BIM and considers many seismic optimisation objectives, with a focus on the ROI criterion. The core of this approach is a model for optimising multiple objectives using discrete dimensional variables. The initial building expense and seismic loss expectation are systematically aligned with the design plan. The initial building expense and seismic loss expectation are regarded as two conflicting optimisation objectives. To achieve an optimal trade-off at the component level, genetic algorithm "Elitist Non-dominated Sorting Genetic Algorithm - II" (NSGA-II) is employed.

Chapter 6 studies the earthquake damage prediction method of multi-scale regional RC frame building based on ANN. Selecting structural parameters that can characterise the whole and partial characteristics of building with considering multiple representative seismic parameters and amplitude modulation coefficients to carry out the building seismic response prediction research. The model's predictive performance and generalisation capacity are evaluated and verified. The prediction result will be utilised to obtain the performance level of the individual building. On one hand, it can form the performance distribution result of the regional building groups. Alternatively, it can display the extent of earthquake damage of individual building, enabling a more thorough assessment of performance and optimisation design.

Chapter 7 is case study and validation. Firstly, take a typical RC frame structure as an example to assess the feasibility of the automated performance evaluation method proposed in chapter 4. Applying the multi-objective optimisation model of discrete size variables in Chapter 5 to carry out "return on investment" trade-off, showing the optimisation result and evaluating it. Then, with the help of the ANN model built in Chapter 6 to predict the structural response of a certain type of building structure under the action of seismic waves, the result of the prediction model is evaluated and its generalisation ability is verified. Finally, take the combination of a certain seismic level and a certain seismic wave record as an example to show the seismic performance distribution of the building groups under this situation, reflecting the idea of multi-scale regional seismic application.

Chapter 8 discusses the study's accomplishments and contributions to current understanding. The study's limitations are also explored. Finally, the future research directions are identified.

CHAPTER 2 LITERATURE REVIEW

This chapter provides a segmented literature review consisting of four sections. Section 2.1 starts by introducing the current research status of seismic design for RC frames and then presents the PBSD theory and FEMA P-58. With significant advancements in the fields of computer science and materials science, the disciplines of civil engineering and natural disaster prevention have also achieved considerable strides in the structural design of buildings exposed to external environmental excitation (Xiong & Huang, 2019). Based on the theory of PBSD, the method named FEMA P-58 was put forward for evaluating the seismic performance of buildings. Section 2.2 explored the existing research using various methods and technologies for seismic performance evaluation, seismic optimisation design and seismic damage prediction respectively. According to the critical analysis of the literature review, the findings and research gaps were identified, providing a clearer understanding of the study's motivation. It increases the need to create a knowledge-based integrated system for seismic evaluation and a multi-objective optimisation algorithm for seismic design in order to predict seismic damage. Therefore, Section 2.3 and Section 2.4 go into the examination of the foundational knowledge and practical use of BIM technologies, semantic web technologies, multi-objective optimal algorithm, and ANN.

2.1 CURRENT RESEARCH STATUS

2.1.1 RESEARCH STATUS OF SEISMIC DESIGN

Inferior RC frame structures constructed only for gravity loads, without adherence to seismic design protocols, have sustained catastrophic damage in prior earthquake occurrences, as corroborated by earlier research (Muho, Kalapodis & Beskos, 2024). Recently, seismic design has been reassessed in response to devastating earthquake globally. Although contemporary seismic design regulations have been implemented in several areas, a lot of assessments reveal that earthquake performance of buildings remains vulnerable to collapses (Tang, Cui & Jia, 2024). Due to the fact that traditional seismic design mostly employs "force-based design" (FBD) methods, which means forces and displacement within elastic limits are computed. The principal aim of most earthquake resilience regulations, for example, GB50011-2010, is for fulfilling the "life safety" design criterion within a certain seismic damage threshold (i.e., a 10% likelihood of exceedance within fifty years). Consequently, while overall structural adequacy may be guaranteed for a particular earthquake intensity, building capacity is often used in a limited number of parts and the majority remain underutilised. Moreover, economic losses might be surprisingly substantial, even when the design scheme effectively guarantees life safety.

The rising requirement for secure and economical earthquake assessment and design, has propelled the advancement of structure seismic design. For example, Monte Carlo simulation technique is used in many building performance assessment studies (Das & Singh, 2023). However, this method often needs high-speed computing and extensive computation time, resulting in considerable extra costs even with the more accurate outcomes for fragility analysis. Compared with FBD, Energy-based seismic design (EBSD) is regarded as a more sophisticated approach. It uses structural hysteretic energy as the primary design indicator, serving as a viable alternative to the maximum value based FBD (Shi et al., 2022). However, this approach does not consider economic losses, but only evaluates the seismic performance from the perspective of structural damage. Moayedi (2020) formulated a displacement-based building performance evaluation and design methodology to satisfy designated displacement-based limitations. Lou et al. (2023) advanced this idea into a loss-based building performance evaluation and design methodology, with the objective of attaining a specified degree of damage costs relative to structural damage through very limited iterations. Although these techniques may reduce the computing requirements linked to recurrent design modification procedures, they are more suitable for early design stages. PBSD is an advancement of displacement-based building performance evaluation method (Monjardin-Quevedo et al., 2022). In contrast to the traditional approach, PBSD articulates design schemes which are associated with building component and IDR to attain specified structural performance targets when buildings encounter designated earthquake intensity. This method facilitates the structure design with a pragmatic and dependable comprehension of the potential risks to life and economic losses associated with future earthquake occurrences (Nia, Moradi & Yang, 2023).

2.1.2 RESEARCH STATUS OF PBSD

The PBSD theory was put forward in the early 1990s, and its seismic precaution concept tries to regulate the seismic performance of structures during future earthquakes, in accordance with a predetermined goal (FEMA, 2018). It is a cyclical procedure that commences with the identification of seismic aims, progresses to the formulation of a preliminary design, evaluates scheme's alignment with the seismic aims, and concludes with redesign and re-evaluation if necessary. Figure 2.1 illustrates the essential phases in the PBSD process. The benefits of PBSD can be concluded as follows:

1.Design individual structure with enhanced assurance.

2.Design individual structure that fulfil specifies performance criteria while minimizing construction expenses.

3. Design individual structure to surpass performance expectations established establishes by current seismic code.

4.Design individual structure that exceed the current seismic code limitations for configuration, materials, and systems.

5.Evaluating the possible efficacy of current seismic code requirements for structures that are newly constructed to inform enhancement to code-based seismic design standards.

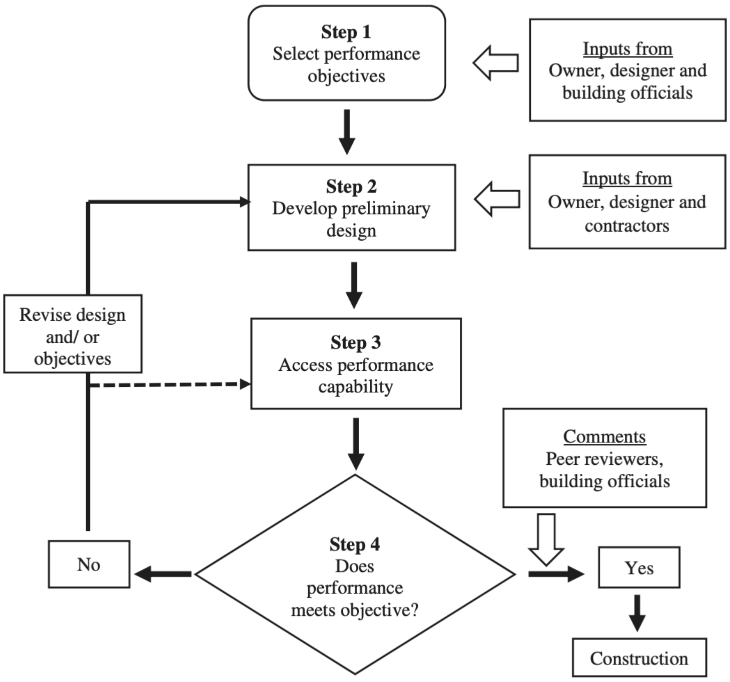


FIGURE 2.1 ESSENTIAL PHASES OF PBSD (NIA, MORADI & YANG, 2023)

Many research institutions have carried out research on the classification and definition of seismic level and performance level. Therefore, the performance target of structures, as a comprehensive reflection of the seismic level and performance level, is no longer limited to the traditional “three level”, but it can achieve higher levels of individualized requirements put forward by the project participants. Based on the conventional norm for seismic precautions, Vision 2000 provides the seismic precaution standard of five magnitude earthquakes for frequent and occasional earthquakes (SEAOC Vision 2000). In traditional design, the intensity values are discrete and a variation of 1 degree in tensity will result in a corresponding change of 1 time in the fundamental seismic acceleration value of the design. To address these issues, certain experts have suggested to directly use ground motion parameters instead of seismic precaution intensity to describe the earthquake impact in the area where buildings are located. “Seismic Code” (CSI, 2016) divided the structural performance level into 5 levels and proposed the IDR as a quantitative index for determination as shown in Table 2.1. However, the indicators regrading performance level developed by these research institutions are not intuitive enough that they are normally not connected with life losses or maintenance costs.

TABLE 0.1 **INTER-STORY DRIFT RATIO USED TO DETERMINE STRUCTURAL PERFORMANCE LEVEL B_I**

Performance level B_i	Inter-story drift ratio
Basically intact B_1	$<x_e$
Minor damage B_2	(1.5-2.0) $<x_e$
Moderate damage B_3	(3-4) $<x_e$
Severe damage B_4	$<0.9x_p$
CollapseB_5	$> x_p$

Note: X_e and X_p represent the structural elasticity and elastoplastic IDR limits respectively, which are taken according to different building types. For reinforced concrete frame structures, the two are 1/550 and 1/50 respectively.

Based on the above theoretical background, FEMA and Applied Technology Council (ATC) cooperated in 2012 to complete the preparation a method for evaluating the building’s seismic performance, named FEMA P-58 (FEMA & ATC,2012). The performance evaluation idea in FEMA P-58 is founded on the full probability model developed by Pacific Earthquake Engineering Research (PEER) Centre. This model considers the various uncertainties in the earthquake action, integrates the influence of structural components, non-structural components and systems on the seismic performance of buildings. It expresses the result as a series of intuitive and easy-to-understand performance indicators such as casualties, maintenance costs and so on. The mathematical expression of the model is shown in Formula 1-1, v(PM) represents the probability of the performance index, the triple integrals from right to left represent: the uncertainty of the ground motion intensity (IM), the uncertainty of structural response using different structure analysis model for a given earthquake intensity (EDP), the uncertainty of structural damage for a given structural response (DS) and the uncertainty of consequences for a given structural damage (PM). In addition, FEMA P-58 has amassed an extensive collection of seismic damage and maintenance data of various building structural systems.

A range of supporting electronic programs and background technical data and other series of products (FEMA, 2018) were also updated, expanded and improved.

$$v(PM) = \iiint G(PM|DS)\,dG(DS|EDP)\,dG(EDP|IM)\,d\lambda(IM) \qquad (2\text{-}1)$$

2.1.2 EVALUATION OF BUILDING SEISMIC PERFORMANCE BASED ON FEMA P-58

FEMA P-58 categorises components with comparable sensitivity and similar outcomes for seismic damage into distinct fragile cluster. It offers over 800 fragility cluster regulations (FEMA, 2018) including fragility cluster serial numbers, identification rules that describe component categorization, fragility curves that describe damage conditions, and consequence functions that describe damage distribution, etc. The fragility cluster is subdivided into a performance cluster by floor, based on structural response criteria such as IDR and Peak Floor Acceleration (PFA). FEMA P-58 proposes three assessment methods that rely on the factors of structural strength, hypothetical situations, and duration, applicable to both newly constructed and pre-existing structures. The evaluation begins by establishing a structure performance model using the performance cluster as the fundamental unit. Subsequently, the seismic analysis result is obtained using seismic hazard analysis. Next, the fragility function is employed to determine the corresponding failed state within the performance cluster, then the consequence function is utilized to calculate the loss associated with each failed state. Finally, the building's aggregate performance index is condensed and presented. Figure 2.2 depicts the performance evaluation process in accordance with FEMA P-58, and the specific procedures are outlined in the following context.

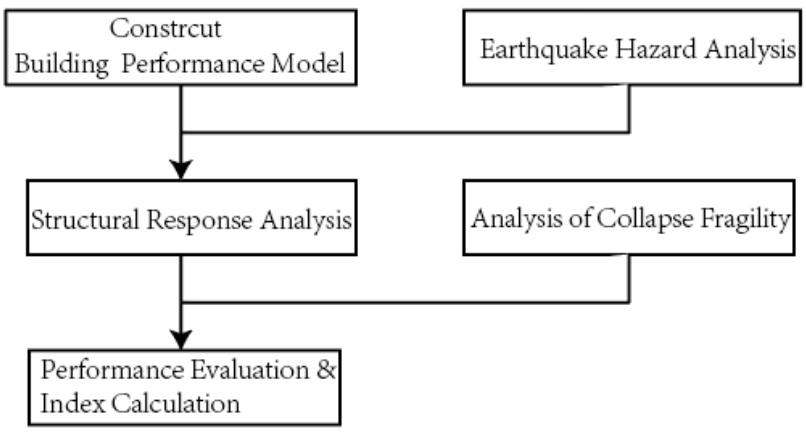


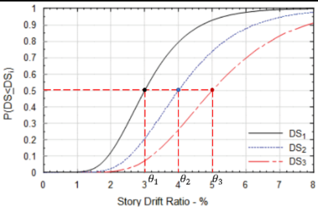
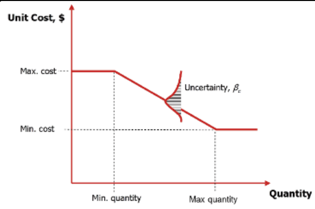
FIGURE 0.2 **BASIC PROCESS OF FEMA P-58'S PERFORMANCE ASSESSMENT**

Step 1: construct building performance model

The model of building performance encompasses various aspects of a building's characteristics that are relevant to earthquake effects. This includes fundamental building information like building's function type, dimensions, and cost of reconstruction. It also incorporates details about the structural and non-structural components, such as the building's location, its susceptibility to seismic responses, potential injuries' type, consequences, and additional information about equipment distribution.

As stated above, FEMA P-58 classifies components into designated fragility clusters, for example, Table 2.2 shows the fragility group rules numbered B1041.032a, in which the standardized classification number is given by NISTIR 6389 (Concepts, Charette & Marshall, 2010). The identification rules describe the characteristics of the components that are categorized under a particular fragility cluster. The fragility curve is a function which explains likelihood of the damage type that may occur when the component undergoes an earthquake response and the seismic demand parameters. It obtained from a large number of experiments, historical earthquake damage experience, expert analysis and statistics, which is a lognormal distribution. This particular category of fragility cluster exhibits three possible damage states, namely DS_1, DS_2 and DS_3 under the action of an earthquake. Each of these states is associated with a distinct distribution function. The median value $_i$ and the discrete value $_i$ of the certain distribution function is stipulated to be given. While each damage state will correspond to a multi-stage consequence function, which represents the possible damage consequences (such as maintenance cost, repair time.) of the component under the damaged state. The number of fragility clusters are estimated and developed by FEMA based on the construction cost of Northern California in the United States in 2011. The regulation will also give the calibration values which is needed to determine each consequence function.

TABLE 0.2 **B1041.032A FRAGILITY CLUSTER REGULATION**

No. of Fragility Cluster	Identification Rules	Fragility Curve	Consequence Function
B1041.032a	The beam-column joint of the ordinary frame, the beam is on one side	 $F_i(D) = \phi(\frac{\ln(D/\theta_i)}{\beta_i})$ $\theta_1 = 0.0175$、$\beta_1 = 0.4$; $\theta_2 = 0.0225$、$\beta_2 = 0.4$; $\theta_2 = 0.0322$、$\beta_3 = 0.4$	 Multi-Stage Consequence Function $Max.cost$、$Max.quantity$ $Min.cost$、$Min.quantity$ β_c

A performance cluster is a collection of components such as beams and columns that conform to a particular fragile cluster and experience identical seismic response, which can be regarded as a subset of the fragility cluster. From the definition, it is evident that performance clusters are typically categorized based on the floors and orientations within the building. As an illustration, the beam-column node that satisfies the identification rules in Table 2.2 is a fragility cluster. Its demand parameter is the IDR, which is a direction-sensitive response index. Different floors and different building directions will have different values. Therefore, the fragile clusters in the north-south and east-west orientations of a certain layer need to be further divided into two types of performance clusters. Acceleration-sensitive components, such as pipeline system whose demand parameter is PFA, are divided into performance clusters based solely on floors, regardless of orientations.

When the building performance model is divided into performance clusters, each type of performance cluster can calculate its damaged condition and consequences under the action of an earthquake according to the fragility curve and consequence function given by the fragility cluster rules.

Step 2: Analysis of earthquake hazard and building's structural response

Earthquake hazard analysis is utilised to give the potential degree of seismic activity that the building may suffer, and it is generally characterised by the intensity of ground motion (Xu, Wu, Feng & Fan, 2021). For the three basic evaluation methods, there are various approaches to ascertain the intensity of ground motion according to different acceleration response spectrum definitions. As the index of ground motion intensity, the response spectrum S_a (T_1) corresponding to Peak Ground Acceleration (PGA) or the spectral acceleration of the structure basic period is utilised. For a certain building structure, the S_a (T_1) index related to the structure period can be adopted as the index, whilst the PGA index can be adopted as the index for the same type of building structure.

After selecting the intensity of the ground motion, FEMAP-58 recommends using nonlinear dynamic time history analysis to analyze structural response as exposed to action of ground motion. If the selected ground motion record closely aligns with the geometry of the target spectrum, only seven ground motion record are required for structural response analysis; otherwise, 11 ground motion records are required (Yan, 2019).

Structural response analysis results are characterised by structural performance parameters, for example, PGA and IDR. The Maximum Inter-story Drift Ratio (MIDR) can provide an overall assessment of the structure's performance. PGA and IDR values for each floor can be used to reflect the damage status of different components, especially for frame structures. The MIDR is the best structural damage discrimination index (Concepts et al., 2010). As a result, the MIDR is used as a benchmark to judge the building's overall performance, while the IDR and PGA serve as the basis for performance evaluation and the response analysis results from a series of ground motion records are formed to generate the response demand state set R, which corresponds to the intensity of each ground motion.

Step 3: Analysis of collapse fragility

The results of the collapse fragility analysis are represented by a collapse fragility function, reflecting the correlation between the likelihood of a building structure collapsing under a particular earthquake intensity and the intensity of ground motions (Abyani, Bahaari & Zarrin, 2019). Incremental Dynamic Analysis (IDA) is a commonly employed method to build the collapse fragility function. The method relies on nonlinear dynamic time history analysis, by continuously increasing the intensity of ground motion and performing the corresponding amplitude modulation on one or more seismic waves applied to the structure. Ultimately, one or more relationship curve of ground motion intensity and structural performance parameter can be obtained (Kita, Cavalagli, Masciotta, Lourenco & Ubertini, 2020).

Step 4: Performance evaluation and index calculation

In order to take into account various uncertain factors, FEMA P-58 utilises the Monte Carlo method to assess performance indicator. Each simulation process is shown in Figure 2.3.

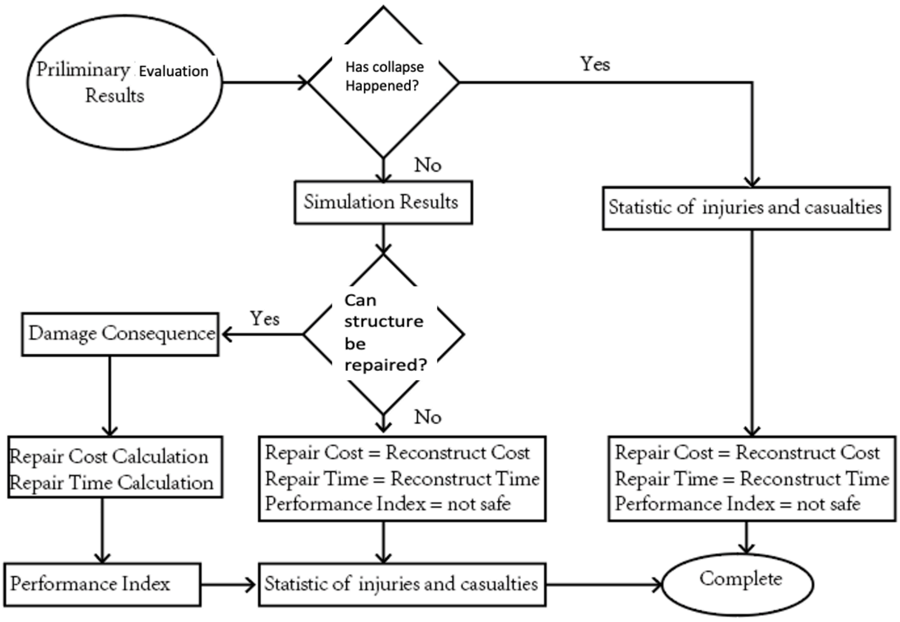


FIGURE 0.3 **A SIMULATION PROCESS OF PERFORMANCE ASSESSMENT**

In the beginning of each simulation, the ground motion is determined, and the probability of structural collapse P_c is calculated according to the collapsed fragility function. Then the process is simulated to produce a random number to determine whether the structure is repairable. When the value of the random number is lower than P_c, the structure is deemed collapsed and the cost to restore it at this time is equivalent to the reconstruct cost. On the contrary, the structure does not collapse. When the residual deformation of the structure exceeds a threshold of 0.01, the probability of being able to repair the structure is normally 50% (Zeng, Deng, Kurata, Duan & Zhao, 2020). Similarly, the simulation generates a random number to judge whether the structure is repairable. When it is judged that the structure is not collapsed and repairable, a possible structural response is randomly generated from the response demand state set R. The damage consequences of the performance cluster are calculated based on this and then consolidated to derive the overall performance index of the building. After one thousand times simulations, one thousand index values are obtained. Then the probability distribution is drawn with a tool to indicate the correlation between the likelihood of achieving a particular performance index and the performance index, which is the performance function.

Step 5: Component-level damage prediction and correction

In actual projects, as different components belonging to the same performance cluster have different repair measures and the data regarding with damage consequence in FEMA P-58 comes from database of Northern California in 2011, it is often necessary to make corresponding adjustments according to the local construction cost pricing specifications when applied to different areas (Xu et al., 2019). These cost data use building components as the basic unit, so it is imperative to implement the damage prediction to the component itself and combine the cost ratio between various damage status in FEMA P-58. This will allow for calculation of the unit maintenance cost associated with a specific damage state of a component. Ultimately, this information will be utilised to obtain the building's overall performance index.

Damage Status DS_n

Because the seismic performance evaluation of structures under the guidance of FEMA P-58 is carried out on a floor-by-floor basis, a certain floor s (s=1, 2, ..., s, ..., N_{story}, where N_{story} represents the total number of building floors) is the research object. All damage states and damage probabilities of various performance groups on this floor are written into the matrix *s_SD[PG]*, as shown in Formula 2-2:

$$s\_SD[PG] = \begin{matrix} \\ DS_1 \\ DS_k \\ DS_{N_{DS}} \end{matrix} \begin{matrix} 1 & p & N_{PG} \\ \begin{bmatrix} 20\% & \cdots & 10\% \\ \vdots & \ddots & \vdots \\ 30\% & \cdots & 40\% \end{bmatrix} \end{matrix} \quad (2\text{-}2)$$

Where

p = 1, 2, ..., p, ..., N_{PG};

k = 1, 2, ..., k, ..., N_{DS};

N_{PG} represents the total number of performance groups of a certain floor;

N_{DS} represents the different damage states of a certain type of performance group number; The k_{th} row and p_{th} column of the matrix represent the probability that the p_{th} performance group of the floor is in the damaged state DS_k and the sum of the value in each column is equal to 100%.

For structural components, they generally belong to multiple performance group categories, the damage state distribution should follow the result of nonlinear time history analysis and the damage state DS_n should be calculated according to Formula 2-3, which is the maximum damage state of the performance group:

$$DS_n = Max(s\_SD[PG_i], s\_SD[PG_j], \ldots) \quad (2\text{-}3)$$

Where

PG_i and PG_j represent the performance group to which the component belongs.

For non-structural components, the damage state distribution generally has a large uncertainty (Dhakal, 2010; Kawaguchi, 2012), so it can be based on the damage status of the performance group and its probability distribution are randomly assigned.

Damage degree *DP*

For non-structural components, assuming that the performance group it belongs to is PG_p, the damage degree *DP* is calculated according to Formula 2-4:

$$DP = \frac{F(PG_p, DS_n)}{F(PG_p, DS_{max})} \tag{2-4}$$

Where

$F(PG_p, DS_n)$ is a function representing the unit maintenance cost of the performance group in the damage state DS_n that calculated according to the consequence function.

$F(PG_p, DS_{max})$ represents the unit maintenance cost of the performance group PG_p in its maximum damage state DS_{max}.

For structural components, there are generally multiple performance group categories, so the damage degree *DP* should be based on Formula 2-4 to sum the unit maintenance costs of each performance group, as shown in Formula 2-5:

$$DP = \frac{\sum_i^N F(PG_i, DS_n)}{\sum_i^N F(PG_i, DS_{max})} \tag{2-5}$$

Consequences of injury *C*

The damage degree *DP* is the damage ratio, which represents the ratio of reconstruction resources to resources required to restore the component from the damage state DS_n to the initial condition. Therefore, the unit damage consequence of the component (Formula 2-6) is expressed as the product of the unit loss data *Unit_Cost* and the damage degree DP, then multiplied by the measurement data *Measure_Data* such as volume, area, length, etc. to get the component damage consequences. The damage consequences of all components are added together to get the performance index of the building.

$$C = DP * Unit\_Cost * Measure\_Data \tag{2-6}$$

2.2 RESEARCH STATUS

2.2.1 BUILDING SEISMIC PERFORMANCE EVALUATION BASED ON FEMA P-58

Yang and Blakeborough et al. (2018) combined the FEMA P-58 approach to predict the seismic damage of a single building. Zeng Xiang et al. (2016) conduct loss prediction for multiple structures in metropolitan settings using the FEMA P-58 guidelines. Cook (2021) evaluate the end-to-end results of the FEMA P5-8 PBSD by using actual data from the 2018 Earthquake in California. To be more specific, they conduct a regional building groups evaluation for about 2.7 million structures during the 2018 earthquake to retroactively estimate economic losses and the quantity of dangerous placards resulting from the disaster. Koohfallah (2024) examines the optimal retrofitting strategy for RC frame school structure using the performance indicator in accordance with the FEMA P-58. A predictive technique for assessing earthquake damage using BIM and FEMA P-58 is proposed by Xu et al. (2019). The aforementioned study utilised strength-based evaluation to evaluate the repair costs of existing buildings in relation to certain earthquake intensity. However, there are three notable shortcomings that should be acknowledged:

1.The segmentation of performance clusters is crucial for establishing building performance models. This process requires the integration of component data, mapping component relationships to performance cluster, and correlating performance clusters with floors. The timely acquisition and transmission of information significantly impact the evaluation optimisation and accuracy. Nevertheless, current research primarily collects and organized component data in tabular format. Designers are then required to manually construct performance models, a process that will consume a lot of time and is susceptible to mistakes. Wemyss et al. (2020) recommended the combination of BIM technology to predict the seismic loss of building under the guidance of FEMA P-58. Xu Zhen et al. (2019) attempted to accomplish the integration of component data by placing the entire evaluation in a particular BIM software environment. Nevertheless, this research still relies on manual identification of performance clusters, and the recognition outcomes are recorded as component attribute factors using BIM interfaces, which leads to a lack of an intuitive relationship between components, performance clusters, and floors. When design schemes or classification regulations are changed, it will result in numerous repetitive adjustments and a failure to accurately represent the logical mapping.

2.The FEMA P-58 provide an approach to assess damage at the component level that is backed by a substantial volume of related damage information. These data are frequently saved in unorganized formats, for example, plain text. The process of linking it to components, which involves establishing a consistent semantic structure to provide compatibility and enable querying of diverse knowledge from various sources, still need further investigation. It is essential to understand the digital administration of the evaluation procedures, as it enables the rapid assessment of construction clusters' performance at the city level.

3.The labour-intensive process of constructing a structure performance evaluation model and the substantial calculations required by FEMA P-58 make it challenging to assess various design schemes. Therefore, a universal framework in place to anticipate the buildings' performance indicators using diverse assessment methods automatically is required. Considering the update in the version of FEMA P-58, it is important that the framework is designed to be easily manageable and maintainable. However, further extensive research is still needed.

2.2.2 SEISMIC OPTIMISATION DESIGN BASED ON "RETURN ON INVESTMENT" CRITERION

The process of designing seismic structures is integrated with the consideration of life cycle costs. The majority of potential costs stem from monetary equivalent costs caused by earthquake occurrences which might happen over the building's lifetime. Typically, prospective expenses are disregarded as they have nothing to do with the seismic structure's resilience. Based on "return on investment" rules, existing research generally regards the combined value of initial building expenses and potential earthquake losses as the overall cost of the building's lifespan. The objective is to seek the most favorable scenario where the overall costs is smallest. The formulation of the overall costs, also known as the objective function, sparks extensive investigation by researchers.

Kanda and Shah suggested a technique to approximate the initial building expense based on the design strength (1997). Some researchers utilized the design intensity factor to quantify the original building expense and put forth a strategy for estimating earthquake losses (Hong & Xie, 1999). Okasha and Frangopol examined the estimated procedure for determining the reliability of structures and conducted different cost calculations using the optimal precautionary intensity (2009). Sarcheshmehpour and Estekanchi (2021) examined the correlation between seismic costs projections and the initial building expenses, taking into account discount factors.

From structural standpoint, the aforementioned research obtains a supposed correlation between the initial building expenses and key features of the structure such as structure reliability, design strength, precautionary intensity. The expected potential seismic loss is often quantified as the multiplication of the initial building expenses and an empirical factor dictated by the building’s structural type. Hence, the initial expenses of constructing the building and the anticipated seismic losses is represented as a function of the particular performance objective. The aim is to achieve performance objective that minimizes the overall costs, for creating the most optimal plan for the buildings. The method prioritizes economic considerations while ensuring safety criteria are met, but it still faces three specific challenges:

- The aforementioned research regards the building seismic optimisation design as a problem with single objective optimisation. Participants, including creators and clients, may passively acquiesce to it. They are unable to make choices that accommodate their preferences, which hinders practical implementation. For example, as the expected potential seismic loss is often quantified as the multiplication of the initial building expense and an empirical factor, the expected potential seismic loss and the initial building expense cannot be treated as the two opposing sub-goals.
- The precision of calculating that initial building expenses is readily influenced by the intricacy of various design schemes and the unpredictable variables in the construction process. The earthquake loss is also influenced by the underlying assumptions of earthquake damage analysis model, subjectivity of damage condition and effects, and the variability of many stochastic indexes. Therefore, due to significant uncertainty and variations, the outcome calculated by adding the two or employing optimisation based on weighting may exhibit a noticeable discrepancy from the anticipated result (Liu, Wen & Burn, 2004), particularly for newly constructed structures during the early design phase.
- There is still room for improvement in optimising specific design schemes. Actually, both the initial building expenses and the seismic loss expectations are intimately correlated with the design plan. As these two expenses are represented as empirical formulas for the performance goals, firstly, the precision of calculation is significantly inadequate and distinction between various design plans is not readily apparent; secondly, there are often multiple design plans that achieve a particular performance goal, resulting in there is no function relationship in the expression to some extent. Therefore, the aforementioned studies regard the structural plan that has the smallest weight as the representative of the “most cost-effective” design for this performance objective. However, it fails to acknowledge - the minimal structural weight usually does not correspond to the smallest original building expenses.

In Min’s research, the initial building expenses is expressed as a function of the structural components’ size and the variable of the reinforcement ratio. The fuzzy decision theory is then used to approximately linear process and estimate the damage caused by earthquakes at various levels of occurrence probability based on the optimal criterion (Min, Burn & Wen, 2010). As for steel frame earthquake design, Ming et al. utilized genetic algorithm to optimise the overall costs during building lifespan, considering the initial building expenses and the expected earthquake expenses as two opposing objectives in the multi-objective optimisation procedures (Liu et al., 2014). A set of alternative designs are produced during the optimisation process. The two research propose the concept of multiple objective optimisation and tries to formulate the initial building expenses as a mathematical function that depends on the variables representing the size of it component. Nevertheless, the current earthquake loss expectation calculation approach is primarily focused on the assessment of the whole structure and does not account for the impact of adjusting the size of individual components on the optimisation outcome. Furthermore, for the convenience of calculation, the existing seismic optimisation design studies treat design variables as continuous variables, whether they are representative structural parameters or component sizes which are not continuous variables. In this case, it is unable to accurately represent the construction and manufacturing needs, therefore rendering it incapable of offering helpful advice for designers. Consequently, further research is still necessary.

2.2.3 EARTHQUAKE DAMAGE PREDICTION

Earthquake damage prediction is generally categorized into two distinct categories according to the sort of response to be predicted. The first category is structural damage state, which can be evaluated either for the entire building or for partial components, such as basically intact or severely damaged. The second category is the response of the structure, namely the Earthquake Demand Parameter (EDP) including IDR and PFA and so on. The factors that affect the prediction include the structure itself and the ground motion records, so structural parameters and seismic parameters should be involved. In general, a combination of multiple factors should be considered to approximate the actual situation as well. Research has confirmed the intricate correlation between seismic motion parameters and structural damage, particularly for seismic parameters. Relying on a single seismic parameter is far from enough (Elenas & Meskouris, 2001). ANN can explain the nonlinear relationship between multiple input parameters and output, which is widely used in the study of earthquake damage response prediction (Paruthi et al., 2022). Using ANN, multiple parameters can be considered to quantify the seismic effect. What is more, ANN make it possible to analyze any quantity and combination of seismic and structural parameters in order to determine the most accurate relationship between these parameters and the building's seismic response (Portillo & Negro, 2022).

Morfidis applied a Multi-layer Feedforward Perceptron (MFP) network, which is a type of feedforward neural consisting of fully connected neurons with a nonlinear kind of activation function, to investigate the behavior of an RC frame under a certain seismic intensity and explored the relationship between the MIDR and 18 parameters including 4 structural parameters and 14 ground motion parameters (2018). It also explored the influence of training algorithms and network structure parameters on the prediction result (Morfidis & Kostinakis, 2018). The parameter selected in this study are adequately reasonable in terms of quantity and qualities and sufficient consideration has been made at the network configuration level, resulting in a relatively optimal prediction outcome. However, since only a certain seismic level is considered, there exist significant constraint in practical applications. Lautour used the MFP network to forecast the comprehensive damage state of the structure of a two-dimensional RC framework with different topologies, stiffness, strength and damping ratio across diverse seismic levels (2009) . As the consideration of different seismic levels is satisfied by the coverage of the selected seismic wave ground motion parameter range as much as possible, the scaling of each seismic wave is not involved. When predicting the damage state of an existing frame exposed to the action of a certain seismic wave after amplitude modulation, the prediction result is not good.

Mackay used a three-layer feedforward neural network and error propagation algorithm to select 19 architectural features and learn the relationship between them and the overall performance level of the building (1992). Jia and Wu pointed out in the study that structural seismic performance itself is directly influenced by the value of the microscopic characteristics of the structure (2021). Therefore, for the RC frame structure, the concrete grade, column size, column reinforcement ratio, story height, lateral span, and height-span ratio beam that have great impact on its seismic performance are selected to establish the mapping relationship with seismic performance. However, this study has not considered issue of the seismic wave amplitude modulation.

In summary, in the existing research, whether it is whole or partial structural parameters, the selection is practical and convincing. Arslan also systematically explores the prediction of the seismic response of various factors (2010). The prediction result, also known as the output parameter, is generally taken as a single index such as IDR, the overall damage state and the component damage state etc. However, none of the existing studies has explicitly addressed the issue of seismic wave amplitude modulation, which results in poor prediction result when subjected to a series of earthquake levels, and it is challenging to extend the model’s application range. Therefore, further in-depth research is required.

2.3 APPLICATION OF BIM AND SEMANTIC WEB TECHNOLOGIES IN SEISMIC PERFORMANCE EVALUATION OF BUILDING

2.3.1 INFORMATION AND KNOWLEDGE MANAGEMENT IN CONSTRUCTION INDUSTRY

The fragmented character of the construction sector is a result of the intricate and close interdependence of the building projects. It always involves multidisciplinary teams such as clients, designers, suppliers and so on. Additionally, it necessitates the utilization of diverse software and hardware instruments at various stages of the entire building's lifespan. Therefore, the construction industry is an information-intensive industry in which information plays a crucial role in disciplinary interactions (Admassie, Ferede, Lema & Ayen, 2022). With the evolution of ICT, the complexity of information management is growing. As the information and knowledge is complex and diverse during the building's lifespan, the efficiency of information management has been developed and implemented.

Interoperability, explained as "*the ability of two or more systems or components to exchange information and to use the information that has been exchange*" (Geraci, Katki, Mcmonegal, Meyer & Porteous, 2002), is a crucial concern for facilitating the sharing of information and knowledge in building projects. During the construction process workflow, drawings and documentations are main elements shared, which require manipulation and revision continuously. Prior to the development of CAD in the early 1980s, documents and drawings were created manually for a long time (Medjdoub, Richens & Barnard, 2001). CAD technology, as being the technique that allows for design data sharing, has a profound influence in the field of construction industry. It established proprietary drawing formats and later developed standards, for example, Data Exchange Format (DXF) which includes measurement, dimensions and layers (Yu et al. 2022).

Data presentation is promoted with the purpose to enhance information transparency sharing among stakeholders in accordance with governments responsibilities and legal legislations (Lisboa & Soares, 2014). Therefore, a wide range of various research topics for different information sharing methods has been developed ever since. To this end, an increasing number of scholars and commercial application developers have begun to establish techniques to shift from document centered methods to computer and modeling integrated methods. In addition, the evolution of knowledge management applications has successfully addressed considerable obstacles in handing and manipulating complex data models. Three stages of knowledge sharing have been summarized in Figure 2.4 (Boddy, Rezgui, Cooper & Wetherill, 2007). In the first stage, the knowledge sharing system is characterised as document-centric management systems. The second stage was developed in response to the recognition by academics and commercial application developers of the need to provide tools for manipulating intricate architectural models. The knowledge managed in this stage is referred to as "knowledge conceptualization and nurturing". BIM and IFCs are products during this time and there has been a noticeable rise in the emergence of ontology. The third stage of knowledge management is defined as knowledge value creation, which is express as "*any process of creating knowledge value, as subjectively perceived by users, out of existing knowledge practices across an organization*" (Rezgui & Miles, 2009).

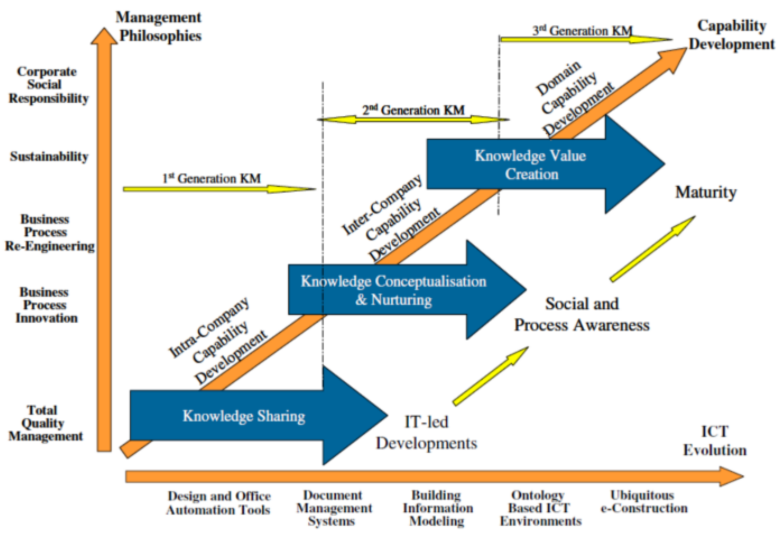


FIGURE 0.4 **GENERATIONS OF KM IN AEC (REZGUI 2007)**

2.3.2 BIM AND IFC STANDARD

The term "BIM" was initially coined in the late 1970s by Professor Charles M. Eastman of the Georgia Institute of Technology, described as follows: "*BIM is a digital visible model of the building process to facilitate exchange and inter-operability of information in digital format*" (Eastman et al., 2008). Facilities Information Council (FIC) defines BIM as follows: BIM is a digital depiction of the physical and functional attributes of facilities and their associated project life cycle data, thereby supporting project decision-making and enhancing project value realization (Ando, Sarlens & Klein, 2019).

As a comprehensive information model, BIM can integrate data, procedures, and resources at various phases of a building life cycle. Moreover, it can provide a complete description of the construction project, applicable to all stakeholders involved (Cursi et al. 2022). In the field of construction engineering, the use of 2D CAD technology is considered as first revolution in design, whereas the advent of BIM technology initiates the second revolution throughout the AEC sector. In comparison to conventional 2D, BIM offers the following benefits:

1. Achieve object-oriented design and visualization of design schemes: BIM make the design process transition from using lines, surfaces, and blocks for abstract representation of design objects to employing components, referred to as “families”, for direct expression of design objects. For example, 2D design pipelines are shown as single lines of various colours, but 3D design use information-based renderable cylindrical pipelines. The rendering effect is more lifelike, resulting in an intuitive and visible physical model that effectively communicates the design idea.

2.Enable collaborative design: the liberty of designers is strictly limited as 3D design is a design process oriented to physical objects. With standardized design specifications, designers from different disciplines can collaboratively operate within a shared 3D environment, enhancing communication and coordination. This approach enables the prompt identification of issues, such as conflicts between pipelines and structural elements, thereby mitigating the challenges associated with later detection.

3.Ablity to carry out complex engineering calculations: by properly leveraging the data contained in the BIM model and integrating it with computational technology, engineering analyses such as collision detection, cost estimation, seismic evaluation can be performed.

4.It can also include the relationship among different entities with the object-level. With the advent of this technology, the information configuration is sufficiently enhanced to perform specific information-processing operations, thereby enabling automated information processing (Jiang, Feng, Zhang & Shi, 2023). This achieves the integration of solutions in one model by exchanging BIM data between various analysis tools.

However, the application of BIM in the design filed still has some disadvantages. The creation of information models requires substantial time and effort. Therefore, this thesis uses Dynamo to transfer diver design schemes into BIM models automatically, which is described in Section 5.2. Moreover, the data exchanging approaches for construction information is crucial, as well as the BIM data management scheme. The most important issue focus on how to filter redundant information from the wide variety of multi-disciplinary life-cycle data to complete specific tasks and to draw only the required information for specific applications. Various formats exist for representation of data in different software leveraged in the construction domain, in this context, exchanging data of different formats directly among different software is very difficult.

To overcome this problem, the application of IFC is proposed by buildingSMART as a set of data exchanging standards in the AEC domain, which have been officially recognised by the International Organisation for Standardisation (ISO) as ISO16739. It can realize the information sharing and exchanging of BIM model data among various professional software for better interoperability (He, 2005). Its core technical content includes the description of engineering information and the acquisition of engineering information (Pinheiro et al., 2018), which can provide an exhaustive depiction of building architectural framework, physical and spatial elements, analytical elements, procedures, resources, controls, participants, and contextual description (International Standards Organization, 2013). The initial version, IFC 1.0, focused on delineating four aspects: architecture, Heating, Ventilating, and Air Conditioning (HVAC), engineering management, and equipment management (Almeida, Chaves, Silva, Carvalho & Caldas, 2023). Subsequently, IFC has evolved and refined data representation within the building domain. Table 2.3 encapsulated the number of the Entities, Types, and property sets present in various IFC versions. Version IFC 4.3 comprises 859 entities, 524 types, and 601 property sets, including eight professional domains including architecture, structural engineering, HVAC, electrical systems, construction management, and operations and maintenance management, which illustrates the comprehensive nature of the IFC standard in conveying building information. Existing research indicates that the IFC standard comprehensively delineates the design schemes, such as aesthetic dimensions, structural analysis, seismic analysis, and cost estimation (Lozano et al.2023).

Table 2.3 The number of entities, types and property sets in different version of IFC

	IFC 2.0	IFC 2x	IFC2x2 ADD1	IFC2x3	IFC2x3 TC1	IFC 4	IFC4 ADD1	IFC4 ADD2	IFC 4.1	IFC 4.2	IFC4.3
Entities	290	379	329	653	653	766	768	776	777	797	859
Types	157	229	313	327	327	391	396	397	432	466	524
Property sets	-	83	312	312	317	408	410	413	487	521	601

Currently, different types of software applications in the construction sector facilitate the input and output of IFC model data. Moreover, some applications, such as Revit (Autodesk), ArchiCAD (Graphisoft), and Tekla Structure (Tekla), have received formal certification from buildingSMART. Therefore, the IFC standard has emerged as a practical data standard for BIM applications in the construction field (Haridas et al, 2017). Nonetheless, due to the richness of the IFC standard semantic expression in the architectural field, the IFC models output by BIM applications often contain extraneous information for particular projects (Zhao, 2017), thereby necessitating the processing of excessive model data and diminishing processing efficiency. Therefore, when this thesis develops the seismic evaluation framework, a series of built-in functions of BIM application are used for preprocessing IFC files to extract only the required data for seismic evaluation in Section 4.2.1.

EXPRESS data specification language is used for composing the conceptual data schema, which focuses on the definition of entities. Entities include data and constraints, where data represents the properties of the entity when it is instantiated and constraints are expressed through rules. Take the IfcObject entity definition in Figure 2.5 as an example to explain the format and content of EXPRESS language.

```
ENTITY IfcObject
    ABSTRACT SUPERTYPE OF(IfcProduct)
    SUBTYPE OF (IfcObjectDefinition);
    ObjectType : OPTIONAL IfcLabel;
    INVERSE
    IsTypedBy : SET [0:1] OF IfcRelDefinesByType FOR RelatedObjects;
    IsDefinedBy : SET OF IfcRelDefinesByProperties FOR RelatedObjects;
    WHERE
    UniquePropertySetNames : ((SIZEOF(IsDefinedBy) = 0) OR IfcUniqueDefinitionNames(IsDefinedBy));
END_ENTITY;
```

FIGURE 0.1 **EXPRESSION OF IFCOBJECT IN EXPRESS**

According to the EXPRESS language format, when the construction project data is instantiated and stored under the corresponding entity category, the information can generally be accessed with the help of standard format files or program interfaces. At present, using IFC physical files in STEP format with the suffix.ifc for information exchanging and sharing is the mainstream approach. The IFC physical file is composed of a header section and a data section, as shown in Figure 2.6.

```
ISO-10303-21;
HEADER;

/*************************************************************************
* STEP Physical File produced by: The EXPRESS Data Manager Version 5.02.0
* Module:                         EDMstepFileFactory/EDMstandAlone
* Creation date:                  Sun Apr 28 10:08:43 2019
* Host:                           DESKTOP-HVRROD7
* Database:                       C:\Users\eCherry\AppData\Local\Temp\{FA
* Database version:               5507
* Database creation date:         Sun Apr 28 10:08:42 2019
* Schema:                         IFC2X3
* Model:                          DataRepository.ifc
* Model creation date:            Sun Apr 28 10:08:42 2019
* Header model:                   DataRepository.ifc_HeaderModel
* Header model creation date:     Sun Apr 28 10:08:42 2019
* EDMuser:                        sdai-user
* EDMgroup:                       sdai-group
* License ID and type:            5605 : Permanent license. Expiry date:
* EDMstepFileFactory options:     020000
*************************************************************************
FILE_DESCRIPTION(('ViewDefinition [CoordinationView]'),'2;1');
FILE_NAME('\X2\987976EE7F1653F7\X0\','2019-04-28T10:08:43',(''),(''),'The
FILE_SCHEMA(('IFC2X3'));
ENDSEC;
```

a) IFC physical file header

```
DATA;
#1= IFCORGANIZATION($,'Autodesk Revit 2016 (CHS)',$,$,$);
#5= IFCAPPLICATION(#1,'2016','Autodesk Revit 2016 (CHS)','Revit');
#6= IFCCARTESIANPOINT((0.,0.,0.));
#9= IFCCARTESIANPOINT((0.,0.));
#11= IFCDIRECTION((1.,0.,0.));
......
#12970= IFCRECTANGLEPROFILEDEF(.AREA.,$,#12969,900.,2000.);
#12971= IFCAXIS2PLACEMENT3D(#6,#17,#19);
ENDSEC;

END-ISO-10303-21;
```

b) IFC physical file data segment

FIGURE 0.2 **IFC PHYSICAL FILE**

Within the scope of the study, this thesis reveals why data exchange is important in the perspective of the information required for building seismic performance evaluation, which would be introduced in detail in Section 4.3.2. As a bridge for data communication between BIM software, the IFC standard can effectively store building information of modeling software and structural analysis software and realize real-time mutual conduction between them to further improve evaluation efficiency. Many scholars have extensively researched the characteristics of data definition methods in BIM building physical models and structural analysis models (Cortes-Perez & Prieto-Muriel, 2020; Lee, Bae & Cho, 2012; Cavalliere, Favia & Lovicario, 2019). They have designed, developed and verified model data conversion interface solutions. The commonly used structural analysis software in China such as YJK and PKPM have also implemented model mutual guidance with BIM software in the form of plug-in. If the BIM model conversion function is introduced in the structural seismic analysis, the efficiency and quality of the evaluation, optimisation and prediction process can be further improved.

In summary, BIM is a collaborative methodology that use a digital presentation of a structure as a dependable data source to facilitate choices relating to building life cycle (Seghier et al. 2025). Due to the efforts of specialists in pertinent domains, the IFC standard has now attained the status of an international standard for the transparency of data and is extensively used in data interaction and exchange within the AEC sector. It employs an object-oriented, formal data definition language, EXPRESS, to describe data and organize data into hierarchical classes. The weaknesses of IFC were generated by an analysis published in a wide range of scholarly publications as:

- The IFC framework has a set of strict definition standards for the description of building information in terms of categories, relationships and levels (Khalili, 2021). Its complex, nested and indirect representation methods require users to spend a certain amount of time and resources to have a systematic understanding of this.

- The limitation of data sharing and exchanging by IFC outside the scope of construction project and facility administration and it has high barriers for extension and update (Chen, Yan, Chen & Li, 2022).

- It is difficult of using IFC to partition the information because of insufficient ability for representing some domain specific data (Wang, Li, Tang & Zhang, 2022).

- As a result of IFC's lack of semantic clarity in mapping entities and relationships, different descriptions of the same definition exist across various federated models (Huh, Ham & Kim, 2023).

The nature of the EXPRESS language underneath the IFC schema is the main cause of the above barriers. The manner in which EXPRESS delineates information dictates that the data in IFC cannot be processed and differentiated in straightforward and clear manner (Ruiz-Zafra, Benghazi & Noguera, 2022). Additionally, there are often multiple ways to express the same information, which make it difficult to unify and reuse resources. Therefore, additional qualitative improvements are needed for effectively managing models defined in EXPRESS (Walle et al., 2023). To this end, the Semantic Web technology is introduced to provide a more appropriate alternative way to solve the interoperability challenges, as it can connect different types of information into one Semantic Web, as well as their underlying meaning (Hardin, 2010).

2.3.3 THE SEMANTIC WEB TECHNOLOGY

2.3.3.1 THE INTRODUCTION OF SEMANTIC WEB

The World Wide Web (WWW) was invented by Tim Berners-Lee more than 30 years ago (Tim, 1994). The enormous chances to pay a visit to various digital files and materials that stored in virtual information environment is therefore offered. To date, the web is working as a crucial platform for billions of geographically dispersed agents to share information. However, as more and more researchers have shown increasingly interests on enhancing the efficiency of knowledge sharing of current web (Lan, Anh & Tran, 1983; Warren & Alsmeyer, 2005), the Semantic Web technologies emerged for such requirement. Berners - Lee et al. proposed the initiative of the Semantic Web in the article by adding a semantic layer to the current web (2018). The W3C describes the Semantic Web as

"The Semantic Web is to provide an infrastructure for the meaningful contents on the web pages, creating an environment where data can be shared and reused across application, enterprise and community boundaries, and providing a platform where machine can quickly retrieval and process the data by using inference and query for sophisticated tasks" (W3C, 2015). In order to achieve the function of the Semantic Web, semantic language which can express machine readable information for documents should be developed. Therefore, it is necessary to construct a brand-new architecture for semantic languages to complete information discovery, access, presentation, and maintenance, as well as the development of Semantic Web applications (Henry, 2004), which is named as "the Semantic Web architecture" (Figure 2.7).

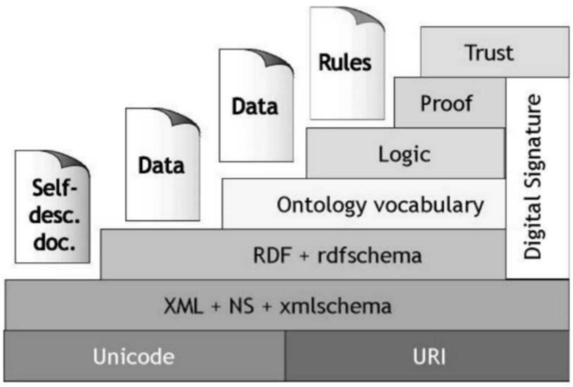


FIGURE 0.7 **A LAYERED APPROACH TO THE SEMANTIC WEB (HENRY, 2004)**

2.3.3.2 ONTOLOGY AND LANGUAGE FOR SEMANTIC WEB

Ontology is recognized as the foundation of Semantic Web development (Taye, 2010). The major impetus behind ontology evolution is to enhance knowledge sharing and reusing across various disciplines (Guarino, 1997). Ontologies can provide a lexicon and framework for representing domain information in a structured format that is readable by both machine and human. The original definition of ontology derives from a philosophical term, which define ontology as

"The study or theory of the explanation of being" (Taye, 2010).

Then the conception of ontology is explained by Lowein that:

"The set of things whose existence is acknowledged by a particular theory or system of thought" (Lowe, 1995).

Then the meaning of ontology was changed from early 1980s as it has drawn the researchers' interests from artificial intelligence community. The definition given by Neches and his colleagues (Neches, 1991) offer a vague guide that describing different tasks for ontology development, for example, identifying classes and relationships between classes:

"An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary".

By the late 1990s, the application of ontology has expanded into various domains and disciplines, including computer science, electronic commerce, knowledge management, etc.(Swartout & William, 1999; Welty & Guarino, 2001). Therefore, the meanings and explanations of ontology have been transferred into various iterations and implementations along with the applications in different areas. Gruber provided a definition of ontology that is frequently cited:

"An ontology is an explicit specification of a conceptualization" (1995).

Based on this definition, Borst (1997) further explained it by focusing on the essence of sharing:

"Ontologies are defined as a formal specification of a shared conceptualization".

In Guarino's theory (2009), an ontology is:

"a set of axioms, i.e., a logical theory designed in order to capture the intended models corresponding to a certain conceptualization and to exclude the unintended ones. The result will be an approximate specification of a conceptualization: the better intended models will be captured, and non-intended models will be excluded".

Regarding the definition of ontology, most Chinese scholars use the definition given by Gruber (2019) as the base and give it different meanings according to the research field and research purpose. For example, Zhang et al. believe that ontology is a set of concepts that describe domain knowledge with relationship between domain objects (2024). Gu Fang et al assert that ontology is a fundamental knowledge system in a certain domain that describes concepts, terms and their relationships in a standardised manner and provides corresponding terms for articulating domain knowledge (2023). While this thesis posits that ontology elucidates concepts and relationships with unified cognition within a certain domain, enabling various entities in the domain to share knowledge and collaborate in a formalised and standardised way. As ontology is being used in various fields and disciplines, the theory is continuously evolving and being implemented, leading to more a more comprehensive knowledge and application of ontology with the academic community. Significantly, the definition of ontology, both domestically and internationally, believes that ontology encompasses several key characteristics such as conceptualization, clarity, formalization, and sharing (Ancione, Ansaldi, Bragatto, Agnello & Milazzo, 2024).

Ontology was initially derived from philosophy and subsequently integrated into AI and computer science to address communication barriers between machines and humans, stemming from the latter's incapacity to comprehend natural language. Domain scholars employed ontological concepts to facilitate mutual understanding of knowledge between humans and machines, thereby promoting efficient information utilisation and sharing. As for the function of ontology, Peleg, Veggiotti, Sacchi and Wilk (2024) think that the primary function of ontology is to facilitate the effective representation and reutilization of knowledge within a certain domain, thereby enabling both domain and related fields to readily access and utilise domain knowledge for research or practical. While Farghaly, Soman and Zhou (2023) believe the ontology is used mainly for the integration of domain knowledge by constructing an ontology with a certain topic. Akcan, Erol and Kose (2023) assert that the important function of ontology is to use its mapping capabilities to comprehensively articulate the semantic relationships among information resources within a certain domain and to uncover significant implicit knowledge. By analysing and synthesising perspectives on the functions of ontology in ontology research literature, it can be concluded that ontology offers significant advantages:

- The interoperability of ontologies makes it possible to link different structured and formalized data to building information models.

- Ontology can define concepts and relationships in target domain flexibly.

- Ontology can provide a consistent and formal taxonomy and classification structure, which enables mapping concepts among diverse fields.

- Ontology language is more user-friendly and allows for better comprehension and updating by users.

- Ontology has the function of reasoning for processing information automatically and supporting decision-making as well. More importantly, the reasoning can be further used by employing semantic rules.

The endeavor of leveraging ontology has been well documented. A comprehensive review of 142 journal articles on ontology evolution in AEC domain is conducted by Farghly (2023). This research is systematically categorised into ten application domains in AEC and also underscores significant shortcomings in existing ontology research. Moreover, Zhong (2019) conducted a scientometric analysis critical review of 199 articles pertaining to the ontologies implemented within the realm of the construction sector. Some representative articles are listed as follows:

- Wetherill et al introduced a prototype ontology applied in the construction industry for managing semantic knowledge as part of the e-COGNOS project, which consists of over 15000 concepts (2002).

- Anumba et al. presented a prototype of semantic web-based information management system (SWIMS), which used an ontology-based method for managing knowledge in a semantic web environment (2008).

- Pauwels et al. indicated the potential of solving the interoperability issue in AEC industry by using ontology approach. The article proposed an ontology-based AEC description framework (2011).

- Abanda, Kamsu-Foguem and Tah created many ontologies, i.e. cost estimation ontology and on-site monitoring ontology, to enhance the building assets maintenance management via their objects, characteristics, and connections (2017).

- There are other scholars who created ontology knowledge base for the building energy consumptions (Wu, Cheng, Wang & Kwok, 2023), road maintenance (Lorvao et al. 2024), object-based CAD information modelling (Li et al. 2023), and seismic evaluation for structures (Naraghi et al., 2024).

There are various ontology construction methods. International research has conducted research on this earlier, yielding key methodologies such as the skeleton method, METHONTOLOGY method (mainly used in the field of chemistry), TOVE method, IDEF-5 method, SENSUS method and Ontology Development 101 method (Le et al. 2019). These ontology construction methods are more applicable to different fields. The TOVE method prioritises ontology assessment but lacks a formalised procedure. While the skeleton method includes an ontology assessment procedure, it just offers guidance for using the skeleton method in ontology construction. There is an absence of defined operational procedures and strategic planning for the operational process. The IDEF-5 method is used for software development and is often utilised in organisations. Its construction method is conducive to ontology reuse, but it cannot be developed in a circular manner. The specific comparison of these methods is shown in Table 2.4.

Table 2.4 The comparison of various ontology construction methods

	the skeleton method	METHONTOLOGY method	TOVE method	IDEF-5 method	SENSUS method	Ontology Development 101 method
the way of construction	mannually	mannually	mannually	mannually	mannually	semi-automatically
the procedure of construction	ontology analysis (defining the meaning and relationship of terms), representation (based on specific language representation), verification	ontology management (clarify the ontology construction plan), ontology construction (specification, conceptualization, formalisation, execution), ontology maintenance (knowledge acquisition, ontology integration, evaluation, documentation, configuration)	motivation design, terminology formalisation, rule formalisation, ontology completion	organization and scope, data collection, data analysis, initialisation of ontology construction, ontology refinement and validation	define terms, connect terms, find root paths, add terms, find domain-specific terms	determine the scope of the ontology, find reusable ontologies, list important terms, define classes, define class attributes, define attribute facets, create instances
the application of method	single domain	Multiple domain	single domain	multiple domain	multiple domain	multiple domain
the application domain	enterprise	chemistry，（ONTO）2Agent	enterprise	enterpise, software develop	electronic science, militray, etc.	Medicine, AEC, etc.

				ment, etc.		
the level of construction detail	little	detailed	little	detailed	normal	detailed
the consideration of resue	no	no	no	no	no	yes

Furthermore, the METHONTOLOGY method is often used in the discipline of chemistry. It has a significant level of maturity and is appropriate for the development of extensive ontologies (Guyo, Hartmann & Snyders, 2023). The drawback lies in the absence of an ontology assessment mechanism. The SENSUS method is often used in the domain of machine translation. This method emphasizes the cultivation of heuristic thinking more than other ways, although it is deficient in the processes of ontology assessment and the documenting of ontology construction results. The Ontology Develop 101 method is an ontology construction method established by Noy et al (2005), from Stanford University. It is widely used by scholars in ontology research and is relatively mature and highly operational. Therefore, this thesis also refers to this method when constructing the ontology.

As the IFC standard relies on the EXPRESS language for information description and transmission through physical file formats, the core for Semantic Web technology lies in Resource Description Framework (RDF) (Nazari & Haydary, 2024). By representing information as a labelled RDF graph, every node in the graph corresponds to a real-world notion or object. The directed line segment between the connected nodes represents the relationship between the objects and a Uniform Resource Identifier (URI) is added to uniquely identify the resource. So far, there are many grammars including RDF/XML, N-Triples, Turtle, etc. that can convert RDF graph into computer-readable text representations, which is convenient for conversion and information sharing between different formats. In order to realize the semantic unification, RDF Schema (RDFS) adds the description of the resource properties and classes on the basis of RDF. Ontology Web Language (OWL) further strengthens the semantic structure by adding type restrictions and complex class expressions. When using OWL and RDFS to represent RDF graphs, new information can be inferred through a standard query method. For example, when a resource belongs to a certain sub-category, it can naturally be inferred that it is also part of the corresponding super-category. The ontology is generally expressed in OWL and the RDF graph described in OWL is referred as OWL ontology.

In order to support more complex logical reasoning, several rule languages with Semantic Web Rule Language (SWRL) as the mainstream have also been developed in Semantic Web domain, that users can use the rule language to express actual needs. The SWRL overcomes OWL's limitations through adopting classes and properties in OWL ontology knowledge base to infer new facts. The syntax of SWRL rule is displayed in Formula 2-7. As a result, the complex real world is depicted as a linked and labeled graph that is easily comprehensible to humans. Flexible and universal information representation can cover any concept, while a significant volume of diverse data from multiple sources can be stored uniformly in computer-readable manner (Chandra & Harel, 1985). The issue such as the lack of information representation entity and the inability to integrate unstructured information in IFC files are solved. It also facilitates information retrieval and query, then further to realized rule-based knowledge reasoning and obtains new knowledge from existing information.

Formula 2-7 syntax of SWRL rule

$A_1, \ldots A_{n-1}, A_n \rightarrow B_1, \ldots B_{n-1}, B_n$

Where A_i and B_i are atomic formulas and each of the atoms could be class, property, instance or built-in in OWL ontology.

The variables used in atoms are indicated using a question mark prefix, for example, A(?x), B(?x,?y),hasProperty(?x,belongsto).

The query language for ontology is developed to meet the need for convenient access to required information parts. Therefore, a number of query languages such as Resource Query Language (RQL), Sesame RDF Query Language (SeRQL) and Protocol and RDF Query Language (SPARQL) have been established for facilitating the information extraction from ontology. SPARQL is the very sophisticated query language (Prud'hommeaux and Seabrone, 2008), which is recommended for querying RDF and become the standard query language (Perez et al., 2009). For instance, the following is an example of a SPARQL query:

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf schema#>PREFIX type:
<http://dbpedia.org/class/yago/>

SELECT ?1b1 ?est

PREFIX prop: <http://dbpedia.org/property/>

WHERE {

    ?country rdfs:label ?1bl

    FILTER(bif:contains(?1bl,"Republic"))

  ?country a type:Country108544813 :

        prop:establishedDate ?est .

  FILTER(?est <"1920-01-01"~~xsd:date)

}
```

As for seismic performance evaluation under the guidance of FEMA P-58, there is a substantial quantity of diverse information from several sources such as building basic geometric information, earthquake intensity information and topological relationships between building components. Ontology can not only provide good hierarchical and relational structural for organization, but also realize explicit mapping between concepts through ontology alignment to improve resource utilization rate. In addition, when all information is fully and consistently expressed in a well-defined hierarchical structure, the ontology can also support reasoning, laying a foundation for the expression of the expression of evaluation logic and the retrieval of semantic information. It can usually be generated based on SWRL rules to match specific patterns of RDF graphs. The subset required by users to realize the reuse and sharing of knowledge (Razavi & Gholizadeh, 2021).

2.3.4 ONTOLOGY-BASED BIM SYSTEM

As introduced in a previous section, BIM has made a great success in visualising, simulating and collaborating of building design in construction domain. It has been leveraged to enhance project collaboration more effectively, as well as to integrate data and manage information for supporting activities through the whole project life cycle (Karan & Irizarry, 2015). As decision making for a certain aim may always be a process which requires across domain information, for example, structure optimal design may consider not only structural resilience knowledge, but also building sustainable knowledge. Therefore, it is still necessary for BIM system to expand its interoperability as to seamlessly include semantic information across different domain. The Semantic Web technologies, in particular ontologies, are progressively integrating with BIM technology as it can provide numerous benefits including semantic clarity in mapping concepts and relationships among various federated models (Radanovic, Khoshelham & Fraser, 2020). Therefore, a change has been led in study and development of the AEC industry (as shown in Figure 2.8).

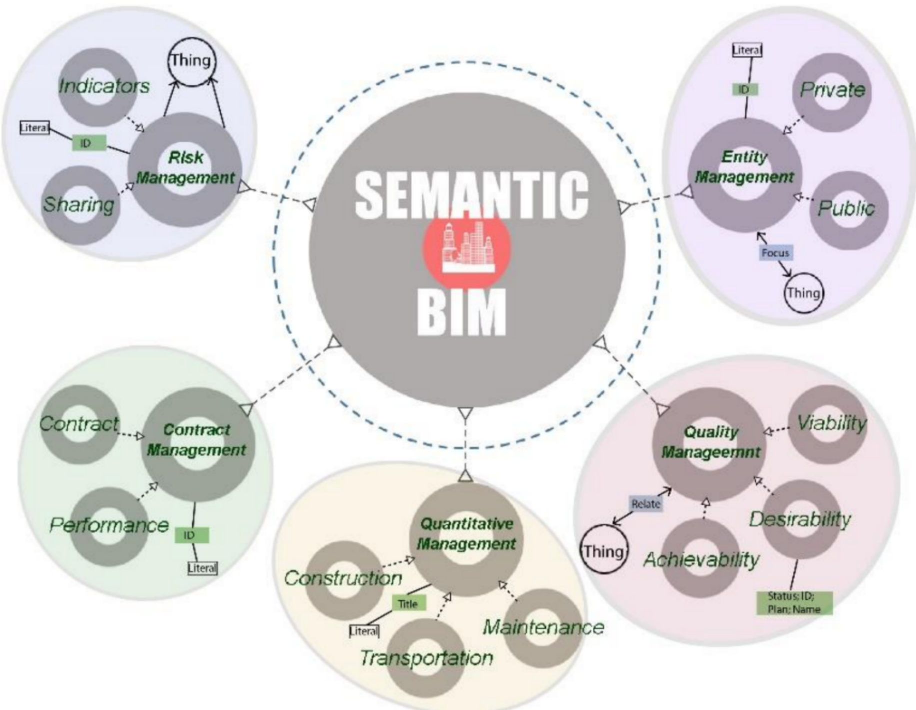


FIGURE 0.8 **BIM ALIGNS WITH SEMANTIC FUNCTION (RADANOVIC, KHOSHELHAM & FRASER, 2020)**

Some recent developments in ontology-based BIM system are presented as follows:

- A semantic information alignment method is proposed by Zhou and El-Gohary (2021). This method can be used to synchronise the depictions utilised in BIM with the depictions utilised in energy rules, which enabled automatic alignment process to support fully automatic energy compliance checking rather than manual or semi-automated process in existing research.

- By retrieving information from BIM models using an ontological knowledge base, Ren et al. constructed a more convincing and efficient method for project and financial management (2021). In practice of Public-Private Partnership (PPP), It provides a more efficient automated processes and reasoning tools to facilitate the better assessments of value for money.

- Zhu developed a BIM model quality inspection ontology and use it to convert inspection clauses into ontology reasoning rules, therefore facilitating model accuracy assessment and enhancing model quality (2022)

- Wang integrated BIM with ontology, using rules articulated in the ontology reasoning language SWRL to amalgamate the IFC model with the ontology knowledge base, and use reasoning to facilitate model parameter evaluation (2023).

- Ni used an ontological approach to model fire safety regulations, transformed IFC files containing building data into RDF format, and utilized SWRL reasoning principle to facilitate information mapping (2024).

The research mentioned above have a commonality in their use of ontology approaches for the processing of knowledge rules or BIM data. When semantic web ontology is used for rules processing, it often involves structured expressions, with the rules being transformed into semantic reasoning rules. In data model processing, it is essential to use the IFC format for conversion prior to using semantic web technologies for data operation. Through the literature review above, it can be concluded that the application of BIM and ontology can provide solutions for follow-up research – BSPEF through the three aspects:

- Use ontology technology to unify the content and processes of various evaluation requirements to have versatility and scalability. When multi-source heterogeneous information is stored in the same semantic environment, real-time correlation and interoperability between information can be carried out to improve knowledge utilization rate.

- Preprocess the IFC File (extract data only required for structure seismic evaluation) to obtain the basic building information from the BIM model and realize the integration of component information. Identify the component topological relationship automatically to make necessary preparations for the automatic establishment of the performance model. With the assistance of the data interaction interface between BIM modelling application and structural analysis application to provide the rapid generation of structural analysis models.

- Utilize the reasoning function of ontology to provide the automatic division and mapping of performance groups and reasonable expression of evaluation logic by expressing the performance group division and mapping rules as SWRL rules. The rapid invocation of semantic information and the real-time retrieval of component level evaluation result are realized with the help of SPARQL.

2.4 COMPUTATIONAL INTELLIGENCE METHODS

It has been discussed in Section 2.3.2 that modeling with high efficiency, information management, and automatic operation completion in the BIM environment cannot be accomplished via traditional BIM procedures without integrating tailored computational methods into the current BIM applications. As a result, the notion of "computational BIM" has arisen, which is a novel methodology that use computational intelligence to deal with various information related to information models.

Computational intelligence generally pertains to a computer program's ability to complete a certain assignment by using real-word facts or empirical findings (Gharehbaghi, Gandomi, Plervris & Gandomi, 2021). To date, scholars and researchers encounter extremely complex and nonlinear problems in the actual world. Conventional modelling techniques falter or perform inadequately when attempting to solve such problems. In this context, computational intelligence methods emerged and has been proven can tackle changeling issues in the sphere of science and engineering.

As stated in Section 1.1, the seismic optimisation design in this thesis is based on the "return on investment" criterion. Therefore, the designers should consider the competing objectives of minimizing the initial building expense and minimizing the expected seismic losses in the future, which are two conflicting goals. Multi-objective optimisation algorithm is well suited for this feature and will be discussed in detail in Section 2.4.1.

Moreover, the calculation process of predicting earthquake damage to building groups is relatively complicated, which means there is no defined formula for calculating the output parameters (indicators that can represent the damage status of building groups) according to the input parameters. The calculation process is hidden in a "black box". ANN can effectively address "black box" dilemma and will be explained in Section 2.4.2.

2.4.1 MULTI-OBJECTIVE OPTIMISATION ALGORITHM

2.4.1.1 THE MOTIVATION OF MULTI-OBJECTIVE OPTIMISATION ALGORITHM DEVELOPMENT

When solving practical problems, people often need to consider multiple goals. For example, both price and quality would be considered when shopping. In manufacturing, multiple goals are usually involved, such as production cost and efficiency. People sometimes will ignore certain goal based on their prior knowledge and past experience or convert these goals into a challenging of optimising a particular outcome. However, in scientific research, if there is no prior knowledge of mutually exclusive and equally important multiple goals, it becomes imperative to treat the problem as a multi-objective optimisation problem, which is referred to as multi-objective optimisation (MOO) (Zhang et al. 2022). Consequently, multi-objective optimisation algorithm is needed to address these problems.

In comparison of single-objective optimisation, multi-objective optimisation is occupied with some advantages. First of all, multi-objective optimisation can process different objective at the same time and can also achieve multiple requirements in one optimisation. By optimisation multiple objectives, even better solutions can be obtained than that of optimizing a single objective. In the data clustering problem, two primary objectives are: minimising the total distance between each sample and its corresponding cluster center, ensuring that the neighbouring samples are classified into one category. The research of xxx reveals that a collection of solution for different trade-offs is located in the compromise of Pareto frontier obtained by optimizing two indicators (Razdar, Adibi & Haleh, 2023). It is challenging to discover the optimal solution with an optimisation procedure that only considers one objective. Secondly, multi-objective optimisation algorithms can provide decision makers with richer selection samples. In this way, decision makers can get the best decision or decisions from multiple choices according to their own preferences. Thirdly, multi-objective optimisation can also provide the distribution of optimal solutions, which is conducive to the mining of novel knowledge and the deeper understanding of the problem. More sensitive targets with the optimal solutions can be studied according to the distribution of the set. Deb et al. designed a circuit layout algorithm based on evolutionary multi-objective optimisation (Deb, Jain, Gupta & Maji, 2015). As two mutually exclusive targets, wiring length and temperature effects caused by device aggregation are considered. The Pareto front distribution obtained by this experiment is relatively steep and from the distribution of the optimal solution, it can be found that the temperature target is more sensitive than the other one. Hence, engineers can mainly consider temperature indicators and secondly consider the wiring length indicator when making a decision.

As one of the latest developments in research methods, multi-objective optimisation has been effectively used to numerous research disciplines, such as control system design (Silva, Fleming, Sugimoto & Yokoyama, 2008), vehicle path planning (Tan, Cheong & Goh, 2007), feature extraction (Y. Zhang & Rockett, 2011), association rule mining (Kaya, 2018), image segmentation (Abdel-Khalek, Ben Ishak, Omer & Obada, 2017), genetic computing (Chaudhary & Kumar, 2019), community detection(Marti, Garcia, Berlanga & Molina, 2016) ,stock trading (Briza & Naval, 2011), etc.. It can be concluded that multi-objective optimisation is very effective in addressing multi-objective problems, especially ones with conflicting goals.

2.4.1.2 TRADITIONAL MULTI-OBJECTIVE OPTIMISATION ALGORITHM

Transforming the sub-objectives into multiple single objectives with mathematical transformation rules is the basic principle of traditional multi-objective optimisation approach. Then, each single objective problem could be solved by using technical approaches, thus to solve the problem of overall objective optimisation (Rabinovich, 2017). Take the Hierarchical sequence method as example, the detailed solution process is shown as follow:

Classify the sub-objectives based on the level of importance, which is denoted as $f_1, f_2, \ldots, f_n$.

Next, generate the best solution x_i of the n objective problems P_i incrementally, as shown in Formula 2-13:

$$(P_i)\min f_i(x) \qquad x \in X \tag{2-13}$$

Where

X is the feasible region and i = 1, 2, …, n. As the number of objectives is n, the final optimisation result is the effective solution X_n.

Traditional multi-objective optimisation methods use the basic principle of single-objective optimisation methods, which have some defects in practical applications, especially in more complex projects. According to literature review (Falcón-Cardona, Hernández Gómez, Coello Coello, & Castillo Tapia, 2021; Contreras, Sanchez & Ramirez, 2018; De & Giri, 2020; Chou & Truong, 2022; Shi et al. 2023), limitations are concluded as follow:

- Complete domain knowledge is required, the weight is not easy to determine, and the concave set problem is challenging to tackle.
- It is necessary to run the optimisation program multiple times for obtaining the most ideal solutions, which is a waste of time. Moreover, if different results are obtained every single time, the decision maker still cannot make a choice which one is the best.

Because of the limitation of traditional multi-objective optimisation algorithms, scholars have proposed a Multi-Objective Genetic Algorithm (MOGA) after continuous research, which will be discussed in the subsequent section.

2.4.1.3 MULTI-OBJECTIVE OPTIMISATION GENETIC ALGORITHM AND ITS DERIVATIVE ALGORITHM

Multi-objective optimisation genetic algorithm

MOGA has developed rapidly in recent years as genetic algorithm has the characteristics of global parallel search for the population, which is suitable for searching in a larger space (Arya, 2022). Multiple solutions can be obtained while running the genetic algorithm once and individuals with certain characteristics can be found through the generation-by-generation combination of chromosomes, which is free from the shackles of traditional algorithms. On the other hand, genetic algorithm's internal mechanism determines that it can readily manage large-scale problems and can search for the global optimal solution without being restricted by the nature of the problem. MOGA is based on the concept of Pareto sorting proposed by Fonseca and Fleming (1993). The algorithm points out that the hierarchical sequence number of all individuals is calculated one by one and the hierarchical sequence number of all non-dominated individuals is assigned as 1 and the hierarchical sequence number of the remaining individuals is 1 more than the number that can dominate it. The selection method of individuals of the same level uses random sampling and a fitness sharing strategy, which is to sort the population according to the level at first, then use the interpolation method to select fitness values for all individuals, and the fitness of the same level individuals also takes same value. Obviously, multiple individuals have the same hierarchical serial number, and it is a very huge task to determine their hierarchical serial number according to their dominance relationship. As MOGA relies too much on the decision maker's choice of sharing function, it is very likely to lead to premature convergence if the choice is not appropriate (Manjhi & Chaturvedi, 2021). In this context, the key issue of MOGA is how to better select fitness functions and assign values to them, and how to maintain the diversity of species.

Non-dominated sorting genetic algorithm

Srinivas and Deb proposed Non-dominated Sorting Genetic Algorithm (NSGA) is a method that uses non-dominated sorting to prioritise individuals with greater fitness for retention in the following generation (Srinivas & Deb, 1994). NSGA adopts a fitness sharing strategy to enable individuals on the quasi-Pareto frontier to be more evenly distributed, maintain the population's variety and prevent premature convergence (Dhabale, Jatti & Singh, 2014). NSGA uses a proportional selection crossover method to produce the next generation and the shortcomings are summarized as follows:

- Developing the Pareto optimal solution set is a time-intensive process. Due to the fact that each iteration of evolution must establish a non-dominated solution set, the algorithm will take a considerable amount of time to execute when the population size is substantial.

- There is no optimal individual retention mechanism, despite the fact that studies have shown that such a mechanism could not only enhance the performance of MOGA but also effectively prevents the loss of outstanding solutions.

- The sharing population's parameters individuals are difficult to determine, hence it cannot be used to adjust the distribution of population individuals.

Non-dominated sorting genetic algorithm -II

NSGA-II was initially suggested by Deb et al. (2002). While possessing the genetic variety, it has the ability to efficiently handle several competing targets involving discrete variables (Park, Hwang & Oh, 2018), which is commonly employed to tackle multi-objective optimisation problems (Wang, Li, Jin, Xiang & Li, 2020).

The NSGA-II algorithm is based on the NSGA and has been improved in three aspects as followed: (a). it reduces the computational complexity by proposing a rapid non-dominated sorting algorithm (Yi et al. 2020). (b). it introduces the elite strategies to increase the sample size (Zhang, Qian & Qian, 2021). By merging the parent population with the offspring population it produces, they jointly complete to generate the succeeding generation population, therefore safeguarding the exceptional individuals from the parent generation, ensuring that some exceptional population individuals will not be eliminated throughout the process of evolution, thereby enhancing the precision of the optimisation result. In addition, the finest individuals would be kept, and the level of population would be quickly increased with all individuals stored hierarchically in the population. (c). It employs the congestion degree comparison operator, which eliminates the need to manually specify the share index in NSGA, sets it as the standard for evaluating individuals within the population (Ke et al., 2021). Therefore, the individuals in the quasi-Pareto domain could be evenly dispersed over the whole Pareto domain, so guaranteeing the diversity in the population.

1.Fast non-dominated sorting

Making an assumption that a population is denoted as P, the procedure is required to compute two indicators, N_p and S_p, for individual p in population P. N_p represents the count of individuals that dominate individuals p in population whereas S_p represents the collection of individuals dominated by individuals p in the population. The algorithm's primary stages are as follows: (a). Identify all $N_p = 0$

population members and add them to the existing collection F_l. (b). For every individual i in the present collection F_l, the set of individuals dominated by it is S_i. Subtract 1 form N_j of each individual j in S_i. If $N_j - 1 = 0$, then move the individual to another set H. (c). As the initial level set of non-dominated individuals, the individual in F_l is optimal. It only dominates other individuals and not dominated by any other individuals. Give each individual in the set a same non-dominated order i_{rank}. Next, repeat the preceding classification procedure for H and assign the relevant non-dominant ranking until each individual has been classified.

2.Calculation of congestion degree

The congestion degree i_d index reflects the population density encompassing a particular point in the population. It is represented by the addition of the length and breadth of the largest rectangle surrounding the individual i but not any other individuals as shown in Figure 2.9. In NSGA-II, the calculation of congestion degree is an important part of ensuring population diversity. Following is an outline of the calculation steps: (a). Initialise the congestion degree i_d of each point to 0. (b). Conduct the population's non-dominated classification for each objective, ensuring that the congestion degree of the two individuals on the boundary is infinite, which is $o_d = I_d = \infty$. (c). Determine the congestion degree of the other individuals:

$$i_d = \sum_{j=1}^{m} (|f_j^{\,i+1} - f_j^{\,i-1}|) \qquad (2\text{-}14)$$

f_j^{i+1} represents the value of the objective function j at the point i+1 and f_j^{i-1} represents the value of the objective function j at the point i-1.

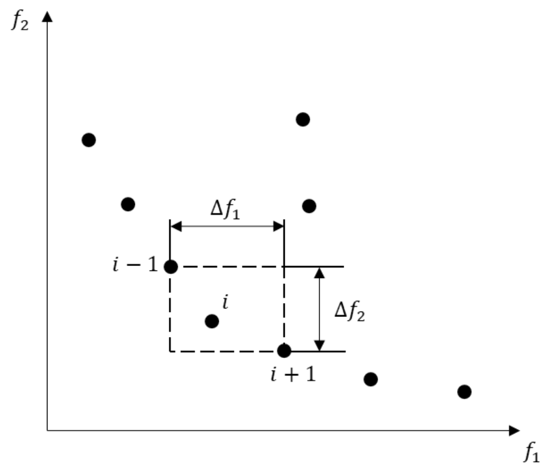


FIGURE 0.3 **THE CROWDING OF INDIVIDUAL I**

3.Congestion degree comparison operator

In step 1 and step 2, two properties of population individual i are obtained, which is non-dominated order i_{rank} and congestion degree i_d respectively. Based on this, the definition of congestion degree comparison operator may be given. when an individual is compared to another, the individual wins if any of the below criteria is met:

(1) if the non-dominated stratum of individual i is superior to that of individual j, it means that

$i_{rank} < j_{rank}$.

(2) if both individual i and j have the same level and the individual j has a greater congestion distance than the individual j, which is $i_{rank} = j_{rank}$ and $i_d > j_d$.

The criteria (1) make sure that the picked individuals are of the superior non-inferior calibre. The criteria (2) choose individuals from the less congested region between two individuals of the same non-inferior level who have no preference regarding their congestion distance. Then winner then proceeds to the subsequent operation.

4.Elite strategy

The NSGA-II employs an elite strategy. By mixing all the individuals of the parent and offspring to perform non-dominated sorting, it is possible to avert the depletion of exceptional parent-generation individuals. The implementation processes are depicted in Figure 2.10.

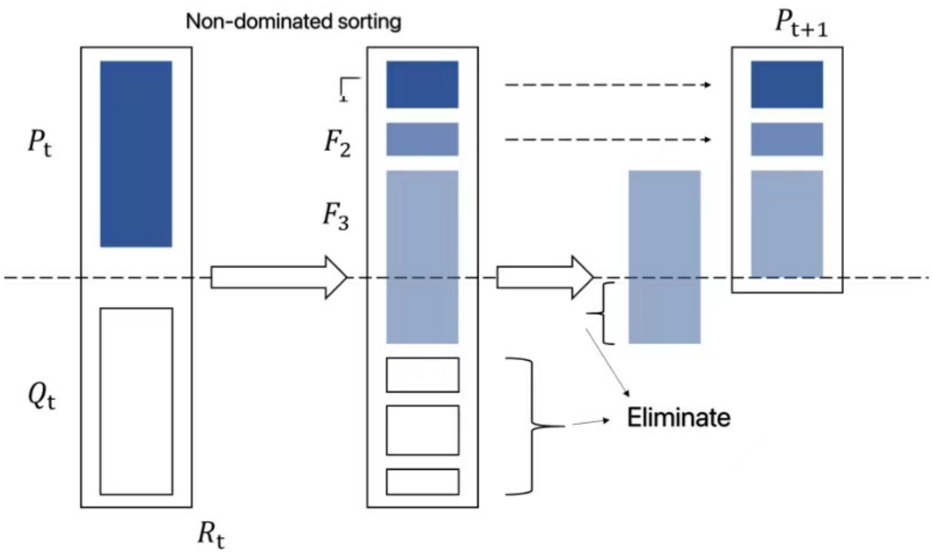


FIGURE 0.10 **PROCEDURES FOR IMPLEMENTATION ELITE STRATEGY**

The new population Q_t and P_t generated in the generation t are merged to create R_t, which has a population quantity of 2N at this time. Conduct non-dominated sorting to R_t and construct a sequence of non-dominated sets F_i to compute the congestion degree. As both progeny and parents are concluded in R_t, the finest individuals in R_t are those from the non-dominated set F_1 after non-dominated sorting. In this case, they are added to the subsequent progenitor population P_{t+1} first. If the size of P_{t+1} compared to N is smaller, add the next-level non-dominated set F_2 into P_{t+1} until population size surpasses N upon adding F_3. Then use the congestion degree comparison operator on the individuals in F_3, take {num(F_3)-(num(P_{t+1})-N)} individuals to ensure that the size of P_{t+1} to equal N. Genetic operators ultimately generate a new progeny population, referred to as Q_{t+1}.

According to literature, the fundamental sequence of NSGA-II is depicted in Figure 2.11 (Deb et al., 2002). In the first phase, perform non-dominated sorting and conventional genetic operations, for example, selection, crossover, and mutation on the initial population to obtain the first-generation subgroup. In the next phase, the parent population is combined with the progeny population to facilitate swift non-dominated sorting beginning with the second generation. Concurrently, calculate the congestion degree of each individual in every non-dominated layer and choose appropriate individuals based on the non-dominated relationship and the congestion degree to establish a new progenitor population. In the last phase, generate a new progeny population using conventional genetic operations, until the program's termination conditions are reached (Razmi, Rahbar & Bemanian, 2019).

In summary, the NSGA-II is one of the most often used multi-objective algorithms, which functions based on non-dominated sorting to effectively provide a diversified Pareto frontier across many goals. Therefore, it can improve the dissemination of solutions throughout the evolutionary process. Multi-objective algorithm represented by NSGA-II has been extensively used to investigate seismic optimal design according to the review in Section 2.2.2. Due to the intricate relation between the two major costs, NSGA-II is the suitable approach to devise an effective method for seismic optimal design to ascertain the Pareto optimal solution set for the construction initial costs and the expected earthquake losses. Chapter 5 will describe how to achieve seismic optimisation design based on "return on investment" criterion in details.

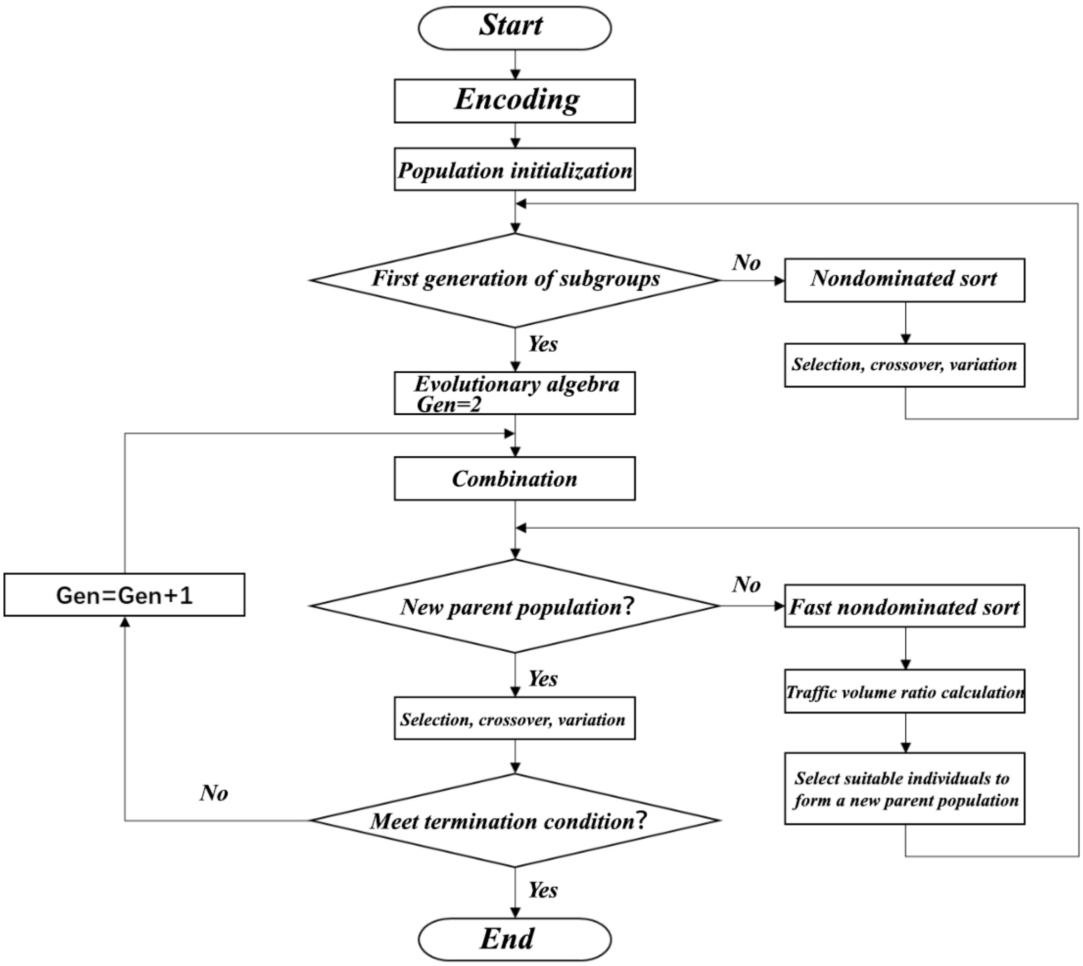


FIGURE 0.11 **FLOW DIAGRAM OF NSGA-II**

2.4.2 ANN

The aspiration to enable computers to emulate human cognition and achieve genuine AI has long been a pursuit of humanity (Gauchi, Bensadoun, Colas & Colbach, 2017). ANN, developed with the help of neuroscience and computer science, are making mankind get closer and closer to this dream. scholars Warren S. McCulloch and Walter Pitts (1943) in University of Illinois proposed a simple neural network model, pointing out that nerve cells in the human brain are essentially a component that can perform logical calculations, and another artificial network can be used to simulate human brain's neural behaviour. Since the 1990s, the computer science's developing progress has been brisk, artificial neural networks uses its superior nonlinear approximation performance to have an irreplaceable role in solving pattern classification, regression, clustering, and optimisation calculations (Ismail, Singh, Shirazian, Albadarin & Walker, 2020). It has been widely used in almost all disciplines and majors such as aerospace (Choudhury & Chandrasekaran, 2020), finance (Choudhury & Chandrasekaran, 2020), machinery (Singh & Abbassi, 2018), agriculture (Saldaña-Robles et al., 2020), medical (Hajder, Kolbusz, Hajder, Nycz & Liput, 2020), civil engineering (Polat, Bingol, Gurgun & Yel, 2016).

As the modelling and imitation object of the neural network is the human brain, human brain's neuron structure is abstracted into a mathematical concept to obtain the fundamental information execution unit of the neural network – the neuron model which is depicted in Figure 2.12. The input signal (x_1, x_2, ..., x_m) is the input of the neuron model, the sum node represents a linear model $u_k = {}_{j=1}^{m}w_{kj}x_j + b_k$, where each input signal x_j is linearly combined according to the corresponding weight w_{kj} and the bias term b_k is brought in. The non-linear activation function f controls whether the neuron sends a signal to the outside by setting a threshold. The output $f(u_k)$ of entire processing unit will be passed to the next neuron model.

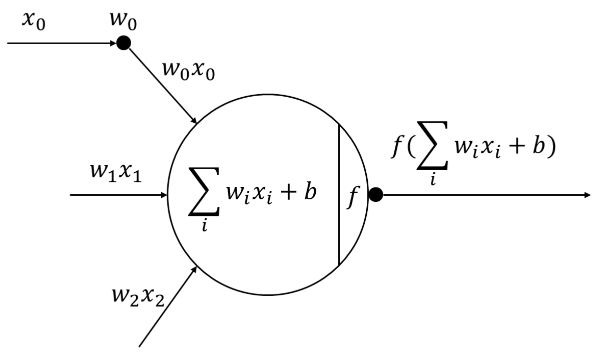


FIGURE 0.4 **A TYPICAL NEURON MODEL**

The different interconnection patterns between multiple neurons constitute a neural network with different properties and functions, such as a feedforward network. According to the order of signal transmission, the feedforward network is composed of multiple "layers" termed input layer, hidden layer and output layer in sequence (Mittal, Devi & Chauhan, 2014). Each layer consists of a certain number of neurons and the neurons in the i-th layer obtain signals only from the neurons in the (i-1)-th layer. As illustrated in Figure 2.13, there is no exchange of signals between the neurons. As the initial neural network layer, each element value of the input vector signal without its own weight value and bias value is received by the input layer. No operations will be performed on the input signal in this layer. A neural network can contain one or more hidden layers. The output layer is the last layer of the neural network. It receives the input from the last hidden layer. The neurons between the two layers are interconnected through weights. The weight value signifies the potency of the link between the units and determines the degree of the input's influence on the output. While the goal of neural network training is to update weight value and bias value to make the output close to the expected output.

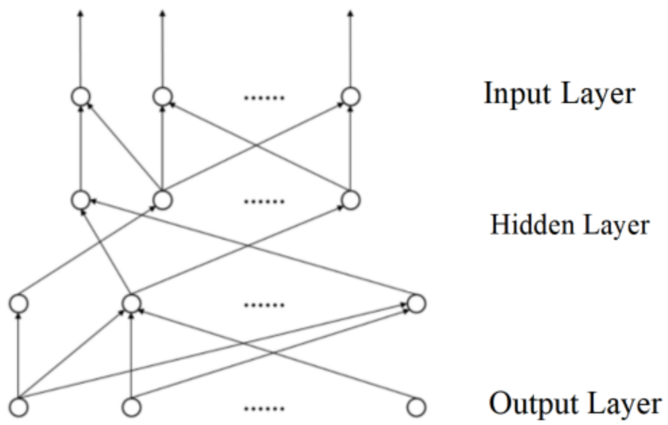


FIGURE 0.13 **STRUCTURE OF FORWARD NETWORK**

The Back Propagation (BP) Neural Network selected in the research, which is proposed by Hornik et al. (1989), is a typical feedforward network and happens to be one of most extensively utilised neural networks. Morfidis et al pointed out that this type of network has good application effects in dealing with Function Approximation (FA) problems and Pattern Recognition (PR) problems (2018). As shown in Figure 2.14, an input layer, several hidden layers and an output layer make up the BP Neural Network. The neurons between each layer are entirely interconnected, which is evidenced in two main processes: the propagation of the input signal forward and the propagation of the learning error backward. When propagating forward, the input signal is transmitted from the input layer to the hidden layer, where it is undergoes processing, and then transmitted to the output layer. If the final output differs from the predicted output, error back propagation will begin. The error will be propagated from the output layer to the input layer via the hidden layer and will be distributed to all units on each layer. This error signal is reflected in the weight correction of each unit so that the network's actual output closely matches the intended output as possible. Constant weight adjustment constitutes the network's learning and training procedure, which is realized with the help of training algorithms.

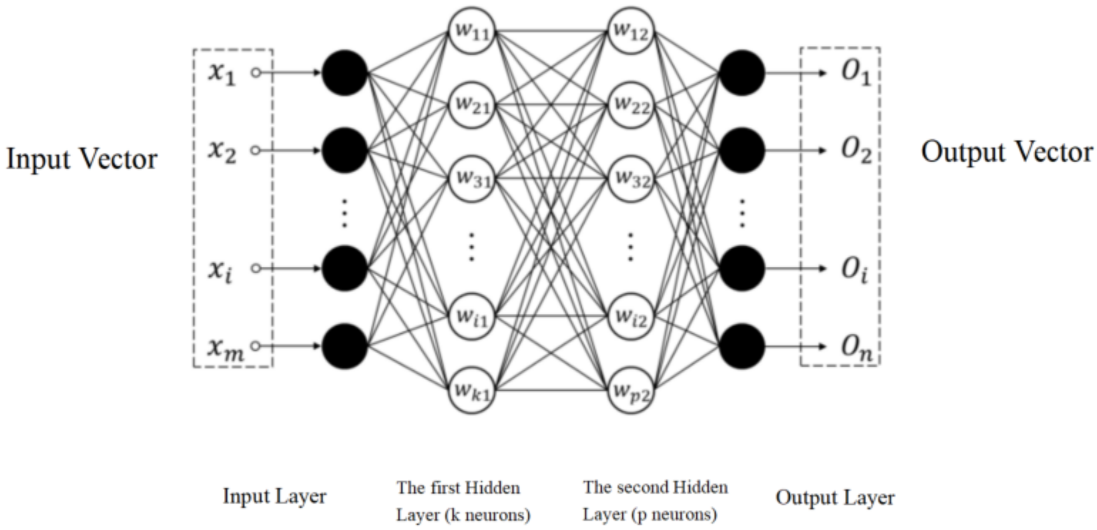


Figure 0.14 **A typical BP neural network model**

Kaushik and Banka proposed a method to reduce the workload of reliability analysis and damage assessment for analyzing the structural seismic reliability (2015). The idea of membership was used to assign training times to network sample and genetic algorithms was also utilized to identify the limit state surface for structural reliability computation in Wang, Chen, Wang and Xiong's research (2022). Zhong, Xie, Qin and Zhang (2022) used uniform design to ascertain the network training set and then developed the BP-based ANN model for assessing the structural reliability. Pei, Liu, Zhang & Chen (2023) used the BP network to substitute the computational process for assessing the structural response, so conserving computation time, and validated its efficacy via illustrative instances.

Based on the existing research on ANN, this method has strong applicability and can be combined with many other methods for improvement and optimisation, so as to solve the various problems in construction industry. Moreover, ANN offers efficient resolutions to both linear and nonlinear complicated issues that conventional mathematics and approaches are unable to adequately address. It establishes a mathematical correlation between input parameters and output parameters via a sequence of numbers as weights and biases, and is extensively used in nonlinear pattern recognition tasks, yielding effective outcomes. During the earthquake damage prediction process of building groups in this thesis, MIDR is used as the output parameter to define the damage status of buildings and different seismic relevant parameters are chosen as input parameter. However, the relationship between the inputs and outputs cannot be expressed in a specific mathematic formula and ANN is suitable for solving the problem.

2.5 SUMMARY

The chapter reviewed the research status of seismic design with an emphasis on seismic performance evaluation based on FEMA-P58, seismic optimisation design based on "return on investment" criterion and earthquake damage prediction. Research gaps and critical evaluation of existing research are concluded. Therefore, the call for establishing an integrated framework derives from the review to manage seismic evaluation information and knowledge, as well as the method of multi-objective optimisation with regard to return on investment criteria based on the component level of building, and consideration of seismic wave amplitude modulation issue for earthquake damage prediction.

Interoperability is the crux of system integration. Hence Section 2.3 commences with the introduction of BIM, then introduces the AEC industry standards IFC. In Section 2.3, semantic web, ontology and ontology-based BIM system are also reviewed. A diverse range of methods and approaches in construction domain are reviewed and through a critical analysis, it can be justified that the BIM-based ontological approach can offer great interoperability for processing both structural design knowledge and seismic information. In this regard, it has been selected for the seismic performance evaluation framework development in the study.

Moreover, Chapter 2 has explored computational intelligence method as well. The multi-objective optimisation algorithm with related definitions and review of different types of algorithms including traditional multi-objective optimisation algorithm, MOGA, NSGA and NSGA-II are introduced. The identification of limitations about these algorithms are concluded. Through a critical analysis, a theoretical basis for choosing NSGA-II to achieve seismic optimisation design are formed. The state of art ANN has been explored as well, then to provide a theoretical basis for proposing earthquake damage prediction method based on it. The methodologies are introduced in Chapter 3.

CHAPTER 3 RESEARCH METHODOLOGY

A precise description of research is expressed by the English Advanced Learner's Dictionary as "*a careful investigation or enquiry, especially through the search for new facts in any branch of knowledge*" (Wehmeier & Hornby, 2000). Furthermore, from defining and refining research problems to reaching and achieving solutions, a series of systematic methods provided solid foundation for research. To conduct research, it is essential to demonstrate the suitable research paradigm and philosophy as well as the appropriate research methods and techniques, which is regarded as the core to complete the research with high quality according to Redman and Mory's research (1923).

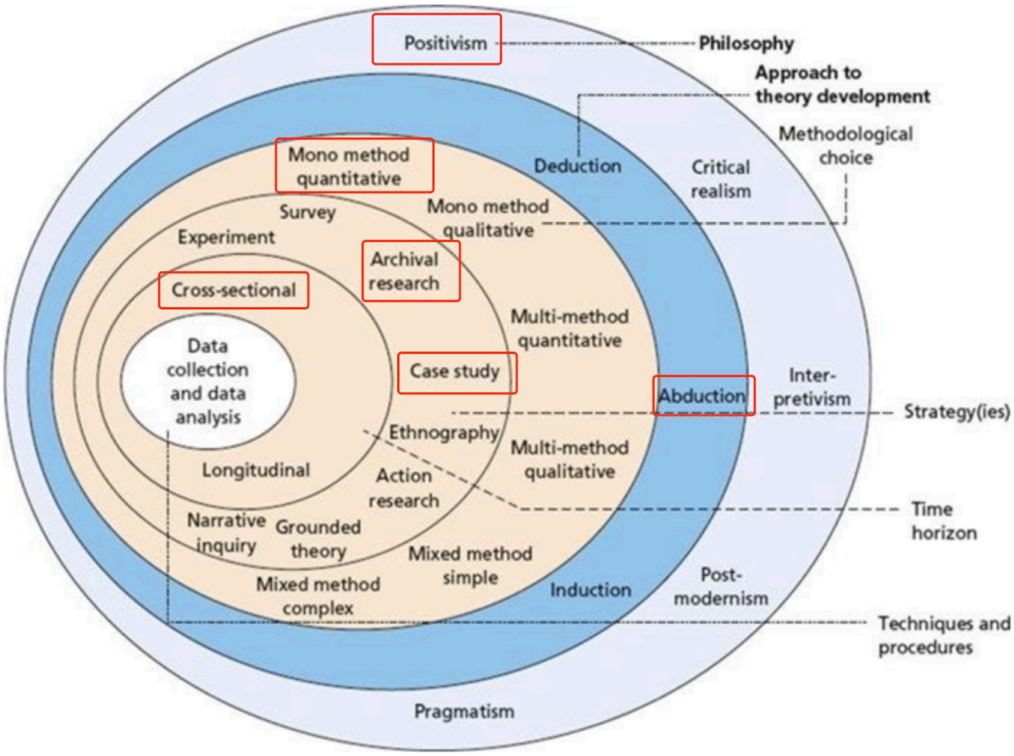


FIGURE 0.1 **METHODOLOGICAL APPROACH UNDERLYING THIS STUDY (IN RED BOXES) BASED ON SAUNDERS' RESEARCH ONION (2000)**

To address the research questions proposed in Chapter 1 of the research, Chapter 3 will express the adopted research methodology to provide solutions to research questions. It will begin with a discussion of the project's overall methodology, which helps define the guiding principles and approaches for this design-based research. For more clarity, Saunders' Research Onion will be referred as a basis for discussion (as shown in Figure 3.1), which is a good method for explaining rigorously the methodological processes of this research (Saunders, Lewis & Thornhill, 2000). Figure 3.1 illustrates that the research process consists of multiple layers. Therefore, the study will begin by describing the research philosophy that underlies it and the research methodological choice that supports it. Next, the research aims of the research design as well as the research approach used would be introduced before stating the chosen research strategies. Finally, the time horizon and the procedures/techniques which are used for collecting and analysing data are stated.

3.1 RESEARCH PHILOSOPHIES

Research philosophy, referred to a structured set of beliefs and hypotheses concerning the growth of knowledge (M. Saunders et al., 2000), is closely associated with research paradigms which is considered as "*the philosophical intent or underlying theoretical framework and motivation of the researcher with regard to the research*" (Mackenzie & Knipe, 2006). The selection of a research philosophy is primarily influenced by the research motivations, the research objectives, and the philosophy typically selected within a discipline (Maceviciute, 2006).

Ontology and epistemology are two central branches of philosophy that deal with the nature of reality and knowledge. They address fundamental questions about what exists and how we can know about what exists.

Ontology is the study of being and existence, which examines the categories of things that exist and how they relate to each other (Phyak, 2022). Ontological questions usually involve topics such as:

- What kind of things exist? Are there physical objects, abstract objects, ideas, and numbers?

- What does it mean to exist? This involves exploring different modes of being. For instance, how does the existence of a physical object differ from that of a thought or a possibility?

- How do entities relate to each other? Ontology also looks at the relationships and hierarchies between different categories of being. For example, how do individual entities relate to larger entities (like a branch to a tree)?

Epistemology is the study of knowledge and belief (Porter, Hutchison & Mathpati, 2019). It explores how we know what we know, the justification of beliefs, and the nature of truth. Typical Epistemological questions include:

- What is knowledge?

- How do we acquire knowledge? This involves examining sources of knowledge such as perception, reason, memory, and testimony.

- What justifies a belief? Epistemologists look into what constitutes sufficient evidence or reason for a belief to be considered knowledge.

While ontology and epistemology are distinct, they are interrelated. Ontological commitments can influence our epistemological views and vice versa. For instance, what we believe exists (ontology) can shape our understanding of what can be known (epistemology). Similarly, how things can be known (epistemology) can affect assumptions about what exists (ontology). They create a comprehensive framework for understanding and interpreting the world (Pan, Li, Wei, Zhang & Luo, 2024).

There are some philosophical stances. Positivism is grounded in the belief that reality is objective and can be discovered through empirical observation and logical analysis (Juan, Jaime, 2023). It emphasizes quantitative methods such as surveys and experiments that can produce statistical data. Positivists seek to formulate general laws and predict patterns of human behavior through measurable evidence and tend to adopt a realist ontology, assuming that reality exists independently of human perceptions, and an objective epistemology, believing that knowledge can be acquired through observation and experimentation without bias (Pavel, 2019; Outhwaite, 2015; Herbert, 2016). Interpretivism suggests that reality is socially constructed, and that understanding human behaviour requires grasping the meaning and experience individuals attach to social phenomena (Philip & Pascal, 2019). This approach typically employs qualitative methods like interviews, focus groups, and ethnography to explore the context and depth of social reality, which aims to understand phenomena from the perspective of participants (Somervile, 2012; Smith, 2006; McKenna, 2020). Moreover, interpretivists often adopt a relativist ontology, which views social reality as multiple and constructed by individuals and a subjective epistemology, acknowledging that researcher biases and perspectives play a role in the creation of knowledge (Aya, Edwards & Rillie, 2017). Realism is the viewpoint that there is a reality independent of human thoughts or beliefs, which asserts that the world exists and possesses properties, regardless of whether or not human observe these properties (Raimund & Schoenegger, 2008). Critical realism acknowledges that human understanding of reality is mediated by perceptions and social conditioning, which accepts that the world and its mechanisms exist independently of knowledge, acknowledge that human perceptions and theories about the world are inherently fallible (Stevens, 2020). Pragmatism is a philosophical tradition centered around the interplay of theory and practice (Dingwall, Cassell & Colin, 2013). It focuses on the practical application of ideas by assessing their truth in terms of the success of their practical consequences.

As stated in Chapter 1, the research objectives are confirmed as follows:

Objective 1: Identify domain knowledge, methodology and current practice of PBSD.

This objective is fulfilled through the critical analysis of literature review, which helps the author comprehend the current status and advantages/disadvantages of existing work. Moreover, relevant seismic codes (such as GB50011-2003) and guidelines (such as FEMA P-58) are explored for identifying the key concepts on the target domain in Chapter 2 and Chapter 4.

Objective 2: Explore information technologies to provide the essential groundwork for the integration of information technology and engineering. Then use them to create a knowledge model that enable a standardized semantic format, allowing for interchange and querying of diverse information and knowledge from several sources.

Information technologies are explored by critical literature review in Chapter 2. A knowledge model for interchanging and querying of seismic assessment evaluation and optimisation design has been developed based on BIM and ontology in Chapter 4. Besides, many other techniques such as IfcOpenShell and pythonOCC are leveraged as well.

Objective 3: A BSPEF would be proposed to realize the automation of building evaluation process.

This objective is fulfilled by utilising BIM and ontology to express the evaluation process and logic through organizing the key concepts in performance evaluation and the relationship between the concepts in Chapter 4.

Objective 4: Establish a Multi-objective Seismic optimisation Design Method (MSODM) for RC frame according to the BSPEF.

This objective is fulfilled by developing a multi-objective optimisation algorithm based on NSGM-II to find the balance point between the two major costs, with BSPED served as the knowledge foundation in Chapter 4 and Chapter 5.

Objective 5: Establish an Earthquake Damage Prediction Method (EDPM) of multi-scale regional RC frame based on ANN and BSPEF.

Input parameters and output parameters are selected based on literature review and the existing theories to train the ANN model for more accurate earthquake damage prediction in Chapter 6. Furthermore, the prediction outcome can be seen intuitively with the visual software.

Objective 6: Validate the application effect of BSPEF, MSODM and EDPM.

The objective is fulfilled by several study cases provided by Sichuan Provincial Government. The application effect of BSPEF and MSODM is justifies by comparison of maintenance cost of different design schemes in Section 7.1 of Chapter 7. The application effect of EDPM is verified by the generalization ability of ANN model in Section 7.2 of Chapter 7.

Therefore, it can be concluded that this thesis lies in both information system and in the seismic engineering domain. The benefit of the currently evolving ICT has been taken initially for establishing methods to develop sophisticated information systems for automated building seismic evaluation, design and damage prediction. Hence, it is essential to comprehend the nature of information systems which comprise appropriate research philosophy.

Information systems is a discipline which relates to "*the development and use of information systems by individuals, groups, organizations and society, where usually those information systems involve the use of computers*" (Kim, Kwon, Heo, Lee & Chun, 2014). The information systems domain has various aspects such as scientific, engineering, technological, managerial and societal aspects (Wood-Harper, Antill & Avison, 1985). As for research regrading with information systems, positivism is commonly adopted as the research approach (Chen & Hirschheim, 2010), where reality is considered to be expressed by "real" objects that exist independently. When the positivist method is adopted, collecting data based on "an observable reality" and trying to find "regularities and causal relationships" are commonly involved, aiming to give a precise and unbiased interpretation of the truth (Johnson & Onwuegbuzie, 2004). The researcher can have their own ideas and independent values for the observed phenomena. Adopting the positivist method is associated with coming up a hypothesis and then testing it, which involves using quantitative analysis of data principally. Therefore, "time- and context-free results" can be obtained and generalized.

However, information systems research does not suit completely into a positivist paradigm as information systems also require management decisions and social processes in some situation. Therefore, interpretivism approach is more and more utilised and accepted in information systems, which emphasises the significance of human involvement in social science (G. Walsham, 1995). As a physical phenomenon, human beings can create meanings which should be interpreted and learned, thus the way of studying human beings and social environments of them should be different from each other (Walsham, 1995). When constructing the domain knowledge base using ontology technology, the necessary building seismic information should be collected by researcher. Moreover, the classes, properties etc. should be defined by researcher to construct the ontology framework. For example, in Section 4.1.2, the application ontology, used to enhance the clarity of the performance evaluation approach and its objective, is proposed based on the concepts of different assessment requirements and classification of assessment results.

Similarly, whereas positivism was once the prevailing philosophy in seismic engineering domain, interpretivism is now extensively utilised as well. Humans generate different meanings and definitions that need interpretation and analysis such as earthquake ground motion. Interpretivists need to join participants' environment and comprehend their point of view. Consequently, interpretivists' values, behaviours also impact the study procedure. For example, the selection of ANN model's input parameters is decided based on both the data analysis of existing works and the author's point of view.

Although it is contended that positivist and interpretivist philosophy may be integrated, while this study leans towards an interpretivism philosophy in some respects, interpretivism is still the most suitable philosophy to embrace. The reasons are explained as follow:

- The nature of the research objective and outcomes need the data justification based on the existing theories and axioms.
- The nature of the research objective necessitates quantitative method to provide a comprehensive grasp of the phenomena examined.
- The method applied includes quantitative data collection.
- The outcomes and findings are intended to be typical of more structure cases and can be generalised.

3.2 RESEARCH APPROACH

There are several research approaches where deduction, induction and abduction are typical three of them. Deduction is a research approach that involves reasoning from the general to the specific, which is often associated with the scientific method and is used to test theories or hypotheses (Taro & Kojima, 2024). The advantages of deductive are a) rigorous testing that can provide a clear and structured method to test hypotheses and validate theories, b) clarity and precision that offers clear guidelines for operationalizing concepts and setting expectations, c) replicability that the structured approach can enhance the possibility of replicating studies and verifying results (Liu et al. 2017; Jacob, 2021; Tan, 2024). Induction is a fundamental method in scientific investigation, focusing on deriving general principles from specific observations (Landino et al., 2014). This approach allows researchers to develop theories that explain patterns and regularities observed in particular data sets. Furthermore, the advantages of induction are a) grounded in data: since the approach originates from specific observations, the resultant theories are often closely aligned with the empirical data, lending credibility and relevance, b) discovery of novel theories: induction allows for the development of new theories or frameworks that may not be evident through deductive reasoning, c) flexibility: as researchers can modify or change their hypotheses based on merging data, making it adaptive to new information (Lu et al., 2019). Formulating a hypothesis according to already existing theories, validating this hypothesis to accept or reject it and revising the theories are required for deduction approach (Gray, 2009; Lawrence, 2005), while induction requires search of patterns that emerge from data collection and analysis, development of theories concluding from the observed results and finally test of theories based on hypotheses (Lawrence, 2000). Additionally, deductive is typically linked with quantitative research while inductive are traditionally to be linked with qualitative research (Brewis & Claire, 2000).

Abduction is a method of reasoning that starts from the observation of a phenomenon and seeks the simplest and most likely explanation for it (Stocker, Wittauer & Ismailidis, 2022). This approach contrasts with deduction, which derives specific conclusions from general premises, and induction, which generalizes from specific observations. Abduction is particularly useful in exploratory research, hypothesis generation, and fields where initial observations are not thoroughly explained.

The research involves a thorough examination of existing systems and approaches for constructing earthquake performance assessment, design, and damage prediction, which cover four main topics. The first topic described the widely used techniques for evaluating a seismic performance of a structure, along with the best practices for seismic design and damage prediction in the building industry. Various approaches, prototypes and methods were examined. It has therefore been concluded that the contribution of current method and technology is very limited and inefficient. The literature review's second section emphasizes information/knowledge management in AEC domain and the advent of BIM and semantic web technologies, which were investigated for their collaboration and interoperability capabilities. Furthermore, this part investigated the possibility of combining BIM and ontology for building seismic performance assessment, which leads to formulating the second research question "How to establish a Building Seismic Performance Evaluation Framework (BSPEF) for the automation evaluation process with the combination of ontology and BIM?". The third part of the literature review explored the multi-objective optimisation algorithm with the review of different types of algorithms and the identification of limitations. Thus, a theoretical basis of design building seismic optimisation framework is formed. At last, the application of ANN in construction domain is reviewed, to explain why ANN is chosen as the method for predicting earthquake damage to regional structures.

The literature review support analysis to achieve research objective described in Section 1.5 by providing a comprehensive information. In this study, both quantitative and qualitative approaches can be stated to be followed for collecting the information and knowledge. In the construction domain, the existing documents of the standards and regulations are considered that were not originally set for automatic process of handling information and knowledge. Based on this analysis, a system for an BIM-based ontological seismic performance assessment and optimisation design for buildings and multi-scale regional seismic damage prediction has been established by researcher, to make transforming the regulatory rules into machine-readable standards become an automatic process, with using multi-objective optimisation algorithm and ANN. At last, the final test of the hypothesis will consist of two case studies: one focusing on the application of performance-based seismic design to individual structures, and the other focusing on the use of multi-scale regional building group earthquake damage prediction.

Therefore, this design-based study developed a new information/knowledge system via conducting a comprehensive examination of literature, analyzing the data, and interpreting the result. In this context, Figure 3.2 depicts the research methodology of the current study, which is primarily a combination of deductive and inductive. Therefore, because of the mixed nature, the research ustilises abductive approach.

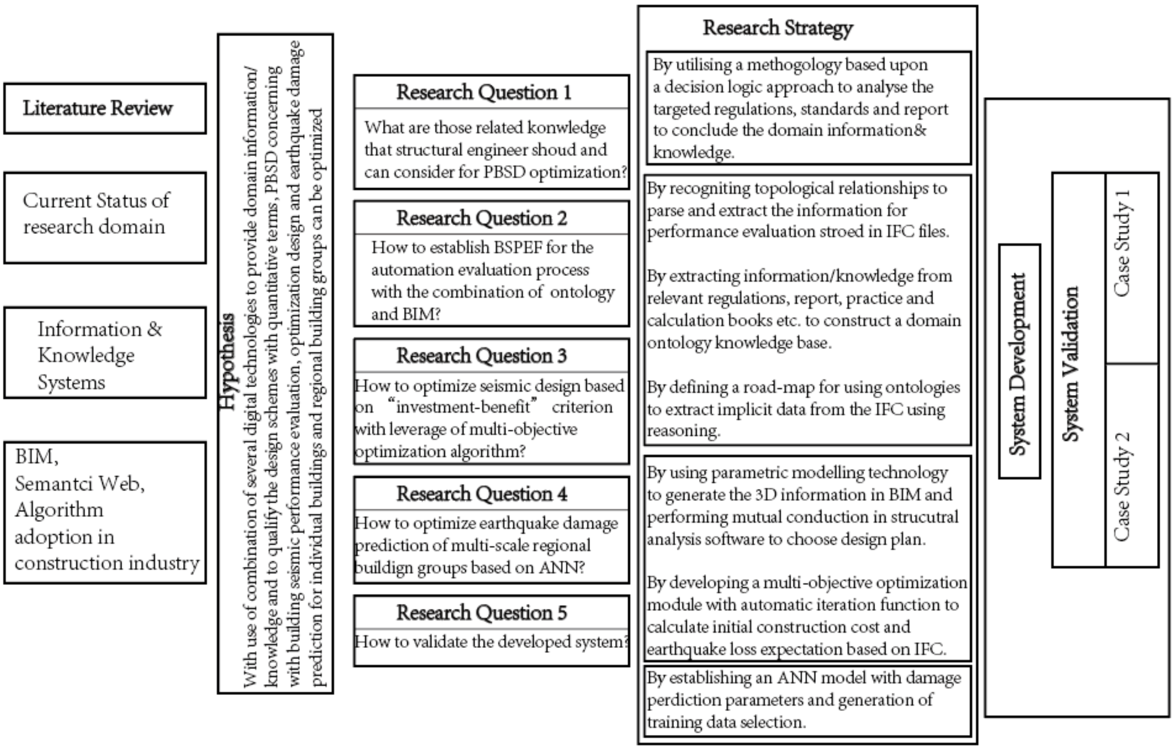


FIGURE 0.2 **THE RESEARCH APPROACH**

3.3 RESEARCH DESIGN AND METHODOLOGICAL CHOICE

This study splits into three phases due to the nature of research objectives (shown in Figure3.3). All these phases are designed to achieve the research objectives, as seen in Figure3.4.

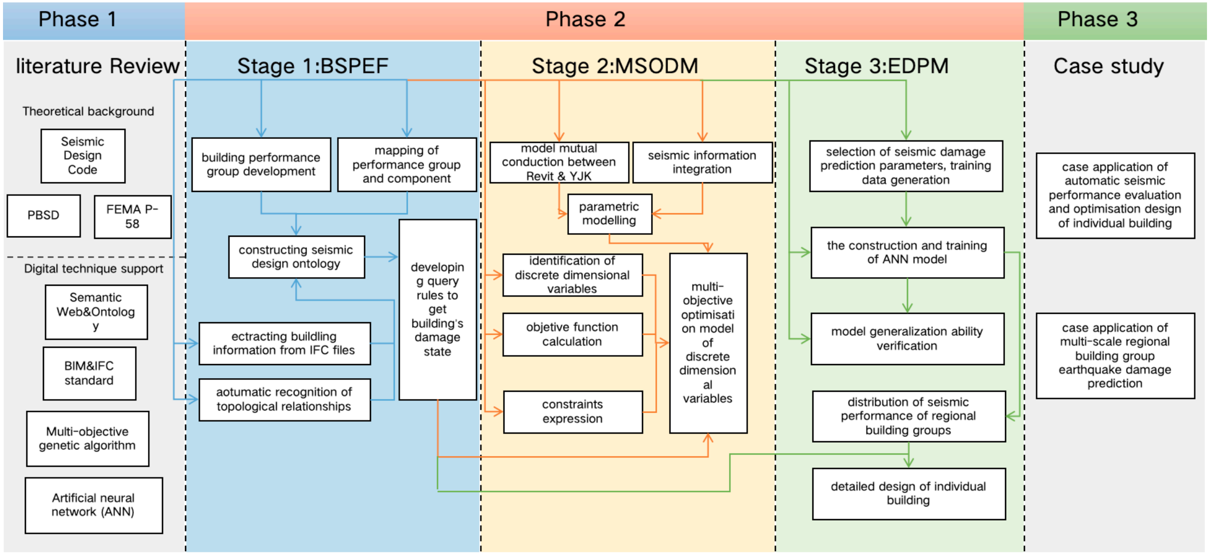


FIGURE 0.3 **THE THREE PHASES OF THE RESARCH**

			Objective 1:Identify domain knowledge, methodology and current practice of PBSD	Objective 2:Explore information technologies to provide the essential groundwork for the integration of information technology and engineering	Objective 3:A BSPEF would be proposed to realize the automation of building evaluation process	Objective :Establish a Multi-objective Seismic optimisation Design Method (MSODM) for RC frame according to the BSPEF	Objective 5:Establish an Earthquake Damage Prediction Method (EDPM) of multi-scale regional RC frame based on ANN and BSPEF	Objective 6:Validate the application effect of BSPEF, MSODM and EDPM
Phase 1		literature review of theoretical background: seismic design code, PBSD, FEMA P-58 etc.	✅					
		literature review of digital tehcnique support: semantic web, ontology, BIM, IFC standard, multi-objective genetic algorithm, ANN etc		✅				
Phase 2	Stage 1	ontology knowledge base development		✅	✅	✅	✅	
		semantic rule development		✅	✅	✅	✅	
		automatic knowledge extraction development			✅	✅	✅	
		information query model development			✅	✅	✅	
	Stage 2	BIM application module development				✅	✅	
		multi-objective optimisation development				✅	✅	
	Stage 3	selection of seismic damage prediction parameters					✅	
		the construction and training of ANN model					✅	
		model generalisation ability verification					✅	
		prediction outcome modelled on visual platform					✅	
Phase 3		Case study of automatic seismic evaluation and optimisation desing of individual building						✅
		Case Study of multi-scale regional building groups earthquake damage prediction						✅

FIGURE 0.4 **METHODOLOGICAL CHOICE FOR RESEARCH OBJECTIVES**

Generally speaking, research methods are quantitative, qualitative or mixed (Tashakkori & Teddie, 2003). For quantitative method, it tends to be linked with a positivist philosophy while for qualitative method, it is traditionally associated with interpretivism philosophy (Polit & Beck, 2010). In fact, there is no necessity to prohibit or prescribe the utilization of a specific method for research paradigms as the mixed methods usually could give researcher a better chance to make research design better constructed in most cases (Liebenau & Lee, 1997). For example, qualitative method can bring up the use of quantitative data to full strength in the case of evaluating human behaviours or attitudes. Both of them could be used in any research paradigm according to Lorleen and Farrugia's research (Lorleen & Farrugia, 2019; Sofaer, 2002). The method is chose based on the research objectives, resources, skills etc. Furthermore, researcher's thought exists not only in the collection and interpretation of the data analysis but also in the choice of the techniques used for collecting data (Kumar, 2014). Therefore, both methods are influenced by researchers' characteristics to some extent.

Different phases and the research method used of each phase are introduced as follows:

Phase 1: In the first phase, the study employs quantitative methods to serve a descriptive and exploratory objective. The literature review is a comprehensive survey and evaluation of the existing research and writings on a particular topic. It can help researchers and scholars identify what has already been studied, debated, and concluded regrading a particular area of interest and this understanding can further provide a foundation for further research. All the domain knowledge, methodology and current research status of PBSD and information technologies are explored to fulfill Objective 1 and Objective 2, which provide the theoretical background and digital technique support knowledge.

Phase 2: In the first phase, the study focuses on prototype development, which is a conventional method often applied in information systems. The suggested theory/notion often results in the creation of a prototype aimed at demonstrating the theoretical framework. The creation of a protype is a suitable assessment strategy throughout the first step of the application development, which seeks to demonstrate some intended functionalities of a system. Details of how these prototypes are developed step by step are introduced in Chapter 4, 5 and 6, which therefore to fulfill objective 3,4 and 5. Quantitative methods are used in Phase 2 as well. For example, during the process of constructing ontology knowledge base, qualitative method is leveraged as FEMA P-58 is looked through systematically to obtain in-depth knowledge about the various concepts regarding with seismic design. Moreover, it is applied during the process of parameterized modelling, multi-objective optimisation modelling and ANN modelling.

Phase 3: Two case studies are applied here to fulfill the Objective 6. Case study in this research to validate the methods developed using quantitative method as well as the justification of methods is relied on the comparison of actual value model outcome.

Consequently, the quantitative method seems to be the most suitable methodological option for the investigation.

3.4 RESEARCH STRATEGIES

The research strategy is referred to as “*a plan of action to achieve a goal*” (Benneworth, 2003). The selection of research strategies, same as that of research methods, is determined not only by research philosophy, method, and purpose but also by the current knowledge, time constraints and available resources as well (Dhanaraj & C., 2006). The most common research strategies are shown as below (Sun, Ju & Lu, 2016; Zhang, 2018; Hara, Kuroda & Nomaguchi, 2020):

- Experiment: Used primarily in scientific research to test causal relationships. It involves manipulating one variable to determine its effect on another variable.
- Survey: A strategy used to collect data from a large population using questionnaires or interviews. It is often used in social sciences.
- Case Study: An in-depth examination of a single instance or event. It is useful for gaining detailed insights and understanding complex issues in real-life contexts.
- Action Research: Involves solving a problem while simultaneously conducting research, which is iterative and collaborative and focuses on practical solutions.
- Grounded Theory: A strategy aimed at generating or discovering a theory through the collection and analysis of data, which is typically associated with qualitative research.
- Archival Research: Involves using existing documents and records as a data source for historical or longitudinal studies.

These strategies can be used individually or in combination, depending on the research objective and questions. The choice of research strategy should align with the research philosophy, approach, and methodological choice determined in the earlier layers of the research onion.

As for this research, it pertains to specific criteria and formal guidelines concerning seismic assessment and multi-objective optimisation design. Moreover, the comprehensive comprehension of research content could be constructed by answering research questions, which indicate the relationship between object domains. For example, “what is building seismic performance evaluation based on FEMA P-58?”, “what methods can be utilised to optimize building seismic design based on ‘return on investment’ criteria?” and “what methods can be leveraged for earthquake damage prediction?” etc. The fulfillment procedure could be specified by finding answers to questions such as “How can domain knowledge regarding to building seismic assessment/design/damage prediction be managed in Semantic Web environment?”, “How to build a parameterized BIM model to provide basic building information then to combine with ontology?”, “How to set a multi-objective optimisation model for building seismic optimisation design based on 'return on investment'?", and "How to develop ANN model for earthquake damage prediction?" etc.. The determination of technology can be specified by finding answers to queries such as "what software could be chosen for implementation?", and "How to choose different parameters?" etc. While the research question assumes that the information/knowledge of building seismic performance assessment/design/damage prediction can be managed using BIM, semantic web, optimisation algorithm and ANN technologies, this hypothesis may be substantiated by addressing questions "what are the existing deficiencies of this research topic?" and "how can use these emerging technologies to be utilised to narrow the gap?".

In order to get the in-depth knowledge in the domain to answer these questions and also consider the reusage of existing ontologies and the utilisation of the existing seismic wave records as the input parameters for earthquake damage prediction, the archival research strategy is therefore selected for the thesis at this stage.

The definition of "case study" may apply to both an analytical unit and a research technique (Yin, 2014). It refers to "*a scenario to which (researchers) have applied their proposed modelling technique, method or program*" as a unit of analysis. Additionally, it represents an comprehensive examination from various perspectives of uniqueness and complexity of a specific undertaking in a "real world" context as a research method (Thomas & G., 2011). It fits into both positivist and interpretivist philosophies, to deductive or inductive approaches and could also contain qualitative and quantitative methods, which is also one of the most used evaluation approaches of developing prototype system. It could contribute to a better comprehension of reality and facilitate the assessment of theories and facts (Aberdeen, 2013). It is also frequently advised to utilise actual cases or data from the real world to evaluate the prototype (Sommerville, 2010).

To be more specific, instrumental case study research strategy is applied for this research, which uses a particular case to gain insights into a broader issue or to refine theory (Li, Lan & Fan, 2019). Two case studies are selected for validation of the proposed methods. The first case study is to justify the feasibility of multi-objective optimisation design method by comparing the maintenance price before and after the optimisation based on a 6-storey RC frame building. The second case study is to test the generalisation ability of ANN model for earthquake damage prediction of building groups based on 30 RC frame buildings. As the development of these method (BSPEF, MSODM and EDPM) are not only developed for these specific 30 RC frames but want to serve as tools for better design of all RC frame buildings, instrumental case study suits the research strategy well.

3.5 RESEARCH TIME HORIZON

Cross-sectional study and longitudinal study are two different approaches of time horizon. Cross sectional study commonly includes the examination of a specific phenomenon at a specific time (Mnk Saunders, Lewis & Thornhill, 2011). While longitudinal studies involve the research covering two or more recurrent assessments of the same sample at different time points (Iacobucci & Churchill, 2010) and it could analyse processes that are ongoing or that undergo changes over a period of time (Chen & Hirschheim, 2010).

In the scope of the study, the collection of information/knowledge regarding building seismic performance evaluation/optimisation design/damage prediction is processed through different sample, the established knowledge base and developed prototype also took part at a particular time but not continuous or changing over time. In this regard, the nature this study is cross-sectional.

3.6 RESEARCH TECHNIQUES: DATA COLLECTION AND ANALYSIS

Data collection plays a crucial role in the research process. As stated in the methodological choice of the research, the source of data and information are from perspectives of both quantitative and qualitative approach. For this study, a plan of data collection was generalized.

Why were data collected

Part of Literature Review

As for domain knowledge relevant with building seismic performance evaluation, optimisation design and earthquake damage prediction, the reason are displayed as below:

- To obtain an overall comprehension of the context with regards to information/knowledge processed during the research procedures.
- To gain the understanding comprehensively not only evaluation criteria, but also the related data demanded to fulfill each criterion.
- To determine the various sources and different methods for gathering necessary data.

As for domain knowledge regarding ICT Technologies (BIM, Semantic Web, multi-objective algorithm, ANN etc.)

- To summarize the existing utilization of emerging technologies to help data interoperability, knowledge development, design optimisation etc.

- To study the possibility of applying these technologies to complement the requirements of building seismic assessment, design and damage prediction.

Part of Prototype Development

As for the BIM parametric modelling

- To generate building seismic design relevant information and provide structure seismic analysis result.

As for Multi-objective optimisation Modelling

- To obtain a balance between initial building costs and seismic anticipated losses

As for ANN Modelling

- To Choose the input parameter for earthquake damage prediction.

What kind of research strategy was adopted associated with data collection

- Analysis of textual files, national policies, and associated standards.
- Carry out methods of both qualitative and quantitative analysis on the required data.
- Transform the content of textual formats into digital formats.
- Put the transformed data into the BIM and Semantic Web Environment.

How were the data gathered and collected

- Based on the content summarized from policies and various documents, textual analysis should be conducted.
- According to the data sourced from internet and other published documents, data analysis should be executed.

This plan constitutes different questions such as “Why were data generated?”, “What kind of research strategy was adopted associated with data collection?”, “How were the data gathered/collected?”. In current research, the data are collected from the published public source such as guidance books or standards by USA, China and other international authority with regards to the target domain knowledge linked to building seismic performance. Moreover, different websites, database and some real building seismic design plan were sought not only to gain domain information/knowledge but also to provide the evidence for choosing parameters when modelling.

3.7 SUMMARY

Chapter 3 has established research methodology for the study, beginning with the underlying philosophy standing to the data gathering and analysis methodologies based on the theory "research onion". Each methodological choice was also explained depended on literature review and best practice of information systems. The theory basis of the prototype development could be provided by the exploratory study. Conversely, the outcome gained from developed prototype system will prove the findings through the critical analysis of exploratory study. Based on this methodology, research procedures of BIM-based ontological BSPEF, multi-objective building seismic optimisation design and ANN-based earthquake damage prediction is detailed in Chapter 4, Chapter 5 and Chapter 6 that follow respectively. Case study evaluation provide the research base for Chapter 7 to examine the prototypical systems respectively.

CHAPTER 4 BUILDING SEISMIC PERFORMANCE EVALUATION FRAMEWORK USING BIM AND ONTOLOGY

This Chapter introduces a framework for building seismic performance evaluation by using BIM and ontology, as depicted in Figure 4.1. The framework is composed of ontology building unit, knowledge extraction unit, semantic rule unit and information request unit, which is used to construct the knowledge base of BSPEF. In knowledge extraction unit, it systematically arranges the essential concepts of performance evaluation and their interrelations to articulate the assessment process and logic, while concurrently storing the diverse, multi-source information necessary for the assessment, encompassing fundamental building data at component level that are extracted from IFC files, structural analysis results, cost and maintenance estimates, and additional earthquake damage information. Based on the information obtained in knowledge extraction unit, four ontologies are developed to store required classes and properties with regards to seismic design in ontology building unit. Simultaneously, ontology alignment and rule reasoning functionalities are used to facilitate entity mapping and information interaction across ontologies, hence automating the assessment process in semantic rule unit. For example, building component are divided into different performance clusters and automatic mapping of performance cluster and component can be realised. Additionally, query rules are also developed for obtaining the structure damage state automatically in information request unit.

Section 4.1 describes how this knowledge base is built. The classification of ontologies is based on origin and kind of information, which includes sketch ontology, aseismic ontology, fragile ontology and application ontology. Section 4.2 states the manner in which ontology efficiently arrange the diverse multi-source data in the knowledge retrieval unit. The information is consistently saved in the format of RDF graphs. Furthermore, in Section 4.3, semantic rules supply the corresponding conceptual foundation for building a bidirectional connection between component and performance cluster. In Section 4.4, the information query module is developed to support semantic query of evaluation information.

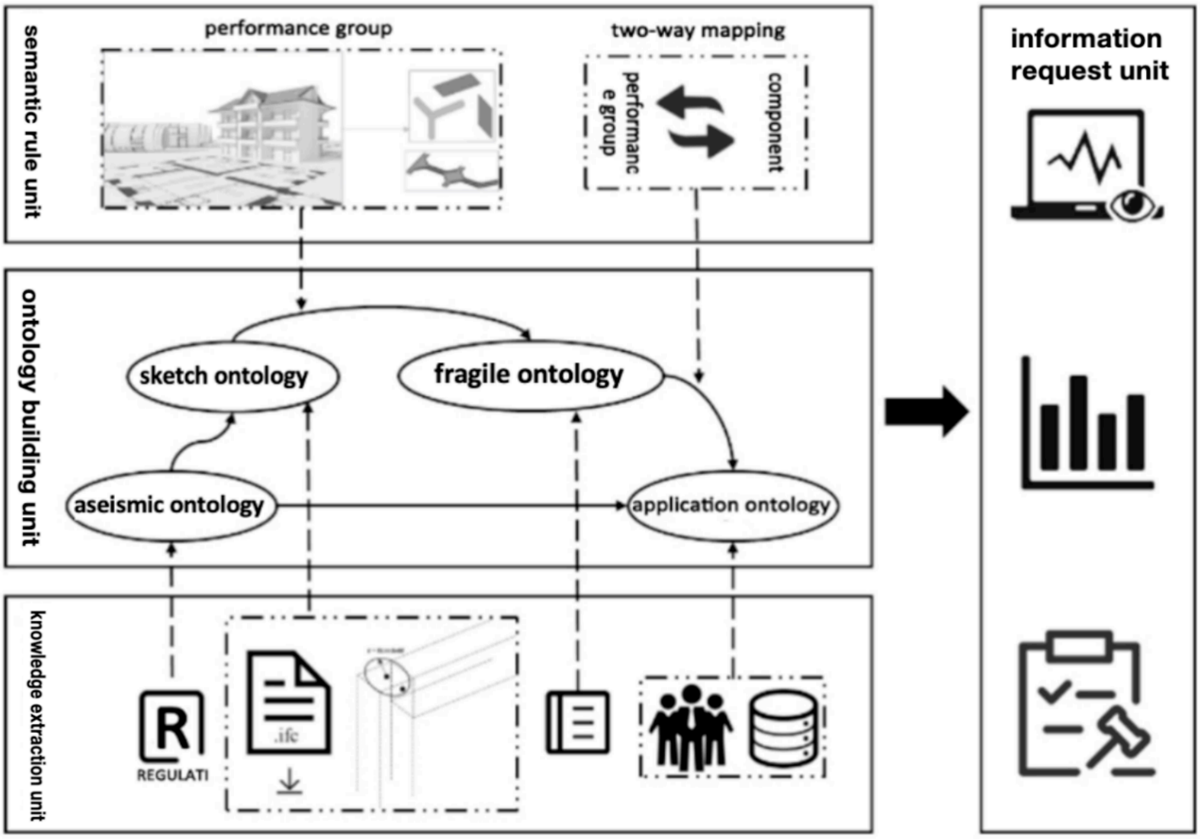


FIGURE 0.1 **FRAMEWORK OF STRUCTURAL SEISMIC PERFORMANCE ASSESSMENT**

4.1 ONTOLOGY BUILDING UNIT

4.1.1 ONTOLOGY DEVELOPMENT METHOD AND EDITING TOOL

Ontology's fundamental components include class, instances and relationships (Shadbolt & Alani, 2011). According to *Ontology construction methods* in Section 2.3.3.2, Ontology Development 101 method proposed by Noy and Mcguinness was selected to develop the ontology (Noy & Mcguinness, 2001). The basic steps are shown in Figure 4.2 (Breitman, Casanova, & Truszkowski, 2007). Protégé, OntoEdit, Ontolingua, OpenCye, WebOnto etc. are usually used as editing tools for ontology development. Protégé is a free and open-source Java-based platform with an intuitive graphical user interface (as shown in Figure 4.3). The input and output formats support ontology description languages such as RDF, UML, OWL, etc., which can be extended through multi-function plug-ins and Java-based APIs. Therefore, Protégé has been selected as ontology editing tool in this research.

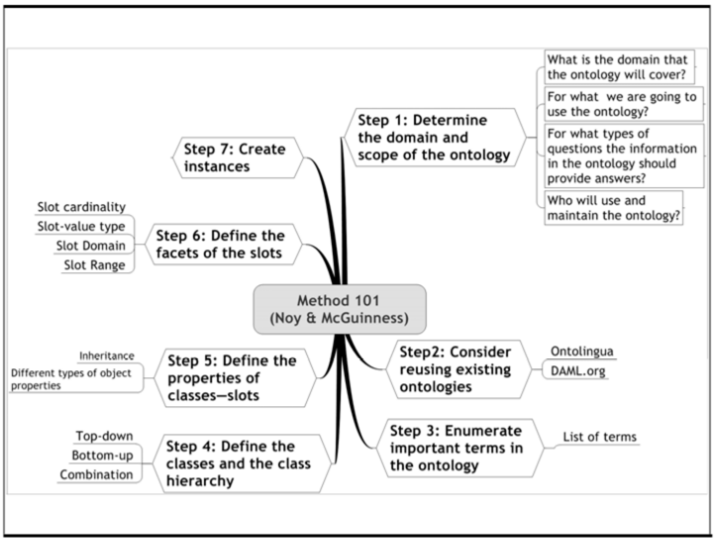


FIGURE 0.2 **ONTOLOGY DEVELOPMENT 101 METHODOLOGY**

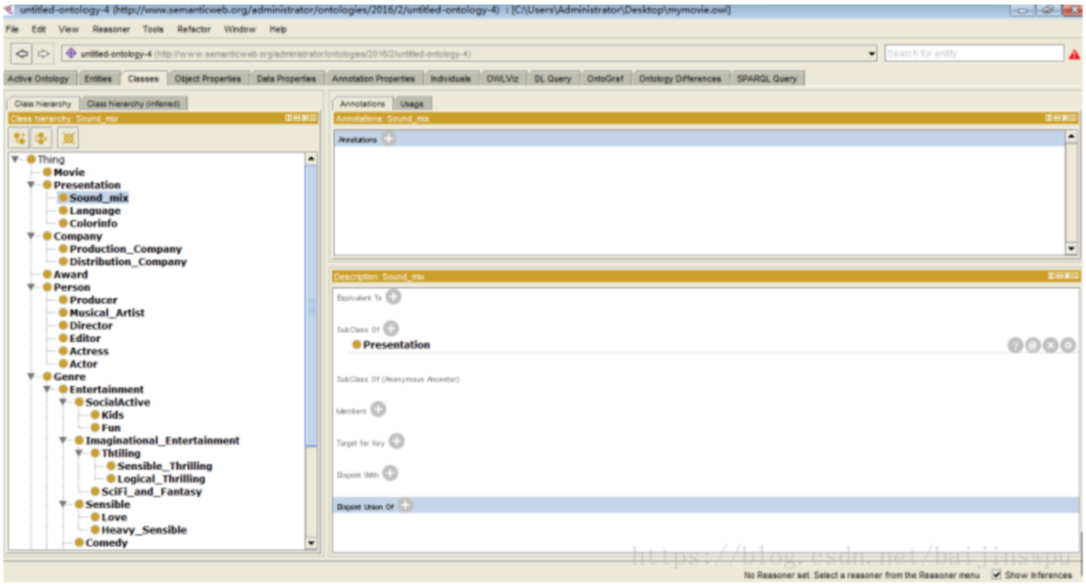


FIGURE 0.3 **PROTÉGÉ SOFTWARE INTERFACE**

For Ontology Development 101, seven major steps are stated as follows:

Step 1: Identify the range and extent of the ontology

Establishing a new ontology involves a procedure to create a knowledge framework of target domain with a specific purpose. Therefore, the first step involves establishing the scope of ontology. To clarify the range and extent of an ontology, a technique of ontology development called competency questions could be used. The ontology engineers can decide what are those concepts should be contained in the ontology by answering competency questions. Some examples of competency questions are given:

- What is the domain that will be defined by ontology?
- What is the application of the ontology?
- Who will be the end user of the ontology?

Moreover, the competency questions are also quite useful to evaluate the developed ontology at the final stage of development.

Step 2: Considering using existing ontologies

It is worth noting that if existing ontologies have already implemented some concepts of target ontology to be developed. Thus, it would be advisable to explore abstracting, extending, or using the complete current ontology to serve this ontology, since it might be a beneficial practice.

Step 3: Enumerate important terms in ontology

It is beneficial to compile terms we want to elucidate or make assertions about to a user. What terms do we want to discuss? What characteristics do the words possess? What observations do we want to make about these terms? For example, significant building-related terminology includes building function, building type, building scale, design information. It is essential to obtain a thorough list of words without concern for conceptual overlap, relationships between the terms and features of the terms.

Step 4: Define classes and class hierarchy

The terms determined in Step 3 should be arranged in a taxonomic hierarchy. Three approaches including the top-down approach, the bottom-up approach and the mixed approach are adopted for organising the class hierarchy (Mike, Uschold, Michael & Gruninger, 1996). The mixed approach is selected for this thesis. Some key prominent notions are delineated firstly and then specialize them as needed. A few top-level terms such as building and a few specific terms such as structural components are established at first. Then these terms can be connected to middle level terms, such as design information. When constructing the class hierarchy, one principle should be followed that the finished taxonomic hierarchy is consistent.

Step 5: Define class properties

In order to support necessary semantic of the ontology, properties should also be defined except the classes. Object property, data-type property and annotation property are three main property types. Object property specifies the connection between different classes; data-type property specifies the connection between data-type values and instances; annotation property offers information for comment of classes in ontology. It should be noting that subclasses inherit all the properties of its superclass.

Step 6: Define the facets of properties

Properties may possess many facets delineating the value type, permissible values, cardinality, and other characteristics of the values that the property many accommodate. For instance, the value of a name property (e.g., "the name of a building") is one string. Specifically, the name represents a property with a value type of String. Conversely, a property has (e.g., "building has design information") may yield different values, with these values being instances of the class "design information". Thus, "has" is a property with a value type instance with "design information" as allowed class.

Step 7: Create instances

Three main tasks are consisted, including the selection of a class, the creation of a specific instance of the class, and the assignment of property values. For example, a specific instance *column A* can be created to represent a specific type of Column. *Column A* is an instance to the Class *Column* representing all columns. It has the following property values defined:

- Building Floor: *the third floor*
- Size: *40mm*60mm*
- Fragile Cluster: *B14403.a*

4.1.2 ONTOLOGY DEVELOPMENT

As stated in Section 2.3, the ontology built in this thesis aims to cover performance evaluation content and express evaluation logic. According to FEMA P-58 and seismic code GB50011-2010, there are mainly three types of information for achieve seismic performance evaluation. The first type is all the relevant structural and non-structural components of a building for the categorization of fragile and performance cluster. Furthermore, it is utilised to characterise the structure data for assessing anticipated seismic damage. The second type of information characterises not only the earthquake itself but also its effects on the building. While the third type of information mainly describes the knowledge of fragile analysis, a crucial part of seismic performance evaluation. Besides, in order to execute the evaluation process, the evaluation process and its associated information should be described as well. Therefore, in order to facilitate the addition, deletion and updating of information and maintenance of the ontology, the sketch ontology, aseismic ontology, fragile ontology and application ontology are built according to the information source and type required for performance evaluation. Each ontology will be expressed in detail one by one.

Sketch ontology

The sketch ontology (Figure 4.4) specifies the structural data set essential for evaluation comprising basic building information and design information. The design information will be obtained from the IFC file from BIM Models and organized according to the structure of "building-floor-component". The dependency relationship between components and floors is defined as "*containedIn_story*" and the topological relationship defined as "*column_LinkedTo_Beam_x*" and "*column_LinkedTo_Beam_y*" between components in this study. Performance clusters are categorised based on floors according to definition. When the structural response is direction-sensitive, it is necessary to distinguish the building direction. While the performance cluster uses component nodes as the fundamental unit, the topological interaction will also be involved. Hence the recognition of topological relationship contains the direction of the component and the recognition of the adjacent relationship between component and component. Therefore, through the establishment of direct contact among building, floor and component, the necessary conditions for the subsequence automatic division of performance groups are provided. Essential structural information encompasses the building's nature, function, size and other relevant details that serve as the primary source of information on the building. Building components have data properties such as Globalld, name, volume, section height, etc.

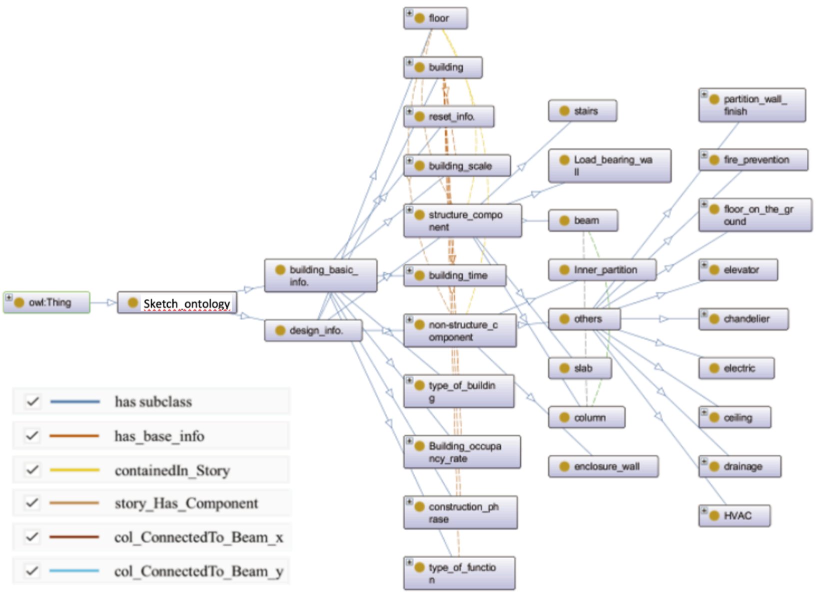


FIGURE 0.4 **CLASSES AND CONNECTIONS WITHIN SKETCH ONTOLOGY**

Aseismic ontology

The aseismic ontology (as shown in Figure 4.5) consists of structural seismic data and the outcomes of structural analysis. It aims to support as a foundation for earthquake hazard assessments and include the computed outcomes of the structural response analysis phase. Structural seismic information contains site category, design seismic groups, seismic precautionary intensity and other building-related seismic information specified by traditional codes (CSI, 2016), as well as performance targets that have been defined by owners in accordance with the specific engineering needs. The structural analysis outcomes are utilised to categorise and retain the seismic response necessary for predicting damage in performance cluster throughout every simulation operation. These outcomes may be retrieved from the calculation book provided by the structural analysis programme.

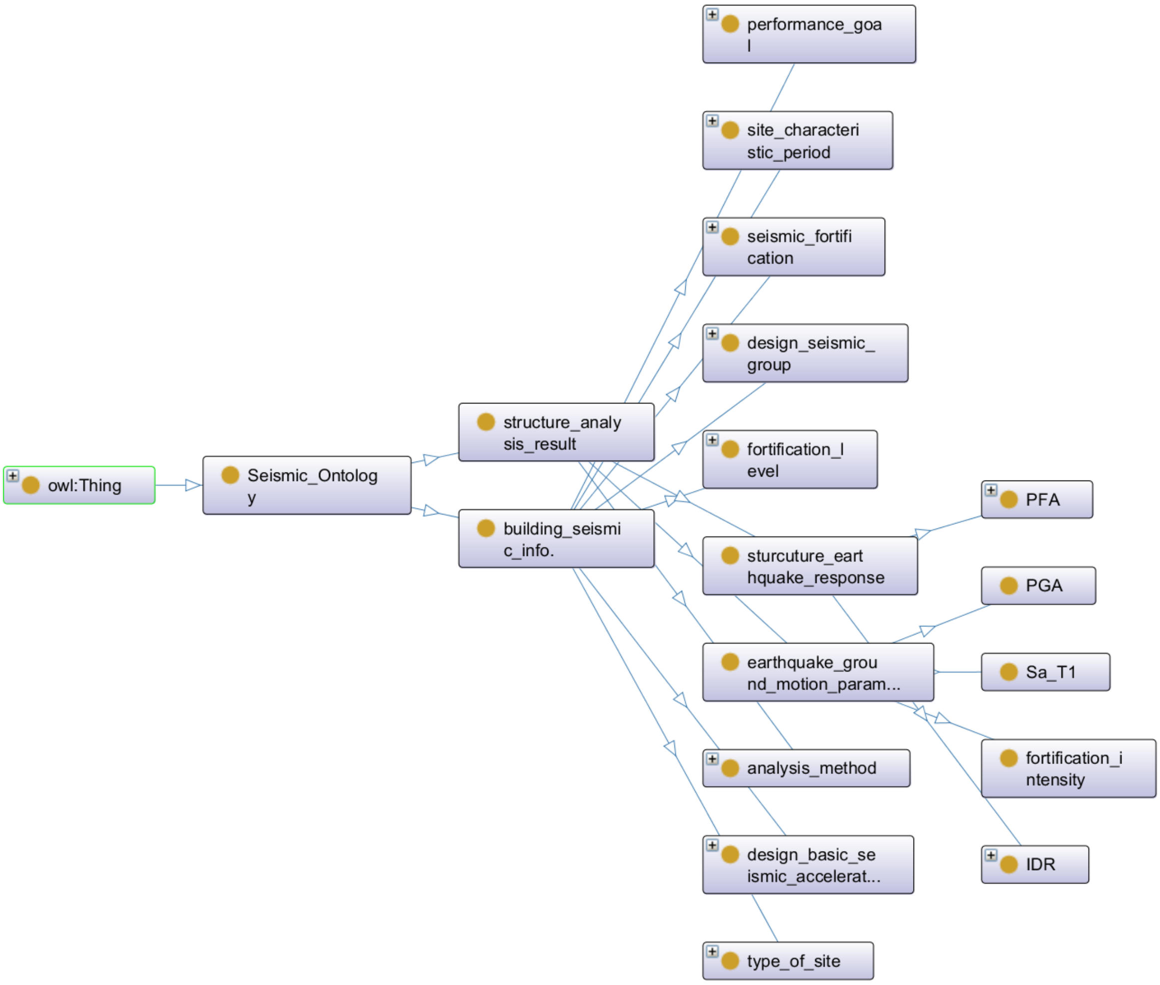


FIGURE 0.5 CLASSES AND CONNECTIONS WITHIN ASEISMIC ONTOLOGY

Fragile ontology

The fragile ontology (as shown in Figure 4.6) encompasses the fundamental assessment information included in FEMA P-58, which includes the two fundamental notions of fragile cluster and performance cluster. The fragile cluster encompasses the definition of fragile cluster category, fragile curve and consequence function. The fragile cluster types involved in common buildings can be listed under its major categories, for example, the “B1041.032a” fragility group belongs to the sub-categories under the “B-shell” major category; each fragility group category is connected with the corresponding fragility curve and consequence function through “*has_fragility_f*”, “*has_cconsequence_f*” relationship. The calibration parameter of each function is also stored in the ontology. The performance cluster, as a subtype of the fragile cluster, is further subdivided according to the performance cluster classification rules. Specifically, when the performance cluster category is direction-sensitive, that is, when the damage parameter is IDR, the fragile cluster will be categorised into distinct performance clusters based on the floor and building's orientation. Otherwise, when the damage parameter is PFA, the fragile cluster of the identical floor should be classified into the identical performance cluster category. Subsequently, the fragile curve in the fragile cluster may be used to determine the loss status and the chance of harm for the performance cluster unit. The desired consequence index which designers are interested in may be calculated by combining the consequence function with the quantity of performance cluster located on the identical floor.

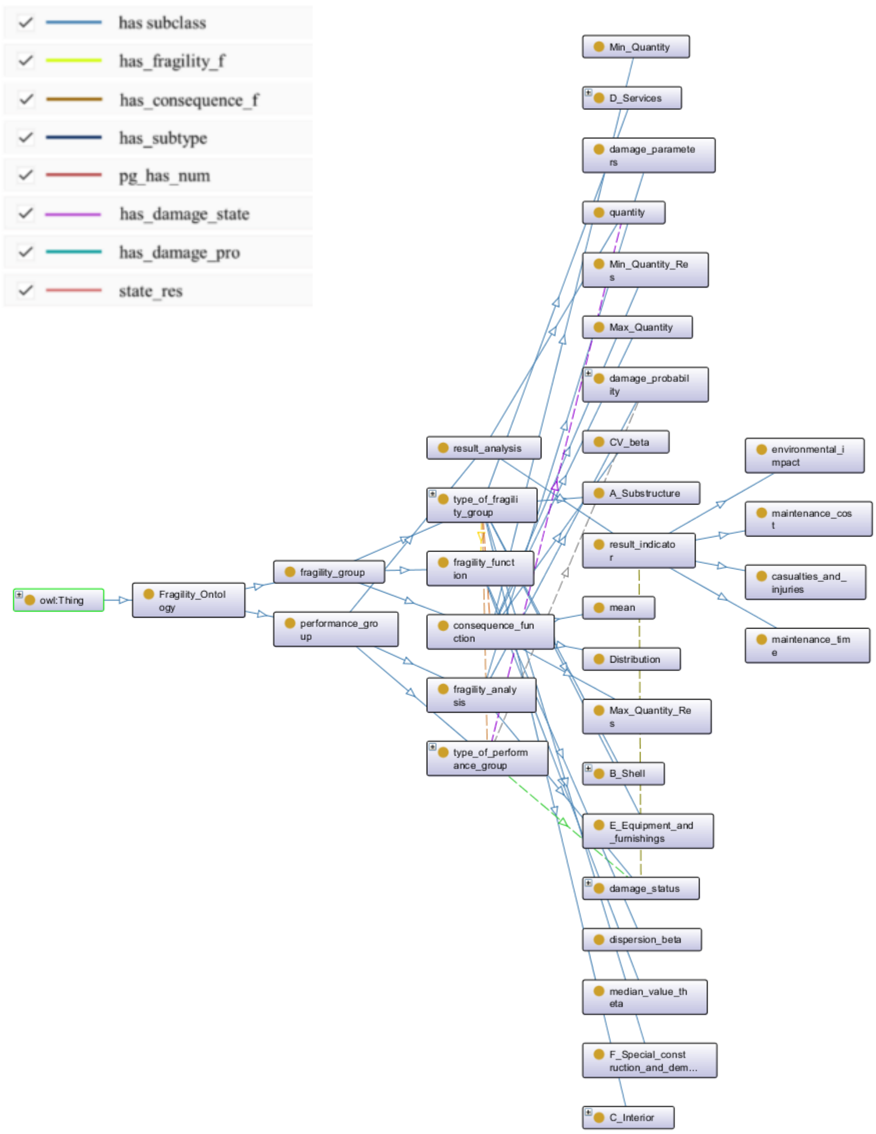


FIGURE 0.6 **CLASSES AND CONNECTION IN FRAGILE ONTOLOGY**

Application ontology

The application ontology (as shown in Figure 4.7) serves the function of elucidating the technique and objective of performance assessment. It integrates the evaluation situation in each simulation. Firstly, it explains the concepts of various evaluation criteria and classification of evaluation results to control the Monte Carlo simulation process. In each simulation, the result of mapping the damage state of the performance group to the component is represented by the fragility information in the damage model and is connected to the fragile component by the "*have_damage_result"* property. The damage state corresponds to the damage status and damage ratio corresponds to damage degree in Section 2.1.2. It is used to combine the damage consequence, the component volume, area and other property data to obtain the damage consequences.

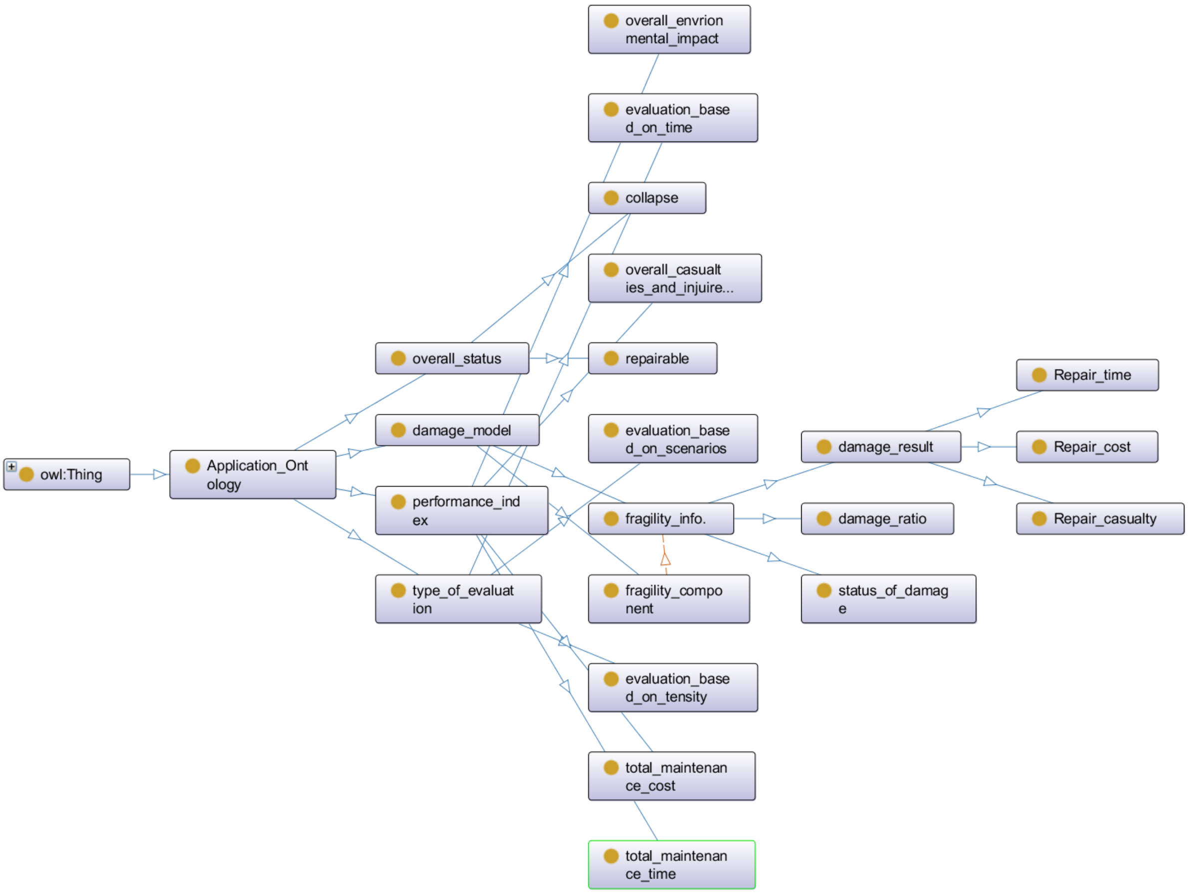


FIGURE 0.7 **CLASSES AND CONNECTIONS IN APPLICATION ONTOLOGY**

A certain relationship needs to be established among the above four ontologies as to facilitate information interaction throughout evaluation process and effectively convey the evaluation principle. As shown in Figure 4.8, the aseismic ontology and the sketch ontology establish a relationship between the building and the building seismic information through the seismic information source property "*has_seismic_info.*". As the foundation for forecasting the harm of the subsequent performance cluster, the seismic response of the building will be linked with floor, for example, the floor is related to the IDR through the property "*has_IDR*". The damage parameter based on the fragility analysis in fragile ontology are derived from the structural response analysis. Therefore, it is related to the seismic response of the structure through the property "*info._from*" in the aseismic ontology. What's more, the components in the sketch ontology are linked with the fragile cluster category in the fragile ontology through the property "*have_component*" and "*has_fragile_info.*", thus linked with the performance cluster category based on the floor information of the component. Finally, it will be included into the damage model which is consistently characterized by fragile component. The object property and data property in the ontology constructed in this study are shown in Figure 4.9.

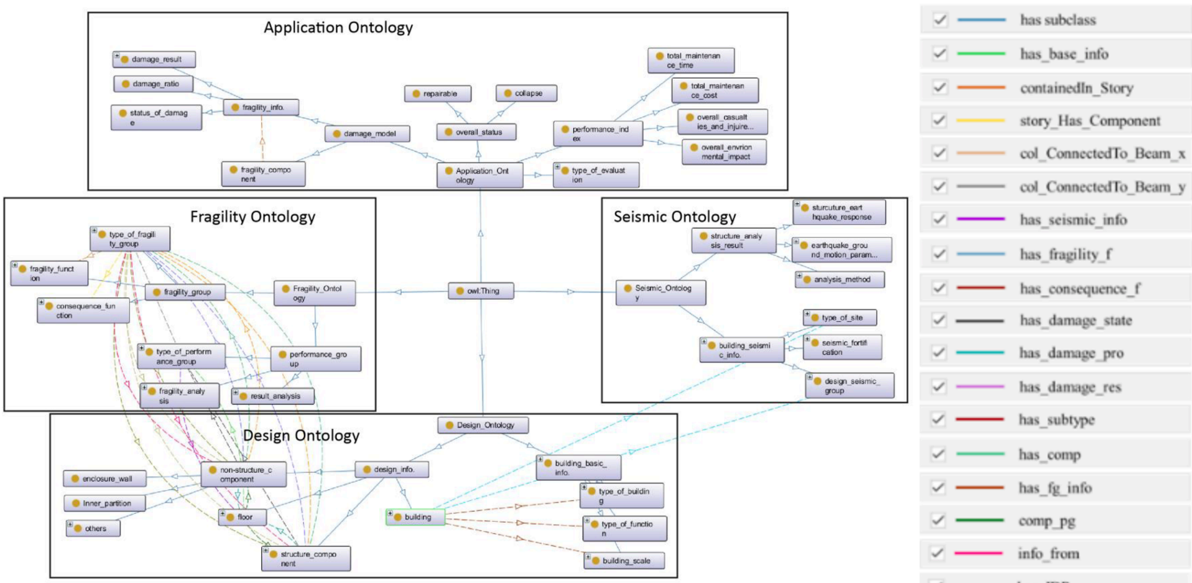
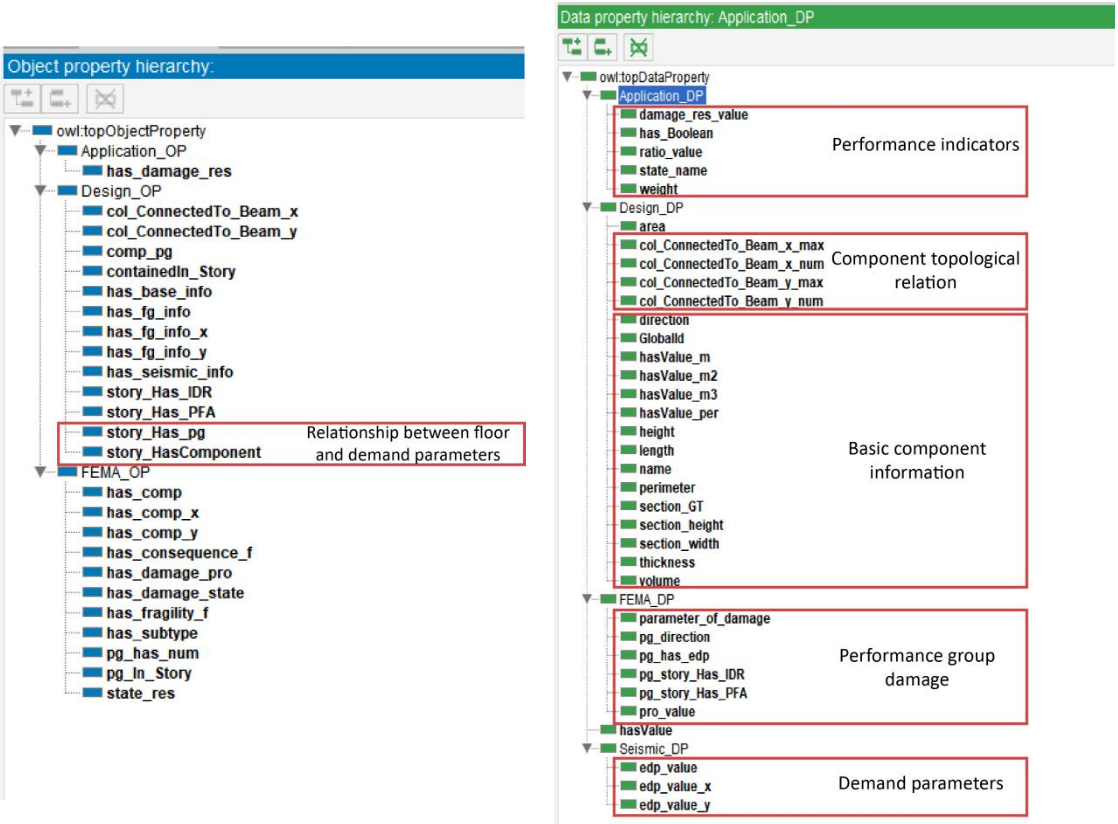


FIGURE 0.8 **PART OF CLASSES AND RELATIONSHIPS IN FOUR ONTOLOGIES**



a) Object properties b) Data properties

FIGURE 0.9 **OBJECT PROPERTIES AND DATA PROPERTIES**

4.2 KNOWLEDGE EXTRACTION UNIT

The knowledge extraction module aims to obtain the information required for performance evaluation in an orderly and efficient manner. The information obtained will be filled in the ontology which has been built in Section 4.1 in the form of ontology instantiation and also used as an information knowledge base in the following evaluation process. This section will explain in detail how to derive the necessary building model information from IFC physical files, BIM-based nonlinear time-history analysis, and the ontology instantiation method.

4.2.1 IFC FILE INFORMATION EXTRACTION AND TOPOLOGICAL RELATIONSHIP RECOGNITION

As the basic unit of performance evaluation, the following information of the component will be taken into consideration: The non-geometric properties such as material and material grade of the component; The geometric size information of the component including height, width, length, volume, area, etc.; The location information of the component, for example, the direction and the topological relationship between the components. The information will be used as the basis for performance group division and damage prediction.

The representation of information in IFC physical files

Consistent with the organization of actual engineering projects, the building model information is organized in the entity structure hierarchy of "project-site-building-floor-building component" in IFC files and two entities are connected by using relationship entity as an intermediate bridge. Specific to a certain type of building component, for example, the IFC expression of beam is expressed as "*ifcBeam"* and its information expression model in IFC is shown in Figure 4.10. To obtain the "*ifcBeam"* entity, the whole entities must be traversed from top to bottom. For example, after obtaining the "*ifcBuilding"* entity, all the "*ifcBuildingStory"* entities contained in the "*ifcBuilding"* entity can be obtained through the "*ifcRelAggregates"* relationship entity. Similarly, the location information of the component is obtained through relationship entity "*IfcRelContainedInSpatialStructure"*.

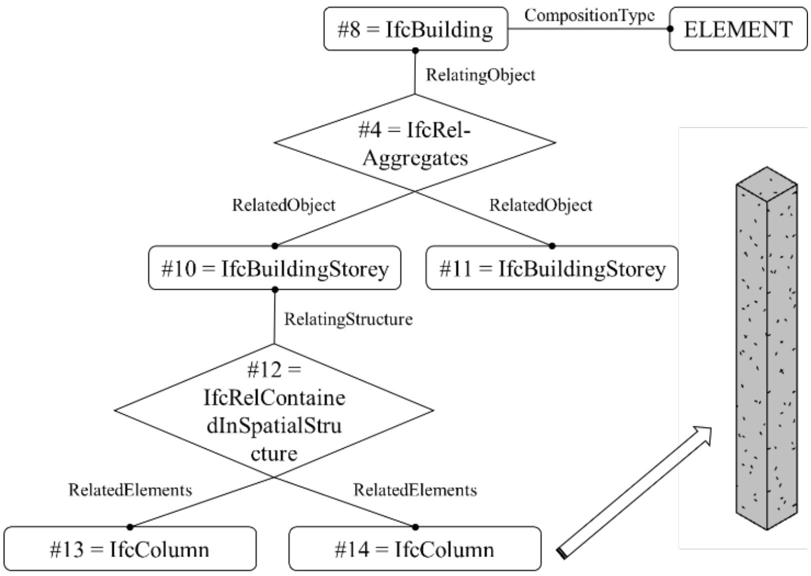


FIGURE 0.10 **INFORMATION EXPRESSION MODEL OF IFCBEAM IN IFC**

After obtaining a certain IFC entity, it is important to analyse the storage of its attribute information for extracting the IFC entity information. As shown in Figure 4.11, the attribute information of an entity is mainly stored in three locations in the IFC files. Taking "*ifcBeam"* as an example, this type of entity predefines attributes through EXPRESS language in the first place and it is directly stored in the IFC statement of the entity instance. As shown in Figure 4.12, in the brackets after the component instance statement (#4454), there are stored information which denoted as info_1 corresponding to the predefined attributes one-to-one, such as GlobalId (the value is "2GfDoqBgn3Yvzam4$_ALsP), Name (the value is "Concrete beam:300*700mm:318509").

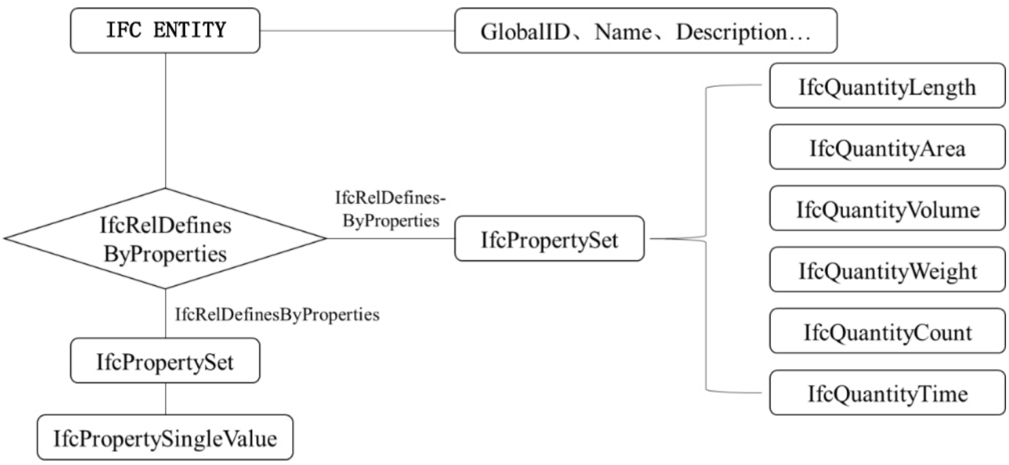


FIGURE 0.11 **STORAGE LOCATION OF ENTITY PROPERTIES EXPRESSED IN IFC**

```
#4454=IFCBEAM('2GfDoqBgn3Yvzam4$_ALsP',#41,' beam :300×700mm:318509',$,' beam
   :300×700mm:315093',#4430,#4452,'318509');
#71802=IFCRELCONTAINEDINSPATIALSTRUCTURE('15Z0v90RiHrPC20066FoKR',#41,$,$,(
#4454,…),#121);
#121=IFCBUILDINGSTOREY('2HskMqVqfAyOC6XNaHOYhO',#41,' 1F ',$,$,#120,$,' 1F ',
.ELEMENT.,3600.);
#90704=IFCRELDEFINESBYPROPERTIES('3y$4YfcH11V9tA55O$M_3w',#41,$,$,(#4454,…),
#4670);
#4670=IFCPROPERTYSET('2GfDoqBgn3Yvzanat_AKzX',#41,$,$,(#4624,#4625));
#4624=IFCPROPERTYSINGLEVALUE('b',$,IFCLENGTHMEASURE(300.),$);
#4625=IFCPROPERTYSINGLEVALUE('h',$,IFCLENGTHMEASURE(700.),$);
```

FIGURE 0.12 **PREDEFINED PROPERTIES**

In addition to the predefined attributes, users will add various additional information to the building components according to their own needs when establishing the BIM model. The three types of information required for performance evaluation discussed above are additional information, of which the expression in IFC is lengthy and complicated as shown in Figure 4.13. First of all, attributes "*relatedObjects*" and "*relatingPropertyDefinition*" are used to correspond with components and "*IfcPropertySet"* entity (#4670) through "*IfcRelDefinesByProperties"* relationship entity (#90704). This relationship entity is regarded as a property set storing a series of component dimensioning entity. Next, the "IfcpropertySingleValue" entity whose name attribute is "h" should be found for obtaining the height information of the beam entity. The "*nominalVaule"* attribute of "IfcpropertySingleValue" entity stores the length value. In addition, it can also be connected through the "*IfcPropertyByProperties*" relationship entity and stored in the *"IfcQuantityLength*", "*IfcQuantityArea*" and other entities according to different attribute types. The size information associated with the component through the *"IfcRelDefinesByProperties*" relationship entity is recorded as info_2. The location information of the component which denoted as info_3 is related to the floor instance (#121) and component through the relationship entity "*IfcRelContainedInSpatialStructure*" (#71802), as shown in figure 4-14. The name property of the floor instance shows the name of the floor where the component is located ("first floor").

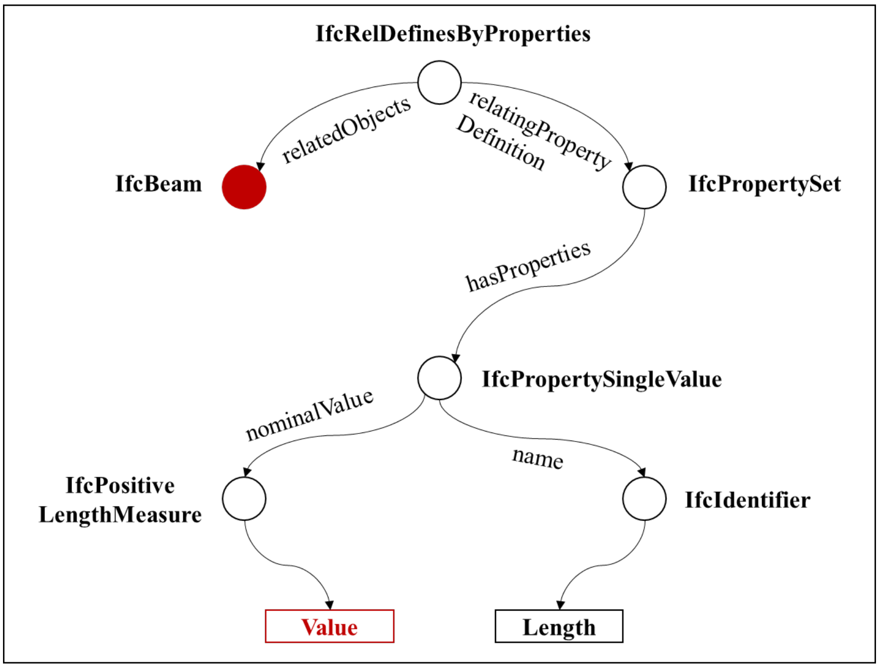


FIGURE 0.13 **RELATIONSHIPS BETWEEN BEAM AND ITS HEIGHT IN IFC**

IFC analysis

IfcOpenShell (2020), an open-source software library, provides a series of built-in functions to preprocess IFC files (explained in Section 2.3.2). As shown in Table 4.1, incoming IFC files are opened using *open* function, then put the class name information of a certain type of entity in the IFC into the *by_type* function, all instance statements of this type of entity can be obtained and stored in a set (entities). For example, for a beam entity, the value of *entity_type* is "*IfcBeam*". At last, all component instances are traversed and info_1, info_2 and info_3 are obtained according to the above analysis of the association components and component information.

TABLE 0.1 **IFC FILE INFORMATION PARSING CODE**

No.	Code	
1.	f = ifcopenshell.open(f)	# Open IFC file
2.	$entities$ = f.by_type(entity_type)	# obtain all sentences concluding certain type of entity
3.	for each $entity$ in $entities$:	# traverse all entities
4.	$info\_1$ = $entity$. __getattr__(Property_1)	# Property_1: GlobalId、 Name...
5.	$info\_3$ = $entity$.__getattr__('ContainedInStructure')[0].__ getattr__('RelatingS-tructure'). __getattr__('Name')	
6.	for $item$ in $entity$.__getattr__('IsDefinedBy'):	

```
if item.wrapped_data.is_a() == 'IfcRelDefinesByProperties':
    if item.__getattr__('RelatingPropertyDefinition').__getattr__('Name’) ==
                                                                'Size Dimension':
        info_2 = item.__getattr__('RelatingPropertyDefinition').__getattr-
                    __('HasProperties')
```

Info_3 is stored by the *“containedInStructure”* property of the instance, the “*RelatedStructure*” property of “*IfcRelContainedInSpatialStructure”* relation entity under the *“containedInStructure”* property points to the floor main body and the “*RelatedElements”* property points to a series of component instance collections located on the floor. The predefined properties of Name and Elevation are stored in the floor instance statement as the information of the floor where the component is located. Similarly, info_2 is stored through the “*IsDefineBy*” property of the instance, which integrates a series of “*IfcDelDefinesByProperties*” relationship entities to store various additional user-defined information. Each “*IfcRelDefinesByPRoperties”* specifies the information category through the Name property and the information category required for performance evaluation is “dimensions”. By judging whether the property value is “dimensioning”, obtaining the property set storing the size information. Then pointing to each “*IfcPropertySingleValue*” property entity through the “*HasPropertyies*” property. “*Name*” is the name of property and “*NominalValue*” presents the value of property.

Automatic recognition of topological relationships

Topological relationship recognition serves as the foundation for the automated categorization of performance clusters. Performance clusters should be categorised based on floors according to definition. When the structural response is direction-sensitive, it is necessary to distinguish the building direction (x/y); while the performance cluster uses component nodes as the fundamental unit, the topological interaction will also be involved, for example, since beams and columns in a reinforced concrete frame system, nodes are categorised into several performance cluster categories based on the dimension and quantity of the node beams. Hence, the recognition of topological relationship contains the recognition of the dependencies between the component and the floor, the recognition of the direction of the component and the recognition of the adjacent relationship between component and component. In order to divide all components on all floors into performance clusters, component should be traversed one by one. Therefore, this thesis proposes a set of automatic recognition algorithms for topological relations and the whole procedure has been seen in Figure 4.14.

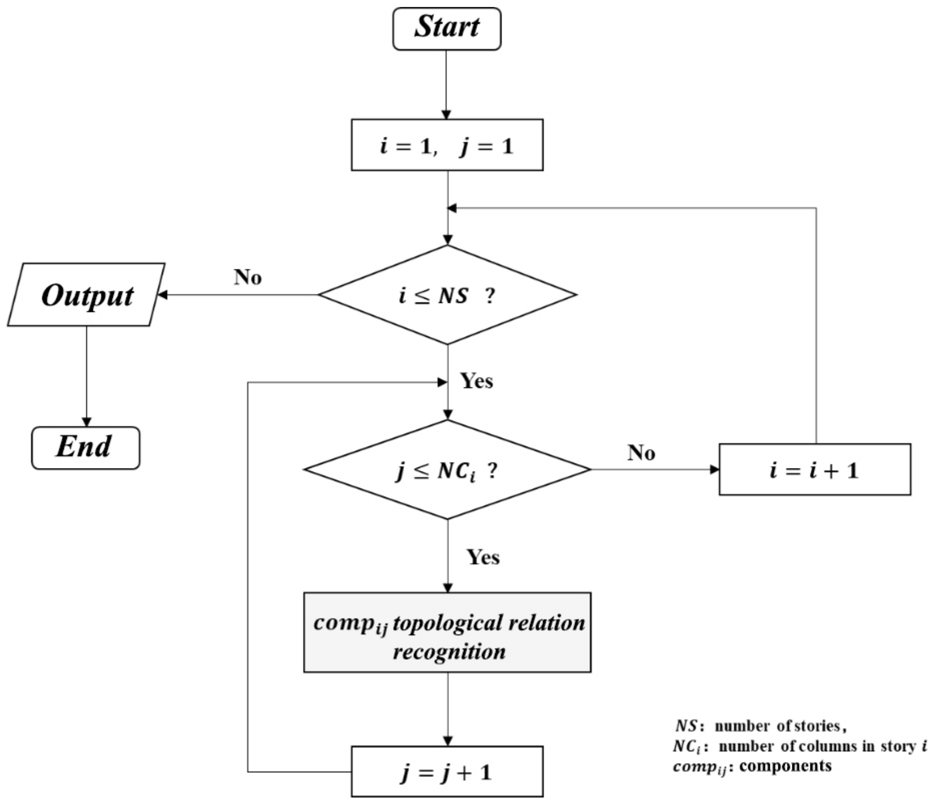


FIGURE 0.14 **FLOW DIAGRAM OF TOPOLOGICAL RELATIONSHIP RECOGNITION ALGORITHM**

Initially, the information of IFC files is extracted to obtain the dependency relationship between the component and the floor. The component is divided by floors and then the component j in each floor i is traversed to identify the topological relationship, which is the component direction recognition and the adjacent relationship recognition between the component and the component. The thesis uses IfcOpenShell and pythonOCC library to proceed in the following three steps (as shown in Figure 4.15) and the code is shown in Figure 4.16.

Step 1: Perform the "face-edge-point" operation on the component to get the end point of the component, denoted as P_k, k=1,2,3,4,5,6,7,8.

Step 2: Calculate the center points p_1 and p_2 of the cross section located at component's both extremities based on the end point information. Take any end point, use P1 as an example, then record the distance from it to other points as D_{1m},. Regarding standard components, if D_{1m} is sorted in descending order, D_{17}, D_{15} and D_{13} are in the first, fourth and fifth positions respectively. Correspondingly, the midpoints of P_1 and P_3, P_5 and P_7 are p_1 and p_2.

Step 3: The component's orientation is dictated by the comparison of the neutral axis vector p_1, p_2 formed by p_1 and p_2 with the building direction. For the recognition of the adjacent relationship of the components, a certain threshold is set. If the length of the end faces' midpoints of the two components is less than the threshold, it can be judged that they are adjacent.

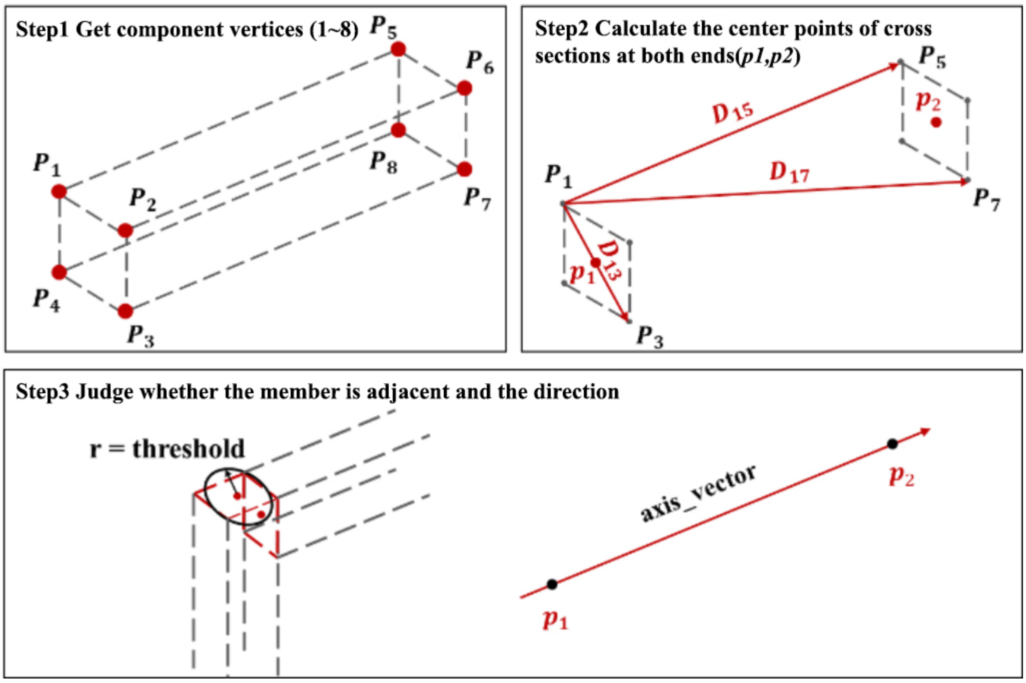


FIGURE 0.15 **COMPONENT DIRECTION AND TOPOLOGICAL RELATIONSHIP RECOGNITION**

```
def get_edges_from_shape(instance_shape):
    """
    Get edges based on faces
    """
    edge_explorer = TopExp_Explorer()
    edge_explorer.Init(instance_shape, TopAbs_EDGE)
    edges = []
    hashes = []
    while edge_explorer.More():
        current_edge = edge_explorer.Current()
        current_item_hash = current_edge.__hash__()
        if not current_item_hash in hashes:
            hashes.append(current_item_hash)
            edges.append(current_edge)
        edge_explorer.Next()
    return edges
```

a) Code snippet 1

```
def mesh_edge(edge):
    """
    Get points based on edges
    """
    edg = topods_Edge(edge)
    assert not edg.IsNull()
    curve_handle, U1, U2 = BRep_Tool_Curve(edg)
    curve = curve_handle.GetObject()
    if curve is None:
        return False
    IS_LINE = curve.IsInstance('Geom_Line')
    points = []
    if IS_LINE:
        p1 = curve.Value(U1)
        p2 = curve.Value(U2)
        points.append(p1.Coord())
        points.append(p2.Coord())
    else:
        U = U1
        nbp = 10
        dU = (U2 - U1)/nbp
        while U <= U2:
            p = curve.Value(U)
            points.append(p.Coord())
            U += dU
    return points
```

b) Code snippet 2

```
#
settings = ifcopenshell.geom.settings()
settings.set(settings.USE_PYTHON_OPENCASCADE,True)

#
instance_shape=ifcopenshell.geom.create_shape(settings,instance).geometry

#
edges=get_edges_from_shape(instance_shape)

#
point_0=(round(mesh_edge(edges[j])[0][0],6),round(mesh_edge(edges[j])[0][1],6),
round(mesh_edge(edges[j])[0][2],6))
```

c) Code snippet 3

FIGURE 0.16 **TOPOLOGICAL RELATIONSHIP RECOGNITION CODE**

In this context, the building model information required for performance evaluation and stored in the IFC file can be parsed and extracted. The automatic recognition of topological relationships based on geometric information also generates topological relationship information of components. This thesis will store this information uniformly in JavaScript data mark format such as JavaScript Object Notation and JSON (as shown in Figure 4.17), which is convenient for viewing information and instantiating ontology.


```
"333": {
    "GlobalId": "3PdLQN5GjCwfz7HiGF7eIf",
    "NameofColumn": "ConcreteColumn:600*600mm:328967",
    "TheFloorInfo.": [
        "TheFifthFloor"
        18000.0
    ],
    "Size": {
        "Volume": 1.296,
        "Length": 3600.0
        "Area": 4.68
    },
    "TopologicalRelationship": {
        "col_ConnetedTo_Beam_x": [
            "3PdLQN5GjCwfZ7HiGF7eWU",
            "3PdLQN5GjCwfZ7HiGF7eXq"
        ],
        "col_ConnetedTo_Beam_y": [
            "3PdLQN5GjCwfZ7HiGF7eZq",
            "3PdLQN5GjCwfZ7HiGF7eZ0"
        ],
        "col_ConnetedTo_Beam_x_num": 2,
        "col_ConnetedTo_Beam_y_num": 2,
        "col_ConnetedTo_Beam_x_max": 450.0,
        "col_ConnetedTo_Beam_y_max": 700.0
    }
}
```


FIGURE 0.17 **BUILDING MODEL INFORMATION STORED IN JSON FORMAT**

4.2.2 BIM-BASED NONLINEAR DYNAMIC TIME-HISTORY ANALYSIS

Nonlinear dynamic time-history analysis refers to the numerical analysis of a system's dynamic response to complex, time-and space-varying loads (Pan et al., 2011). This analysis usually involves material or geometric nonlinearities, such as plastic deformation, contact problems, and large displacements or rotations. Compared with linear analysis, nonlinear dynamic analysis can more accurately simulate real-world physical phenomena. For earthquake nonlinear dynamic analysis, seismic waves are used as input to perform dynamic analysis on the structure. IDR and PFA are two important representations of structural dynamic performance (Kappos & Eng, 2010).

A complete BIM model contains a variety of building information, which makes modeling be a time-consuming process. However, Nonlinear dynamic time-history analysis only requires for structural information. Therefore, BIM models are converted into structural analysis models to reduce the workload of modeling. The commonly used Chinese structural design programme YJK offers the REVIT-YJKS interface for converting Revit data. In addition, it is the officially designated software of Shanghai Urban Construction Design and Research Institute. Therefore, this thesis uses this interface to realize the direct generation of model of structure from the Revit programme. The process is shown in Figure 4.18.

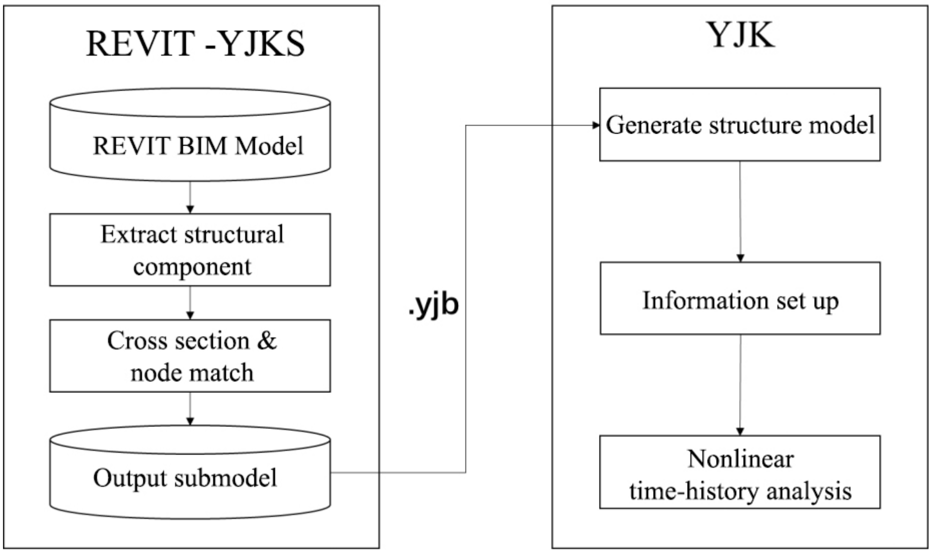


FIGURE 0.18 **FLOW OF YJK STRUCTURE MODELS GENERATED FROM REVIT MODELS**

REVIT-YJKS module in Revit is used to select the structural component that need to be converted into structural analysis model. Their section and nodes will match the pre-defined sub-models in YJK. As the name of family parameters in Revit may be different from the name of definition parameters in YJK, when sub-models are transformed and extracted, the transformation between the family parameters used by Revit components and definition parameters used by YJK should be carefully considered. The associated cross-sectional shape, geometric parameters, material types, etc. must be matched during transforming as shown in Figure 4.19. Then these sub-models are exported in .ydb file format and loaded into YJK for the purpose of generating models of structure. Finally, operations such as floor generation, load layout, parameter setting of nonlinear dynamic time-history analysis are performed in YJK programme.

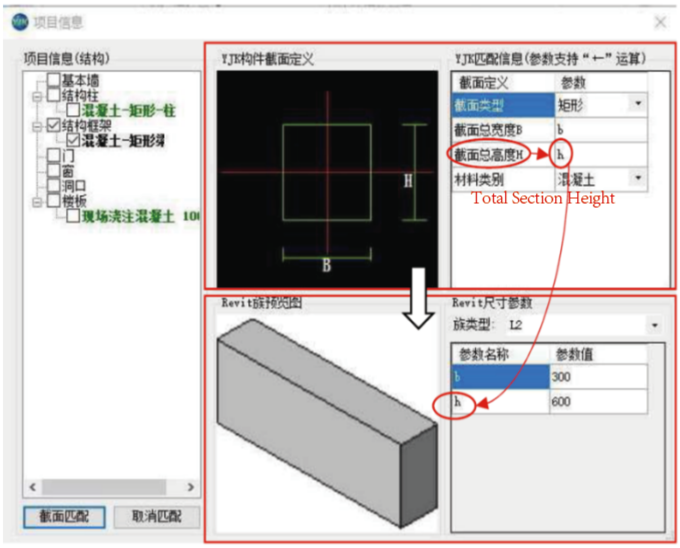


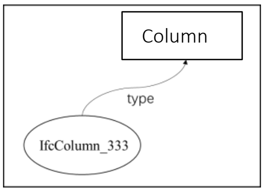
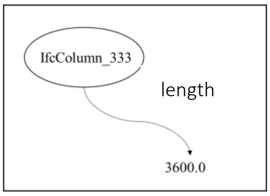
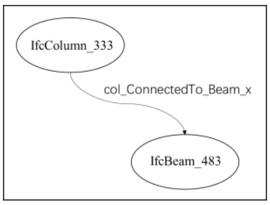
FIGURE 0.19 **REVIT-YJK PARAMETER CONVERSION**

The YJK structure seismic analysis yield reaction findings, namely the IDR and PFA, which are extracted from the calculation book. The results are then included into the structure seismic response concept via ontology instantiation. It can be associated with the floor as the foundation for predicting the harm of the follow-up performance cluster.

4.2.3 ONTOLOGY INSTANTIATION

Instantiation refers to the association of abstract concepts in an ontology with concrete entities in the real world. Instantiation allows for the creation of particular instances of concepts in an ontology. The RDF graph described by OWL is called an OWL ontology. As the RDF graph can be understood as a triplet with a URI identification of the node, the instantiation of the ontology can be processed by writing the RDF triplet. RDFLib is a Python library for processing RDF. It is widely used in the development and instantiation of ontology due to its clarity and easiness of comprehension. This thesis uses the RDFLib library to read information for importing into each ontology that has been built. The code and its description are shown in Table 4.2. Each element in the triple (s, p, o) is created or referenced in the ontology by declaring the corresponding URI and naming with the help of the built-in URIRef function in RDFLib library.

TABLE 0.2 **ONTOLOGY INSTANTIATION CODE**

Code	Explanation
g = rdflib.Graph()	# RDF graph initialization
g.add((s, p, o))	# Add RDF triplet
s = rdflib.URIRef(URIRef_s + individual_name) p = rdflib.URIRef(URIRef_p + 'type') o = rdflib.URIRef(URIRef_o + 'NamedIndividual')	# Instance Creation # URIRef_s: Instance Namespace prefix
s = rdflib.URIRef(URIRef_s + individual_name) p = rdflib.URIRef(URIRef_p + 'type') o = rdflib.URIRef(URIRef_o + class_name)	# Specify the type pf instance 
s = rdflib.URIRef(URIRef_s + individual_name) p = rdflib.URIRef(URIRef_p + data_property) o = value	# Create the data property of instance 
s = rdflib.URIRef(URIRef_s + stratnode) p = rdflib.URIRef(URIRef_p + object_property) o = rdflib.URIRef(URIRef_o + endnode)	# Create object property of instance # stratnode: name of instance # endnode: name of instance 

4.3 SEMANTIC RULE UNIT

When multi-source heterogeneous data is effectively organized and stored uniformly in the form of simple and intuitive RDF graphs, the association and interoperability of information are mostly shown in two aspects. On the one hand, there is the need for information interaction within or between ontologies. As described in the section of ontology development, the damage parameter in the fragile ontology is derived from the seismic reaction of the structure in the aseismic ontology, hence the instance information will be shared between these two different concepts. On the other hand, the two-way mapping between performance clusters and components, as the core of the entire performance evaluation, is achieved through mutual reference and association between floor information, component information, fragile cluster information and performance cluster information. The function of information interaction in the format of logic or established principles is named as ontology reasoning. In this study, the semantic rule module develops and applies reasoning rules.

Generally speaking, reasoning rules are described with the help of an ontology-based rule language. SWRL rules is a rule language that allows users to modify according to their own needs and can work with different rule engines to meet the different needs of users. SWRL rules consist of an *assumption (body)* and an *inference (head)* which is generally represented as "*assumption->inference*". Both the assumption and the inference are composed of many components represented by symbol "^" and each component is stated in the format of p ($term_1$, $term_2$, …, $term_n$) where p can represent the ontology class, property or SWRL intrinsic function, term may refer to a variable, example or numeric information.

Semantic rules are mainly used for performance cluster discrimination and performance cluster mapping (Gauchi, Bensadoun, Colas & Colbach, 2017). The performance discrimination consists of two stages. In the initial phase, the fragile cluster is categorised. In the next phase, on the basis of the fragile cluster identification result, the performance cluster is categorised according to the floor and orientation data of the fragile cluster. The categorisation standards for fragile clusters may be classified into two categories: the first category requires the integration of fundamental information about structure's component. In this case, it is necessary to merge the second category with recognition outcomes of the component's topological relationship. For example, the weakened flange joint of the beam with welded web in the steel frame is used as the basic unit of the performance cluster and is classified as two distinct types "B1035.0-0.1" and "B1035.011" based on the quantity of beams connected to the joint.

Table 4.3 displays the overarching semantic principles developed in this study which are applicable to the above two types of classification. The term "*Compnent_Class" in the context of fragile cluster classification rules refers to the many classes of components such as beams, columns, slabs, and others*. For the instance *"?comp"* in the component class, its corresponding property value such as fundamental data and topological connections are acquired through the relationship "*have_properties*", "*have_value*". Next, the property value is compared with the limit value of the fragile cluster category using the SWRL built-in function "*swrlb:lessThan"* to divide it into different fragility groups. For component instance in different fragile cluster category, the floor and orientation information where they are located are acquired through "*contained_In_Story"* and "*orientation_info."* in the performance cluster division rules. When they are matched with performance cluster category in the identical floor and direction, the relationship between them is developed through property "*have_pc_info*" and "*pc_have_comp*" to realise the discrimination of component performance clusters ultimately.

TABLE 0.3 **GUIDELINES FOR SEPARATING INTO FRAGILE AND PERFORMANCE CLUSTERS**

Type of classification	SWRL rules
Fragile cluster classification	Component_Class(?comp) ^ has_property(?comp, ?property) ^ hasValue(?property, ?value) ^ swrlb:lessThan(?value, limit_value) ^ Fragile Cluster(?fc) ^ -> has_fc_info(?comp, ?fc) ^ has_comp(?fc, ?comp)
Performance cluster classification	type of fragile cluster(?fc) ^ has_comp(?fg, ?comp)^ containedIn_Story(?comp,?story) ^ orientation_info(?comp, direction) ^ type of performance cluster(?pc) ^ pc_story(?pc, ?story) ^ pc_orientation(?pc, orientation) -> has_pc_info (?comp, ?pc) ^ pc_has_comp(?pc, ?comp)

Performance cluster mapping involves transferring the mapping the fragility analysis results for the performance cluster at the component level. The corresponding SWRL rules can be found in Table 4.4. The *assumption(body)* obtains the damage state *"?damage_state"* and damage probability "*damage_pro*" of the performance group category instance *"?pg"* by comparing the floor information and orientation information of the component. The inference executes the associations between the component instances and the above two types of information through property "*have_damage_res*" to realise the correlation of the harm condition and harm probability of each fragility component and its performance cluster. Then the ultimate damage condition and damage ratio are computed according to the component-level damage prediction formula proposed in Section 2.2 of this study.

TABLE 0.4 **RULES FOR PERFORMANCE CLUSTER MAPPING**

The type of classification	SWRLrules
Performance cluster mapping	type of performance cluster(?pc) ^ pc_story(?pc, ?story) ^ pc_orientation(?pc, orientation) ^ has_damage_state(?pc, ?state) ^ has_damage_pro(?state, ?pro) ^ pc_has_comp(?pg, ?comp) ^ containedIn_Story(?comp, ?story) ^ comp_direction(?comp, direction) -> fragility component(?comp) ^ has_damage_res(?comp, ?state) ^ has_damage_res (?comp, ?pro)

This research chooses SPARQL language to construct information request unit. SPARQL is a query language specially created for RDF. It allows for easy retrieval of information from an RDF graph by selecting and extracting the relevant knowledge. SPARQL statements usually consist of a SELECT part and a WHERE part. The SELECT part clarifies the information variable that needs to be queried, which is represented by "?". The variable range is limited in the form of triples in the WHERE part. The querier binds the variables of the matched triple set to the corresponding part of each triple and finally returns the result.

According to the performance cluster classification principles in Section 4.3, the performance clusters and floors have been directly related through the "*pc_story*" property. Meanwhile, a correlation has been established between seismic reaction of the building and the floor. The fragility curve, which determines the point at which the performance cluster fails, is linked to the fragile cluster category as well. Thus, the performance cluster instance *"?pc*" of all categories *"?fc"* on a specific floor *"?story*" and a specific orientation "*?orientation*" can be obtained by the query statements shown in Table 4.5. The IDR *"?IDR*" the median value of the fragility curve *"?theta*" and the discrete value of the fragility curve *"?beta*" which are damage data related the performance group unit can be obtained in real time. When the ultimate evaluation outcome is semantically associated with the fragility component, this simple and easy-to-write query method can also be quickly retrieved.

TABLE 0.5 **ILLUSTRATION OF RETRIEVING COMPONENT DAMAGE DATA USING SPARQL**

Query Content	SPARQL SELECT rules
Damage information of component	SELECT ?story ?orientation ?fc ?pc ?IDR ?theta ?beta WHERE { ?fc has_subtype ?pc. ?pc pc_story ?story. ?pc pc_orientation?orientation. ?story has_IDR ?idr. ?idr hasValue ?IDR. ?IDR orientation ?orientation. ?fc has_fragility_DS1_theta ?theta. ?fc has_fragility_DS1_beta ?beta. }

4.5 SUMMARY

This chapter proposes a BSPEF based on BIM and ontologies, as outlined in the P58 performance assessment material. The core aspect is the development and reasoning of ontology. The development of ontology realises the expression of the evaluation content by organizing the key concepts in the seismic performance assessment process of the building and the relationship between the concepts. It is suitable for the prediction of various performance indicators of the building under any evaluation type. It is versatile and the knowledge is easy to modify and expand. By preprocessing the IFC files, the basic building information from different BIM software can be obtained and topological relationship of the components can be automatically identified as a necessary condition for the automatic division of performance groups. The transfer of models between BIM software and structural analysis software may enhance the efficiency and quality of structural response analysis. The structural response information and other assessment information such as damage data in FEMA P-58 are stored in the same ontology semantic environment to realise the association and interoperability of knowledge. Moreover, it uses simple and comprehensible expression to facilitate the understanding of the assessment process by non-professionals. Ontology reasoning uses SWRL rules to realize the automatic two-way mapping between performance groups and components. It can also realize the rapid development of performance models, improves the evaluation efficiency and strengths the evaluation logic. Besides, the rules are easy to be managed, maintained and can be synchronized with the fragility group regulations for additions, deletions and modifications at the same time. Finally, using SPARQL language for semantic query, retrieval requirements can be expresses intuitively and more detailed component-level damage information can be obtained efficiently.

The research findings in this chapter can realise the component-level damage prediction of a single building under a certain earthquake level, that is, the final outcome is expressed as the damage distribution with the component as the basic unit. According to FEMA P58, the expected earthquake loss can be calculated based on the damage status of components. Additionally, the expected earthquake loss is one of the indicators for the structural seismic optimal design. Therefore, it can serve as the research foundation for the subsequent chapter’s optimisation of the structural seismic design at the component level.

The multi-objective seismic optimisation design method is proposed in this chapter is displayed in Figure 5.1. After the stakeholders establish the performance goals that the building must achieve, the discrete values are organised into many sets of design plan sets of the first strategy according to the design variable range given by the designer, which are inputted into downstream unit from customer unit. Subsequently, the BIM application unit utilises parametric modelling method to create a 3D information model and perform mutual conduction using structural analysis programme to determine if individual design plan fits the predetermined criteria. If the predetermined criteria are met, the IFC file including comprehensive model data will be generated; otherwise, the design plan will be sent back to customer unit, removed from the plan set and reselected. Once a plan from the plan set meets the design requirements and the quantity of plans surpasses the quantity of iteration populations, the automated iteration function in the optimisation unit will compute initial building expenses and anticipated seismic loss by using IFC files. Next, NSGA-II algorithm will be utilised to assign weights to conflicting aims. During the whole iterative process, the role of two upstream units enables the production of design plan modelling and the assessment of performance specifications.

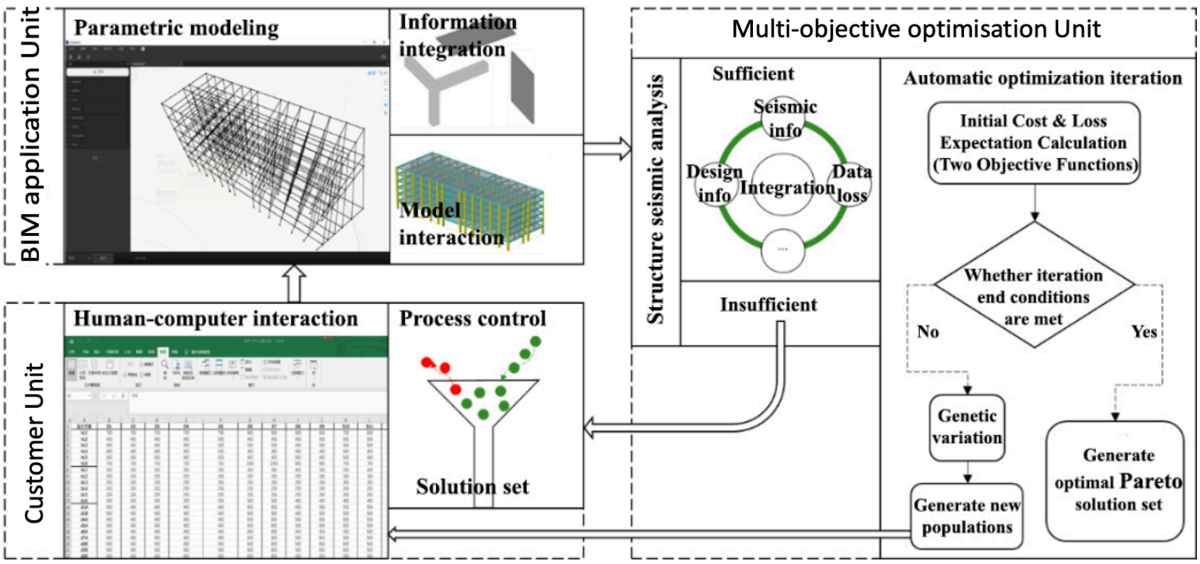


FIGURE 0.1 **MULTI-OBJECTIVE SEISMIC OPTIMISATION DESIGN TECHNIQUE USING BIM FOR RC FRAME**

The core of this method lies in a collection of multi-objective optimisation models with discrete size factors, whose purpose is achieving a harmonious equilibrium between the initial building expenses and seismic loss expectations. While the BIM application unit seeks to make the whole optimisation design process using BIM to become more efficient.

5.1 BIM APPLICATION UNIT

The BIM application unit is based on the BIM platform and provides three major functions during the optimisation process: parametric modeling, information integration and model interaction. Information integration allows for obtaining component data using the IFC files to enable automatic computation of the initial building expense and seismic loss expectation, which has been already explained in detail in Chapter 4. For model interaction, the interface facilitates the interaction between Revit programme and YJK programme in the part of model mutual interaction of Chapter 4 as well. Structure analysis information can be imported to Revit with a single click. This section will focus on the parametric modelling process using BIM.

Throughout the optimisation iteration process, a multitude of diverse design schemes will be produced. While the rapid generation of building models, as a prerequisite for objective function calculation and seismic analysis, significantly influences optimisation efficiency. Dynamo, as a built-in plug-in of the Revit software, can be used for visual programming modeling (as shown in Figure 5.2). It utilises the robust internal node library to provide a range of intricate geometric operations and interactive functions that allow integration with other data application tools like Excel and Zoho. In Revit, the process of creating a model involves encapsulating the codes that realize different functionalities and representing them as nodes that are clearly comprehensible. In addition, users can also use the Python node to call Application Programming Interface (API) for customized nodes and secondary development.

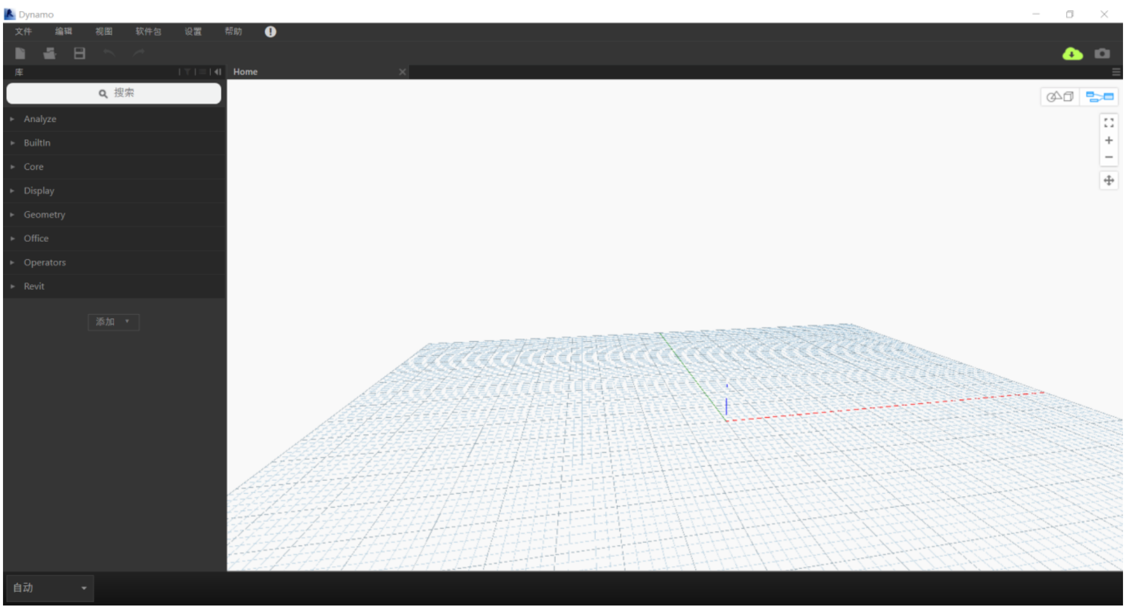


FIGURE 0.2 **VISUAL WORKING INTERFACE IN DYNAMO**

The chapter will use the dynamo for quick parametric modelling for individual design plan. Take the design scheme in Figure 5.3 as an example, the process from modeling sequence of “grid positioning point – component positioning line – component” to realize the generation of $column_{1A}$, $column_{2A}$ and $beam_{12A}$ in the design scheme is shown.

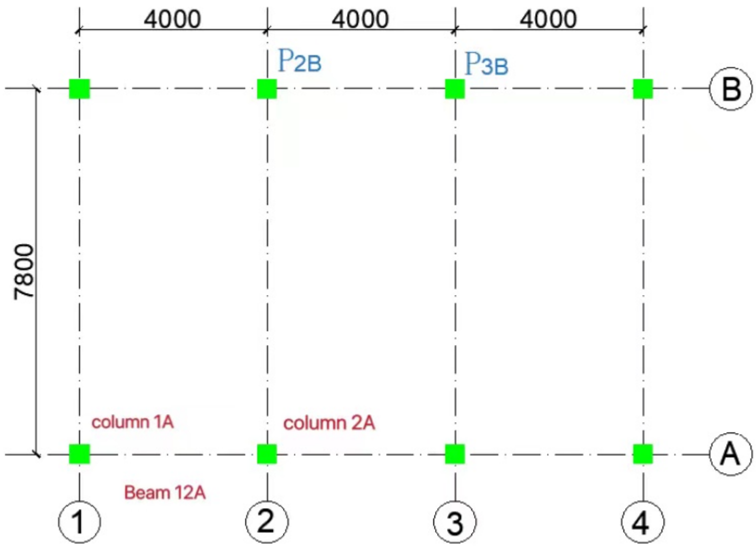


FIGURE 0.3 **AN EXAMPLE DIAGRAM OF A DESIGN CASE**

Step1: Grid positioning point generation

Coordinate information of the intersection of each grid was entered in the code block and connect to the Python script node, then the information flow will be used as input variables IN [0], IN [1], IN [2]. Figure 5.4 shows the detailed content of the Python script node. By receiving the input from the upstream node and calls the *Point.ByCoordinates* node from the Dynamo Geometry node library at the same time to realize the function of “using the given 3 Cartesian coordinates to form a point”. Finally, the point set is output as an OUT variable. After the above information processing flow runs, the grid positioning points on the right part of Figure 5. will be generated automatically, which is the necessary geometric elements for the generation of component positioning line.

```python
import clr

clr.AddReference('ProtoGeometry')
from Autodesk.DesignScript.Geometry import*

#The input will be stored as a list of IN variables
grid_x = IN[0]  # Receiving grid - x coordinate point
grid_y = IN[1]  # Receiving grid - y coordinate point
elevation = IN[2]  # Receiving elevation
p_lists = [ ]
for k in range(len(elevation)):
   for j in range(len(grid_y)):
      for i in range(len(grid_x)):
         x_ij = 1000*grid_x[i]
         y_ij = 1000*grid_y[j]
         z_ij = 1000*elevation[k]
         p_ij = Point.ByCoordinates(x_ij, y_ij, z_ij)  #function node
#Output content is assigned to OUT variables
OUT = p_lists
```

FIGURE 0.4 **GENERATE LOCATING POINT OF AXIS NETWORK" PYTHON SCRIPT**

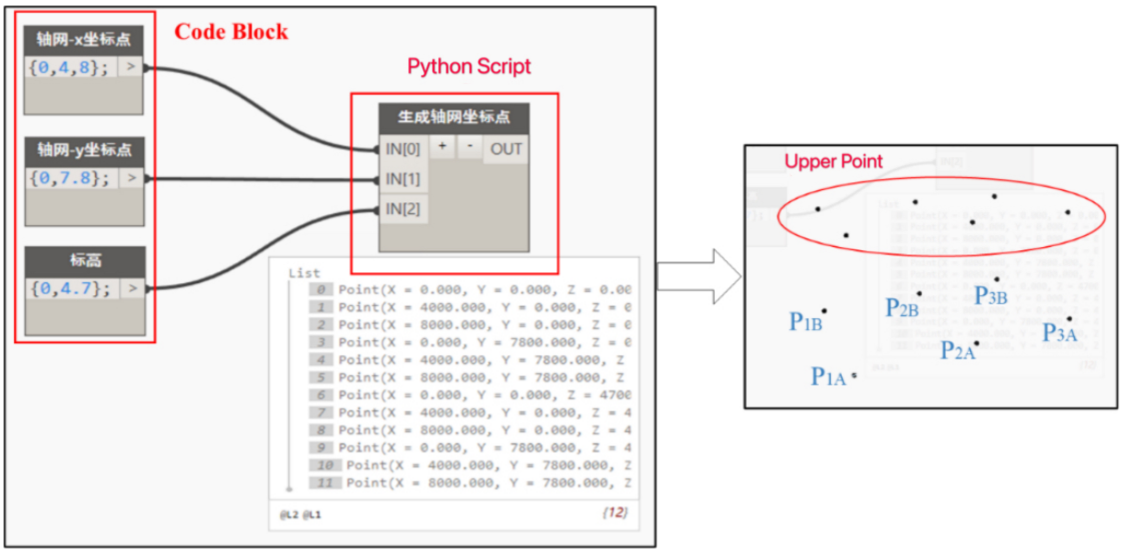


FIGURE 0.5 **LOCATING POINT GENERATION OF AXIS NETWORK**

Step 2: Component positioning line generation

In the Revit project, the positioning line of the column is the central axis, and the positioning line of the beam is the top axis by default, as shown in Figure 5.6. The grid coordinate point set was gathered in step 1 and input it to the downstream node, so it can be indexed according to the arrangement position of the two ends of the component positioning line in the set. When obtaining the positioning points at both ends of the positioning line, the *Line.ByStartPointEndPoint* node in Dynamo Geometry node library supports the creation of a straight line between two input points. Finally, the component positioning line is generated as shown in Figure 5.7.

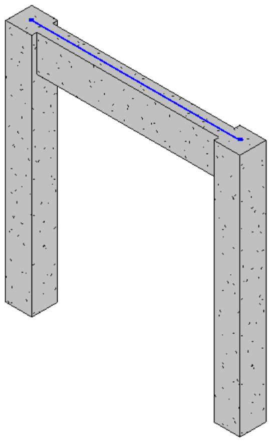


FIGURE 0.4 **POSITION LINE OF COMPONENT**

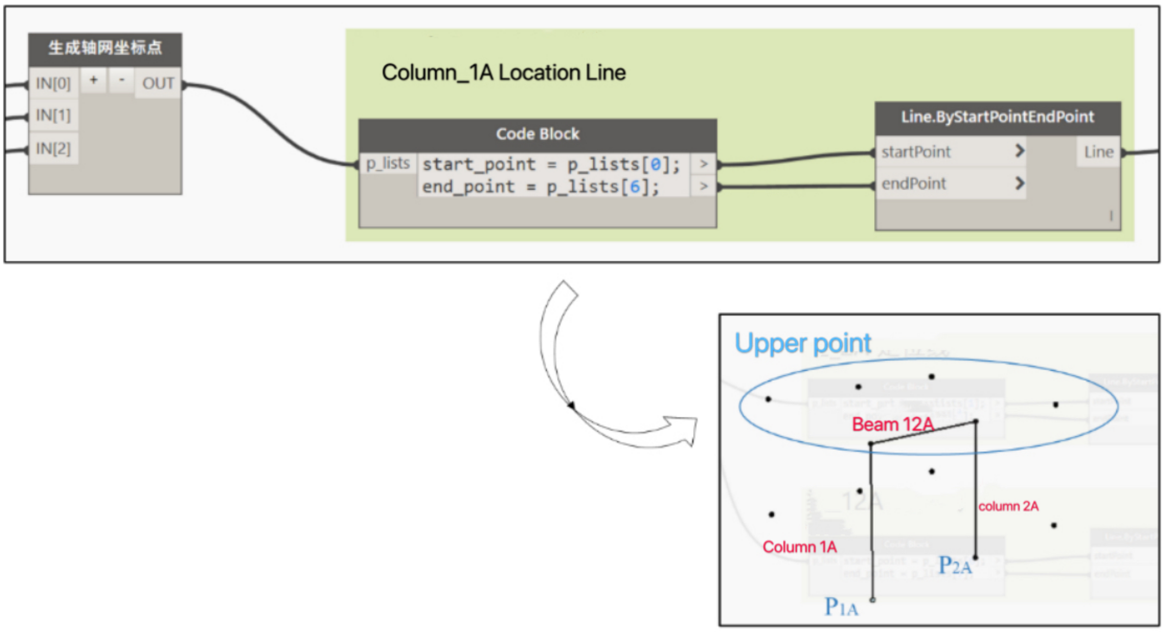


FIGURE 0.5 **POSITION LINE GENERATION OF COMPONENT**

Step 3: Component generation

Taking column components as an example, the Dynamo Revit node library provides the *StructuralFraming.ColumnByCurve* node, which is used to generate column components based on component positioning line, story height and family type information. The component positioning line has already been obtained in Step 2. Story height is obtained by selecting the existing elevation level in the Revit project document through the Levels function node. The component type is specified by the Family Types node which can support access to all available family types in the project document.

When the above function group is executed, the corresponding family instance will be automatically generated in the Revit project. The generation of frame beam component is displayed in Figure 5.8.

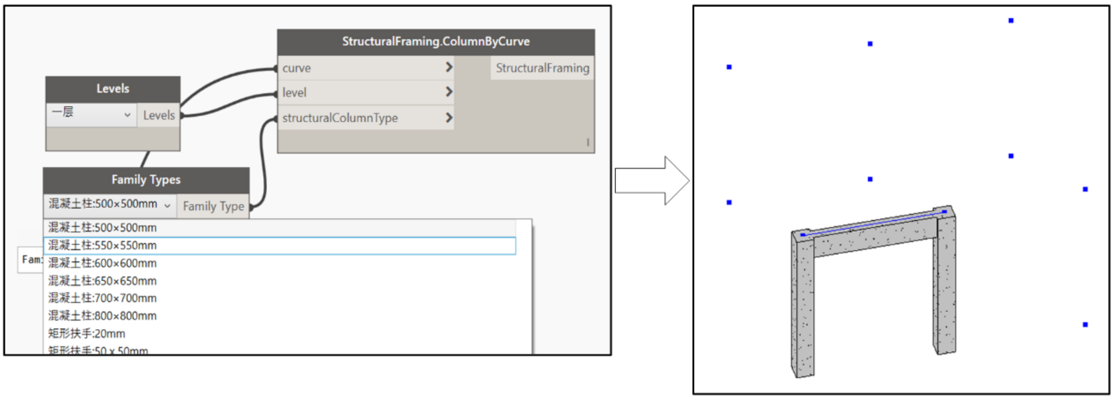


FIGURE 0.6 **COMPONENT GENERATION**

The above modeling process uses grid positioning, component type and location as driving parameters, users can only need to modify the corresponding values at the parameter nodes according to the design plan to realize the generation of the model, which greatly improves the modeling efficiency and quality. As Dynamo supports interaction with Excel software, the parameter values of the design scheme can also be stored in an Excel file in advance and then read into the parameter node for running.

5.2 MULTI-OBJECTIVE OPTIMISATION UNIT OF DISCRETE SIZE VARIABLES

The study presents a discrete-dimensional model for multi-objective optimisation with discrete dimensional variables, aiming to determine the structural seismic optimisation plan as a decision-making issue with several objectives at the component level. Using the multi-objective genetic algorithm, based on the condition of fulfilling the specified criteria and achieving the performance goals (CSI, 2016), the final result is conveyed by a set of optimal design schemes that strike an equilibrium between the initial building expense and the expected earthquake damage. The model is formulated using mathematical notation, as seen in Equation 5-1 (Lisboa & Soares, 2014):

$$\begin{aligned} & find\ x = [v_1, v_2, \dots, v_i, \dots, v_n]^T \quad i = 1,2, \dots, n \\ & V - \min f(x) = [C_0(x), C_1(x)]^T \\ & s.t. PO(x) \le [PO] \\ & g_j(x) \le 0 \quad j = 1,2, \dots m \\ & v_i \epsilon \{v_{i1}, v_{i2}, \dots, v_{ip}, \dots, v_{iq}\}\ p = 1,2, \dots q \end{aligned} \qquad (5\text{-}1)$$

X in the Formula 5-1 represents the collection of design schemes, named the design plan x. The design plan x is characterised by a sequence of cross-sectional dimension variables V_i that indicate structure component. The initial construction expense and anticipated seismic expense have a one-to-one correspondence with the design plan and are used as two independent objective functions for minimum optimisation. While meeting the traditional specification design requirements, each design plan must satisfy the performance requirements put forward by all stakeholders. Simultaneously, the values of the dimensional variables are also constrained to discrete values that meet the conventional criteria.

This model primarily encompasses the identification of discrete size variables, the computation of objective functions, the formulation of constraints, and the choice of optimisation algorithms, which will be introduced in the following context.

5.2.1 INDENTIFICATION OF DISCRETE DIMENSIONAL VARIABLES

For the RC frame structure during the preliminary design stage, the cross-sectional size of the structural member is considered as a design variable and assume that the topological structure of the building, the material of the component and the size of the non-structural component are set to be known and fixed. Then reinforcement of the components can be automatically reinforced by the structural analysis programme.

Taking the standard RC frame building architecture in Figure 5.9 as an illustrative instance, this study assumes that the concrete component have rectangular section and the feature of these sections may be described by their size (M, 2001). As the design plan is axisymmetric, the columns on axis-a are exactly the same as the columns on axis-c and the beams on axis-1 are exactly the same as the beams on axis-6. For the frame column, it is assumed that the height of each floor is fixed, and all columns' cross-section are square, the frame column may be categorised into 6 groups based on the component's plane position and the load condition. The width of the section will be utilized as the design variables, which will be defined based on the component's axis code. For instance, d_{1A} represent column's cross-section width variables where axis 1 and axis A intersect. Frame beams are categorised into 4 groups based on their span and orientation, assuming the span is already fixed. Width and height of rectangular section serve as design variable that are denoted by the category number, such as b_{L4} and h_{L4} indicate the cross-section width and height variables of the 4th category of beam. Then the design scheme of the frame is displayed using fourteen sample design variables as below.

$$x = [d_{1A}, d_{2A}, d_{3A}, d_{1B}, d_{2B}, d_{3B} | b_{L1}, b_{L2}, b_{L3}, b_{L4} | h_{L1}, h_{L2}, h_{L3}, h_{L4}]^T \quad (5\text{-}2)$$

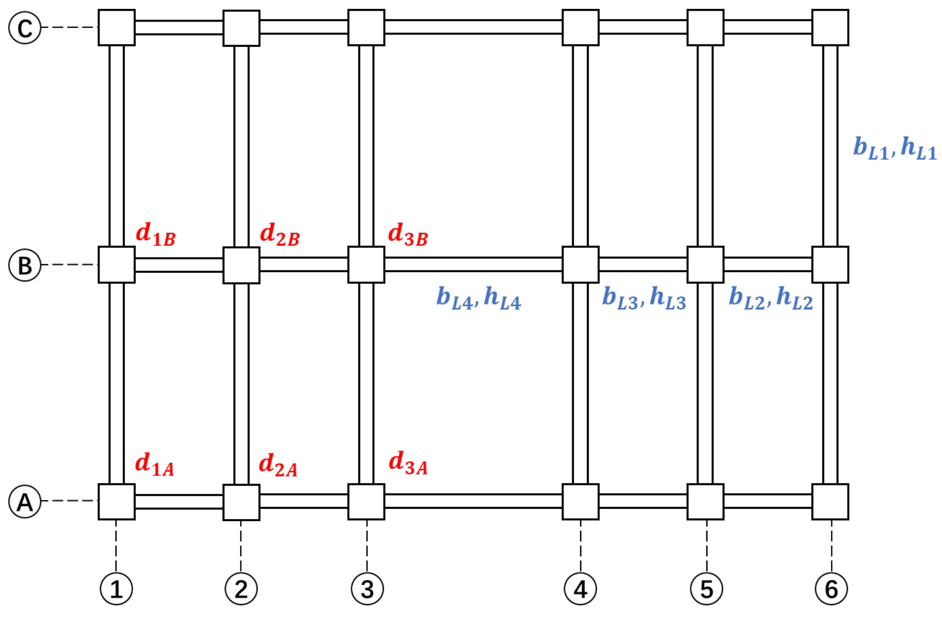


FIGURE 0.7 STANDARD BUILDING ARCHITECTURE OF RC FRAME

In the primary selection of the traditional structural section design, an estimate may be made about the dimensions of the column section according to $A_c \geq \frac{N_c}{[\mu_N] f_c}$. $[\mu_N]$ is the maximum value of axial compression ratio, which is selected according to the structure type and seismic grade in the seismic regulation. f_c Is the design value of the axial compressive strength of reinforced concrete. is the estimated design value of the column axial force. The initial section height of the beam section is usually 1/15 ~1/10 of the span and the initial section width of the beam section is usually 1/3 ~1/2 of the section height. Based on this, designers can preliminarily determine the range of discrete design variables.

5.2.2 OBJECTIVE FUNCTION CALCULATION

Initial building expense calculation

The integrated unit price technique is to use the unit prices of each part of the project as the total cost unit price, which is generated after comprehensive calculation and includes direct costs, indirect costs, profits and taxes (the calculation formula is shown in *Consequences of injury C* in Section 2.1.2). On February 17, 2003, the Ministry of Construction in China promulgated the *"Construction Project Integrated Unit Pricing Specifications" (GB50500-2003)*, which took effect on July 1, 2003. Since then, the integrated price technique has been formally adopted in the construction engineering cost sector. As the research backdrop of this thesis is situated in China, the integrated unit price technique is applicable for determining the initial building expense when the model information is obtained by analysing the IFC. As this study assumes that everything remains unchanged except for structural components, the cost associated with non-load-bearing interior and exterior walls, floor slabs, and other non-structural elements can be ignored and solely included in frame beams C_{beams} and frame columns $C_{columns,}$ shown in Formula 5-3.

$$C_1 = C_{columns} + C_{beams} \tag{5-3}$$

$$C_{fra} = C_C \sum_{i=1}^{N} (V_C - V_S)_i + C_S \gamma_S \sum_{i=1}^{N} (V_S)_i + C_f \sum_{i=1}^{N} (A_f)_i \tag{5-4}$$

NOTE: the cost of frame beams and columns should be included in the cost of concrete, reinforcement and formwork, which is expressed in C_{fra} uniformly. As shown in Formula 5-4: C_c, C_s and C_f are the cost of concrete per unit volume, the cost of steel bar per unit weight and the cost of formwork per unit area respectively. Y_s is the bulk density of the steel bar. V_c and V_s are the concrete volume and the steel bar volume respectively, where the steel bars include longitudinal steel bars, stirrups, etc. A_f is the surface area of the formwork and N is the number of components.

Anticipated seismic loss calculation

The damages resulting from earthquake activity may often be categorised as direct losses and indirect losses. The study just considers the maintenance costs associated with direct losses, which is the expenses incurred for repair or replacement of structural and non-structural elements (Min et al., 2010). The approach described in Chapter 4 may be used to determine the extent of damage to a structure under a specific earthquake level. If all potential future earthquake actions are taken into account, this study defines earthquake loss expectation as the product of seismic losses and the likelihood of earthquake occurrence. This method does not take a whole lifecycle cost into account.

1. Seismic hazard curve

The seismic hazard curve represents the correlation between the intensity of ground motion and its yearly average probability of exceeding. Cornell et al. approximates the expression of the seismic hazard curve shown in Formula 5-5 (Cornell, Jalayer, Hamburger & Foutch, 2002):

$$H(x) = k_0 x^{-k} \tag{5-5}$$

K0 and k can be fitted according to the relevant parameters of Design Based Earthquake (DBE) and Maximum Considered Earthquake (MCE), shown in Formula 5-6 and 5-7.

$$k = \frac{\ln(v_{DBE}/v_{MCE})}{\ln(IM_{MCE}/IM_{DBE})} \tag{5-6}$$

$$\ln k_0 = \frac{\ln(IM_{DBE})\ln(v_{MCE}) - \ln(IM_{MCE})\ln(v_{DBE})}{\ln(IM_{DBE}/IM_{MCE})} \tag{5-7}$$

v_{DBE} and v_{MCE} represent the annual exceeding probability of DBE and MCE respectively. IM_{DBE} and IM_{MCE} represent the index of DBE and MCE respectively.

The fitted seismic hazard curve of ground motion intensity and annual exceeding probability is displayed in Figure 5.10, from which the annual exceeding probability corresponding to a certain ground motion intensity can be calculated.

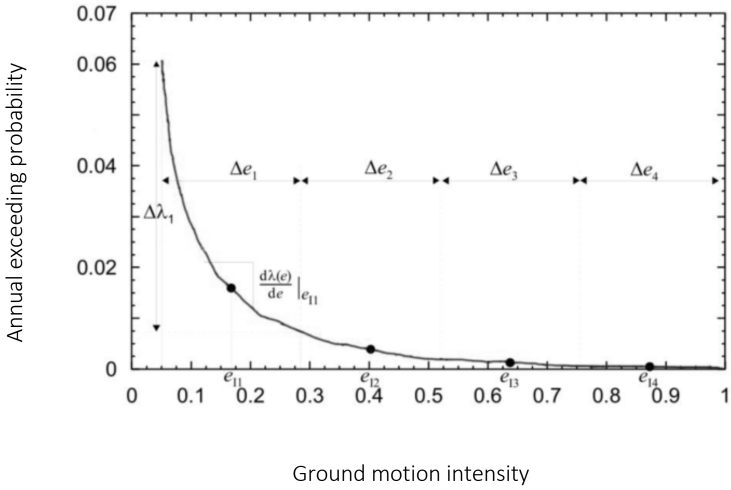


FIGURE 0.8 SEISMIC HAZARD CURVE

2. Average annual repair cost

Theoretically, according to the seismic hazard curve in Figure 5.10, the total repair cost annual exceeding probability curve in Figure 5.11 can be obtained by calculating the repair cost at each feature point. The area enclosed by the curve is the average annual repair cost C_{annual} caused by the potential earthquake.

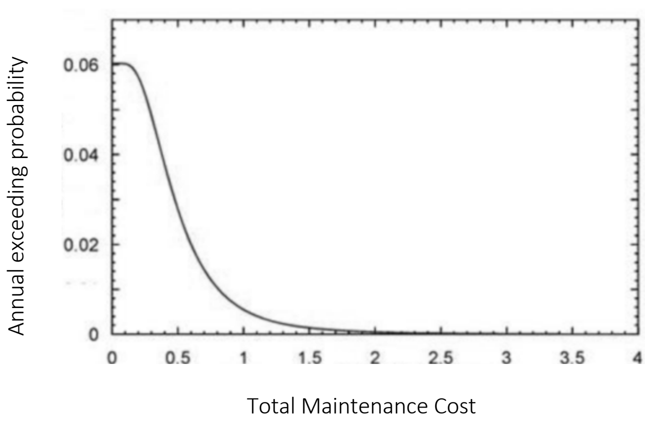


FIGURE 0.9 **PERFORMANCE CURVE**

In FEMA P-58, C_{annual} is used as the benchmark when calculating the average cost for repairing of a structure within its service life. It is computed according to the net present value of the equivalent future expenditures each year, so the earthquake loss expectationshould be:

$$C_1(x,T,\lambda) = C_{annual}(\frac{1-\frac{1}{(1+\lambda)^T}}{\lambda}) \quad (5\text{-}8)$$

X represents the structural design plan, which will affect the value of C_{annual}; T is the structural design service life, which is 50 years for ordinary houses and structures in China as the research background is settled in Sichuan, China; λ is the discount rate. According to the research results of the (Razavi & Gholizadeh, 2021), when the design service life is within 50 years, the fixed discount rate can be taken as 0.035 for earthquake repair loss calculation.

5.2.3 EXPRESSION OF CONSTRAINTS

During the design optimisation process, the design plan must satisfy the criteria of seismic performance, structure and other relevant factors (Xiong & Huan, 2023). Therefore, the value of the design variable is constrained both directly and indirectly.

Indirect constraints refer to design variables' implicit function must fulfil specific criteria. The output of such constraints requires structural analysis, and the manual judgement is made in conjunction with the results given by the software. Specifically, it refers to satisfying the criteria of conventional code for component strengths, reinforcement ratio and total building displacement under the action of earthquakes. Additionally, it involves achieving the desired performance goals in PBSD. The IDR is often utilised a quantitative measure to assess if RC frame constructions satisfy the predetermined performance criteria across various earthquake intensities displayed in Table 1.1.

Direct constraints are to explicit limitations on the permissible values of design variables. The designer establishes these restrictions based on the unique state of the building and must consider the feasibility of future manufacturing and construction. For example, the physical dimensions of component size is generally selected as a multiple of 50mm or 100mm (Park, Hwag & Oh, 2018), the ratio of beam height to width is limited to 1.5-2.5, keep component types as few as possible.

5.2.4 THE APPLICATION OF NSGA-II

The effectiveness of the optimisation problem is highly dependent on the application of the optimisation algorithm. Therefore, it is crucial for qualified experts who have the target domain knowledge to choose the optimal optimisation algorithm according to the specific circumstances of the issue. According to the description of Formula 5-1, the optimisation problem studied in this section is a multi-objective optimisation issue of discrete variable sets. The problem of multi-objective is to obtain a series of optimal solution sets, which is Pareto optimal solution.

When applying the NSGA-II algorithm for the discrete size variable multi-objective optimisation issue suggested in this section, these parameters should be given first: the size of starting population, crossover likelihood, mutation probability, clear plan penalties and termination conditions. The penalties are imposed when the structure fails to fulfil the specific restrictions expressed in Section 5.3.3, then the resolution will be subjected to disciplinary measures, and the punishment procedure removes the design scheme (Eleftheriadis et al., 2018). The termination condition of this study is set as the occurrence of every individual design plan in 3 consecutive generations, with all of them having occurred in the prior generations. As the customer has the ability to manipulate the iterative process, it is unnecessary to predefine the maximum iteration frequency. To generate new populations, designers must exercise artificial judgements based on the outcome of nonlinear time-history analysis obtained from the YJK application. A design plan that satisfies the constraints may be included as one of the schemes in the new population. Otherwise, it should be discarded, and other design plans created simultaneously will be chosen instead. The new design schemes will be verified again until the total number of populations is satisfied. The final optimisation result constitutes the Pareto optimal solution, representing as a series of better design schemes with the component size as a variable.

5.3 SUMMARY

The core of the multi-objective seismic optimisation design technique of RC frame using BIM proposed in this chapter lies in a discrete multi-objective optimisation model of dimensional variables. By treating the structural seismic optimisation design guided by performance-oriented thinking as a multi-objective optimisation problem, the initial building expense and the seismic loss expectations are treated as two conflicting goals in term of expenses. For the specific design plan, use the size of representative load component of RC frame structure and establish one-to-one correspondence between the two major costs and the design plan. Iterative optimisation of the NSGA-II algorithm is used to obtain a series of Pareto solution sets for stakeholders to make trade-off decisions. Additionally, the separation of the user module and the subsequent optimisation module allows the technical staff to focus on the design itself without having the corresponding theoretical knowledge of algorithm, which means the application threshold of this method is relatively low. The application of BIM technology, on the one hand, helps designers avoid manual complex and multiple iterative modeling and repetitive modeling during the structural analysis with the help of parametric modeling function. On the other hand, it realizes rapid prediction of cost at the component level based on IFC standards. It improves the efficiency and quality of optimized design to a certain extent.

CHAPTER 6 RESEARCH ON EARTHQUAKE DAMAGE PREDICTION OF MULTI-SCALE REGIONAL RC FRAME BASED ON ANN

As stated in the research background part of Chapter 1, this thesis is derived from a research project which aimed to study how to predict earthquake damage for regional building groups and then to improve seismic performance of the individual building of building groups that have suffered serious damage. In previous Chapter 4 and Chapter 5, BSPEF and MSODM are developed to improve structure seismic performance through performance evaluation and multi-objective optimisation design. This chapter will then illustrate how to predict the level of earthquake damage for regional building groups and how to identify the buildings that have suffered serious damage. Subsequently, the approaches developed in Chapter 4 and Chapter 5 can be applied to these seriously damaged buildings for seismic performance improvement.

This chapter proposes a prototype to predict the earthquake damage for multi-scale regional RC frame using ANN technology (ANN technology has been introduced in Section 2.4.2). The selecting method of damage prediction parameters and the generation of training data are introduced step by step in Section 6.1 and Section 6.2 respectively. Section 6.3 describes how to establish an ANN model. The model of earthquake damage prediction based on ANN is developed in Section 6.4 and how to apply the model for multi-scale regional RC frame is stated in Section 6.5.

6.1 SELECTION OF EARTHQUAKE DAMAGE PREDICTION PARAMETERS

6.1.1 INPUT PARAMETERS

Structure-related parameters and earthquake-related parameters are among the factors influencing the results of earthquake damage. When selecting the structural parameters, it is considered not only that the selected parameters should be the main factor that affects the load characteristics of the RC frame structure, but also the principle of easy access to obtain the parameters. Therefore, the purpose of rapid prediction can be achieved. The seismic parameters should reflect the characteristics of seismic waves. If the data features are highly correlated, it may lead the model to overfit, which denotes the occurrence when a model excels on training data but underperforms on novel data. Overfit results in a model with poor generalisation ability and cannot be effectively applied to practical problems. Therefore, the input parameters should not be correlated or have little influence on each other (Lee, 2018).

Structural Parameters

Structural parameters needed are those which characterise better seismic properties of the building structure during an earthquake. Morfidis and Kostinakis (2018) pointed out that the more critical structural parameters include: total height, planar layout, floor height layout, structural system, structural eccentricity, grade of concrete and steel bars, size and reinforcement of structural elements, foundation system and soil classification. A six-story RC frame structural is provided by Sichuan Provincial Government as the object of study and 10 representative structure parameters are chosen by considering the overall and fragmentary characteristics of the structure, which are shown in Table 6.1. These parameters are extensively used in established methodologies for assessing the fragility of RC buildings (Guo & Li, 2021; Lu et al., 2020; Morfidis & Kostinakis, 2018; Su & He, 2018; Ye, Zhang, & Zhu, 2019) and have been recognized by contemporary seismic codes as the factors that significantly influence the earthquake damage status of RC buildings (e.g., Eurocode 8, BS EN 1998-5, and GB50011-2010)

TABLE 0.1 **INFORMATION OF 10 STRUCTURAL PARAMETERS**

No.	Parameter	Unit	No.	Parameter	Unit
1	Ground Floor Height	m	6	Reinforcement ratio of Beam	—
2	Floor Height	m	7	Concrete Strength	MPa
3	Size of Column	mm	8	Reinforcement Strength	—
4	Height of Beam	mm	9	Horizontal side span	mm
5	Reinforcement ratio of Column	—	10	Horizontal midspan	mm

Note: multi-layer frame structure is commonly found in teaching buildings, office buildings, hospitals, etc., because of functional requirements, the height of the ground floor is often higher than that of other floors.

Seismic Parameters

The seismic parameters represent the impact of different ground motion recordings on the structure's seismic response. Morfidis and Kostinakis (2018) selected 14 seismic parameters and evaluated their influence on the neural network prediction accuracy of the MIDR in literature. The conclusions are stated as: The inclusion of more than 5 seismic parameters as input parameters results in increased prediction accuracy and the optimal combination of ground motion parameters yields a minimum mean squared error of 0.044 and comprises 13 seismic parameters (as shown in Figure 5.1). Lautour et al. (2009) selected Peak Ground Acceleration (PGA), Peak Ground Displacement (PGD), Spectral Intensity (SI), characteristics frequency and effective duration as parameters to depict the properties of ground motion. Du and Padgett (2020) selected the spectral acceleration that matches basic period median value of case sample as the ground motion Intensity Measure (IM) Index. The research has shown that these input parameters have good predictive effects on earthquake prediction. Therefore, 13 seismic parameters will be selected as input parameters and take all these parameters into account.

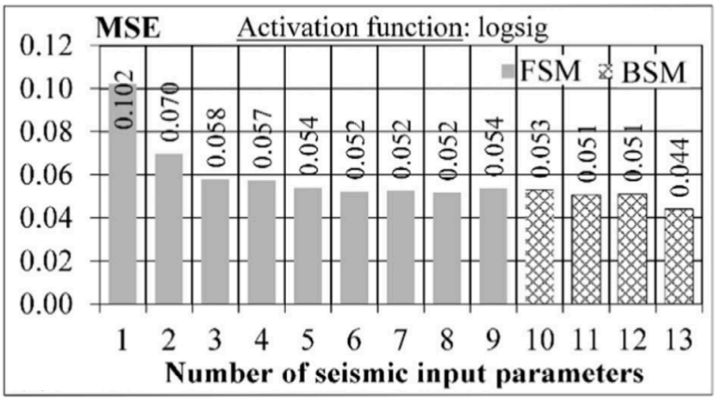


Figure 5.1 MSE OF DIFFERENT NUMBER OF SEISMIC INPUT PARAMETERS

Ground motion characteristics include three elements: amplitude, frequency spectrum and duration (Cortés-Pérez, Cortés-Pérez & Prieto-Muriel, 2020). There are generally four types of earthquake duration definitions: bracket, consistent, important and effective duration (De & Giri, 2018). Except for important duration, the remaining duration parameters will change with the change of ground motion amplitude, which can refer to either the duration of the original ground motion record or the duration of the ground motion record after amplitude modulation. The important duration is the time between the different percentage of the cumulative energy during the ground motion acceleration time history and the total ground motion input energy, which does not change with the amplitude of the ground motion and has no obvious correlations with PGA, PGV and $S_a(T_1)$. The cumulative energy is widely applied in practical research as 5%-75% (D_{s5-75}). Jianping et al. (2020) select D_{s5-75} based on Arias intensity as the duration parameter of ground motion. Arias intensity calculation is shown in Formula 6-1:

$$I_A = \frac{\pi}{2g}\int_0^{t_{max}} a^2(t)dt \tag{6-1}$$

In this formula, a(t) represents the time history of the ground motion recording acceleration, t_{max} represents the total recording time and g is the gravity acceleration.

The definition of $D_{s5\text{-}75}$ based on Arias intensity is shown in Figure 6.2. I_{A0} and I_{A1} represent 5% and 75% of the accumulated energy respectively and the time difference between the two corresponding times is the duration.

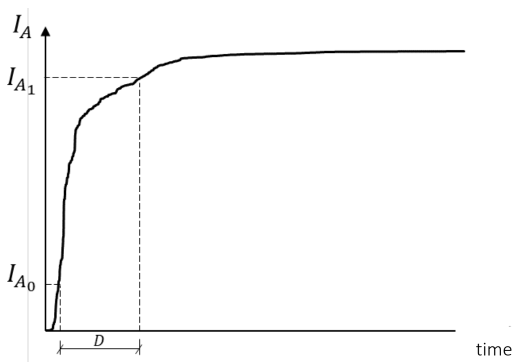


FIGURE 0.2 **THE DEFINITION OF D_{s5-75} BASED ON ARIAS INTENSITY**

Earthquake prediction needs to consider the amplitude modulation factor, due to the abruptness and dynamic nature of seismic waves of seismic waves events. Earthquakes are abrupt phenomena that are transient and change rapidly. When this dynamic action is converted into an equivalent static action for design, the actual bearing capacity of the component in an earthquake is higher than when it is designed according to static forces. To account for this disparity in bearing capacity, an adjustment coefficient, referred to as the amplitude modulation factor, is included into seismic design. Therefore, the selection of seismic parameters for this thesis should take into account the two aspects below. Firstly, try to select those parameters that can effectively describe the characteristics of ground motion. Secondly, consider the issue of seismic wave amplitude modulation. Therefore, the 13 seismic parameters in Table 6.2 are selected. PGA, PGV PGD, I_a, SED, CAV, ASI, HI, EPA, V_{max}/A_{max} will all be scaled during amplitude modulation while PP and effective duration D will remain unchanged.

TABLE 0.2 **INFORMATION OF 13 SEISMIC PARAMETERS**

Parameter	Calculation Equation	Unit	Explanation
Peak Ground Acceleration: PGA	$\max\lvert a(t)\rvert$	g	Seismic parameters (amplitude characteristics) determined by time history records
Peak Ground Velocity: PGV	$\max\lvert v(t)\rvert$	cm/s	
Peak Ground Displacement: PGD	$\max\lvert d(t)\rvert$	cm	
Arias Intensity: I_a	$I_a = (\frac{\pi}{2g})\int_0^{t_{tot}}(a(t))^2dt$	m/s	
Specific Energy Density: SED	$SED = \int_0^{t_{tot}}(v(t))^2dt$	cm^2/s	
Cumulative Absolute Velocity: CAV	$CAV = \int_0^{t_{tot}}\lvert a(t)\rvert dt$	cm/s	
Acceleration Spectrum Intensity: ASI	$ASI = \int_{0.1}^{0.5}S_a(\xi = 0.05, T)dT$	g • s	Seismic parameters determined by response spectrum (spectral characteristics)
Housner Intensity: HI	$HI = \int_{0.1}^{2.5}\mathrm{PSV}(\xi = 0.05, T)dT$ Pseudo-velocity spectrum PSV	cm	
Effective Peak Acceleration: EPA	$EPA = \left(\frac{1}{2.5}\right)\{\overline{S_a}(\xi = 0.05, T)\}_{0.1}^{0.5}$	g	
Vmax/Amax(PGV/PGA)	$\max\lvert v(t)\rvert/\max\lvert a(t)\rvert$	s	
Predominant Period:PP	$PP = T[maxS_a(\xi = 0.05, T)]$	s	
Effective Duration:D	D_{s5-75} based on Arias strength	s	Seismic parameters based on earthquake duration (duration characteristics)
Target-PGA	Maximum acceleration by time history analysis	g	Amplitude modulation information

In order to realise the prediction of the bridge fragility curve of the IDA method (this method has been explained in *Analysis of collapse fragility* in Section 2.1.2), Du and Padgett (2020) selected 12 seismic waves to scale 30 times which is 360 seismic wave series in total, covering each IM value interval, which is time-consuming but necessary. This study takes PGA as the amplitude modulation index and adds Target-PGA seismic parameters, scaling each selecting series of seismic waves by as many discrete values as possible. After the trained network is utilised to build the fragility function, fix the PGA value and input the remaining seismic parameters. The lower limit of the Target-PGA parameter is generally a small value such as 0.02g. The upper limit of the value will be computed according to the basic seismic data of the building structure under consideration. It should be greater than the PGA value when all structures collapse.

According to the selection of structural parameters and seismic parameters, this paper selects a total of 23 input parameters including 10 structural parameters and 13 seismic parameters to form the input vector x shown in Formula 6-2:

$$x = [x_{struct}|x_{seism}]^T$$

$$x_{struct} = [H_1|H_2|b|h|\rho_c|\rho_b|f_c|f_s|l_1|l_2]^T \quad (6\text{-}2)$$

$$x_{seism} = [PGA|PGV|PGD|I_a|SED|CAV|\mathrm{ASI}|\mathrm{HI}|\mathrm{EPA}|\mathrm{PGV/PGA}|\mathrm{PP}|\mathrm{D}|\mathrm{Target-PGA}]^T$$

6.1.2 OUTPUT PARAMETERS

The output parameters determine the prediction type, but also affect the selection of neural network type, network configuration, training algorithm and its evaluation index. The MIDR is a global, structural and deterministic index (Kappos & Eng, 2010). It is a reliable index reflecting the overall structure and non-structural damage of RC buildings (Elenas & Meskouris, 2001; Naeim, 1989). It characterises the overall damage condition of the building and is also crucial in establishing the fragility function. The output parameters shown in Table 6.3 are formed according to the research purpose.

TABLE 0.3 **INFORMATION OF OUTPUT PARAMETER**

Output parameter	Abbreviation	Numerical Type	Prediction Type
Maximum inter-story Drift Ratio	MIDR		Function Approximation(Boddy et al.)

Since MIDR is a real number greater than zero, this type of prediction problem belongs to function approximation (FA) problem. The output value o $\in R^1$ is predicted by the input vector x $\in R^{23}$ and the mathematical expression is (Saldaña-Robles, 2020):

$$f(x): R^{23} \rightarrow R^1 \tag{6-3}$$

Therefore, in the following network configuration part of ANN model development (Section 6.3.1), the performance indicators can be selected as Mean Square Error (MSE) and correlation coefficient R, which are common performance indicators for FA problems.

After the input and output parameters selected in Section 6.2 are assigned values from a series of real cases analysis, the training data used to build the ANN prediction model can be created. For structural parameters, this study considers the premise of covering different parameters combinations as much as possible and select a series of architectural case design plans to generate corresponding structure plans to generate corresponding structural parameter values.

For seismic parameters, this study adopts the new version of NGA-WEST2 strong motion record database released by PEER Center (PEER, 2003), which is widely recognized as complete and reliable ground motion parameters. A total of 21,324 records with complete distance, site condition information and a damping ratio of 5% are selected. The distribution of seismic distance and magnitude are shown in Figure 6.3, which can be seen that the strong motion record database provides many near-field's (100-1000km) strong earthquake records.

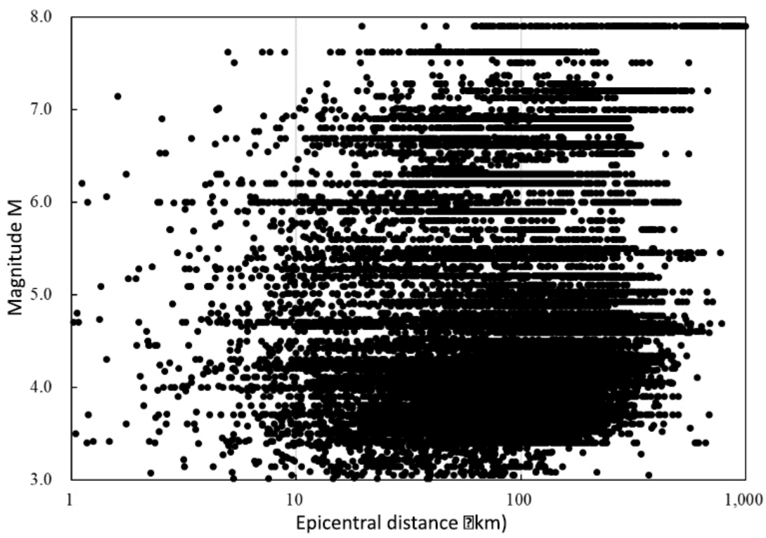


Figure 0.3 **Distance-magnitude distribution of earthquake waves**

This thesis combines the seismic wave selection method in the existing research, and proposes to screen seismic waves from four aspects (Ji, 2018): magnitude, seismic distance, site conditions and recorded amplitude modulation. Seismic parameter values are calculated according to the formula in Table 6.2.

Step 1: Preliminary Selection Conditions of Magnitude Parameters

The current accumulation of strong earthquake records determines that when selecting actual records, most of them need to linearly modulate the record amplitude. Therefore, PGA is acceptable within a certain range. Ji (2018) recommends using the intensity attenuation relationship to calculate the magnitude ranges of earthquake events corresponding to different target intensities of 10-200km and gives the recommended magnitude ranges for each precautionary intensity, as shown in Table 6.4.

TABLE 0.4 **PRIMARY FILTER OF MAGNITUDE PARAMETER**

Precautionary Intensity	VI	VII	VIII	IX
Range of Earthquake Level	[5.0-7.5]	[5.0-8.0+]	[6.5-8.0+]	[7.0-8.0+]

Step 2: Preliminary Selection Conditions of Seismic Distance Parameters

Numerous studies and verifications have been conducted on the role of distance in the primary selection conditions. It is generally believed that the correlation between distance and structural nonlinear response is lower than the magnitude (Ji, 2018). The study selects the reference range in Table 6.5 according to the epicenter distance of each design group.

TABLE 0.5 **PRIMARY FILTER OF SEISMIC DISTANCE PARAMETER**

Seismic distance	Group		
Precautionary Intensity	1st group	2nd group	3rd group
VI	[0,12]	[12,25]	[25+]
VII	[0,15]	[15,30]	[30+]
VIII	[0,18]	[18,40]	[40+]
IX	[0,20]	[20,50]	[50+]

Step 3: Site Selection Condition

Due to the complexity and discreteness of the site response itself, its priority in the preliminary selection conditions is lower than the magnitude and distance. If the number of final candidate records is sufficient, further screening can be combined with the site type of the target site. Station records that are significantly inconsistent with the target site type should not be used and records that are the same or different from the target site category should be used. Ji (2018) recommends using the conversion relationship between the 30m borehole shear wave velocity V_{s30} of the station site and the site classification shown in Table 6.6 to determine the site category.

TABLE 0.6 **SITE FILTER CONDITION**

Site Type	suggested value
IV	$V_{S30} \leq 160m/s$
III	$160 \leq V_{S30} \leq 260m/s$
II	$260 \leq V_{S30} \leq 550m/s$
I	$V_{S30} > 550m/s$

Step 4: Amplitude Modulation Record

In the practice of wave selection, the upper limit is the primary component that affects the amplitude modulation coefficient, rather than the lower limit. After Steps 1-3 are screened for seismic wave, Wen et al. (2019) adopts the input PGA of rare earthquakes under different precautionary levels as the target value to inversely calculate the corresponding recorded amplitude modulation coefficient. [0.2 – 5] is selected as the initial amplitude modulation interval of NGA-West2 and the upper limit of amplitude modulation can be floated to 10. The use of IDA analysis will involve amplitude modulation within a sufficient range of smaller and larger PGA values. Therefore, the seismic waves screened in step 3 need to be limited by amplitude modulation coefficients for different PGA values.

Finally, for each design scheme, the seismic wave is selected to use for seismic analysis, set relevant parameters in YJK software according to the code requirement and perform nonlinear time history analysis to acquire the IDR response as the output parameter. Then the output parameter will form the training data together with the input parameter corresponding to the design plan, which is utilised to build the prediction model. For individual nonlinear dynamic time history analysis, when the frame's IDR exceeds 0.1, numerical instability will appear (Goda & Tesfamariam, 2015). In this case, the training samples with MIDR 0.1 are removed from the training data.

6.3 ANN MODEL DEVELOPMENT

According to the introduction of ANN principles and basic models in Section 6.1, if a BP neural network with better performance for a specific problem is trained, it is necessary to optimise and adjust the number of hidden layers, the number of neurons in each layer, activation function and training data partition on the basis of clear performance indicator. Finally, the optimally configured prediction model can be obtained. For a certain configuration network, the optimal performance of different training algorithms will also be different.

6.3.1 NETWORK CONFIGURATION

Step 1: Performance Indicators

For FA problems, common performance indicators are Mean Square Error (MSE) and correlation coefficient R. The MSE is the average of the sum of squares of each data deviation from the true value, as illustrated in Formula 6-4 (Xiong, Li & Lu, 2020). The definition clearly indicates that a lower MSE value corresponds to a superior prediction effect.

$$MSE(Y,\hat{Y}) = \frac{1}{N}\sum_{i=1}^{N}(\hat{Y} - Y)^2 \qquad (6\text{-}4)$$

Y represents the true value, $\hat{Y}$ represents the predicted value and N is the amount of data.

The correlation coefficient R measures the linear relationship between the predicted value $\hat{Y}$ and true value Y, shown in Formula 6-5 (Falcón-Cardona et al., 2021). The definition clearly indicates that a higher R value corresponds to a superior prediction effect.

$$R(Y,\hat{Y}) = \frac{Cov(Y,\hat{Y})}{\sqrt{Var(Y)Var(\hat{Y})}} \qquad (6\text{-}5)$$

$Cov(Y,\hat{Y})$ is the covariance of Y and $\hat{Y}$, Var(Y)is the variance of Y, $Var(\hat{Y})$is the variance of $\hat{Y}$.

Step 2: Number of Hidden Layers

The theory demonstrates that a single-hidden-layer feedforward network can map all continuous functions. Lee (2018) believes and verifies that a single-layer feedforward neural network can predict functions more accurately and its efficiency has also been confirmed in many related studies. Liu (2010) also pointed out that a three-layer BP neural network, consisting of a hidden layer, has the capability for estimating given continuous function, provided that number of neurons in the hidden layer may be freely chosen. In this context, the number of hidden layers is configured as 1, which is a three-layer BP neural network.

Step 3: Number of Neurons

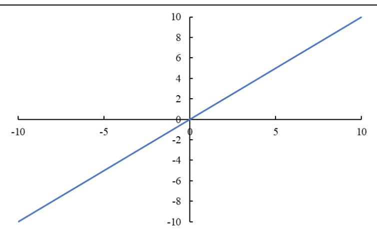
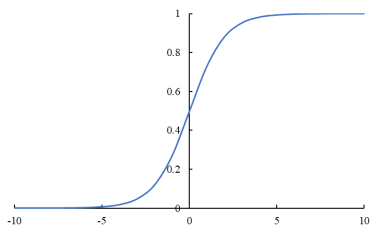
The number of neurons varies according to the nature and specific circumstances of the problem being studied. There is currently no direct method to determine the number. The optimal number is generally established via the process of trial and error.

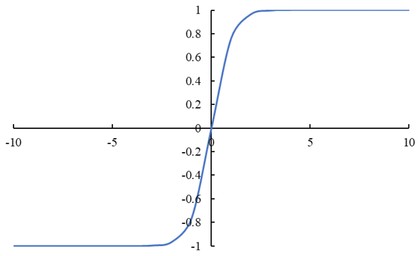
Step 4: Activation Function

The activation function introduces nonlinear characteristics to the neural network, which can make the neural network approximate complex functions at will. Typical activation functions include sigmoid function, tanh function, Relu function, as well as its enhanced versions such as Leaky-ReLU, P-ReLU, R-ReLU, and so on (Das & Singh, 2019).

Table 6.7 shows the three common types of activation function information (FEMA P-58, 2017). For hidden layer, the FA problem can use the logsig function or the tansig function. While for output layer, the purelin linear function is selected as the output value is any positive real number, without being restricted to the range of [-1, 1] or [0, 1].

TABLE 0.7 **INFORMATION ABOUT THREE TYPES OF EXCITATION FUNCTIONS**

Activation Function	Image	Function Expression	Function Scope
purelin		$f(x) = x$	$(-\infty, +\infty)$
logsig		$f(x) = \frac{1}{1+e^{-x}}$	$(0,1)$

tansig		$f(x) = \dfrac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$	$(-1,1)$

Step 5: Training Data Division

The purpose of sample set division is to guarantee the network's ability to generalise and prevent overfitting. Training samples are generally partitioned into training data, testing data and validation data based on a certain ratio. For each network configuration and training algorithm model, the training data is used to train the weight parameters and does not affect the established hyperparameters such as algorithms and configurations. After the training data has been used to determine each set of network models, the verification data is then used to test whether the model is accurate. In this process, the verification data will not affect the weight parameters obtained by training, but it can be used to adjust the hyperparameters such as the number of neurons and the number of iterations according to the testing results of verification data in different models. The model with highest accuracy will be selected. However, the optimal model under the action of the validation data is not certainly smallest for other data that has not yet appeared. Therefore, a validation data set that has not been trained at all is required to assess the correctness of the model and corroborate the generalisation ability of the network.

To prevent over-fitting, the training algorithm internally uses the validation data set to check the conditions for the termination of training, but the results obtained using the validation data set cannot draw definitive conclusions about the performance of the neural network, so the result of validation data will be ignored in this study.

Step 6: Normalization Function

Generally, input parameters need to be normalized, which is necessary for optimizing training effects (Rafiq, Bugmann & Easterbrook, 2001). Normalisation is to de-dimensionalize different parameters and reduce the numerical difference. Besides, it is to make the network converge quickly (Liu, 2020). As a common normalization method, linear normalization has two forms. Generally, the normalization form to be used is selected according to the activation function used by the network (Hornik, Stinchcombe & White, 1989). When the activation function uses the logsig function with a value range of (0,1), then use $y = \frac{x-min}{max-min}$ to normalize the data to the interval [0.1]. When the activation function uses the tansig function with a value range of (-1,1), use $y = \frac{2(x-\min)}{max-min} - 1$ to normalize the data to the interval [-1,1]. y and x are the values before and after normalization. Max and min represent the maximum and minimum values of the parameters before normalization.

Similarly, the target output vector also needs to be normalised, but in order to get the final true value, the normalised output vector needs to be reverse-transformed.

6.3.2 TRAINING ALGORITHM

Suppose there is the prediction model f(x,θ), where θ represents a vector θ with m unknown parameters $\theta = (\theta_0, \theta_1, \ldots, \theta_{m-1})^T$, the corresponding vector $x = (x_0, x_1, \ldots, x_{n-1})^T$, the corresponding target value $\hat{f} = (\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_{n-1})^T$. The neural network training process is to continuously adjust the parameters to minimise the discrepancy between the predicted value and the target value, even if the cost function is minimised. Generally, the MSE is used as the function expression as shown in equation 6-6. Solve the unknown parameters by making the partial derivative of the cost function equal to zero in equation 6-7, where $r_i(\theta)$ is called residual.

$$L(\theta) = (\hat{f} - f(x,\theta))^T (\hat{f} - f(x,\theta)) = \sum_i^n (\hat{f}_i - f(x_i,\theta))^2 = \sum_i^n (r_i(\theta))^2 \tag{6-6}$$

$$\nabla L(\theta) = \frac{\partial L(\theta)}{\partial \theta} = \sum_i^n \frac{\partial(\theta)}{\partial \theta} 2r_i(\theta) = 0 \tag{6-7}$$

Since the solution of the above unknown parameters belongs to the nonlinear least squares problem, it is necessary to rely on an optimized iterative algorithm.

This study will adopt Levenberg-Marquardt (LM) algorithm and Scaled Conjugate Gradient (SCG) algorithm for FA problem training (Yue et al, 2023). The LM algorithm is a practical and efficient calculation method. Applying the BP neural network using this algorithm to the earthquake damage prediction model can effectively deal with the problem of parameter redundancy. The model's fitting speed is faster and the accuracy is better. The SCG algorithm is an effective method for solving large-scale linear and nonlinear equations (Ohtsuka, Teshima, Matsumoto & Hikita, 2006).

6.4 ANN-BASED EARTHQUAKE RESPONSE PREDICTION MODEL

6.4.1 TRAINING PROCESS

For the FA problem, the research develops a three-layer BP neural network model. The input layer contains 23 nodes, composed of 10 structural parameters and 13 seismic parameters after normalization. The selection of normalization function is according to the hidden layer activation function. The number of hidden layers is 1; the number of neurons is chosen using trial and error method; two activation functions include logsig and tansig are provided; the output layer contains 1 node which is the MIDR after denormalization and select the purelin linear function as the activation function. The sample data in Section 6.3 is partitioned into training data, test data and verification data with a certain ratio. Two types of algorithms, LM and SCG, are used as training algorithms.

For each training algorithm, the number of hidden layer neurons in the network configuration and the activation function in the hidden layer are considered as variable factors. Meanwhile, MSE or correlation coefficient R is used as the performance indicators to determine the best neural network configuration, considering the optimal situation of total data set, training data set and test data set under the two performance indicators respectively, as shown is Table 6.8. Additionally, since the division of training data, test data and verification data are random, the models generated for each training under the same configuration will be different. It is iterative process for each model configuration.

TABLE 0.8 **EVALUATION CRITERIA**

	MSE	R
Total data set（Data_All）	All-mse-min	All-R-max
Train data set（Data_Train）	Train-mse-min	Train-R-max
Test data set（Data_Test）	Test-mse-min	Test-R-max

6.4.2 MATLAB CODE IMPLEMENTATION

MATLAB (2018b version) is used to establish the BP network model for this research. Take a certain model configuration (training algorithm: LM algorithm; hidden layer activation function: logsig; normalization function: [0,1] type; number of hidden layer neurons: m) as an example to introduce the training process. The code is shown in Figure 6.4.

```
%{ The network configuration
        Training algorithm—trainlm
        Hidden layer excitation function—logsig
        Normalized function—[0,1]
        Number of hidden layer neurons—m
}%

% Load raw training data
load parameters.mat
load results.mat
temp = randperm(size(parameters,1));        % Random number generation
i_training = parameters(temp(1:n),:)';      % Random generation of training data (n = total number)
o_training = results(temp(1:n),:)';

% Normalized processing
[inputs, is_input] = mapminmax(i_training,0,1);      % Training set input normalization
[targets, os_output] = mapminmax(o_training,0,1);    % The training set output is normalized

% Initialize the network configuration
hiddenLayerSize = m;               % Specify the number of hidden layer neurons
net = fitnet(hiddenLayerSize);     % Generate an initialization network

% Data set partitioning(ratio1+ratio2+ratio3=1)
net.divideParam.trainRatio = ratio1;
net.divideParam.valRatio = ratio2;
net.divideParam.testRatio = ratio3;

% Other Network Configurations
net.trainFcn = 'trainlm';                     % Specify training algorithm
net.layers{1}.transferFcn = 'logsig';         % Specify the hidden layer excitation function
net.layers{2}.transferFcn = 'purelin';        % Specify the output layer excitation function

% Training
[net, tr] = train(net, inputs, targets);
% Training result 1 - Connect weights and bias
weight1 = net.iw{1,1};                        % Input - Hidden layer weight 20*23
weight2 = net.lw{2,1};                        % Hidden layer - Output weight 1*20
bias1 = net.b{1};                             % Input - Hidden layer bias 20*1
bias2 = net.b{2};                             % Input - Hidden layer bias 1
% Training result 2- Prediction result
bodyfatOutputs = net(inputs);                 % Total data set prediction results
trOut = bodyfatOutputs(:,tr.trainInd);        % The training set predicts the results
vOut = bodyfatOutputs(:,tr.valInd);           % Validation sets predict results
tsOut = bodyfatOutputs(:,tr.testInd);         % Test sets predict results

% Inverse normalization(Take the test set for example)
Out_sim_ts = mapminmax('reverse', tsOut, os_output);             %Renormalization on test sets results
tsTarg = targets(:,tr.testInd);                                  %The original test set after normalization
Out_sim_targ_ts = mapminmax('reverse', tsTarg, os_output)        %Renormalization  on original test sets

% Performance index calculation (Take the test set for example)
ts_mse = mse(net, tsTarg, tsOut);             % Mean square error of test set MSE
ts_R = corrcoef(tsOut, tsTarg);               % Test set correlation coefficient R
```

FIGURE 0.4 **MATLAB CODE**

Step 1: Loading Data

Training data *in.mat* format are stored and loaded with the help *of.mat* method in MATLAB. The input parameter matrix is named *parameter.mat* and the output parameter matrix is named *results.mat*. The random permutation function *randperm* receives the size of the row vector and return the row vector after the numbers are randomly arranged. The newly generated row vector performs an index function, shuffling the sequence of samples based on the original training data to obtain new training data for subsequent data set division.

Step 2: Normalization

The *mapminmax* function processes the matrix by normalizing the minimum and maximum values of each row to [YMIN, YMAX]. According to the selected hidden layer activation function logsig, the [0,1] type is adopted here which means [YMIN, YMAX] = [0,1]. After processing, the normalized input matrix input $\in R^{23*n}$ and output matrix output $\in R^{1*n}$ are obtained. *Is_input* and *os_output* represent the process settings that allow consistent processing of other matrices for denormalization of subsequent output value.

Step 3: Network Configuration Initialization

The number of neurons in the hidden layer is taken as the input of the fitnet function, a 3-layer initialized BP neural network is generated in MATLAB. The brief code example is demonstrated: net = fitnet (minmax(input_data), [H1 H2], {'logsig' 'puerlin'}

Step 4: Data Set Division

The randomly arranged training data has been generated by the randperm function in step 1. It is the only need to specify the division ratio of the data set which are trainRatiov, valRatiov and testRatio respectively in this step. Then the training data, validation data and test data may be produced from the normalized input and output of training data.

Step 5: Other Network Configuration

The LM algorithm is selected as the training algorithm in this example, which is represented by trainlm in MATLAB. The hidden layer activation function is logsig and the activation function in the output layer is purelin.

Step 6: Training and training results

The normalized input matrix is input and matrix is output into the initialization BP neural network for network training.

On the one hand, the training results are expressed as the weights and biases of the neurons connected to each layer. On the other hand, it is expressed as the prediction outcomes of the training data. The weight from the input layer to the hidden layer (weight 1 $\in$ R^{24*23}) is obtained through the iw$\{1, 1\}$ property of the network. The weight from the hidden layer to the output layer (weight 2 $\in$ R^{1*23}) is obtained through the lw$\{2, 1\}$ property of the network. The bias between the two neurons (bias 1 $\in$ R^{24*1}, bias 2 $\in$ R^1) is obtained through the b property of network b{1}.

According to the division ratio in Step 4, the prediction outcomes of the training data, the verification data and the test data can be obtained from the prediction outcomes of the total data set *bodyfatoutputs.*

Step 7: Denormalization

Taking the test data as an example, the denormalization of the data still uses the *mapminmax* function in Step 2, the difference is that use "reverse" as the marked keywork. The process setting *os_output* represents the same processing method as the previous normalization to obtain the prediction result of the test data and the denormalized result of the original test data.

Step 8: Performance index calculation

Taking the test data as an example and combined with the prediction results obtained in the above steps, the MSE and the correlation coefficient R index of the test data are calculated by calling the *mse* and *corrcoef* functions.

Therefore, MATLAB software is used to build ANN models under different network configurations and training algorithm combination and the prediction effect is evaluated according to the indicators in Table 6.8. The optimal prediction model then can be selected for subsequent analysis.

6.4.3 MODEL GENERALISATION ABILITY VERIFICATION

The generalisation ability of an ANN model pertains to the capacity to effectively predict and categorise unknown data subsequent to acquiring the characteristics of a specific data set. It is one of the important indicators for measuring the quality of an ANN model, indicating its capacity to adapt to new data. On this context, this section verifies the generalisation capability of the optimally configured network obtained in Section 6.4.2. The data set has not yet appeared is relative to the original training sample. As there are two types of input parameter including structure parameters and seismic parameters, therefore considering the following three scenarios:

Scenario 1: only the structure parameter x_{struct} is different from the original sample set and structural cases that have not yet appeared in the original sample set can be considered.

Scenario 2: only the seismic parameter x_{seism} is different from the original sample set. On the one hand, the Target-PGA parameter can be adjusted. On the other hand, the parameter can be fixed and seismic wave sequence that has not appeared in the original sample set can be selected.

Scenario 3: Neither the structural parameter x_{struct} not the seismic parameter x_{seism} has ever appeared, which means the scenario 1 and 2 should be comprehensively considered here.

The above new data set still uses the nonlinear time history analysis function in the YJK programme to obtain the true value. Apart from this, the collapse fragility function of the unknown cases is considered to be developed for the first scenario. Through the comparison of the function under the ANN prediction with the actual function, the reliability of the prediction result can be explored. Meanwhile, the influence of the introduction of Target-PGA parameters on the training results of other common seismic wave parameters can also be tested.

According to the ANN earthquake response prediction model trained in Section 6.4, this research proposes a multi-scale regional earthquake damage prediction method, as shown in Figure 6.5.

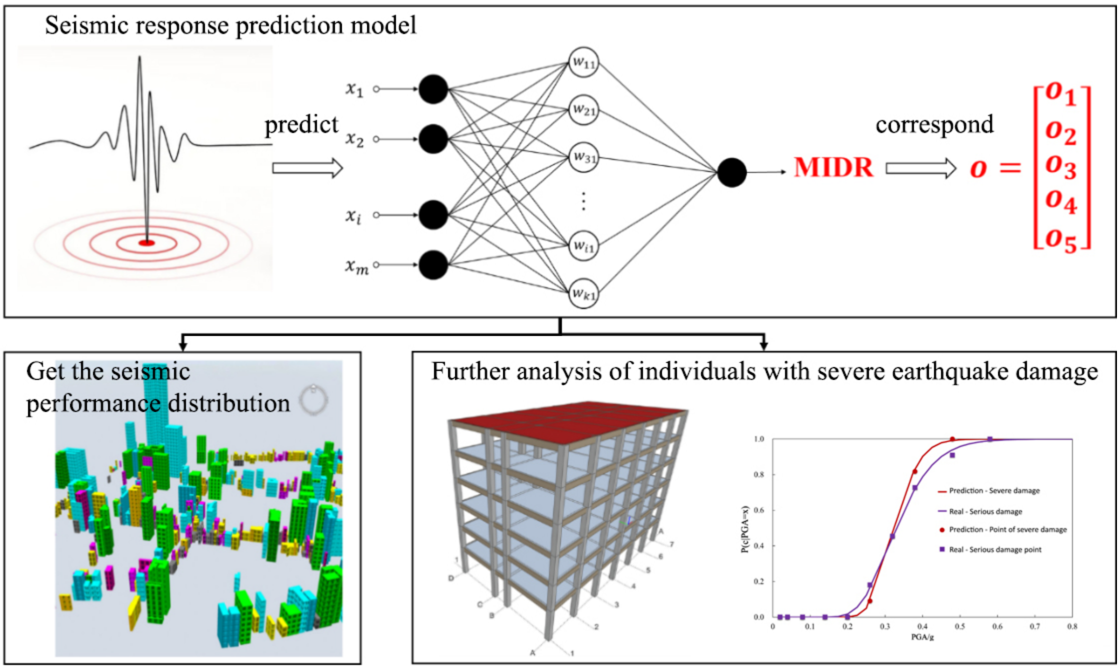


FIGURE 0.5 **METHOD OF MULTISCALE REGIONAL SEISMIC DAMAGE PREDICTION**

Stage 1, based on the ANN model developed in this chapter, for a certain area of RC building group, the seismic response prediction model can quickly obtain the MIDR of each building unit under a certain seismic level and a certain seismic wave record action.

Stage 2, since the response value is a reliable indicator reflecting the overall structure and non-structural damage situation of the RC frame building, the general damage status of the regional building group can be quickly obtained by professional designers. It can also be used as the basis for adjusting the regional precaution target and guiding post-earthquake disaster relief. According to the existing research on rules of the corresponding relationship between structural performance level and structural IDR displayed in Table 1.1, the performance level indicators of each building unit can also be quickly obtained, therefore non-professionals can also understand the distribution of earthquake damage of the entire area, such as the proportion of slightly damaged building and their locations.

Stage 3, While grasping the overall earthquake damage information of the area, urban seismic planners can also locate buildings with severe earthquake damage based on actual needs. For example, based on the research results in Chapter 4, economic performance evaluations can be performed as the basis for post-earthquake repairs and restorations. It can also be combined with the research results of Chapter 5 to comprehensively consider economy and safety in the pre-earthquake precaution and carry out more detailed performance optimisation design. In addition, as MIDR is a crucial factor in establishing the collapse fragility function of the structure, the response value can also be used as a rough substitute for the collapse fragility analysis in the case of insufficient computing resources to achieve collapse analysis of a large number of single structures.

In summary, the earthquake response prediction model based on ANN is able to realise multi-scale regional seismic design at the overall and partial levels, providing effective guidance for pre-earthquake precaution and post-earthquake disaster relief.

6.6 SUMMARY

With the help of ANN, this chapter takes the six-story RC frame structure as the research object. After comprehensively considering the overall and partial characteristics of the building structure, 10 representative structural parameters are selected. Taking into account the characteristics of ground motion, including amplitude, frequency spectrum and duration, and seismic wave amplitude modulation factors, 13 seismic parameters were selected. The effect of the aforementioned 23 parameters on the MIDR prediction is studied. Furthermore, different ANN network configurations (number of hidden layer neurons, activation function, and normalization function) and a combination of two training algorithms (LM and SCG) are adopted in order to obtain a more ideal prediction effect. The training obtains a model that enables prediction of structural response of the RC frame building to any seismic wave after a series of amplitude modulations and verifies the generalisation ability of the model with greater accuracy. Finally, a method for multi-scale regional earthquake damage prediction is proposed. This method combines the established ANN earthquake response prediction model, starting from the overall and partial levels of the region, on the one hand, it can efficiently assess the extent of earthquake damage to a cluster of buildings. on the other hand, it realises the rapid analysis for the collapse situation and adopts more detailed performance design for the more severely damaged building unit.

CHAPTER 7 CASE STUDIES

As this thesis is originated from a research project cooperated by Shanghai Urban Construction Design and Research Institute and Sichuan Provincial Government for the aim to study how to predict earthquake damage in Sichuan area, two case studies (will be introduced in detail in the following Section) which are all provided by Sichuan Provincial Government to determine whether the proposed methods work as intended and whether the ontology knowledge base can yield meaningful results. Moreover, all building cases existing in this thesis are real projects in Sichuan Province. In Sections 7.1 and 7.2, respectively, two structural design implementations for individual buildings and building groups are demonstrated.

7.1 CASE APPLICATION OF PERFORMANCE-BASED SEISMIC DESIGN OF INDIVIDUAL BUILDING

This section takes a typical RC frame structure as an example, on the first step, seismic performance evaluation approach using BIM and ontology proposed in Chapter 4 is applied to the performance evaluation process under the guidance of FEMA P-58, conducting four intensity-based assessments under the action of earthquake. Then, BIM-based multi-objective seismic optimisation design method of RC frame proposed in Chapter 6 is applied to this case and the seismic optimisation design from the component level is realized.

7.1.1 PROJECT OVERVIEW

A new building is a six-story cast-in-place RC frame structure, which functions as an office building. It is a real project whose information is provided by Sichuan Provincial Government. The specific design information is as follows: the project covers an area of 960.96 m^2, the total floor area is about 5765.76 m^2, the total height is 22.7m, the floor height is 3.6m except the ground floor height is 4.7m and the plane size is 18.3m * 52.0m. Every level of the building is designed to be a typical floor and the layout of the building is shown in Figure 7.1. The horizontal AB axis and CD axis frame beam's cross section is 300mm * 700mm, the longitudinal 7-8 axis frame beam's cross section is 300mm *700mm, the longitudinal 5-6 axis, 6-7axis, 8-9 axis and 9-10 axis frame beam's cross section is 250mm *400mm, the rest of the frame beams' cross section are all 250mm * 450mm and frame columns' cross section are all 500mm * 500mm. Frame beams, columns, floors and roof slabs are all cast-in-situ, slab thickness is 100mm, concrete adopts type C30, it is assumed that all steel bars are all the same everywhere that longitudinal reinforcement adopts type HRB335 and stirrup adopts type HPB235, outer wall is 240mm thick infill wall, inner wall's thickness is 120mm, non- accessible roof's dead and live loads are 4.38 kN/m^2 and 0.5 kN/m^2 respectively while the dead and live loads on the floor are 3.24 kN/m^2 and 2.5 kN/m^2 respectively. The floor slab is a load-bearing transverse element of a building, supporting the floor load and lateral forces. It functions to absorb and convey seismic during an earthquake. Nonetheless, as comparison to longitudinal structures such as beams and columns, its load-bearing capacity is relatively small and its damage will not have a significant impact on the overall reliability of the building structure. Therefore, slabs are not modelled and considered in this thesis.

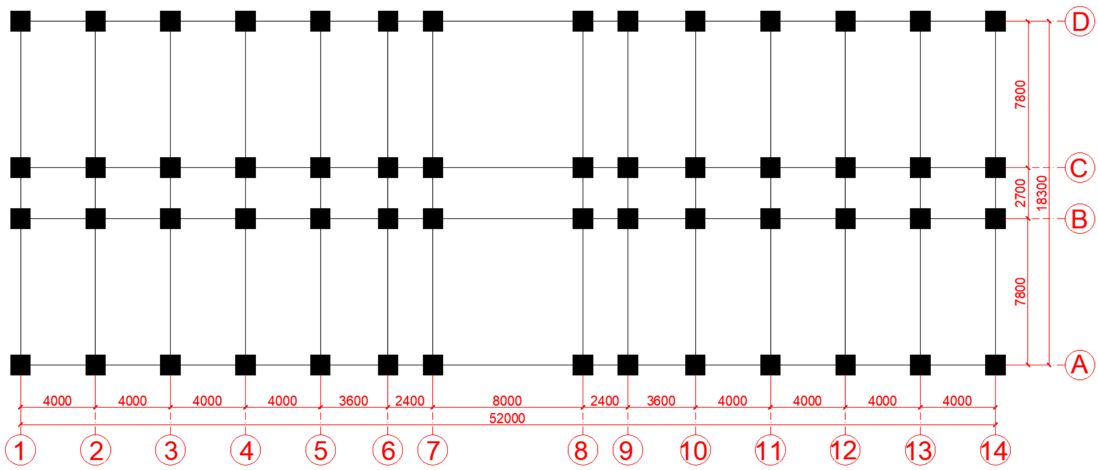


FIGURE 0.1 **BUILDING'S DESIGN SCHEME LAYOUT**

The studied building belongs to category C structure and the detailed seismic information is expressed as follows: the seismic precautionary intensity is 7 degrees, the design basic seismic acceleration value is 0.1g, the site category is category II, the design seismic group is classified as the second group, and the site characteristic period is 0.4s. The seismic analysis of the structure is conducted in the YJK software, the period reduction method is used to consider the impact of the infill wall on the seismic resistance of the frame, and the reduction factor of 0.6 is assumed (Guo, 2012). The IDR under frequent earthquakes is 1/777, the maximum axial compression ratio of the bottom column is 0.49, and the basic period of the structure's first mode is 1.0843s, all of which meet the requirements of the regulations.

7.1.2 AUTOMATED SEISMIC PERFORMANCE EVALUATION

Structure's Performance Model Development

1. acquisition of fundamental building information

Firstly, develop a three-dimensional information model of the building in Revit 2016, including frame beam, column components, floor slabs, retaining walls and internal partition walls. Then IFC files is exported. In the python 3.7 environment, preprocess the IFC file to extract the fundamental details of the component, the floor information, and then determine the neighbouring connection between the beam and the column. Generate the aforementioned data in a unified JSON format, which is displayed in the column information of Figure 7-2a) and in the inner partition wall information of Figure 7-2b).


```
"333": {
    "GlobalId": "3PdLQN5GjCwfz7HiGF7eIf",
    "NameofColumn": "ConcreteColumn:600*600mm:328967",
    "TheFloorInfo.": [
        "TheFifthFloor"
        18000.0
    ],
    "Size": {
        "Volume": 1.296,
        "Length": 3600.0
        "Area": 4.68
    },
    "TopologicalRelationship": {
        "col_ConnetedTo_Beam_x": [
            "3PdLQN5GjCwfZ7HiGF7eWU",
            "3PdLQN5GjCwfZ7HiGF7eXq"
        ],
        "col_ConnetedTo_Beam_y": [
            "3PdLQN5GjCwfZ7HiGF7eZq",
            "3PdLQN5GjCwfZ7HiGF7eZO"
        ],
        "col_ConnetedTo_Beam_x_num": 2,
        "col_ConnetedTo_Beam_y_num": 2,
        "col_ConnetedTo_Beam_x_max": 450.0,
        "col_ConnetedTo_Beam_y_max": 700.0
    }
}
```


a) Parsed column information


```
"46": {
    "GlobalId": "0a26ZMNbXErREbybazuaFg",
    "NameofWall": "BaseWall: Internal - InfillWall - 120mm:339506",
    "TheFloorInfo.": [
        "TheFirstFloor"
        3600.0
    ],
    "Size": {
        "Volume": 3.1104,
        "Length": 7200.0
        "Area": 25.92
    },
    "Structure": {
        "IsExternal": false
        "LoadBearing": false
        "ExtendToStructure": false
        "Reference": "Internal - InfillWall - 120mm"
    },
    "Direction": "x"
}
```


b) Parsed interior partition information

FIGURE 0.2 **RESULTS OF IFC FILE PREPROCESSING**

The structure model information is parsed and stored in the ontology through instantiating, with taking the keyword "key" in JSON format as the property name and taking the value as the property value.

According to FEMA P-58's assessment of the potential elements and contents found in a standard office building, this study derives an estimated inventory of the categories and amounts of components other than the known components in the selected cases and instantiates them in the ontology. The instantiation of non-structural component information like plumbing, heating, electricity, flooring and pendant lights is depicted in Figure 7.3.

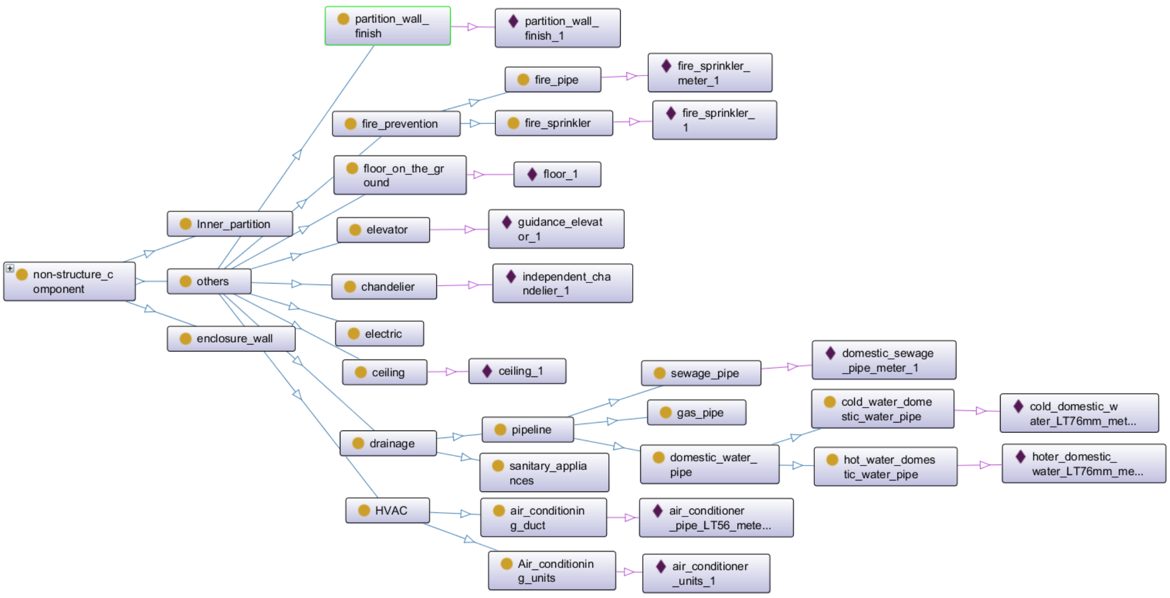


FIGURE 0.3 **INFORMATION ABOUT NON-STRUCTURAL COMPONENTS**

Fragility group classification and other necessary information assumptions of the components in the new building, including structural components (beam-column nodes), non-structural components (floor, external retaining wall, internal partition wall, partition wall finish) and equipment (water, heating and electricity engineering project), etc., summarized as shown in Table 7.1.

TABLE 0.1 **INFORMATION ABOUT FRAGILITY GROUPS IN THE CASE**

No. Of fragility group	description	Unit	EDP
B1041.031a			
B1041.031b			
B1041.032a	Beam-column joint of ordinary frame	/	
B1041.032b			
B1041.033a			IDR
B1041.033b			
B2011.101	Non-structural exterior wall	9.29m^2	
C1011.001a	Lightweight gypsum partition wall	30.48m	
C3011.001a	Partition wall finish	30.48m	
C3027.001	floor	9.29 m^2	
C3032.001a	Suspended ceiling, only vertical support	23.225 m^2	PFA
C3034.001	Independent lighting chandelier	/	

No. Of fragility group	description	Unit	EDP
D1014.011	Traction elevator	/	
D2021.011a	Cold water pipe	304.8m	
D2022.011a	Hot water pipe with small diameter	304.8m	
D2031.011b	Domestic sewage pipeline	304.8m	
D3041.011a	HVAC pipes	304.8m	
D4011.021a	Fire pipes	304.8m	
D4011.031a	Fire sprinkler	/	
D3052.011a	Air conditioner group	113.27m³	

2. performance cluster division

The categorisation of performance clusters is carried out by means of SWRL rules, taking the beam-column node as an example shown in Figure 7.4. Based on the fragile cluster regulations in FEMA P-58, the beams and column components of RC frame are divided by nodes. The node column component named *ifcColumn_185* is classified to fragile group B1041.031a and B1041.031b respectively in the x-direction and y-direction based on the quantity and dimensions of neighbouring beams. Subsequently, it is categorized into several performance clusters according to the floor information where the component is located, and the performance cluster information of nodal beam is the same as those of the column component. Table 7.2 displays the corresponding SWRL rules. The RDF graph of the division result obtained by reasoning in the ontology is shown in Figure 7.5. The column instance and the corresponding performance group are directly related through the *has_pg_info* and *pg_has_comp* object properties to make necessary preparations for the subsequent component-level damage prediction.

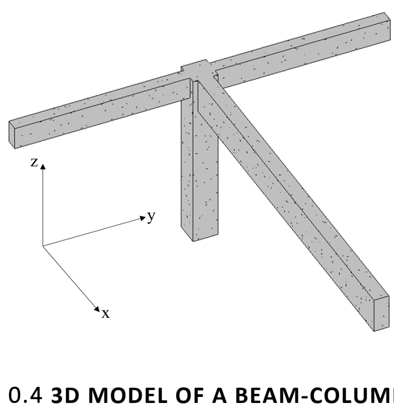


FIGURE 0.4 **3D MODEL OF A BEAM-COLUMN NODE**

TABLE 0.2 **SWRL RULES FOR FRAGILE CLUSTER AND PERFORMANCE CLUSTER CLASSIFICATION**

// B1041.031a fragile cluster classification
column(?col)^col_ConnectedTo_Beam_x_num(?col,"1"^^xsd:integer)^ col_ConnectedTo_Beam_x_max(?col,?beam_size)^col_ConnectedTo_Beam_x(?col,?beam_x)^ swrlb:lessThan(?beam_size,"609.6"^^xsd:double)^B_Shell(?fg)^hasValue(?fc,"B1041.031a"^^xsd:string) ->has_fc_info_x(?col,?fc)^has_fc_info_x(?beam_x,?fc)^has_comp_x(?fc,?col)
column(?col)^col_ConnectedTo_Beam_y_num(?col,"2"^^xsd:integer)^ col_ConnectedTo_Beam_y_max(?col,?beam_size)^col_ConnectedTo_Beam_y(?col,?beam_x)^ swrlb:lessThan(?beam_size,"609.6"^^xsd:double)^B_Shell(?fc)^hasValue(?fc,"B1041.031b"^^xsd:string) ->has_fc_info_y(?col,?fc)^has_fc_info_y(?beam_y,?fc)^has_comp_y(?fg,?col)
//performance cluster classification- x direction
B1041_031a (?fc) ^ has_comp_x(?fc, ?comp)^containedIn_Story(?comp,?story) ^ Type of performance cluster (?pc) ^ pc_story(?pc, ?story) ^pc_direction(?pc, "x"^^xsd:string) -> has_pc_info (?comp, ?pc) ^ pc_has_comp(?pc, ?comp)
//performance cluster classification- y direction
B1041_031b (?fc) ^ has_comp_x(?fg, ?comp)^containedIn_Story(?comp,?story) ^ Type of performance cluster (?pc) ^ pc_story(?pc, ?story) ^pc_direction(?pc, "y"^^xsd:string) -> has_pc_info (?comp, ?pc) ^ pc_has_comp(?pc, ?comp)

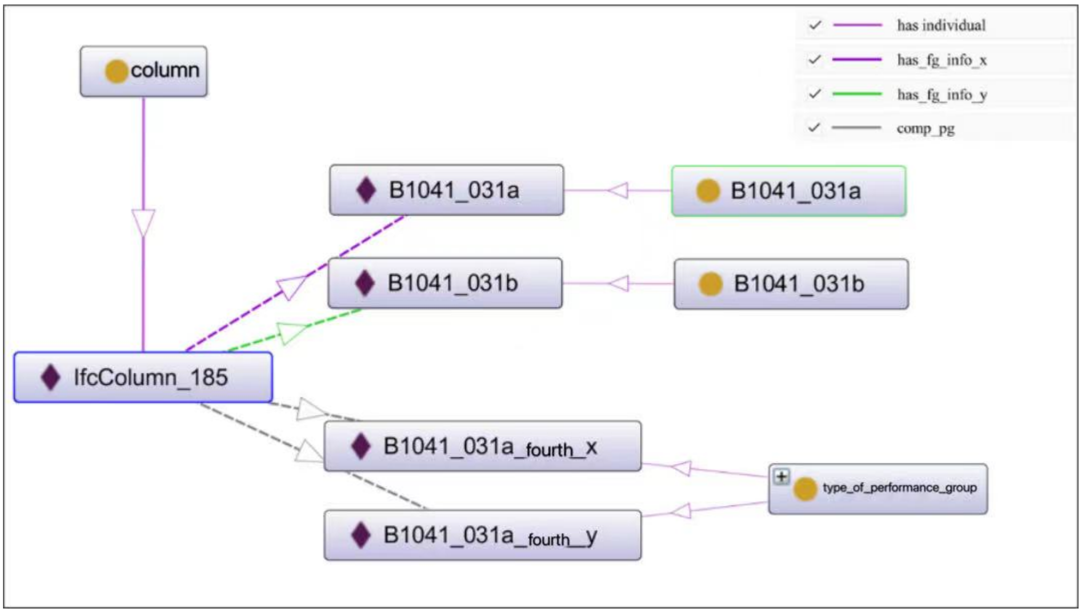


FIGURE 0.5 **RESULTS OF PERFORMANCE CLUSTER CLASSIFICATION**

Earthquake Risk Analysis

1. determination of target response spectrum

Based on the "*China Earthquake Parameter Zoning Map*" (*GB 18306-2015*), this case study has determined the four intensity levels of ground motion parameters for frequent earthquakes, basic earthquake, rare earthquakes and extremely rare earthquakes in conjunction with the characteristics of the construction site. The acceleration response spectrums are drawn corresponding to the four-magnitude earthquake level, as shown in Figure 7.6. The horizontal segment of the acceleration response spectrums represents the short-period part, in which the response of the structure to earthquake is mainly controlled by acceleration. The end point of the horizontal segment is defined as characteristic period.

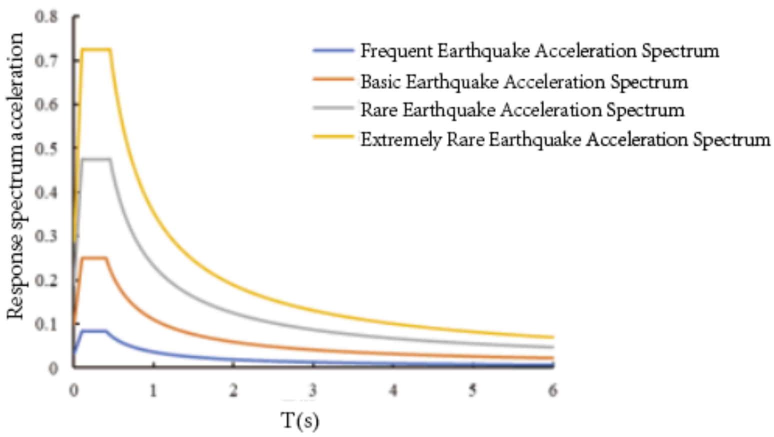


FIGURE 0.6 **ACCELERATION SPECTRUM**

2. seismic wave selection and amplitude modulation

In this case, $S_a(T_1)$ under 5% damping is chosen as the ground motion intensity index and the reference number of ground motion amplitude modulation.

In the FEMA P-695 report of ATC-63, there are 50 strong earthquake records suitable for structural response analysis, which are widely used in the evaluation of structural collapse resistance and IDA analysis (Yan, 2019). This case takes the design response spectrum as the desired spectrum, matches the desired response spectrum with the primary period of the structure and selects 11 ground motions for further investigation of response. The eleven pieces of ground motion record information and acceleration response spectrum are displayed in Table 7.3 and Figure 7.7.

TABLE 0.3 INFORMATION OF 11 PIECES OF EARTHQUAKE WAVES

No.	File Name	location/time/earthquake name
EQ1	169_IMPVALL.H\H-DLT262	Delta, 10/15/1979, Imperial Valley-06
EQ2	1116_KOBE\SHI000	Shin-Osaka, 1/16/1995, Kobe, Japan
EQ3	725_SUPER.B\B-POE270	Poe Road (temp), 11/24/1987, Superstition Hills-02
EQ4	752_LOMAP\CAP000	Capitola, 10/18/1989, Loma Prieta
EQ5	960_NORTHR\LOS000	Canyon Country- W Lost Cany, 1/17/1994, Northridge-01
EQ6	829_CAPEMEND\RDL270	Rio Dell Overpass- FF, 4/25/1992, Cape Mendocino
EQ7	953_NORTHR\MUL009	Beverly Hills- 14145 Mulhol, 1/17/1994, Northridge-01
EQ8	68_ SFERN\PEL090	LA- Hollywood Stor FF, 2/9/1971, San Fernando
EQ9	1158_KOCAELI\DZC180	Duzce, 8/17/1999, Kocaeli, Turkey
EQ10	1244_ CHICHI\CHY101-E	CHY101, 9/20/1999, Chi-Chi, Taiwan
EQ11	1602_ DUZCE\BOL000	Bolu, 11/12/1999, Duzce, Turkey

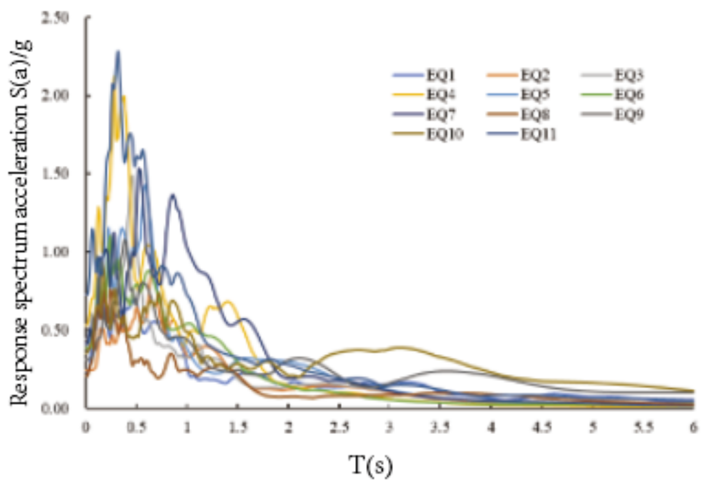


FIGURE 0.7 **ACCELERATION RESPONSE SPECTRUMS OF 11 PIECES OF EARTHQUAKE WAVES**

The target spectrum and the spectral acceleration value S_a (T_1) of individual seismic wave response spectrum are obtained during structure's basic period, and then each seismic wave amplitude is modulated according to the single-point amplitude modulation technique (Lv, Liu, LEE, & Yu, 2018). Figure 7.8 displays the contrast between the average response spectrum after amplitude modulation and the standard response spectrum under the unusual earthquake level. The average response spectrum of ground motion closely resembles the form of the standard design spectrum based on visual assessment, so the selected ground motion is reasonable and can be used for follow-up analysis.

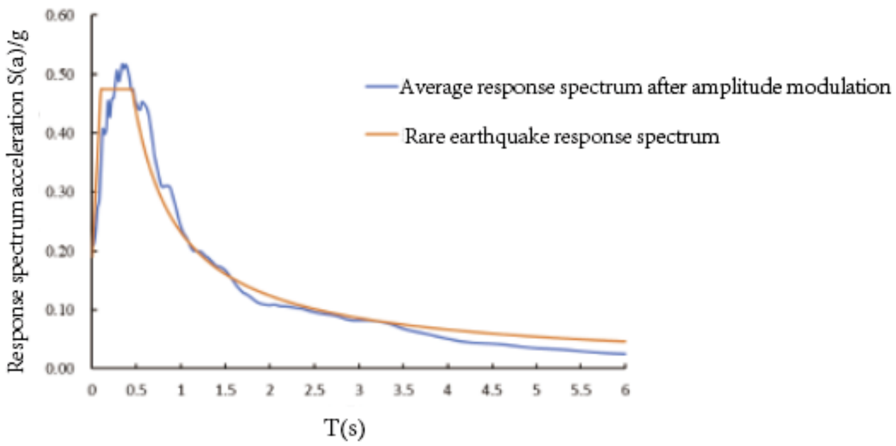
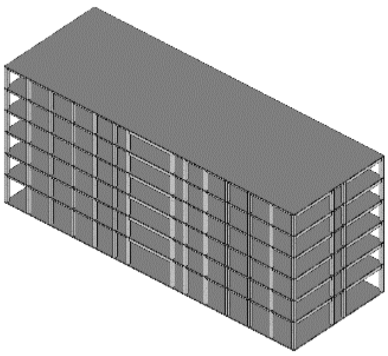


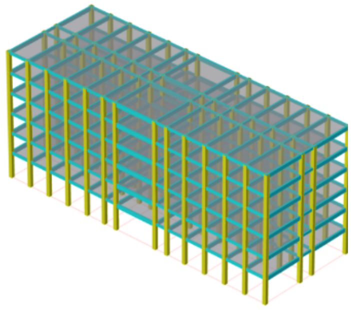
FIGURE 0.8AVERAGE ACCELERATION SPECTRUM CURVE FROM AMPLITUDE MODULATION TO RARE EARTHQUAKE LEVEL

Structural Response Analysis

Select structural elements like beams and columns in the Revit structural model in Figure 7.9 a), use the REVIT – YJKS module to export the sub-model .ydb file and import it into the YJK model programme to create a structural analysis model in Figure 7.9 b).



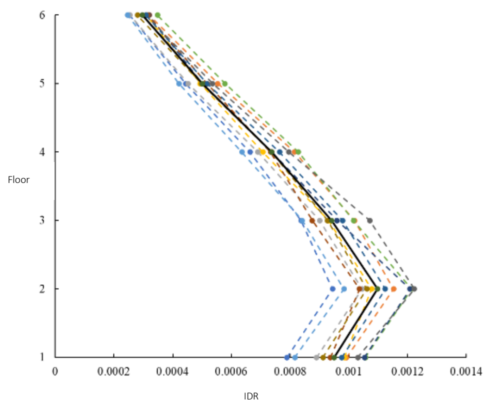
a) Structure model in Revit



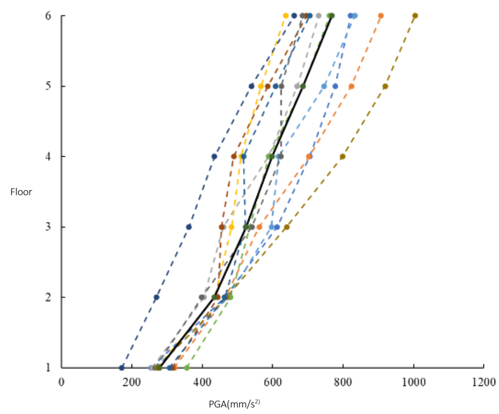
b) Structural analysis model in YJK

FIGURE 0.9 **REVIT PHYSICAL MODEL AND YJK ANALYSIS MODEL**

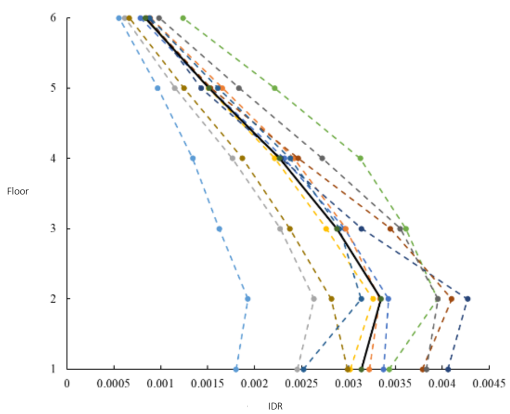
Subsequently, eleven ground motions were respectively adjusted to four seismic intensity degrees. Elasto-plastic time history analysis was conducted on the practical model in YJK programme to determine the demand parameters including MIDR and PFA. The distribution of 11 ground motion responses and the average response distribution are depicted in Figure 7.10, where dotted dashed lines represent the structural response envelope induced by a specific ground motion, whereas the black lines represent the average response of all chosen ground motions. For frequent occurrence of earthquakes, MIDR and PFA are relatively small and the values of them are concentrated which are close to the mean value. Therefore, structural components and enclosure/ partition members of frame structure may show commendable seismic performance in this case. However, for extremely infrequent seismic events, the values of MIDR and PFA increase significantly that the beam-column joints of the structure may began to be damaged.



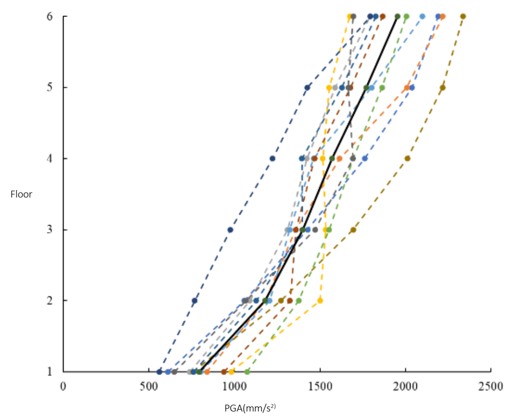
a) Maximum inter-story drift ratio in frequent earthquake



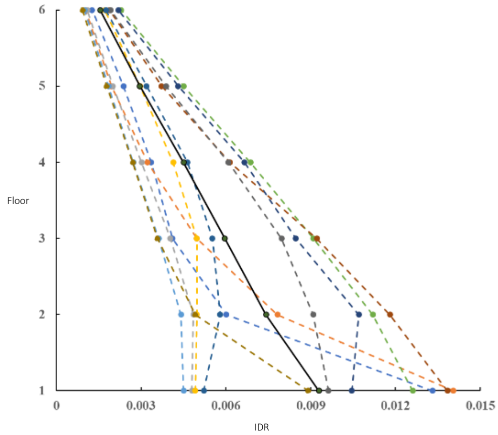
b) Peak floor acceleration in frequent earthquake



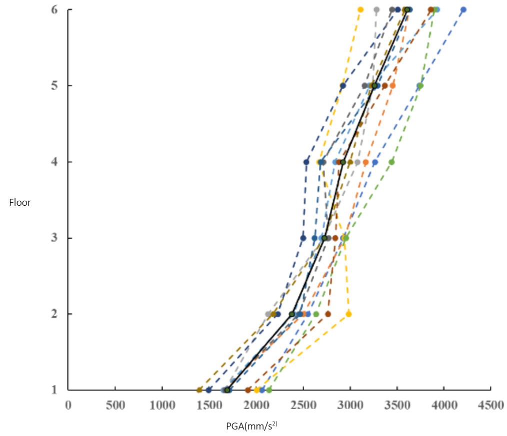
c) Maximum inter-story drift ratio in basic earthquake



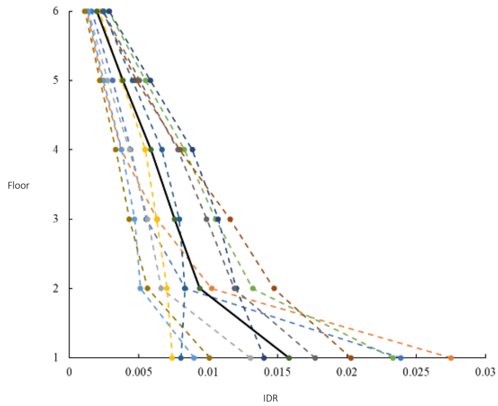
d) Peak floor acceleration in basic earthquake



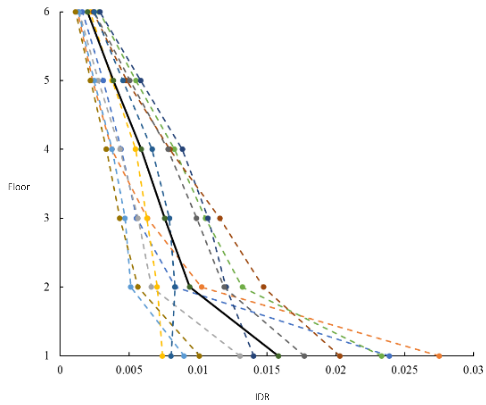
e) Maximum inter-story drift ratio in rare earthquake



f) Peak floor acceleration in rare earthquake



g) Maximum inter-story drift ratio in extremely rare earthquake



h) Peak floor acceleration in extremely rare earthquake

FIGURE 0.10 **SEISMIC RESPONSE DISTRIBUTION**

Analysis of the Overall Structure Fragility

The present requirements describe a link between performance levels and the MIDR between floors for determining the total performance level of RC frame building. The performance levels may be categorised as follows: basically intact with MIDR 1 in 550, slight damage with MIDR 1 in 250, moderate damage with MIDR 1 in 120, serious damage with MIDR 1 in 50. When the IDR between floors is greater than 1/50, the structure is considered to be destroyed and collapse occurs. The IDA approach is utilised to calculate the probability distribution of the structure with different performance levels, so as to produce the performance curve with different performance levels. Specifically, for each seismic wave, combined with the selection of ground motion intensity in this case, the amplitude is modulated to $S_a(T_1)$ to be 0.02g, 0.034g, 0.05g, 0.075g, 0.102g, 0.15g, 0.215g, 0.275g, 0.329g, 0.35g, 0.45g, 0.55g, 0.65g, 0.75g. Input the amplitude-modulated ground motion into the structural analysis model for elasto-plastic time history until the structure collapses (the MIDR reaches

1/50) and the IDA curve of $S_a(T_1)$ and the IDR are obtained. Repeat 11 seismic waves to get 11 IDA curves, as shown in Figure 7.11.

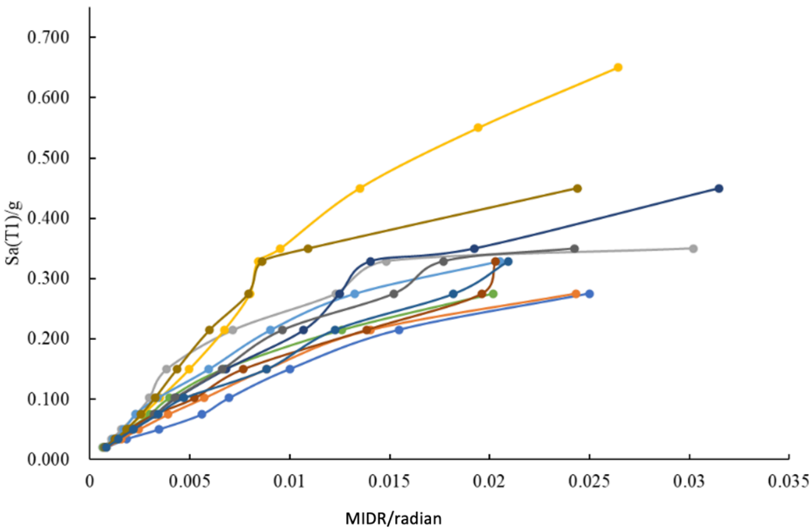


FIGURE 0.11 **IDA CURVE**

Based on the IDA curve obtained in Figure 7.11, the collapse fragility tool is used for probability statistical analysis and the fragility function of each performance level with a log-normal distribution is fitted, which represents the function of the probability of reaching each performance level and the earthquake intensity $S_a(T_1)$, as shown in Figure 7.12. The fitting parameters of the normal distribution are shown in Table 7.4. Subtract the exceeding probability values corresponding to the ground motion from the four curves in turn, the likelihood of a building collapsing due to a certain ground motion may be obtained. It can also be directly regarded as the probability value corresponding to the severe damage curve.

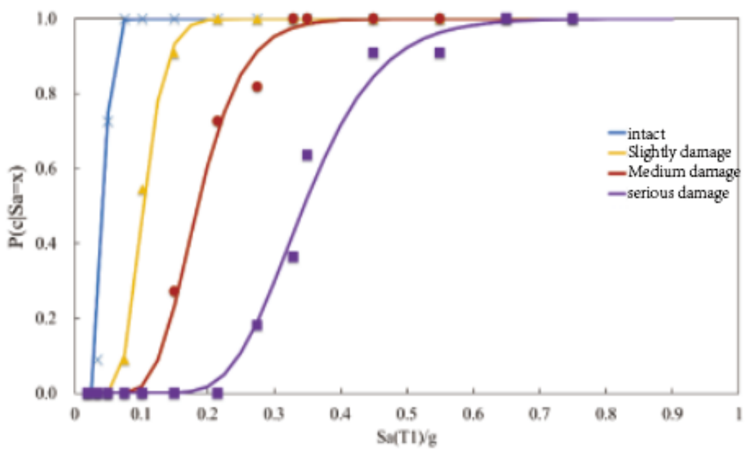


FIGURE 0.12 **FRAGILITY CURVE**

TABLE 0.4 **THE NORMAL DISTRIBUTION FITTING PARAMETERS OF FRAGILITY CURVES AT EACH PERFORMANCE LEVEL**

Basically intact		Slightly damage		Medium damage		Serious damage	
Mean(g)	Standard deviation	Mean(g)	Standard deviation	Mean(g)	Standard deviation	Mean(g)	Standard deviation
0.044	0.19	0.103	0.25	0.185	0.29	0.345	0.26

Strength-based Assessment

This case study analyzes the real loss data, which does not consider the residual displacement fragility and only considers the maintenance loss cost caused by the earthquake. Four simulations of maintenance cost evaluation under the intensity of ground motion for the selected cases are carried out and the number of simulations is taken as 1000. For each simulation, the Monte Carlo method of generating artificial EDP matrix in the Yang's research (2009) is used to randomly generate a possibility from the set of structural responses under the action of 11 seismic waves.

1. Prediction of component-level damage

After each possible building structure response demand value is generated, the damage status, damage probability and damage consequences of various performance clusters can be calculated according to the fragility curve and consequence function in the fragile cluster regulations. Then, the fragility analysis results are reversely mapped back to the component through the performance cluster reverse mapping. It is incorporated into the fragility component of the application ontology and combined with the component-level damage prediction formula proposed in this research, the ultimate failure state and failure ratio of the component are obtained. The final damage consequence of the component is expressed as the product of the damage ratio, the unit loss data (repair valuation) and the unit of measurement. The performance index of the building (total maintenance cost) is obtained by summing up all damage consequence value. The performance groups reverse mapping is expressed and inferred through the SWRL rules proposed in Section 4.3.3 and the final mapping result is displayed in Figure 7.13.

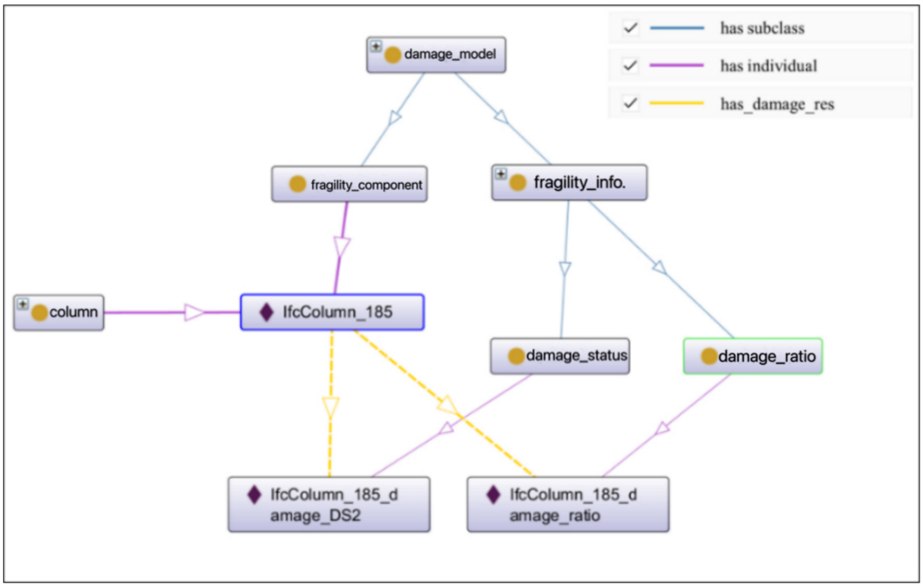


FIGURE 0.13 **DAMAGE RESULTS OF VULNERABLE COMPONENTS**

2. Repair pricing information

The component damage ratio refers to the proportion of the cost required to restore a component to its original condition, in relation to the overall construction cost, so the unit loss data corresponds the construction cost. As stated in *initial building expense calculation part* in Section 5.3.2, Construction cost is calculated using the integrated unit price method ((N. Zhang, 2015), which includes all costs such as material costs, labor costs, machinery costs, regulatory fees, profits and taxes, as shown is Table 7.5 and 7.6 (CNY is the symbol of Chinese yuan). For structural components (beams and columns), new construction *cost a = concrete unit price * component volume + steel unit price * steel weight + formwork unit price * component surface area*. For non-structural components such as walls, new construction *cost a = new component unit price* number of components contained in the component cluster*. Considering the uncertainty of the cost information, the other components not shown in the table are priced according to the repair cost in FEMA P-58 (in US dollars) and converted to CNY based on the US dollar exchange rate in December 2024 (1 US dollar = 7.33 CNY, 1 GBP = 9.15 CNY).

TABLE 0.5 **INTEGRATED UNIT PRICES OF STRUCTURAL COMPONENTS**

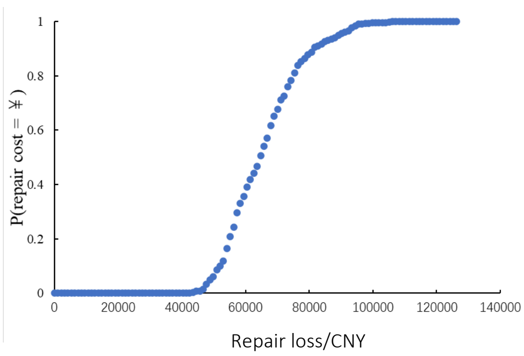
Component Type	Concrete project（CNY/m³）			Concrete pumping（CNY/m³）	Rebar project（CNY/t）		Form work project（CNY/㎡）
	C30	C35	C40		HPB300	HRB400	
Beam	455	465	475	20	5116	5156	50
Column	500	510	520	20	5116	5156	50

TABLE 0.6 **COMPREHENSIVE UNIT PRICES OF NON-STRUCTURAL COMPONENTS**

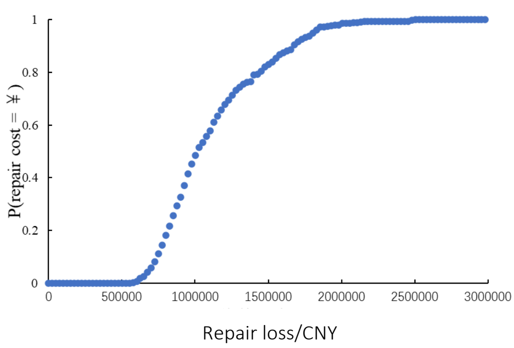
Project	Infill wall（CNY/m³）	elevator（CNY）
Comprehensive unit price	423	200,000

3. evaluation results and analysis

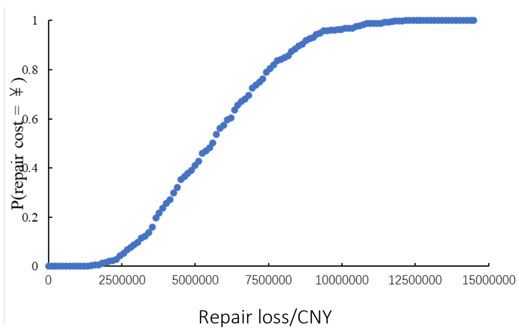
This study conducts four simulations of maintenance cost evaluation under the ground motion intensity for the selected cases, the number of simulations is 1000 and the result are displayed in Figure 7.14. The cost of loss and the exceeding probability show a normal distribution curve. In order to explore the loss distribution of individual performance cluster, take the particular loss value that corresponds to the P = 50% point as a sample. Then, compare the expenses associated with repairing structural components and non-structural components when subjected to various action levels, as shown in Figure 7.15.



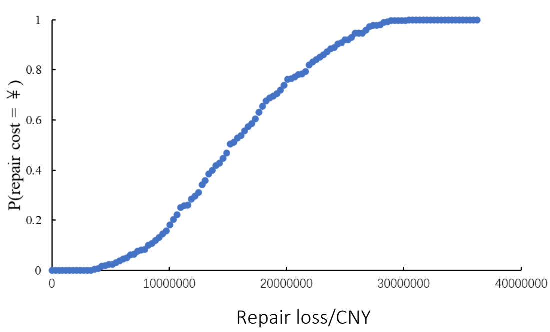
a) Performance function of frequent earthquake



b) Performance function of basic earthquake



c) Performance function of rare earthquake



d) Performance function of extremely rare earthquake

FIGURE 0.14 **PERFORMANCE FUNCTIONS OF MAINTENANCE COST UNDER FOUR GROUND MOTION INTENSITIES**

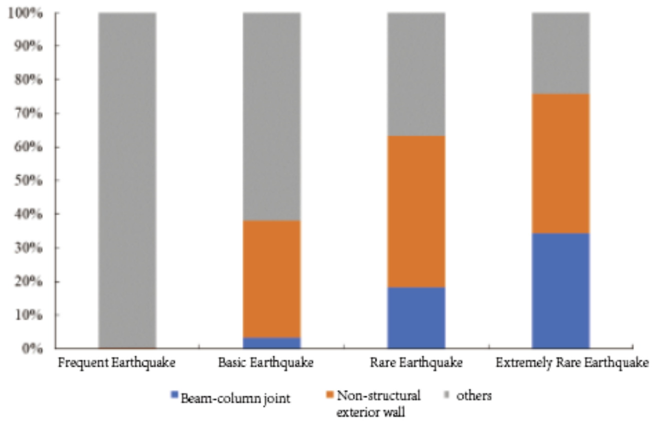


FIGURE 0.15 **PERFORMANCE CLUSTERS' LOSS RATIO**

This study divides all component into three parts according to their functions: structural components (beam-column nodes), enclosure/partition components (interior and exterior walls) and others (floor, ceiling and auxiliary components, pipelines, large equipment, elevators, etc.). The structural components and enclosure/partition members of frame structure in this case showed commendable seismic performance due to frequent occurrence of earthquakes. The maintenance expense was incurred due to damage to other components such as decorative equipment and the total maintenance cost was 64565 CNY. Due to the seismic activity, the enclosure/partition components experienced degradation as well, resulting in a maintenance expense of 35% and a total expense of 1,014,858 CNY. The beam-column joints of the structure began to be damaged as a result of a rare earthquake, and the total maintenance cost was 5583789 CNY. This phenomenon occurs due to the escalation of ground motion intensity, where structural components begin to exert influence, resulting in more severe damage. On the contrary, the proportion of repair costs for non-structural members is getting smaller. Under the extremely infrequent seismic events, the cumulative expenditure for maintenance amounted to around 15.19 million CNY.

In this instance, the expense of the multi-story building frame construction is calculated to be 1800 CNY/m^2, then the replacement costs are about 5765.75 * 1800 = 10.38 million CNY. It can be seen that the loss ratio (the ratio of earthquake loss to replacement cost) under the action of the four earthquake levels is 0.62%, 9.78%, 53.59% and 146.34% respectively. According to the damage loss ratio of structure summarised by Zheng's research, it explains that the overall seismic performance of the structure is that the "building is basically intact due to frequent earthquake, building is slightly damaged under the action of basic earthquake, building is severely damaged due to rare earthquake, and building is completely damaged under the action of extremely earthquake.” (Zheng, Xiang & Zheng, 2016). At the same time, the minimum requirements of “three levels” in regulations should be met. The outcomes are displayed in Table 7.7.

TABLE 0.7 **LOSS RATIO INFORMATION**

Earthquake level	Loss Ratio	Loss Ratio Range	If meet the requirement
Frequent Earthquake	0.62%	[0, 5%)	yes
Basic Earthquake	9.78%	[5%,40%)	yes
Rare Earthquake	53.79%	[40%, 70%)	yes
Extremely Rare Earthquake	146.34%	—	—

7.1.3 MULTI-OBJECTIVE SEISMIC OPTIMISATION DESIGN

Seismic Related Information

1. seismic information

Based on the fundamental seismic data of this instance and taking into account the requirements of other interested parties, the designers preset that the structure should adhere to the minimal seismic performance standards of “three levels”: building would be kept in original condition when small earthquake occurs; building can be repaired when moderate earthquake occurs; and building would not collapse when big earthquake occurs. Judging from the performance evaluation results in Section 7.1.2, the extremely rare earthquakes cause excessive maintenance losses, which easily affect the loss levels under other earthquake levels. In addition, due to the low possibility of earthquake occurring at this magnitude compared to other types of earthquakes, they are not considered in the optimisation process.

Due to the difference structural periods of the same type of buildings with different configurations, individual seismic level of the section should be characterized by the PGA and determined based on the “*China Earthquake Parameter Zoning Map*” (GB 18306 – 2015). The annual probability of exceeding at each level is calculated according to Formula 7-1:

$$P_{annual} = 1 - (1 - P)^T \qquad (7\text{-}1)$$

T represents the design base period; P represents the probability of exceeding in the design base period. If there are frequent earthquake, the probability of exceeding in 50 years is 63%. The relevant information of each earthquake level is listed in Table 7.8.

TABLE 0.8 **EARTHQUAKE LEVEL INFORMATION**

Earthquake Level	PGA	Exceed Probability	Annual Exceed Probability
Frequent Earthquake	0.033g	50 years 63%	1.97%
Basic Earthquake	0.100g	50 years 10%	0.21%
Rare Earthquake	0.190g	50 years 2%	0.04%

According to the description in Section 5.3.2, the performance curve encloses the region representing the average yearly repair expense caused by potential earthquake disasters. Due to the limited computing resources, this study proposes to approximate the area encircling the curve with three seismic levels of frequent earthquakes, precautionary earthquakes and rare earthquakes as the base points, which is displayed in Figure 7.16. The approximate area consists of three parts, calculated as follows:

$$S = S_1 + S_2 + S_3 \qquad (7\text{-}2)$$

$$S_1 = C_1 P_{annual1} \qquad (7\text{-}3)$$

$$S_2 = \frac{1}{2}(P_{annual1} + P_{annual2})(C_2 - C_1) \qquad (7\text{-}4)$$

$$S_3 = \frac{1}{2}(P_{annual2} + P_{annual3})(C_3 - C_2) \qquad (7\text{-}5)$$

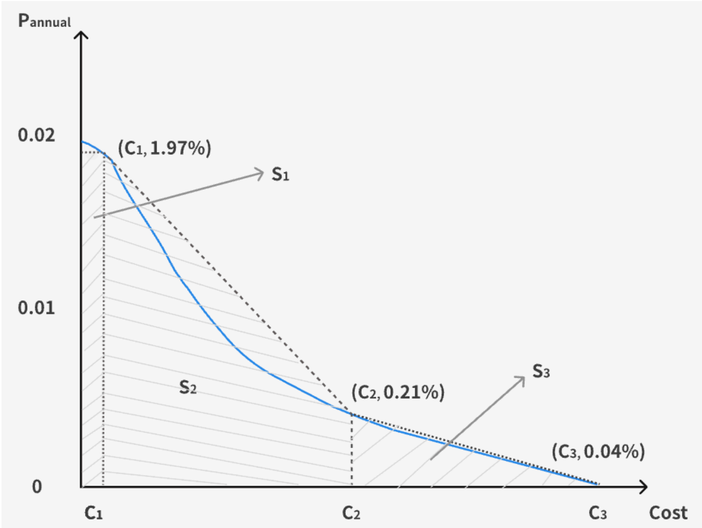


FIGURE 0.16 **APPROXIMATE CALCULATION OF AVERAGE ANNUAL RESTORATION COST**

$P_{annual1}$, $P_{annual2}$, $P_{annual3}$ are the annual probability of exceeding at the three seismic levels, which are 1.97%, 0.21% and 0.04%. C_1, C_2, C_3 represents the seismic maintenance losses under the three seismic levels. When the earthquake level is less than the frequent earthquakes, the annual probability of exceeding is relatively great, the loss cost is relatively low, and the curve range is relatively small. It can be approximated by the maintenance loss caused by frequent earthquakes, represented by S_1. When the seismic level is between frequent earthquakes and fortified earthquakes, the curve section is concave and relatively long. It is approximated by the trapezoidal area enclosed by two seismic level situations, represented by S_2. When the seismic level is between the fortified earthquake and the rare earthquake, the average annual repair cost can also be estimated by the trapezoidal area enclosed by the two seismic level situations. As the annual probability of exceeding is small, the concavity of the curve section becomes smaller and almost a straight-line decline. This area is named as S_3. When the earthquake level is greater than that of a rare earthquake, considering that the probability is extremely small, the area enclosed by curve is relatively small which can be ignored when calculated.

Then the average annual repair cost is converted to earthquake loss expectation according to Formula 5-8.

2. seismic wave selection

The equal-weighted full-period matching scheme has strong performance throughout a broader period segment, but has obvious deviations in the long period segment, with the maximum relative error reaching 45%, and its discreteness is double that of other period segments. Therefore, to improve the accuracy of structural analysis, this study uses the optimised weight function selection method proposed by Ji to choose eleven ground motions from the ATC-63's FEMA P-695 series of earthquakes for further examination of their reaction (Ji, 2018). This method overcomes the difference in ground motion characteristics under different seismic levels and the deficiencies of different structures selected for the same site.

The 11 seismic waves information is displayed in Table 7.9. Figure 7.17 depicts the response spectrum record of each seismic wave, the average response spectrum after amplitude modulation and the design response spectrum under rare earthquakes. The resemblance between the two is evident. Therefore, the selected ground motion is reasonable and can be used for subsequent analysis.

TABLE 0.9 **INFORMATION ABOUT 11 PIECES OF EARTHQUAKE WAVES**

No.	File Name	Location/Record Time/Earthquake Name
EQ1	169_IMPVALL.H\H-DLT262	Delta, 10/15/1979, Imperial Valley-06
EQ2	174_ IMPVALL.H\H-E11140	El Centro Array #11, 10/15/1979, Imperial Valley-06
EQ3	1165_ KOCAELI\IZT090	Izmit, 8/17/1999, Kocaeli, Turkey
EQ4	292_ ITALY\A-STU000	Sturno (STN), 11/23/1980, Irpinia, Italy-01
EQ5	960_NORTHR\LOS000	Canyon Country- W Lost Cany, 1/17/1994, Northridge-01
EQ6	126_ GAZLI\GAZ000	Karakyr, 5/17/1976, Gazli, USSR
EQ7	1605_ DUZCE\DZC180	Duzce, 11/12/1999, Duzce, Turkey
EQ8	68_ SFERN\PEL090	LA- Hollywood Stor FF, 2/9/1971, San Fernando
EQ9	721_ SUPER.B\B-ICC000	El Centro Imp. Co. Cent, 11/24/1987, Superstition Hills-02
EQ10	1244_ CHICHI\CHY101-E	CHY101, 9/20/1999, Chi-Chi, Taiwan
EQ11	1063_ NORTHR\RRS228	Rinaldi Receiving Sta, 1/17/1994, Northridge-01

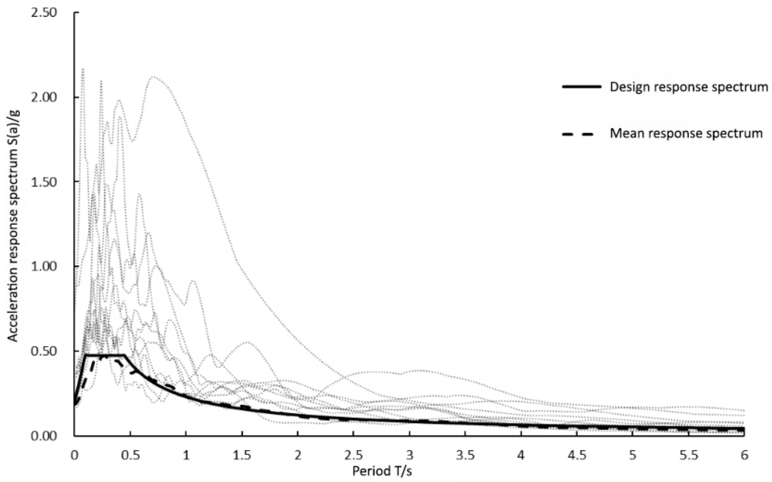


FIGURE 0.17 **EARTHQUAKE WAVE INFORMATION**

Initial Setting

1. dimensional variable setting

According to the structural layout of the practical application in Figure 7.1, the designer first selects twelve height and width dimension variables of the six frame beam categories, which have varying spans in the standard layer. Additionally, ten width dimension variables of the ten frame column categories with different loading conditions in different plane positions. These twenty-two design variables constitute a structure plan, as displayed in Table 7.10. The naming of each design variable complies with the rules given in Section 5.3.1, the variable range and optional number are given by the designer in combination with experience and actual project requirements.

TABLE 0.10 **DESIGN VARIABLES**

Component Type	Design Variables	Range of Design Variables	Optional Quantity
Frame Beam	h_{L1}	650,700,750,800,900	5
	h_{L2}、h_{L3}	450,500	2
	h_{L4}、h_{L5}	400,450,500	3
	h_{L6}	700,750,800,900,1000	5
	b_{L1}	300,350,400,450	4
	b_{L2}、b_{L3}、b_{L4}	250,300	2
	b_{L5}	200,250,300	3
	b_{L6}	300,350,400,450,500	5

Component Type	Design Variables	Range of Design Variables	Optional Quantity
Frame Column	d_{1A}、d_{1B}、d_{5A}、d_{6A}、d_{6B}	400,450,500,550,600,650,700,800	8
	d_{4A}、d_{7A}、d_{4B}、d_{5B}、d_{7B}	450,500,550,600,650,700,800	7

2. BIM application process

After inputting the structural component type and size information of the design plan into the Dynamo parameter node, the frame model shown in Figure 7.18 can be generated.

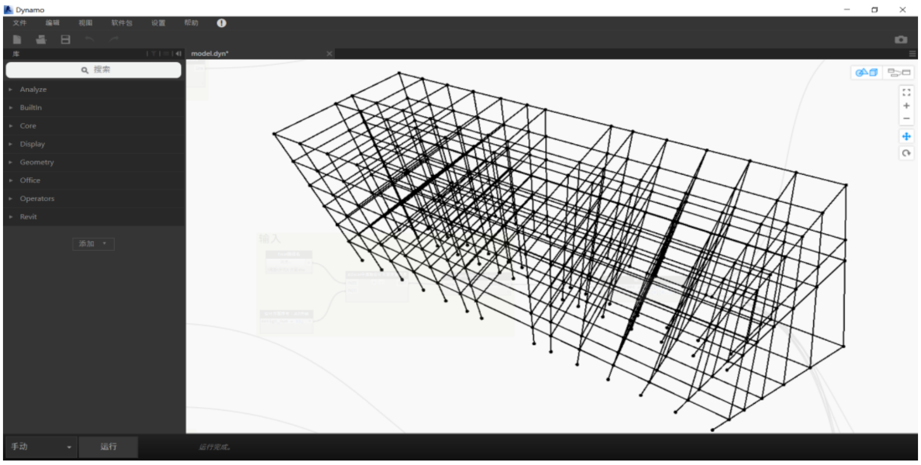


FIGURE 0.18 **MODEL PREVIEW DURING THE RUN IN DYNAMO**

In the subsequent iteration process, on the one hand, the Revit model of each design scheme uses the Revit-YJK interface to import the structural model into the YJK software with one click and performs nonlinear time history analysis. Then the result would be judged by designers whether the constraint conditions are met. On the other hand, when the design plan meets the constraint conditions, the Revit model can export the IFC files and processes the component information based on the IFC analysis in Chapter 2 to achieve the calculation of the initial building expense objective function and the seismic loss objective function. The comprehensive unit price information for case cost calculation is shown in Tables 7.5 and 7.6 of Section 7.1.2, earthquake loss expectation is calculated according to Formula 5.8, Formula 7.2 and Formula 7.5.

3. optimized algorithm parameter setting

Once the design variables of study application have been chosen, establish the starting population size N to be 24, the crossover probability C to be 0.8, the mutation probability M to be 0.1, and assign the design variable values of individual plan of the initial population as the initial input of the NSGA-II algorithm. When choosing the initial plan, it should try to satisfy fewer types based on diversity and rationality, so as to facilitate production and construction. Subsequently, following the procedure outlined in Figure 5-16, the NSGA-II algorithm is used to iteratively optimise the system while ensuring that the pre-defined seismic performance objectives are satisfied, until the termination criteria for iterations are fulfilled. Therefore, the initial building expense and anticipated earthquake loss can be minimized. The ultimate outcome is represented as the Pareto optimal collection of solutions.

Optimisation Iteration and Outcome Analysis

At the end of the iteration of this case study, Gen = 17 (hereinafter referred to as F_{17}). Figure 7.19 displayed the objective function created during individual iteration, starting with the parent generation and continuing until the 17th filial generation. The red dot represents the Pareto optimum solution set and they are not dominated by any other design plan. They constitutes the viable area with other points.

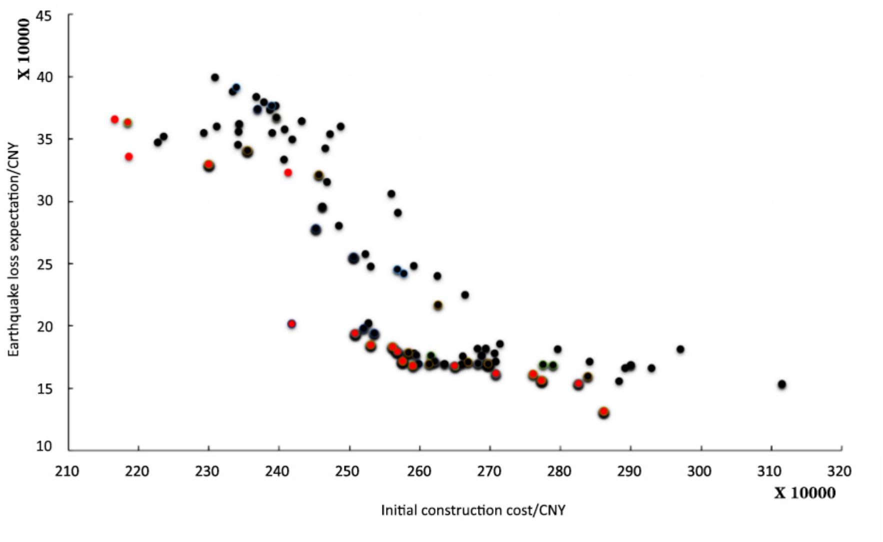


FIGURE 0.19 **ITERATION SITUATION**

As displayed in Figure 7.20, the objective function distribution for the starting population may more effectively include the design situation in the center as well as at both extremes. Obviously, the initial building expense increase with the decreasing of the seismic loss expectation, suggesting that it is required and justifiable to consider it as a conflicting target for their separate optimisation (Min et al., 2010). Similarly, the Pareto solution set has the same trend in the distribution of initial building expense and seismic loss expectation. The two goal expenses have achieved a pronounced optimisation impact when making comparison with the starting population. Assuming that individual point representing a design solution in the diagram may be denoted by (*initial building expense, anticipated earthquake loss*), notably, as at point 2, the starting population is represented by (*2591110, 248571*) and the Pareto optimal solution set is (*2583454, 179010*). After optimisation, the two major costs have been both reduced and the reduction in earthquake losses is greater. Take another example in point 1, the Pareto optimal solution set consists of the coordinates (*2452064, 194026*), while the starting population is represented by the coordinates (*2412529, 323074*). It means an extra 1.63% of the initial building expense and can reduce the anticipated earthquake loss by 39.95%. In addition, it can be seen from the minimum interval of the two major costs that the calculation accuracy is greatly improved.

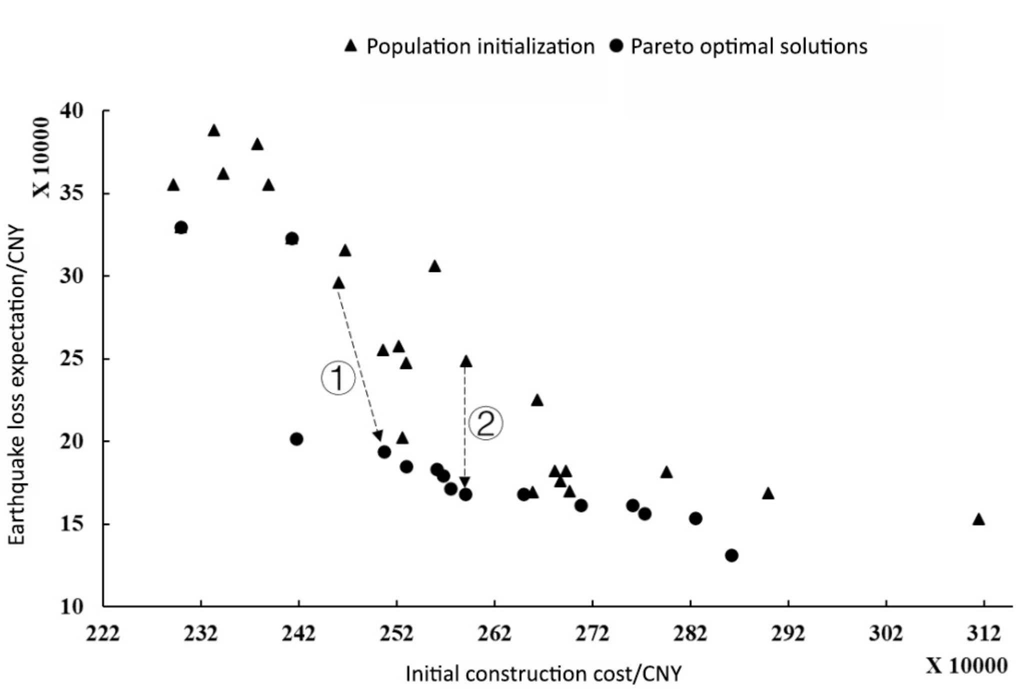


FIGURE 0.20 **OBJECTIVE FUNCTION DISTRIBUTION OF SCHEME SET BEFORE AND AFTER OPTIMISATION**

Figure 7.21 and Figure 7.22 respectively show the Pareto optimal solution set and the distribution of the MIDR response difference in the primary and secondary directions of the building under the action of the initial population under moderate earthquakes. The comparison reveals that the optimised displacement difference is drastically reduced, indicating that the building's cost distribution is improved when it responds symmetrically in the two seismic directions.

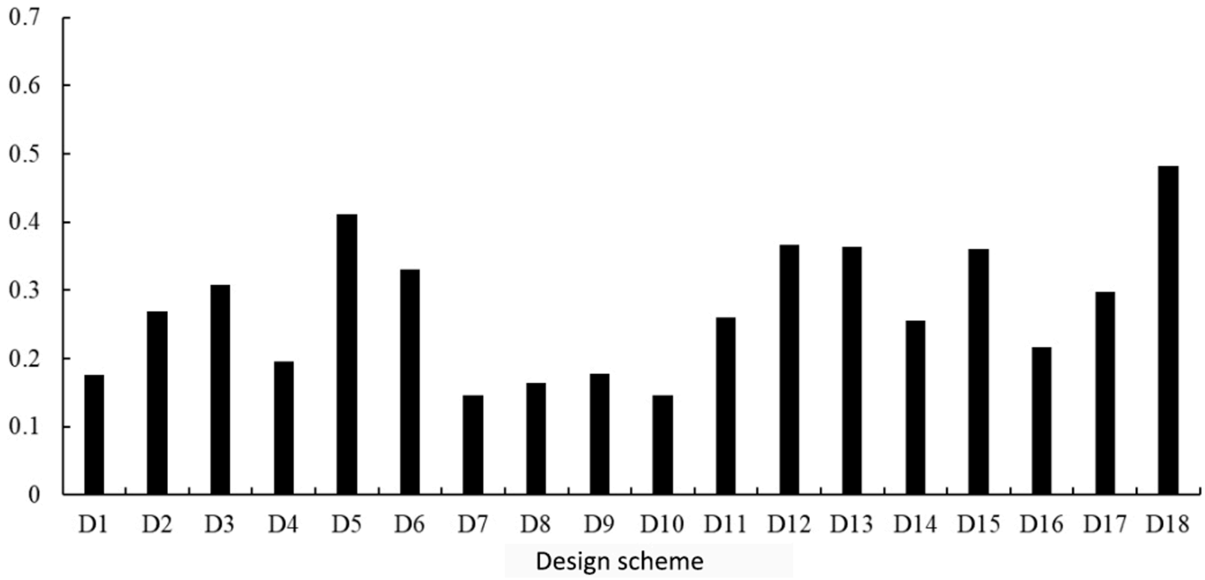


FIGURE 0.21 **THE MAXIMUM VALUE DISTRIBUTION OF RESPONSE DIFFERENCE IN PRIMARY AND SECONDARY DIRECTIONS OF PARETO OPTIMAL SOLUTION SET UNDER THE ACTION OF MEDIUM EARTHQUAKE**

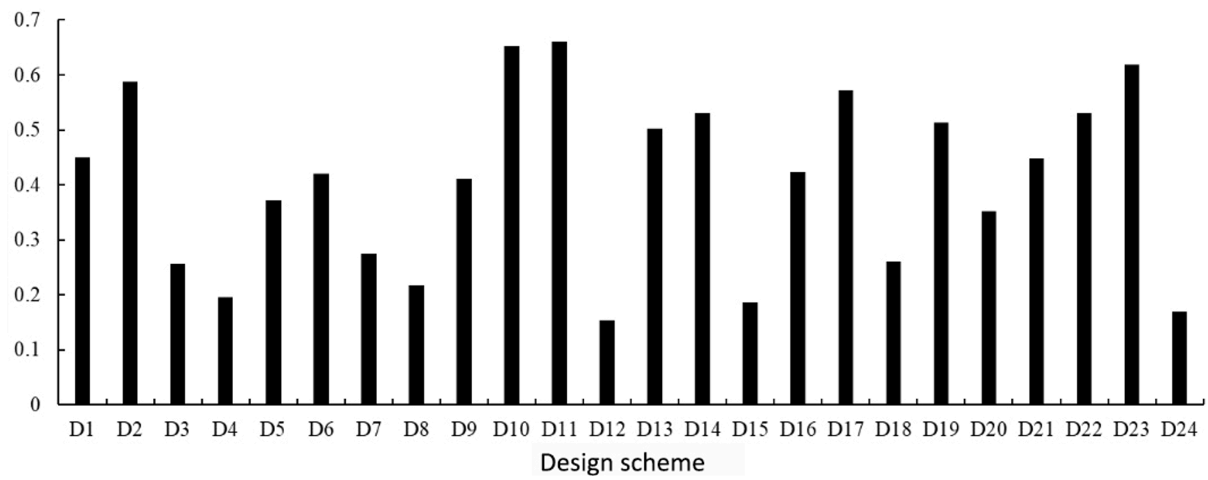


FIGURE 0.22 **THE MAXIMUM VALUE DISTRIBUTION OF RESPONSE DIFFERENCE IN PRIMARY AND SECONDARY DIRECTIONS OF INITIAL SCHEME SET UNDER THE ACTION OF MEDIUM EARTHQUAKE**

Figure 7.23 and Figure 7.24 respectively show the Pareto optimal solution set and maximum value of the response difference distribution of initial population under the action of small earthquakes in the building main direction in the linear elastic analysis and elastoplastic analysis. Obviously, the response difference after optimisation is largely reduced. This is because the elastic-plastic analysis in the optimisation iteration is based on a certain series of seismic records and the elastic analysis depends on the characteristics of the building itself. It can be seen that the optimisation trend appears that the building characteristics are consistent with the seismic records used.

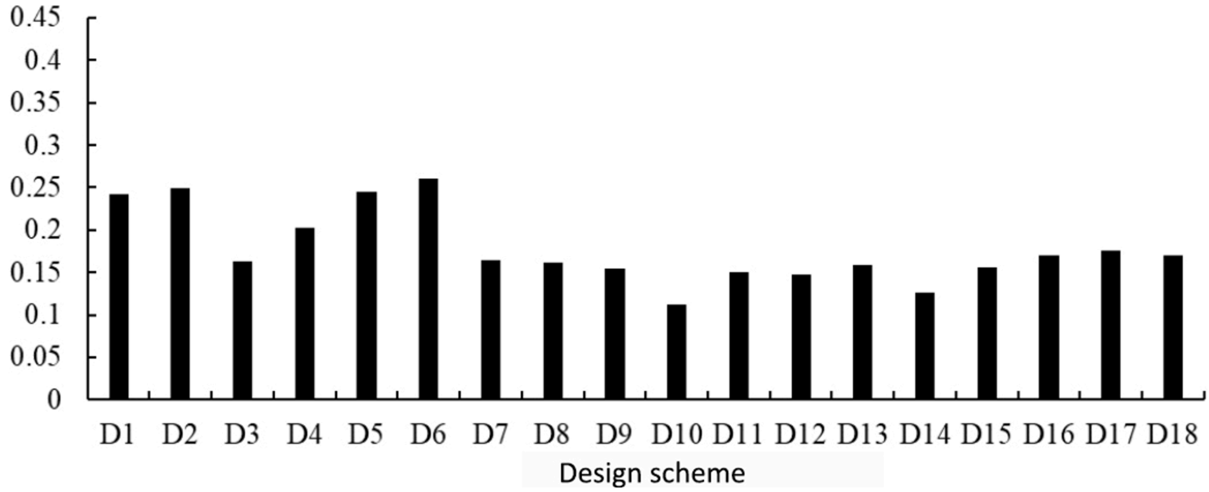


FIGURE 0.23 **THE MAXIMUM VALUE DISTRIBUTION OF RESPONSE DIFFERENCE IN PRINCIPAL DIRECTION ELASTIC AND ELASTOPLASTIC OF PARETO OPTIMAL SOLUTION SET UNDER THE ACTION OF SMALL EARTHQUAKE**

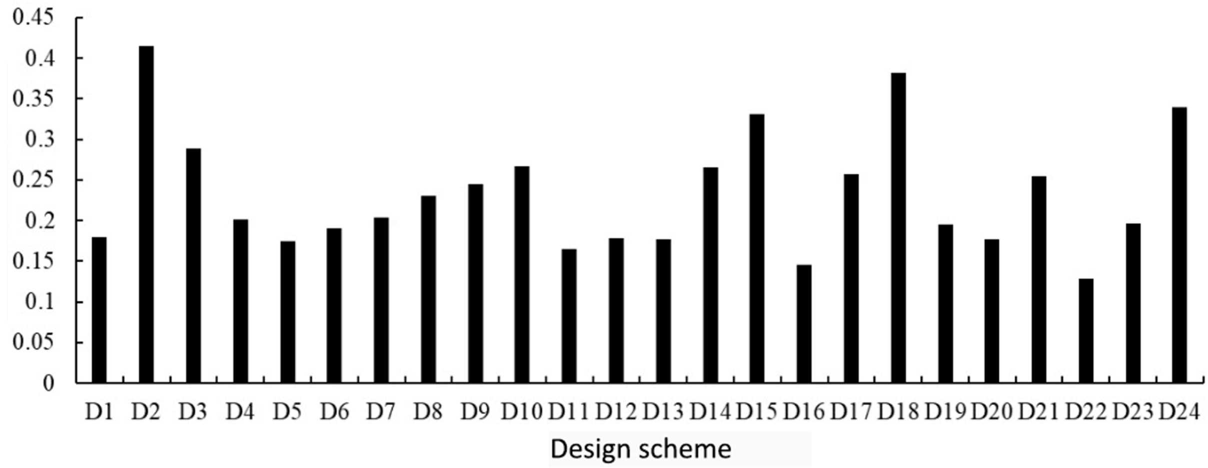


FIGURE 0.24 **THE MAXIMUM VALUE DISTRIBUTION OF RESPONSE DIFFERENCE IN PRINCIPAL DIRECTION ELASTIC AND ELASTOPLASTIC OF INITIAL SCHEME SET UNDER THE ACTION OF SMALL EARTHQUAKE**

Figure 7.25 show the distribution of the average loss ratio (expected earthquake loss/initial construction cost) of each building in the plan set before and after optimisation. As assuming that everything else is unchanged except for structural components, the cost of non-structural components is not taken into account when calculating the initial construction costs. While the cost of structural components, non-structural components and equipment is considered when calculating earthquake loss expectations. In order to obtain the loss ratio in the general sense, the cost of RC frame structural components is considered to account for 1/6 of the total cost generally, which is calculated by multiplying the loss ratio by 6. It can be concluded from the figure that optimisation has a more obvious effect on reducing the average loss.

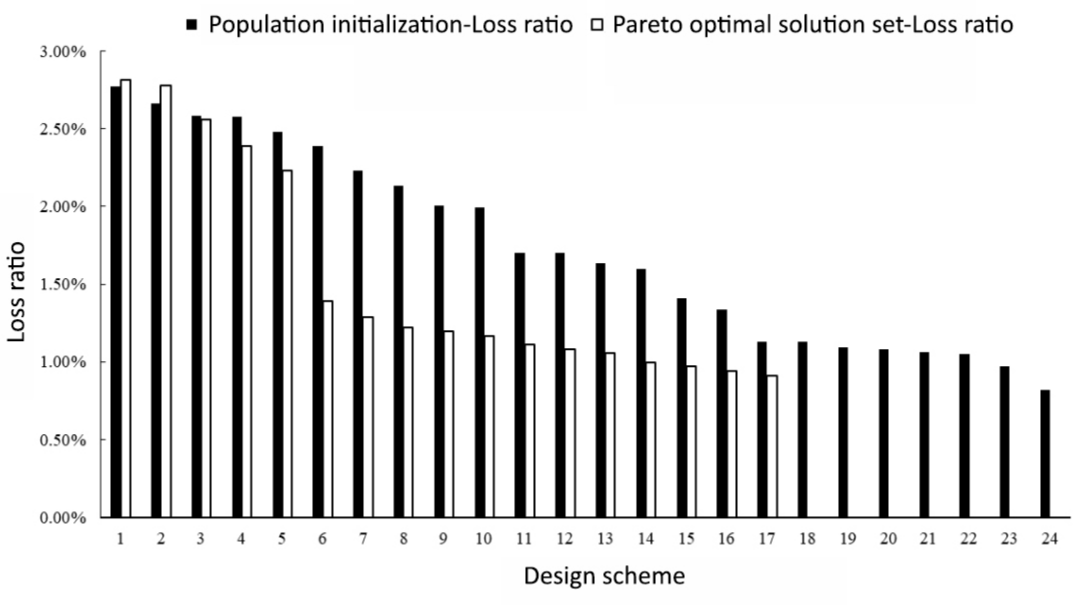


FIGURE 0.25 **LOSS RATIO DISTRIBUTION BEFORE AND AFTER OPTIMISATION**

7.2.1 CASE TRAINING SAMPLES

30 RC Frame Building Cases

According to the selected structural parameters in Section 6.2.1, this study takes as much as possible to cover different parameter combinations as the premise and forms the 30 RC frame building case training samples which are also provided by Sichuan Provincial Government (as shown in Table 7.11).

TABLE 0.11 **STRUCTURAL PARAMETERS OF TRAINING SAMPLES**

No.	Height of ground floor/m	Height of other floors/m	Size of column/mm	Height of beam/mm	Reinforcement ratio of column	Reinforcement ratio of beam	Strength grade of concrete/MPa	Strength grade of steel	Horizontal side span/mm	Horizontal midspan/mm
D1	4.5	3.5	600	600	0.012	0.013	35	335	6800	2400
D2	4.2	3.6	500	650	0.016	0.0087	40	335	6000	3000
D3	4.2	3.6	400	650	0.02	0.0097	35	400	6000	3000
D4	4.5	3.6	600	550	0.014	0.015	35	335	6800	2400
D5	4.5	3.5	550	650	0.015	0.008	30	400	6000	2700
D6	4.2	3.6	600	700	0.013	0.01	35	335	6800	2400
D7	4.2	3.6	500	700	0.018	0.009	40	335	6000	2400
D8	5.5	3.6	700	700	0.022	0.01	30	335	6800	2400
D9	5.5	3.6	600	550	0.019	0.011	40	400	6000	3000
D10	3.6	3.6	550	600	0.011	0.0105	30	400	6800	2400
D11	4.5	3.6	550	600	0.015	0.009	40	400	6000	3000
D12	4.2	3.6	550	700	0.01	0.008	40	400	6000	3000
D13	3.6	3.6	500	650	0.015	0.008	35	400	6000	2400
D14	3.6	3.6	550	550	0.01	0.014	30	400	6800	2400
D15	4.2	3.6	500	600	0.012	0.008	35	400	6000	3000
D16	4.2	3.6	450	750	0.014	0.0085	40	400	6000	2400

No.	Height of ground floor/m	Height of other floors/m	Size of column/mm	Height of beam/mm	Reinforcement ratio of column	Reinforcement ratio of beam	Strength grade of concrete/MPa	Strength grade of steel	Horizontal side span/mm	Horizontal midspan/mm
D17	4.5	3.6	450	750	0.022	0.0095	35	400	6000	3000
D18	3.6	3.6	550	750	0.01	0.009	35	400	6000	2400
D19	5.5	3.6	450	750	0.03	0.011	35	335	6000	2700
D20	4.2	3.6	550	700	0.015	0.011	40	335	6000	3000
D21	3.6	3.6	550	600	0.018	0.013	30	335	6000	2400
D22	4.2	3.6	700	500	0.0124	0.012	35	400	6000	3000
D23	4.5	3.6	450	700	0.0174	0.01	40	400	6000	2700
D24	4.5	3.6	550	700	0.017	0.0114	40	335	6000	2400
D25	5.5	3.6	700	600	0.012	0.012	35	400	6800	2400
D26	4.2	3.6	400	750	0.023	0.0099	35	400	6000	2400
D27	4.5	3.6	600	500	0.0247	0.016	30	335	6000	2700
D28	4.5	3.5	600	600	0.0205	0.0134	35	335	6800	2400
D29	3.6	3.6	500	500	0.0275	0.0167	30	335	6800	2400
D30	3.6	3.6	500	750	0.0218	0.018	30	335	6000	2400

Except for the 10 selected structural parameters, the other assumptions remain unchanged, including:

- a six-story building, six longitudinal spans and each span is 4.8m, three horizontal spans and the side spans are the same. The frame beams, columns, floors and roof panels are all cast-in-place, with a thickness of 150mm. All steel bars including stirrups are of the same grade. The hollow blocks with a thickness of 200mm and a bulk density of 10kN/m^3 are evenly arranged along the axis of each beam. The dead and live loads of the roof are 5kN/m3 and 2kN/m3 respectively. The dead and live loads of the floor are 5kN/m3 and 2kN/m3 respectively.

- Class C structure, seismic precautionary intensity is 7 degrees. The design basic seismic acceleration value is 0.1g, the site category is Class II, the design seismic group belong to the second group, the site characteristic period is 0.40s and the damping ratio is 5%.

Selection of 60 Seismic Waves

First, according to the four steps of seismic wave screening in Section 6.3, this case study combines basic seismic information to screen 136 seismic wave records form the NGA-West2 database. The screening result is summarized as shown in Table 7.12.

TABLE 0.12 **FILTER CONDITIONS IN ALL STEPS**

Filter Conditions	Range	Earthquake wave quantities after selection
Earthquake Level（Epicenter Distance range of 10-200km）	[5.0-8.0]	6181
Epicentre Distance	[15,30]	555
	[260,550]	377
Amplitude Modification Coefficient	PGA∈[0.16， 0.95]	136

For the screening conditions of the amplitude modulation coefficient, the amplitude modulation coefficient for the target PGA=0.19g (rare earthquake) is limited to [0.2-5], which means 0.038≤ PGA ≤ 0.95g. While for the target PGA = 0.8g, the amplitude modulation coefficient is limited within 5 which means PGA is screened in the range of [0.16, 0.95].

According to the screening condition in Table 7.12, 132 seismic waves with complete seismic wave information can be searched from the PEER database from the 136 seismic waves. In order to save computing resources as much as possible, for the PP parameters of each seismic wave during the characteristic period, the limited range is [0.2s,0.65s], 79 seismic waves are remained. Then the number of steps NTPS is limited to [0,15000], 68 seismic waves are remained. Finally, 60 seismic waves were finally screened out based on the comparison between the average design spectrum after amplitude modulation and the design response spectrum of rare earthquakes. According to the calculation method of the seismic parameters in each region in Table 6.2, the coverage of the ground motion parameters of 60 seismic waves are obtained except Target-PGA, as shown in Table 7.13. Based on the *Seismic parameter* part in Section 6.1.1, the amplitude modulation parameters are selected according to Table 7.14 for this case study.

TABLE 0.13 **GROUND MOTION PARAMETERS VALUE OF 60 EARTHQUAKE WAVES**

Earthquake Motion Index	Unit	Maximum Value	Minimum Value
PGA	g	0.151	0.714
PGV	cm/s	5.258	100.112
PGD	cm	0.184	47.513
Ia	m/s	0.090	6.028
SED	cm^2/s	8.370	10388.514
CAV	cm/s	149.001	2560.562
ASI	g • s	0.106	0.609
HI	cm	13.298	301.938
EPA	g	0.105	0.597
Vmax/Amax	s	0.026	0.207
PP	s	0.204	0.649
D_{s5-75}	s	0.733	18.074

TABLE 0.14 **VALUES OF AMPLITUDE MODULATION PARAMETER**

No.	Target-PGA/g	NO.	Target-PGA/g
1	0.02	6	0.26
2	0.04	7	0.32
3	0.08	8	0.38
4	0.14	9	0.48

5	0.20	10	0.58

Structural Response Analysis

According to the above information, there are a total of 30 design schemes and 600 (60*10) seismic waves. For each design scheme, 200 seismic waves were randomly allocated, relevant parameters were set in the YJK software according to the specifications. Nonlinear time history analysis was then conducted for obtaining 6000 training data for prediction model construction. In the subsequent neural network model building and training process, 5600 sample data formed by 28 design schemes are selected and the remaining two design schemes are used as test sets to test the prediction effect of the model. Furthermore, the raining samples with MIDR0.1 is removed and finally get 5468 training data.

7.2.2 CASE RESPONSE PREDICTION MODEL

Training Process

Figure 7.26 depicts this case study's training procedure. The 5468 data samples in section 7.2.1 are divided into training data, test data and verification data at the ratio of 60%, 20% and 20%. The number of hidden Layer neurons is set from 10 to 60, which are 51 situations in total. The rest configurations are set according to the default settings in section 6.5.1 and loop 50 times for each combination of network configuration and training algorithm. Considering the above factors, a total of 62200 (102*2*6*50=62200) training sessions are required.

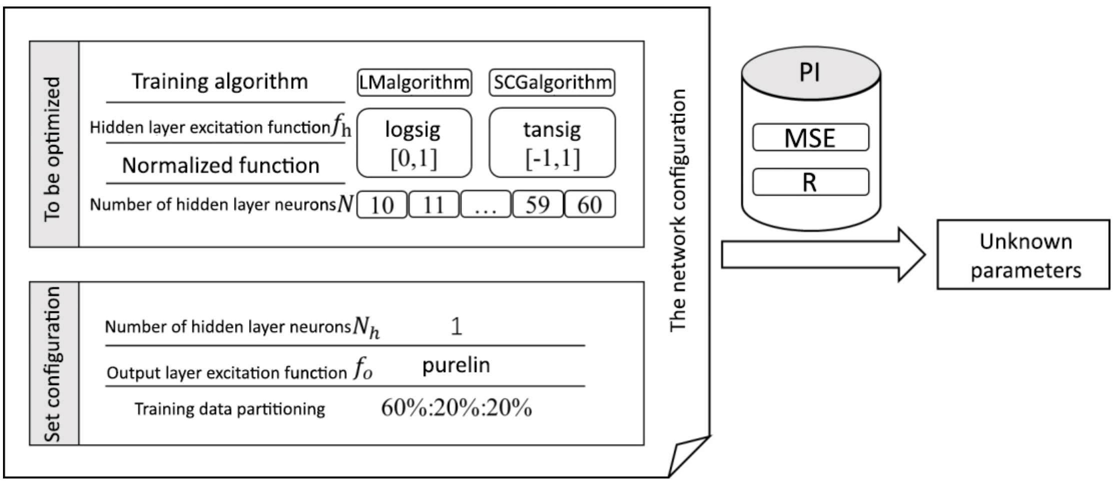


FIGURE 0.26 **TRAINING PROCESS OF FA PROBLEM**

Training Result

Step 1 to step 8 in section 6.5.2 clarify the process of training the FA problem in this study with a certain model configuration in MATLAB. Users only need to change the training algorithm (LM/SCG), the number of hidden layer neurons (10-60) and the hidden layer activation function (logsig/tansig) to drive training in another network configuration. The 62200 training results are sorted according to the evaluation template in Table 6.8 to obtain data in Table 7.15.

TABLE 0.15 **PERFORMANCE EVALUATION RESULTS**

Performance Evaluation Standards		Training Algorithm/Activation Function of Hidden Layer			
		LM/tansig	LM/logsig	SCG/tansig	SCG/logsig
min(MSE)	All data set	0.0070	0.0019	0.0138	0.0035
	Training data set	0.0020	0.0006	0.0109	0.0028
	Testing data set	0.0099	0.0026	0.0132	0.0034
max(R)	All data set	0.9546	0.9500	0.9079	0.9069
	Training data set	0.9875	0.9841	0.9266	0.9243
	Testing data set	0.9303	0.9293	0.9050	0.9029

The basic conclusions derived from the above table are:

- Regardless of the performance evaluation criteria adopted, the training effect of the training algorithm LM is superior to that of the SCG algorithm, which is reflected in the smaller MSE value and the larger correlation coefficient R value.
- When the MSE is used as the evaluation index, the performance of the hidden layer activation function when taken as logsig is better than that of tansig. When applying the correlation coefficient R as the evaluation indicator, the tansig function outperforms the logsig function in terms of performance.

According to the above six performance evaluation criteria, the optimal number of neurons in the hidden layer is sorted out as shown in Table 7.16.

TABLE 0.16 **OPTIMAL NUMBERS OF NEURONS IN THE HIDDEN LAYER**

Performance Evaluation Standards		Training Algorithm/Activation Function of Hidden Layer			
		LM/tansig	LM/logsig	SCG/tansig	SCG/logsig
min(MSE)	All data set	33	20	53	43
	Training data set	45	50	53	43
	Testing data set	12	24	46	25
max(R)	All data set	33	20	53	43
	Training data set	45	50	53	43
	Testing data set	18	23	46	25

The basic conclusions derived from the above table are:

- When the LM algorithm is selected as the network training algorithm, for the total data set and training data set, the optimal number of neurons is not relevant with the performance index. For validation data set, the optimal number of neurons is correlated with the performance index.
- When the SCG algorithm is selected as the network training algorithm, the optimal number of neurons is solely dependent on the data set type and is not related with performance indicators.

A Certain Training Result Under the Optimal Configuration

From Table 7.15 and Table 7.16, it demonstrates that the network configuration with the smallest MSE for validation data set adopts LM algorithm, logsig hidden layer activation function and 24 hidden layer neurons, which is named as “LM-logsig-24”. This section will show the results of a certain simulation under this “LM-logsig-24” configuration, as shown in Figure 7.27.

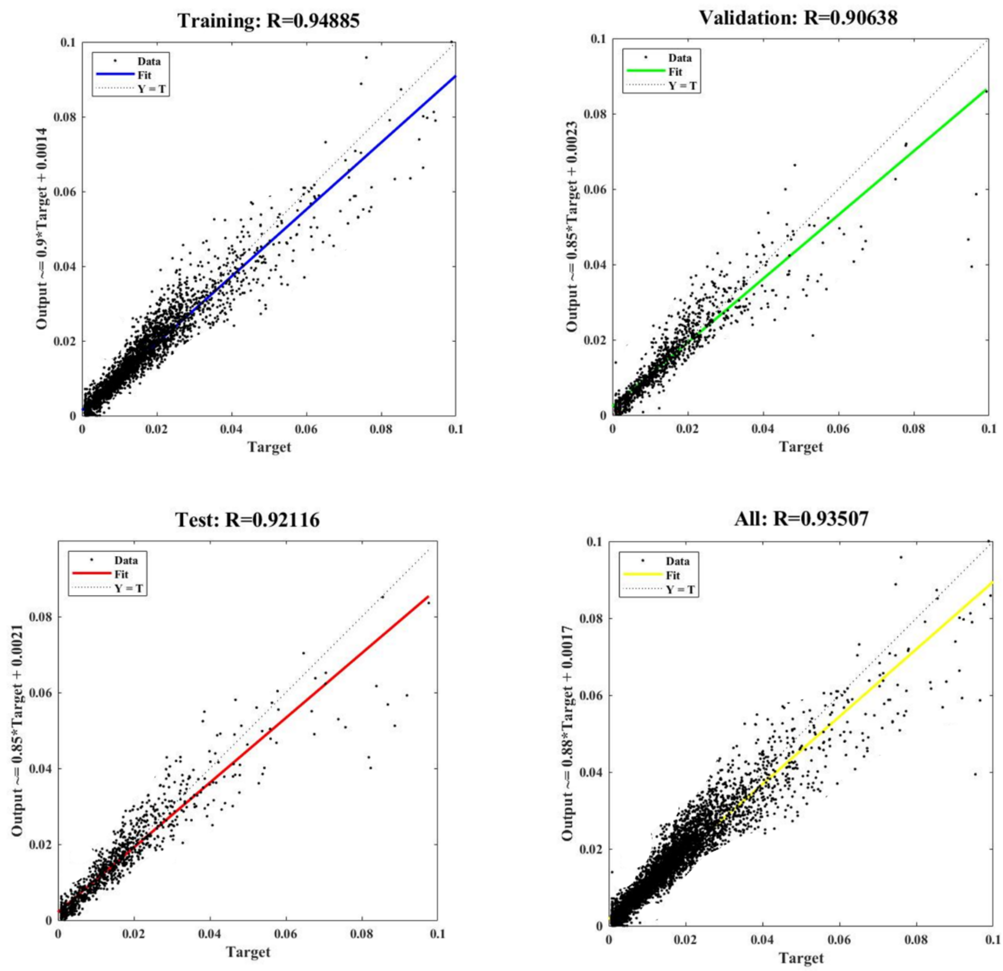


FIGURE 0.27 **"LM-LOGSIG-24" SIMULATION RESULT**

In the Figure 7.27, the little black dots represent the sample data, the horizontal axis and vertical axis coordinate values are the actual and predicted values of the sample, respectively. In the four cases of training data set, validation data set, verification data set and total data set, the true value and the predicted value are fitted to obtain four regression curves. Ideally, the network output should be identical to the expected output which is the dashed line. The solid line depicts the linear regression of the data. Generally speaking, the correlation coefficient during fitting is above 0.9, which indicates that the neural network has good performance (Du & Padgett, 2020). The simulated correlation coefficient R values of the four curves shown in the figure are 0.9489, 0.9064, 0.9212, 0.9351, which are all greater than 0.9. The data sample points are in good accord with the curve, which demonstrates that the model fits well.

It can also be seen from the figure that when the MIDR is in the (0,0.025] interval, the model prediction effect is better, and the degree of dispersion is smaller. When the MIDR exceeds 0.025, the degree of dispersion of the model prediction increases. This is because the building structure has reached the collapse limit at this time, the damage is severe, the structure response does not have regularity and the prediction result has little impact on the establishment of the collapse fragility curve, so the prediction effect is still within the acceptable range.

Judgement of Model Superiority and Inferiority

Although there is a good simulation effect, it can only represent the model performance of a certain simulation under the "LM-logsig-24" network configuration. In this section, the coefficient of determination R^2 is a mathematical indicator used for judging the overall pros and cons of model. The calculation formula of this indicator is displayed below:

$$R^2 = \frac{\left(l\sum_{i=1}^{l}\hat{y}_i y_i - \sum_{i=1}^{l}\hat{y}_i\sum_{i=1}^{l}y_i\right)^2}{\left(l\sum_{i=1}^{l}\hat{y}_i^{\,2} - \left(\sum_{i=1}^{l}\hat{y}_i\right)^2\right)\left(l\sum_{i=1}^{l}y_i^{\,2} - \left(\sum_{i=1}^{l}y_i\right)^2\right)} \tag{7-6}$$

Y_i(i=1,2,…,n) is the true value of the ith data, is the predicted value of the ith data, and l is the total number of data samples, which is 5468 for this case study. The original data were simulated 50, 100 and 500 times respectively. The numerical results of the model obtained by the simulation were calculated using the coefficient of determination and the statistical results are shown in Table 7.17.

TABLE 0.17 **DATA STATISTICS OF DETERMINATION COEFFICIENT R^2 UNDER VARIOUS SIMULATION TIMES**

Times of simulation	50 times	100 times	500 times
Mean	0.8550	0.8564	0.8592
Var	0.0007	0.0007	0.0006
Std	0.0255	0.0261	0.0239

According to the data in the table above, the average value of the coefficient of determination R2 stabilized at [0.85, 0.86] as the number of simulations increased. Under the conditions of each number of simulations, the values of the difference and the standard deviation are both less than 0.1. These data indicate using this network model has good simulation results and stable performance. In this context, the model can ensure better fitting accuracy when multiple random data is drawn to from the training data set and the test data set. It also has a higher universality.

Generalization Ability Verification

This study has verified the generalization capacity of the better configuration network (LM-logsig-24) obtained. Then, the following context will explain the situation of the new sample data in various situations and select the ANN network model obtained after a certain training under the "LM-logsig-24" configuration to predict the new sample data in various situations. The prediction results are displayed and the analysis of the results are given. The network models of all kinds of new sample data are all the same in this case.

1.new sample data generation

(1). ANN prediction of unknown structure under known seismic waves

Select building structure "D29" as the study case and select 11 seismic waves from the existing 60 seismic wave series that can be closer to the response spectrum of the case target after amplitude modulation. Each seismic wave is performed 10 times of amplitude modulation as shown in Table 7.14. In YJK software, the MIDR of the building under each earthquake situation in Table 7.18 is analyzed and calculated by nonlinear time history. 100 (11*10) sample data to be verified are obtained.

TABLE 0.18 **110 EARTHQUAKE SCENARIOS**

No. of earthquake waves	640; 2391; 391; 611; 3746; 4136; 803; 4132; 4117; 4130; 549
Amplitude Modification Index	[0.02,0.04,0.08,0.14,0.20,0.26,0.32,0.38,0.48,0.58]

(2). ANN prediction of known structures under unknown seismic waves

For D1, the first step is to randomly select 3 seismic waves that are not included in the 60 seismic waves in the training sample from the 132 seismic waves with complete seismic wave information as described in Section 7.2.1. Each seismic wave is subjected to 10 amplitude modulations in Table 7.14, which is called "the known amplitude modulation parameters". The second step is to perform the amplitude modulation of 60 seismic wave series of training samples under 3 new Target-PGA VALUES, which is called "the unknown amplitude modulation parameters".

Situation 1: the known amplitude modulation parameter sub-case

The parameter information for the selected 3 new seismic waves is shown in Table 7.19. In YJK, the MIDR of the building under each earthquake situation is analyzed and calculated and 30 (3*10) sample data to be predicted and verified are obtained.

TABLE 0.19 **PARAMETERS INFORMATION OF 3 NEW EARTHQUAKE WAVES**

No. of earthquake wave Earthquake index	Unit	125	766	4861
PGA	g	0.335466897	0.345423296	0.338127065
PGV	cm/s	26.39452844	37.44013943	31.21161854
PGD	cm	4.888105746	13.24091143	12.67647163
Ia	m/s	1.011514638	1.266832945	2.027255262
SED	cm^2/s	439.8922945	1328.29563	3333.669209
CAV	cm/s	625.096731	853.399283	1488.331971
ASI	g • s	0.253317659	0.269511676	0.175155161
HI	cm	215.9740458	269.2634247	163.5616917
EPA	g	0.253641657	0.267854971	0.177515146
Vmax/Amax	s	0.078679979	0.108389156	0.092307365
PP	s	0.815843122	0.684105255	0.74993333
D_{s5-75}	s	2.52739985	2.29749864	10.33373118
Target-PGA	g	[0.02,0.04,0.08,0.14,0.20,0.26,0.32,0.38,0.48,0.58]		

Situation 2: the unknown amplitude modulation parameter sub-case

This sub-case is for the 60 seismic waves in the training sample. On the basis of the situation 1, 3 new Target-PGA values ae additionally considered for each seismic wave. The three major seismic levels in the specification are chosen for the case structure, which are 0.033g, 0.10g and 0.19g. In YJK, the MIDR of the building under each earthquake situation is analyzed and calculated. 180 (60*3) sample data to be predicted and verified are obtained.

(3). ANN prediction of unknown structure under unknown seismic waves

For D30, 3 seismic waves are randomly chosen that are not included in the 60 seismic waves in the training sample from the 132 seismic waves with complete seismic wave information. The first step is to perform amplitude modulation to each seismic wave under 3 new Target-PGA values. The second step is to perform amplitude modulation in Table 7.14 to each seismic wave, then the MIDR of the building under each earthquake situation is analyzed and calculated in YJK. 39 (3*3+3*10) sample data to be predicted and verified are obtained.

2. ANN prediction results and analysis

First, construct a certain ANN training network "*net*" under the "LM-logsig-24" configuration according to Section 6.5.2. Then the input parameters such as structural parameters and seismic parameters of various new samples are normalized and input into "*net*" respectively. The predicted value can be obtained after denormalization.

Plot the prediction results under each situation, Scene 1: ANN prediction of unknown structure under known seismic waves, Scene 2-1: ANN prediction of known structure under unknown seismic waves and known amplitude modulation parameters, Scene 2-2: ANN prediction of the known structure subject to the unknown seismic wave and unknown amplitude modulation parameter, Scene 3: ANN prediction of the unknown structure subject to the unknown seismic wave are shown respectively in Figure 7.28. The performance indicators predicted in each situation are shown in Table 7.20.

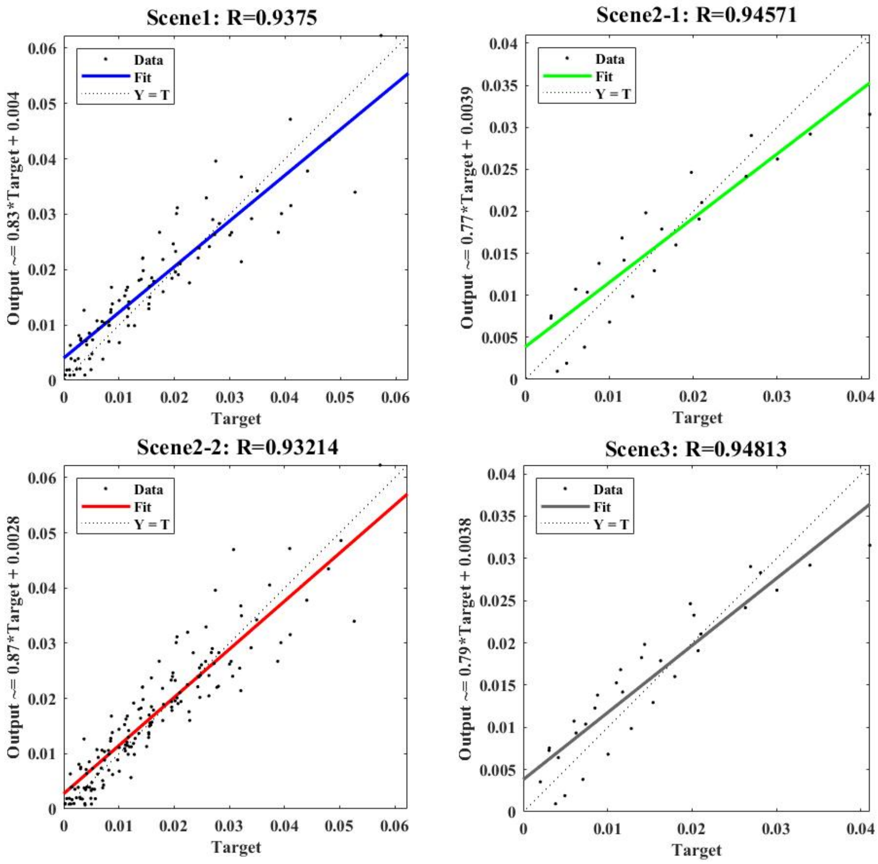


FIGURE 0.28 **PREDICTION RESULTS OF NEW SAMPLE DATA**

TABLE 0.20 **PERFORMANCE INDICATORS PREDICTED FOR EACH SCENARIO**

	Scene1	Scene2-1	Scene2-2	Scene3
MSE	0.0027	0.0018	0.0022	0.0001
Relationship coefficient R	0.9375	0.9457	0.9321	0.9481

The foregoing findings provide the following conclusions:

- The trained ANN model can predict new sample data more accurately and has a better fit.
- It is more consistent with the training results in section 3 that the prediction accuracy under the new sample data also shows the law that the MIDR has a smaller dispersion in the (0,0.025] region and a larger dispersion after it exceeds 0.025.
- The distribution of prediction points consists of across different situations, indicating that the generalization ability of the ANN training model for unknown structural parameters and unknown seismic parameters is the same, which in turn verifies the rationality of parameter selection.

According to the establishment steps of the collapse fragility function in Section 7.1.2, the real output (YJK analysis result) and the ANN predicted output of the 110 new sample data (11 seismic wave records * 10 amplitude modulation) in the case (a) are used to construct the fragility function. The comparison result is shown in Figure 7.29. The figure shows that the two severe damage performance function curves are basically consistent in the first half (PGA =0.20g). In the middle section (0.20g <PGA <0.60g), they are discrete on both sides near the 50% probability point. In the latter half, they are basically coincided. As can be concluded, the ANN model is more effective in predicting the collapse fragility function. To some extent, it can also explain that the Target-PGA parameters can reflect the characteristics of seismic waves together with the other 12 seismic parameters.

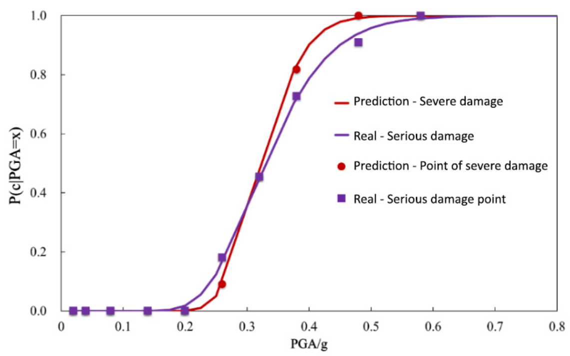


FIGURE 0.29 **COMPARISON DIAGRAM OF COLLAPSE VULNERABILITY FUNCTION**

From the verification effect of the above new data set, the generalization ability of the ANN model trained in this chapter is good.

7.2.3 MULTI-SCALE REGINAL EARTHQUAKE DAMAGE PREDICTION APPLICATION CASE STUDY

Based on the optimal ANN response prediction model obtained in Section 7.2.2, the multi-scale regional earthquake damage can be carried out in accordance with the method in Section 6.6. Taking the following scenarios as an example for application explanation.

The designated seismic level is 0.008g, which means the Target-PGA is set to 0.08g. Select the 4860-Chuetsu-oki seismic wave and use a certain ANN training network under the "LM-logsig-24" configuration to predict the MIDR of 30 building cases. Then according to the judgment criteria of the general performance level of the RC frame building, obtain the seismic resistance of each individual structure performance level. The thesis uses blue, green, yellow, orange and red to represent buildings damage status that are basically intact, slightly damaged, moderately damaged and severely damaged respectively. The performance level distribution of the regional building complex is displayed on the visualisation platform, as shown in Figure 7.30. It is evident from the image that influenced by this earthquake, the seismic performance level of the structure is between slight and moderate damage based on the color of buildings which are blue or green.

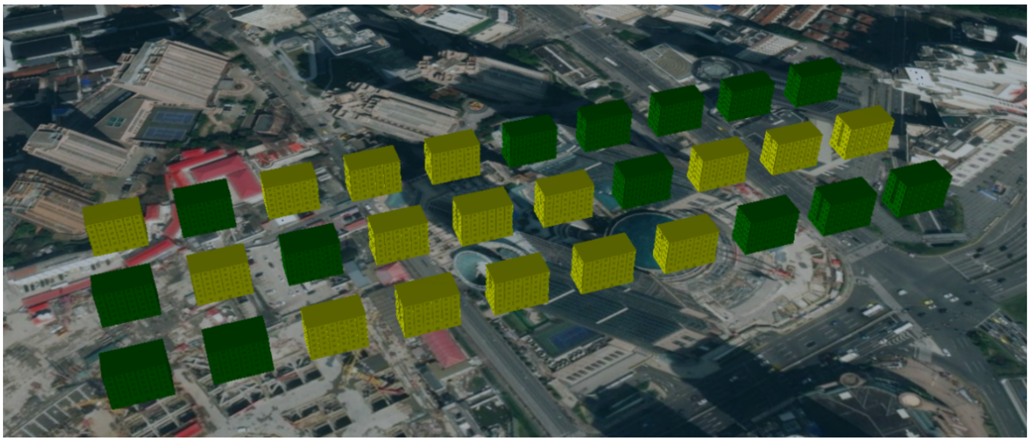


FIGURE 0.30 **THE DISTRIBUTION OF SEISMIC PERFORMANCE LEVEL FOR REGIONAL BUILDING GROUPS**

Similarly, regional seismic precautionary personnel can also perform performance distribution predictions based on other seismic levels and seismic wave records. A series of prediction results will provide effective guidance for the formulation and adjustment of seismic performance targets for the entire area. Additionally, the index of the MIDR between floors can be combined to evaluate the collapse of an individual building according to the collapse fragility function construction method in the generalization ability verification in Section 7.2.2. For individual structures that are extensively compromised, a comprehensive performance assessment and optimisation design are conducted, integrating the case study presented in Section 7.1, while reconciling the primary performance criteria of safety and cost-effectiveness.

7.3 SUMMARY

The chapter examines the feasibility of the methods proposed in Chapter 4 to Chapter 6 through case studies.

Judging from the application of case study 1, the seismic performance evaluation method of structure using BIM and ontology can provide more detailed evaluation information while reserving design space for new buildings. It can quickly and accurately predict performance indicators under various earthquakes. The evaluation efficiency is greatly improved. In addition, this method is according to FEMA P-58 and realizes the component-level damage prediction, which means the final result is expressed as the damage distribution with the component as the basic unit. It fully prepared for the subsequent multi-objective seismic optimisation design in combination with the initial building expense. While the application of the multi-objective optimisation model of discrete size variables matches the two major costs with the design plan by comprehensively considering the overall and local seismic characteristics of the structure. In this way, it improves the calculation accuracy while ensuring the reliability of the optimisation plan. The model optimisation results also indicate an enhance level of optimisation.

Judging from the application of case study 2, the input parameters selected in this study to affect seismic damage analysis are reasonable for the common RC frame structure types in the existing research. It not only considers the structural characteristics of the overall and partial levels of the building, but also considers the characteristics of seismic waves and amplitude modulation factor. While ensuring a higher prediction effect, it also expands the application range of the model, which realizes the earthquake response prediction of group buildings under more complete earthquake conditions. The prediction results and generalization capabilities of the model are also ideal, providing support for the rapid assessment and prediction of urban-level building groups.

CHAPTER 8 CONCLUSION AND FUTURE WORKS

The notion of PBSD may effectively address the constraints of conventional standard design, for example, having only one precautionary objective and difficulties in satisfying certain demands. This concept has significant research significance. Given the ongoing progress of information technologies, it is crucial to implement digital transformation in the building engineering sector. Consequently, this subject launches a sequence of investigations into the PBSD of buildings. It combines BIM, semantic web, and AI technologies to enhance the digital proficiency of practical cases while solving theoretical problems. The research has produced a structure seismic performance assessment system of process automation and expandable content, providing technical assistance for the engineering implementation of PBSD. The study focuses on optimising the design for earthquake events and predicting earthquake damage, which builds upon prior research and takes into account both the macro and micro properties of the building and seismic characteristics. Thus, solutions with higher level of optimisation and a broader variety of applications for prediction. Furthermore, the research offers valuable information for mitigating earthquake damage in places of varying scales, focusing on both individual building and larger clusters of buildings.

8.1 CONCLUSION

Based on the theoretical background of performance-based design, this research uses BIM, semantic web, and AI technology to carry out seismic performance design research for individual buildings and regional buildings, and completes the six objectives stated in Section 1.5:

Objective 1: Identify domain knowledge, methodology and current practice of PBSD.

First of all, the concept of seismic design is introduced in the literature review. The disadvantages of traditional seismic design method FBD is explained, which calls for the rising requirement for secure and economical earthquake assessment and design. Therefore, various evolved design methods such as Monte Carlo simulation method, EBSD method, DBD method and PBSD are explored. Compared to other methods, PBSD has some obvious advantages which is clarified based on existing research.

Then the essential phases and methodology of PBSD are explained. FEMA P58 is the best practice of PBSD, which considers the various uncertainties in the earthquake action, integrates the influence of structural components, non-structural component and systems on the seismic performance of buildings. It expresses the seismic performance result as a series of intuitive and easy-to-understand performance indicators. Moreover, the detailed procedure of PBSD based on FEMA P-58 is illustrated as well.

At last, the current research status of PBSD based on FEMA P-58 are introduced. Some notable research gaps are therefore identified.

Objective 2: Explore digital technologies and use them to create a knowledge model that enable a standardized semantic format, allowing for interchange and querying of diverse information and knowledge from several sources.

Digital technologies such as BIM, ontology, MOGA and ANN are explored in the Section 2.3 and Section 2.4. Furthermore, a BIM-based ontology knowledge model is developed in Chapter 4. According to the seismic codes and FEMA P-58, all the relevant information required from seismic evaluation are collected. Based on the types of the seismic evaluation information, four ontologies are established. Sketch ontology specifies the structural data set essential for evaluation comprising basic building and design information. Aseismic ontology consists of structural seismic data and the outcomes of structural analysis. Fragile ontology encompasses the fundamental assessment information included in FEMA P-58. The terms of performance cluster and the rules of its division are defined in the Fragile ontology. Application ontology serves the function of elucidating the technique and objectives of performance assessment. In addition, a certain relationship is established among four ontologies as to facilitate information interaction throughout evaluation process and effectively convey the evaluation principle. This knowledge model also serves as the knowledge base for the automatic seismic assessment and multi-objective optimisation design.

Objective 3: A BSPEF would be proposed to realize the automation of building evaluation process.

The buildings' seismic performance evaluation will involve multi-source heterogeneous data including basic building information, structural analysis results, and maintenance cost valuation. This thesis first uses the three-dimensional digital model provided by BIM technology to store actual engineering information so that it can be parsed and transmitted in a structured format. Then, based on the ontology technology, organize key concepts and the relationships between concepts to form a performance evaluation framework under the guidance of FEMA P-58. Data/information in calculation books and other textual format was put in a unified semantic environment with building model information. Reasoning and query language was used to realize the association and interoperability of these information. Eventually, automated BSPEF will be realized under the entire framework organization.

Objective 4: Establish a Multi-objective Seismic optimisation Design Method (MSODM) for RC frame according to the BSPEF.

On the basis of BSPEF, combined with "return on investment" criterion, this research regards the structural seismic optimisation as a multi-objective optimisation problem that seeks to strike an equilibrium between initial building expense and anticipated earthquake loss. Starting from the overall and partial levels of structural seismic characteristics, a collection of multi-objective optimisation models with discrete dimensional variables is developed. The model expresses the two major costs as the corresponding function of the design plan with the component size as the variables, and selects the NSGA-II algorithm that can effectively deal with multiple conflicting targets of discrete variables for optimisation iteration. Optimisation result is a set of optional design plan and show the optimisation trade-off between the two major costs.

Objective 5: Establish an Earthquake Damage Prediction Method (EDPM) of multi-scale regional RC frame based on ANN and BSPEF.

In view of the time-consuming analysis of structural response, high computational cost, and the difficulty of popularization in regional building groups, this research takes the common RC frame structure as the research object, and establishes an alternative model for seismic response analysis of this type. On the one hand, as there are multiple parameters describing the seismic characteristics and seismic wave characteristics of the structure, an ANN model that can handle multi-variable and nonlinear relationships is selected. On the other hand, representative parameters that can reflect the macro and micro characteristics of the structure are selected. Moreover, the amplitude modulation coefficient is introduced into the seismic parameters to expand the application range of the prediction model. After the training has obtained a structural response model that can better predict this type of building under the action of any seismic wave considering the amplitude modulation factor, this study proposes to efficiently determine the seismic performance distribution of group buildings, and finally achieve multi-scale regional earthquake damage prediction.

Objective 6: Validate the application effect of BSPEF, MSODM and EDPM.

First, actual engineering cases are provided by Sichuan Provincial Government and perform automated under each seismic level. Then, the case is subjected to a multi-objective seismic optimisation design guided by the “return on investment” criterion to obtain a series of better design schemes that comprehensively consider safety and economic requirements. Then, based on ANN training on 5468 sample data, the earthquake response prediction model of a certain type of RC frame structure group building under the action of earthquake leveling within any consideration range is obtained, and the prediction effect and generalisation capabilities of the model are verified. Finally, take the combination of a certain seismic level and a certain seismic wave record as an example to show the seismic performance distribution of the building groups under this circumstance, reflecting the concept of multi-scale regional seismic precaution.

8.2 CONTRIBUTION

From the perspective of academic community, the contribution of the thesis mainly includes the following four points:

Contribution 1: A domain knowledge base of seismic assessment evaluation and multi-objective optimisation design has been established in this thesis. This ontology can be reused by other researchers for the relevant topics.

Contribution 2: A theoretical framework for seismic performance evaluation of buildings has been formed. It could achieve (1) expressing the process and logic of evaluation process by organizing the key concepts and the relationship between the concepts during the performance evaluation; (2) storing the multi-source heterogeneous information from different software such as Revit to provide basic building information, YJK to provide structural analysis results, and so on; (3) facilitating the automated evaluation of actual seismic performance of buildings.

At present, the research on building seismic performance evaluation has problems such as poor versatility of evaluation need to be improved. This thesis systematically sorts out the evaluation content and logic and proposes a structure seismic performance evaluation framework using BIM and ontology. The framework meets the evaluation requirements of various performance indicators in the existing FEMA P-58. The effectiveness and excellence of assessment is vastly enhanced, which provides support for the rapid prediction of urban-scale building clusters in the future. Additionally, the framework is extensible, which is not only reflected in the design space reserved for new buildings, but also in that the framework is easy to modify when the FEMA P-58 version is changed.

Contribution 3: A multi-objective optimisation model of discrete size variables is developed.

The seismic optimisation design of buildings based on the “return on investment” criterion often regards single variables such as precautionary intensity and reliability as optimisation parameters, and the minimum sum of the two objective functions of initial construction expense and seismic loss expectation is regarded as the optimisation criterion, leading to limited optimisation results which is confined to the overall level of building but not component level. The multi-objective optimisation model proposed in this study establishes a one-to-one correspondence between the design scheme and the two major costs with structural size as variables. It realizes the micro-level optimisation design and improves the optimisation level. With the help of the multi-objective optimisation algorithm NSGA-II, the two major costs are regarded as conflicting goals for multi-directional optimisation. The outcome is shown as a Pareto optimal design solution set for scientific decision-making by project stakeholders.

Contribution 4: a seismic response prediction model considering the macro and micro characteristics of the structure and the seismic wave amplitude modulation coefficient was established.

With the aid of ANN technology, this research takes the reinforce concrete frame structure as the research object and comprehensively considers the structural characteristics of the building and seismic wave characteristics. Then, the structural parameters and seismic parameters that affect the seismic response are reasonably selected and trained to obtain a model capable of predicting the structural response of a specific type of reinforced concrete frame structure due to a set of amplitude modulation seismic waves. Therefore, the earthquake damage distribution prediction of group buildings under more complete earthquake conditions and the rapid establishment of the collapse vulnerability function of individual buildings are realized.

From the perspective of practical application, there are four contributions concluded as follows:

Contribution 1: Design variables are treated as discrete variables but not continuous variables, which can accurately represent the construction and manufacturing needs as the size of component is us rally multiples of 50 or 100 for ease of production. Therefore, designers can obtain the most ideal design variables intuitively.

Contribution 2: Amplitude modulation factor has been taken into account for earthquake damage prediction in this thesis, so the designers can get more accurate structural earthquake response. This is because earthquake are abrupt phenomena that are transient and change rapidly. When this dynamic action is converted into an equivalent static action for design, the actual bearing capacity of the component in an earthquake is higher than when it is designed according to static forces.

Contribution 3: The performance level distribution of the regional building complex is displayed on the visualization platform and marked with different colors to represent the different performance level. Therefore, the designers can distinguish the performance level of individual buildings very intuitively.

Contribution 4: For performance evaluation and optimisation design, this thesis not only consider about structural stability and safety, but also consider the economy as well. It is more in line with the actual situation and therefore is more practical for designers.

8.3 FUTURE WORK

As the constraint of the research scope and time, a PhD project could only emphasize a specific object of a target domain. Limitations of this research are concluded as follow:

Limitation 1: the scope of this research is relatively narrow.

PBSD is a comprehensive theoretical framework that allows design to include considerations of safety, cost-effectiveness, and environmental preservation by fully assessing performance. This research only examines the objective of optimising building maintenance costs as a representative for estimating earthquake-related losses. Moreover, this thesis only focuses on the RC frames, which is limited by the type of research cases provided by Sichuan Provincial Government.

Limitation 2: the efficiency of traversing IFC files is relatively low.

In the procedure of extracting information from IFC files for ontology instantiation, all data in IFC files should be traversed which is time-consuming.

Limitation 3: there is no user-friendly interface for designers.

In this research, different prototype systems are developed to answer the research question. To prove an original concept, the establishment of these systems are conducted in different software environment such as protege-OWL, Revit etc. It is full of challenges for customers with low or no expertise of the programme and its plug-ins.

According to limitations stated above, future work of this thesis can be concluded:

Future work 1: expand the scope of this research.

As society's need for structure's performance continues to rise, the assessment of sustainability factors such as greenhouse gases emission will have practical importance as well. Additionally, in the context of analysing seismic damage, it is advantageous to acquire several seismic damage indexes for structures, such as the IDR, PFA and other multi-response predictions. This approach proves to be beneficial in enhancing performance assessment and optimising efficiency in design. Further research and theoretical verification are urgently needed.

Moreover, this research only takes RC frame building as research objectives, more building types can be further regraded as research objectives to represent more comprehensive situation in practice.

Future work 2: research on methods of obtaining the exact data from IFC files.

As stated above, the IFC files are traversed in order to get the required data for ontology instantiation in this thesis. In order to improve the efficiency of the automatic seismic evaluation and optimisation design process, new methods or approaches should be studied.

Future work 3: develop an interface.

It is necessary to develop a user-friendly interface to improve the usability for customers with low or no expertise of the programme and its plug-ins. The easy-to-access platform can facilitate the decision-making process for different construction projects.

Future work 4: apply the methodology for other domains.

The methodology of establishing building seismic performance evaluation framework could be applied for ontology development in construction-related research domains, for example, precautionary construction management, sustainable construction, and supply chain management. Consequently, the ontology built in the thesis can be reused as a semantic resource for numerous other applications in ACE industry. It provides an approach to store and use multi-source heterogeneous data and information. The interoperability and extensibility of ontology can also facilitate the acquisition of solutions that integrate all phase of the building's life cycle and enable stakeholders to make decisions with a comprehensive and holistic view.

REFERENCE

Abanda, F. H., Kamsu-Foguem, B., & Tah, J. H. M. (2017). BIM – New rules of measurement ontology for construction cost estimation. *Engineering Science and Technology, an International Journal, 20*(2), 443-459. doi:https://doi.org/10.1016/j.jestch.2017.01.007

Abanda, F. H., Tah, J. H. M., & Keivani, R. (2013). Trends in built environment semantic Web applications: Where are we today? *Expert Systems with Applications, 40*(14), 5563-5577. doi:https://doi.org/10.1016/j.eswa.2013.04.027

Abdel-Khalek, S., Ben Ishak, A., Omer, O. A., & Obada, A. S. F. (2017). A two-dimensional image segmentation method based on genetic algorithm and entropy. *Optik, 131*, 414-422. doi:https://doi.org/10.1016/j.ijleo.2016.11.039

Aberdeen, T. (2013). Case study research: Design and methods (4th Ed.). *canadian journal of action research*.

Anumba, C. J., Pan, J., Issa, R., & Mutis, I. (2006). Construction project information management in a semantic web environment. *Engineering Construction & Architectural Management, 15*(1), 78-94.

Anumba, C. J., Pan, J., Issa, R., & Mutis, I. (2008). Construction project information management in a semantic web environment. *Engineering Construction & Architectural Management, 15*(1), 78-94.

Arslan, M. H. (2010). An evaluation of effective design parameters on earthquake performance of RC buildings using neural networks. *Engineering Structures, 32*(7), 1888-1898. doi:https://doi.org/10.1016/j.engstruct.2010.03.010

Beetz, J., Leeuwen, J. V., & Vries, B. D. (2009). IfcOWL: A case of transforming EXPRESS schemas into ontologies. *Artificial Intelligence for Engineering Design Analysis & Manufacturing, 23*(01), 89-101.

Benneworth, P. (2003). Doing regional research in a devolving state: Research methodology and strategy in the post-devolution United Kingdom. *Regions Magazine, 243*(1), 12-13.

Boddy, S., Rezgui, Y., Cooper, G., & Wetherill, M. (2007). Computer integrated construction: A review and proposals for future direction. *Advances in Engineering Software, 38*(10), 677-687. doi:https://doi.org/10.1016/j.advengsoft.2006.10.007

Borst, W. N. (1997). Construction of Engineering Ontologies for Knowledge Sharing and Reuse. *universiteit twente*.

Bouška, R. (2016). Evaluation of Maturity of BIM Tools across Different Software Platforms. *Procedia Engineering, 164*, 481-486. doi:https://doi.org/10.1016/j.proeng.2016.11.648

Breitman, K. K., Casanova, M. A., & Truszkowski, W. (2007). *Semantic Web: Concepts, Technologies and Applications*: Springer-Verlag London.

Brewis, & Claire. (2000). Social Research Methods: Qualitative and Quantitative Approaches. *International Journal of Social Research Methodology*.

Briza, A. C., & Naval, P. C. (2011). Stock trading system based on the multi-objective particle swarm optimization of technical indicators on end-of-day market data. *Applied Soft Computing, 11*(1), 1191-1201.

Brooks, T. J. (2014). *Building Information Modeling: Beyond design, commissioning and construction.* Clemson University.,

Cao, Y., & Zhao, Z. (2009). *Improved structural reliability analysis BP neural network response surface method.* Paper presented at the The second structure international forum on New Engineering Progress.

Cardoso, J. B., Almeida, J. R. D., Dias, J. M., & Coelho, P. G. (2008). Structural reliability analysis using Monte Carlo simulation and neural networks. *Advances in Engineering Software, 39*(6), 505-513.

Cavalliere, C., Dell'Osso, G. R., Favia, F., & Lovicario, M. (2019). BIM-based assessment metrics for the functional flexibility of building designs. *Automation in Construction, 107*, 102925. doi:https://doi.org/10.1016/j.autcon.2019.102925

Chandra, A. K., & Harel, D. (1985). Horn clause queries and generalizations. *Journal of Logic Programming, 2*(1), 1-15.

Chaudhary, D., & Kumar, B. (2019). Cost optimized Hybrid Genetic-Gravitational Search Algorithm for load scheduling in Cloud Computing. *Applied Soft Computing, 83*, 105627. doi:https://doi.org/10.1016/j.asoc.2019.105627

Chen, W. S., & Hirschheim, R. (2010). A paradigmatic and methodological examination of information systems research from 1991 to 2001. *Information Systems Journal, 14*(3), 197-235.

Cheng, J., Li, Q. S., & Xiao, R.-c. (2008). A new artificial neural network-based response surface method for structural reliability analysis. *Probabilistic Engineering Mechanics, 23*(1), 51-63. doi:https://doi.org/10.1016/j.probengmech.2007.10.003

Choudhury, B., & Chandrasekaran, M. (2020). Electron beam welding of aerospace alloy (Inconel 825): A comparative study of RSM and ANN modeling to predict weld bead area. *Optik, 219*, 165206. doi:https://doi.org/10.1016/j.ijleo.2020.165206

Concepts, R., Charette, R. P., & Marshall, H. E. (2010). UNIFORMAT II Elemental Classification for Building Specifications, Cost Estimating, and Cost Analysis. *NIST Interagency/Internal Report (NISTIR) - 6389*.

Cornell, C. A., Jalayer, F., Hamburger, R. O., & Foutch, D. A. (2002). Probabilistic Basis for 2000 SAC Federal Emergency Management Agency Steel Moment Frame Guidelines. *Journal of Structural Engineering, 128*(4), 526-533.

Cortés-Pérez, J. P., Cortés-Pérez, A., & Prieto-Muriel, P. (2020). BIM-integrated management of occupational hazards in building construction and maintenance. *Automation in Construction, 113*, 103115. doi:https://doi.org/10.1016/j.autcon.2020.103115

Dawood, N. (2009). Development of 4D based performance indicators in construction industry. *Engineering Construction & Architectural Management, 17*(2), 210-230.

de Lautour, O. R., & Omenzetter, P. (2009). Prediction of seismic-induced structural damage using artificial neural networks. *Engineering Structures, 31*(2), 600-606. doi:https://doi.org/10.1016/j.engstruct.2008.11.010

Deb, K., Jain, P., Gupta, N. K., & Maji, H. K. (2015). Multiobjective placement of electronic components using evolutionary algorithms. *IEEE Transactions on Components & Packaging Technologies, 27*(3), 480-492.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation, 6*(2), 182-197.

Del Gobbo, G. M., Williams, M. S., & Blakeborough, A. (2018). Seismic performance assessment of Eurocode 8-compliant concentric braced frame buildings using FEMA P-58. *Engineering Structures, 155*, 192-208. doi:https://doi.org/10.1016/j.engstruct.2017.11.016

Dhabale, R., Jatti, V. S., & Singh, T. P. (2014). Multi-objective Optimization of Turning Process During Machining of AlMg1SiCu Using Non-dominated Sorted Genetic Algorithm. *Procedia Materials Science, 6*, 961-966. doi:https://doi.org/10.1016/j.mspro.2014.07.166

Dhanaraj, & C. (2006). Research Methodology in Strategy and Management. *Academy of Management Review, 31*(2), 497-500.

Ding, L. Y., Zhong, B. T., Wu, S., & Luo, H. B. (2016). Construction risk knowledge management in BIM using ontology and semantic web technology. *Safety Science, 87*, 202-213. doi:https://doi.org/10.1016/j.ssci.2016.04.008

Du, A., & Padgett, J. E. (2020). Investigation of multivariate seismic surrogate demand modeling for multi-response structural systems. *Engineering Structures, 207*, 110210. doi:https://doi.org/10.1016/j.engstruct.2020.110210

Eastman, C. M. (2008). BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers.

Eleftheriadis, S., Duffour, P., Greening, James, . . . Mumovic. (2018). Investigating relationships between cost and CO2 emissions in reinforced concrete structures using a BIM-based design optimisation approach. *Energy & Buildings*.

Elenas, A., & Meskouris, K. (2001). Correlation study between seismic acceleration parameters and damage indices of structures. *Engineering Structures, 23*(6), 698-704. doi:https://doi.org/10.1016/S0141-0296(00)00074-2

Falcón-Cardona, J. G., Hernández Gómez, R., Coello Coello, C. A., & Castillo Tapia, M. G. (2021). Parallel Multi-Objective Evolutionary Algorithms: A Comprehensive Survey. *Swarm and Evolutionary Computation, 67*, 100960. doi:https://doi.org/10.1016/j.swevo.2021.100960

Ferner, H., Wemyss, M., Baird, A., & Beer, A. Seismic performance of non-structural elements within buildings.

Fonseca, C. M., & Fleming, P. J. (1993). *Multiobjective genetic algorithms.* Paper presented at the IEEE Colloquium on Genetic Algorithms for Control Systems Engineering.

Gauchi, J. P., Bensadoun, A., Colas, F., & Colbach, N. (2017). Metamodeling and global sensitivity analysis for computer models with correlated inputs: A practical approach tested with a 3D light interception computer model. *Environmental Modelling & Software, 92*, 40-56. doi:https://doi.org/10.1016/j.envsoft.2016.12.005

Geiß, C., Aravena Pelizari, P., Marconcini, M., Sengara, W., Edwards, M., Lakes, T., & Taubenböck, H. (2015). Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques. *ISPRS Journal of Photogrammetry and Remote Sensing, 104*, 175-188. doi:https://doi.org/10.1016/j.isprsjprs.2014.07.016

Geraci, A., Katki, F., Mcmonegal, L., Meyer, B., & Porteous, H. (2002). *IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries.* Paper presented at the IEEE Std.

Gharehbaghi, S., Gandomi, M., Plevris, V., & Gandomi, A. H. (2021). Prediction of seismic damage spectra using computational intelligence methods. *Computers & Structures, 253*, 106584. doi:https://doi.org/10.1016/j.compstruc.2021.106584

Goda, K., & Tesfamariam, S. (2015). Multi-variate seismic demand modelling using copulas: Application to non-ductile reinforced concrete frame in Victoria, Canada. *Structural Safety, 56*, 39-51. doi:https://doi.org/10.1016/j.strusafe.2015.05.004

Gray, D. E. (2009). Doing Research in the Real World. *Sage*.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies, 43*(5), 907-928. doi:https://doi.org/10.1006/ijhc.1995.1081

Guarino, N. (1997). Understanding, building and using ontologies. *International Journal of Human Computer Studies, 46*(2-3), 293-310.

Guarino, N. (2009). *What Is an Ontology?* : Building Ontologies With Basic Formal Ontology.

Guo, H. (2012). Influence of infill wall on seismic performance of reinforced concrete frame structure. *Architecture of Shanxi*(36), 38-39.

Hajder, M., Kolbusz, J., Hajder, P., Nycz, M., & Liput, M. (2020). Data Security Platform Model in Networked Medical IT Systems based on Statistical Classifiers and ANN. *Procedia Computer Science, 176*, 3682-3691. doi:https://doi.org/10.1016/j.procs.2020.09.018

Handl, J., & Knowles, J. (2007). An Evolutionary Approach to Multiobjective Clustering. *IEEE Transactions on Evolutionary Computation, 11*, 56-76.

Hardin, S. (2010). Tim Berners-Lee: The semantic Web-Web of machine-processable data. *Bulletin of the American Society for Information Science & Technology, 31*(3), 12-13.

He, H. S. (2005). IFC-based framework for evaluating building envelope performance. *Building Civil & Environmental Engineering*.

Henry, L. (2004). Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential. *Online Information Review, 28*(3), 242-243.
Hietanen, J. (2006). IFC model view definition format.
Hong, F., & Xie, L. (1999). Decision analysis of structure seismic fortification. *Earthquake Engineering and Engineering Vibration*(02), 9-14.
Hornik, K. M., Stinchcomb, M., & White, H. (1989). Multilayer feedforward networks are universal approximator. *Neural Networks, 2*(5).
Huang, C., & Huang, S. (2020). Predicting capacity model and seismic fragility estimation for RC bridge based on artificial neural network. *Structures, 27*, 1930-1939. doi:https://doi.org/10.1016/j.istruc.2020.07.063
Hubert, G., Shen, X., Stephen, G., & Victoria, F. (2014). Considering the Feasibility of Semantic Model Design in the Built-Environment. *Buildings, 4*(4), 849-879.
Hui, L., & Zhang, Q. (2009). Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II. *IEEE Transactions on Evolutionary Computation, 13*(2), 284-302.
Iacobucci, D., & Churchill, G. A. (2010). Marketing Research: Methodological Foundation,(with Qualtrics Card). *Instructors Manual for Marketing Research Methodological Foundations, 14*(1), 131.
Ismail, H. Y., Singh, M., Shirazian, S., Albadarin, A. B., & Walker, G. M. (2020). Development of high-performance hybrid ANN-finite volume scheme (ANN-FVS) for simulation of pharmaceutical continuous granulation. *Chemical Engineering Research and Design, 163*, 320-326. doi:https://doi.org/10.1016/j.cherd.2020.09.002
Jeong, Y. S., Eastman, C. M., Sacks, R., & Kaner, I. (2009). Benchmark tests for BIM data exchanges of precast concrete. *Automation in Construction, 18*(4), 469-484. doi:https://doi.org/10.1016/j.autcon.2008.11.001
Ji, K. (2018). *strong ground motion selection for multiple levels of seismic fortification demand in China.* Institute of Engineering Mechanics,
Jia, D.-W., & Wu, Z.-Y. (2021). Seismic fragility analysis of RC frame-shear wall structure under multidimensional performance limit state based on ensemble neural network. *Engineering Structures, 246*, 112975. doi:https://doi.org/10.1016/j.engstruct.2021.112975
Jianping, H., Shiyan, C., Xiaohui, Y., & Dagang, L. (2020). Influence of Earthquake Sustainment on Vulnerability and Seismic Performance of RC Frame Structure. *Journal of Building Structures, 42*, 116-127. doi:10.14006/j.jzjgxb.2019.0677
Johnson, R., & Onwuegbuzie, A. (2004). Mixed Methods Research: Paradigm Whose Time Has Come. *Educ.Res., vol.33,no.7,pp.14-26*.
Kanda, J., & Shah, H. (1997). Engineering role in failure cost evaluation for buildings. *Structural Safety, 19*(1), 79-90. doi:https://doi.org/10.1016/S0167-4730(96)00039-2
Kappos, L. A. J., & Eng, D. (2010). Seismic damage indices for RC buildings: evaluation of concepts and procedures. *Progress in Structural Engineering & Materials, 1*(1), 78-87.
Karan, E. P., & Irizarry, J. (2015). Extending BIM interoperability to preconstruction operations using geospatial analyses and semantic web services. *Automation in Construction, 53*, 1-12. doi:https://doi.org/10.1016/j.autcon.2015.02.012
Kaya, M. (2018). *Multi-objective genetic algorithm based approaches for mining optimized fuzzy association rules.* Paper presented at the Springer-Verlag.
Khalili, A. (2021). An XML-based approach for geo-semantic data exchange from BIM to VR applications. *Automation in Construction, 121*, 103425. doi:https://doi.org/10.1016/j.autcon.2020.103425
Kim, H., Kwon, S., Heo, J., Lee, H., & Chung, M. K. (2014). The effect of touch-key size on the usability of In-Vehicle Information Systems and driving safety during simulated driving. *Applied Ergonomics, 45*(3), 379-388. doi:https://doi.org/10.1016/j.apergo.2013.05.006
Kita, A., Cavalagli, N., Masciotta, M. G., Lourenço, P. B., & Ubertini, F. (2020). Rapid post-earthquake damage localization and quantification in masonry structures through multidimensional non-

linear seismic IDA. *Engineering Structures, 219*, 110841. doi:https://doi.org/10.1016/j.engstruct.2020.110841

Koo, B., & Shin, B. (2018). Applying novelty detection to identify model element to IFC class misclassifications on architectural and infrastructure Building Information Models. *Journal of Computational Design and Engineering, 5*(4), 391-400. doi:https://doi.org/10.1016/j.jcde.2018.03.002

Kubba, S. (2017). Chapter Five - Building Information Modeling (BIM). In S. Kubba (Ed.), *Handbook of Green Building Design and Construction (Second Edition)* (pp. 227-256): Butterworth-Heinemann.

Kumar, R. (2014). Research Methodology: a Step-By-Step Guide For Beginners. *Sage*.

Lan, Anh, & Tran. (1983). A Semantic Web Primer. *Library Hi Tech*.

Lawrence, N. (2005). Social research methods : qualitative and quantitative approaches. *Social Research Methods Qualitative & Quantitative Approaches, 39*(3), 447-448.

Lawrence, N. W. (2000). Social Research Methods: Qualitative and Quantitative Approaches. *Teaching Sociology, 30*(3).

Lee, S.-I., Bae, J.-S., & Cho, Y. S. (2012). Efficiency analysis of Set-based Design with structural building information modeling (S-BIM) on high-rise building structures. *Automation in Construction, 23*, 20-32. doi:https://doi.org/10.1016/j.autcon.2011.12.008

Lee, X. (2018). *residential house seismic research based on BP neural network.* Institute of Disaster Prevention,

Li, Y., Liu, J., Wang, Z., Jin, C., Hao, J., & Li, H. (2021). Investigation of seismic performance of RC column with electrochemical chloride extraction-strengthening by MPC-CFRP. *Engineering Structures, 247*, 113228. doi:https://doi.org/10.1016/j.engstruct.2021.113228

Liebenau, J., & Lee, A. S. (1997). *Information Systems and Qualitative Research*: Information Systems and Qualitative Research.

Lima, C., Diraby, T. E., & Stephens, J. (2005). Ontology-based optimisation of knowledge management in e-Construction. *Electronic Journal of Information Technology in Construction, 10*, 305-327.

Lisboa, A., & Soares, D. (2014). E-government Interoperability Frameworks: A Worldwide Inventory. *Procedia Technology, 16*, 638-648. doi:https://doi.org/10.1016/j.protcy.2014.10.012

Liu, M., Wen, Y. K., & Burns, S. A. (2004). Life cycle cost oriented seismic design optimization of steel moment frame structures with risk-taking preference. *Engineering Structures, 26*(10), 1407-1421. doi:https://doi.org/10.1016/j.engstruct.2004.05.015

Liu, X. (2010). Research on data normalizaton of input layer of BP Neural Network. *Mechanical Engineering and Automation, 000*(003), 122-123.

Lorleen, & Farrugia. (2019). WASP (write a scientific paper): The ongoing process of ethical decision-making in qualitative research: Ethical principles and their application to the research process. *Early human development*.

Lv, D., Liu, T., LEE, S., & Yu, X. (2018). Analysis of the influence of target spectrum and amplitude modulation method on selection of ground motion. *Earthquake Engineering and Engineering Vibration, 38*(04), 021-028.

M, C. C. (2001). Optimal lateral stiffness design of tall buildings of mixed steel and concrete construction. *The Structural Design of Tall Buildings, 10*(3), 155-177.

Maceviciute, E. (2006). Book review: Researching information systems and computing. *Professor T.d.wilson*.

MacKay, & Da Vid, J. C. (1992). Bayesian Interpolation. *Neural Computation, 4*(3), 415-447.

Mackenzie, N., & Knipe, S. (2006). Research dilemmas: Paradigms, methods and methodology. *Issues in Educational Research, 16*(2), 193-205.

Martí, L., García, J., Berlanga, A., & Molina, J. M. (2016). A stopping criterion for multi-objective optimization evolutionary algorithms. *Information Sciences, 367-368*, 700-718. doi:https://doi.org/10.1016/j.ins.2016.07.025

Mcculloch, W. S., & Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *biol math biophys*.

Medjdoub, B., Richens, P., & Barnard, N. (2001). Building Services Standard Solutions. *Springer Netherlands*.

Mike, Uschold, Michael, & Gruninger. (1996). Ontologies: principles, methods and applications. *Knowledge Engineering Review, 11,93-136*.

Min, L., Burns, S. A., & Wen, Y. K. (2010). Multiobjective optimization for performance-based seismic design of steel moment frame structures. *Earthquake Engineering & Structural Dynamics, 34*(3), 289-306.

Mittal, A., Devi, G., & Chauhan, P. (2014). *Application of ANN to Predict Liquefaction Potential of Soil Deposits for Chandigarh Region, India*: Proceedings of the Third International Conference on Soft Computing for Problem Solving.

Morfidis, K., & Kostinakis, K. (2018). Approaches to the rapid seismic damage prediction of r/c buildings using artificial neural networks. *Engineering Structures, 165*, 120-141. doi:https://doi.org/10.1016/j.engstruct.2018.03.028

Naeim, F. (1989). *The Seismic Design Handbook*: The Seismic Design Handbook.

Neches, R. (1991). Enabling Technology for Knowledge Sharing. *AAAI AI Mag, 12*.

Ohtsuka, S., Teshima, T., Matsumoto, S., & Hikita, M. (2006). *Relationship between PD-induced electromagnetic wave measured with UHF method and charge quantity obtained by PD current waveform in model GIS.* Paper presented at the Electrical Insulation and Dielectric Phenomena, 2006 IEEE Conference on.

Okasha, N. M., & Frangopol, D. M. (2009). Lifetime-oriented multi-objective optimization of structural maintenance considering system reliability, redundancy and life-cycle cost using GA. *Structural Safety, 31*(6), 460-474. doi:https://doi.org/10.1016/j.strusafe.2009.06.005

P., Pauwels, and, ., D., Van, Deursen, . . . and. (2011). A semantic rule checking environment for building performance checking. *Automation in Construction, 20*(5), 506-518.

Park, H. S., Hwang, J. W., & Oh, B. K. (2018). Integrated analysis model for assessing CO2 emissions, seismic performance, and costs of buildings through performance-based optimal seismic design with sustainability. *Energy and Buildings, 158*, 761-775. doi:https://doi.org/10.1016/j.enbuild.2017.10.070

Park, H. S., Hwang, J. W., & Oh, B. K. (2018). Integrated analysis model for assessing CO_2 emissions, seismic performance, and costs of buildings through performance-based optimal seismic design with sustainability. *Energy & Buildings, 158*(PT.1), 761-775.

Pazlar, T., & Turk, Z. (2008). Interoperability in practice : geometric data exchange using the IFC standard. *Electronic Journal of Information Technology in Construction, 13*, 362-380.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry, 25*(13), 1605-1612.

Pinheiro, S., Wimmer, R., O'Donnell, J., Muhic, S., Bazjanac, V., Maile, T., . . . van Treeck, C. (2018). MVD based information exchange between BIM and building energy performance simulation. *Automation in Construction, 90*, 91-103. doi:https://doi.org/10.1016/j.autcon.2018.02.009

Polat, G., Bingol, B. N., Gurgun, A. P., & Yel, B. (2016). Comparison of ANN and MRA Approaches to Estimate Bid Mark-up Size in Public Construction Projects. *Procedia Engineering, 164*, 331-338. doi:https://doi.org/10.1016/j.proeng.2016.11.627

Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies, 47*(11), 1451-1458.

Rabinovich, Y. I. (2017). Universal procedure for constructing a Pareto set. *Computational Mathematics and Mathematical Physics, 57*(1), 45-63.
Rafiq, M. Y., Bugmann, G., & Easterbrook, D. J. (2001). Neural network design for engineering applications. *Computers & Structures, 79*(17), 1541-1552. doi:https://doi.org/10.1016/S0045-7949(01)00039-6
Razavi, N., & Gholizadeh, S. (2021). Seismic collapse safety analysis of performance-based optimally designed reinforced concrete frames considering life-cycle cost. *Journal of Building Engineering, 44*, 103430. doi:https://doi.org/10.1016/j.jobe.2021.103430
Razmi, A., Rahbar, M., & Bemanian, M. (2019). PCA-ANN integrated NSGA-III framework for dormitory building design optimization: Energy efficiency, daylight, and thermal comfort. *Applied Energy, 305*, 117828. doi:https://doi.org/10.1016/j.apenergy.2021.117828
Ren, G., Li, H., Liu, S., Goonetillake, J., Khudhair, A., & Arthur, S. (2021). Aligning BIM and ontology for information retrieve and reasoning in value for money assessment. *Automation in Construction, 124*, 103565. doi:https://doi.org/10.1016/j.autcon.2021.103565
Rezgui, Y., & Miles, J. (2009). *Transforming SME strategies via innovative transient knowledge-based alliances in the construction sector.* Paper presented at the IEEE International Conference on Industrial Informatics.
Ribeiro, L. d. A., Soares, A. d. S., Lima, T. W. d., Jorge, C. A. C., Costa, R. M. d., Salvini, R. L., . . . Gabriel, P. H. R. (2015). Multi-objective Genetic Algorithm for Variable Selection in Multivariate Classification Problems: A Case Study in Verification of Biodiesel Adulteration. *Procedia Computer Science, 51*, 346-355. doi:https://doi.org/10.1016/j.procs.2015.05.254
Saldaña-Robles, A. L., Bustos-Gaytán, A., Diosdado-De la Peña, J. A., Saldaña-Robles, A., Alcántar-Camarena, V., Balvantín-García, A., & Saldaña-Robles, N. (2020). Structural design of an agricultural backhoe using TA, FEA, RSM and ANN. *Computers and Electronics in Agriculture, 172*, 105278. doi:https://doi.org/10.1016/j.compag.2020.105278
Sarcheshmehpour, M., & Estekanchi, H. E. (2021). Life cycle cost optimization of earthquake-resistant steel framed tube tall buildings. *Structures, 30*, 585-601. doi:https://doi.org/10.1016/j.istruc.2021.01.038
Saunders, M., Lewis, P., & Thornhill, A. (2000). *Research Methods for Business Students (2nd edition)*: Research methods for business students /.
Saunders, M., Lewis, P., & Thornhill, A. (2011). Research Methods for Business Students. *Qualitative Market Research, 3*(4), 215 - 218.
Schmitz, M., Ramírez, K., Mazuera, F., Ávila, J., Yegres, L., Bezada, M., & Levander, A. (2021). Moho depth map of northern Venezuela based on wide-angle seismic studies. *Journal of South American Earth Sciences, 107*, 103088. doi:https://doi.org/10.1016/j.jsames.2020.103088
Shadbolt, N., & Alani, H. (2011). *Handbook of Semantic Web Technologies*: Handbook of Semantic Web Technologies.
Silva, V. V. R., Fleming, P. J., Sugimoto, J., & Yokoyama, R. (2008). Multiobjective optimization using variable complexity modelling for control system design. *Applied Soft Computing, 8*(1), 392-401.
Singh, S., & Abbassi, H. (2018). 1D/3D transient HVAC thermal modeling of an off-highway machinery cabin using CFD-ANN hybrid method. *Applied Thermal Engineering, 135*, 406-417. doi:https://doi.org/10.1016/j.applthermaleng.2018.02.054
Sofaer, S. (2002). Qualitative research methods. *Int J Qual Health Care, 14*(4), 329-336.
Sommerville, I. (2010). Software Engineering. *software engineering*.
Srinivas, N., & Deb, K. (1994). Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. *Evolutionary Computation, 2*(3), 221-248.
Swartout, & William. (1999). Ontologies. *IEEE Intelligent Systems & Their Applications*.
Tan, K. C., Cheong, C. Y., & Goh, C. K. (2007). Solving multiobjective vehicle routing problem with stochastic demand via evolutionary computation. *European Journal of Operational Research, 177*(2), 813-839. doi:https://doi.org/10.1016/j.ejor.2005.12.029

Tashakkori, & Teddie. (2003). Handbook of Mixed Methods in Social & Behavioral Research. *Sage Publications*.
Taye, M. M. (2010). Understanding Semantic Web and Ontologies: Theory and Applications. *computer science*.
Thomas, & G. (2011). A Typology for the Case Study in Social Science Following a Review of Definition, Discourse, and Structure. *Qualitative Inquiry, 17*(6), 511-521.
Tim, B.-L. (1994). Tim Berners-Lee: WorldWideWeb.
Venugopal, M., Eastman, C. M., & Teizer, J. (2015). An ontology-based analysis of the industry foundation class schema for building information model exchanges. *Advanced Engineering Informatics, 29*(4), 940-957. doi:https://doi.org/10.1016/j.aei.2015.09.006
Verstraeten, R., Pauwels, P., Meyer, R. D., Meeus, W., & Lateur, G. (2009). *IFC-based calculation of the Flemish Energy Performance Standard*: eWork and eBusiness in Architecture, Engineering and Construction - ECPPM 2008.
Walle, R., Deursen, D. V., Pauwels, P., Verstraeten, R., Campenhout, J. V., Roo, J. D., & Meyer, R. D. A semantic rule checking environment for building performance checking.
Walsham. (1995). Interpretive case studies in IS research: nature and method. *European Journal of Information Systems, 4*(2), 74-81.
Walsham, G. (1995). The Emerigence of Interpretivism in IS Research. *Information Systems Journal, vol.6, no.4,pp 376-394*(4), 307-309.
Wang, Y., Li, C., Jin, X., Xiang, Y., & Li, X. (2020). Multi-objective optimization of rolling schedule for tandem cold strip rolling based on NSGA-II. *Journal of Manufacturing Processes, 60*, 257-267. doi:https://doi.org/10.1016/j.jmapro.2020.10.061
Warren, P., & Alsmeyer, D. (2005). *Applying semantic technology to a digital library: a case study*: Semantic Web Technologies: Trends and Research in Ontology-based Systems.
Wehmeier, & Hornby. (2000). *Oxford advanced learner's dictionary of current English*: Oxford advanced learner's dictionary of current English.
Welty, C., & Guarino, N. (2001). Supporting ontological analysis of taxonomic relationships. *Data & Knowledge Engineering, 39*(1), 51-74. doi:https://doi.org/10.1016/S0169-023X(01)00030-1
Wen, R., Ji, K., & Ren, Y. (2019). Review on selection of strong motion input for structural time-history dynamic analysis 结构时程分析中多层次强震动记录输入选取研究综述. *39*, 18. doi:10.13197/j.eeev.2019.05.1.wenrz.001
Wetherill, M., Rezgui, Y., Lima, C., & Zarli, A. (2002). Knowledge management for the construction industry: the e-cognos project. *Electronic Journal of Information Technology in Construction, 7*.
Wix, J., & Liebich, T. Industry Foundation Classes Ifc 2x.
Wood-Harper, Antill, & Avison. (1985). *Information systems definition : the multiview approach*: Information systems definition : the multiview approach.
Xiong, M., & Huang, Y. (2019). Novel perspective of seismic performance-based evaluation and design for resilient and sustainable slope engineering. *Engineering Geology, 262*, 105356. doi:https://doi.org/10.1016/j.enggeo.2019.105356
Xu, Z., Zhang, H., Lu, X., Xu, Y., Zhang, Z., & Li, Y. (2019). A prediction method of building seismic loss based on BIM and FEMA P-58. *Automation in Construction, 102*, 245-257. doi:https://doi.org/10.1016/j.autcon.2019.02.017
Yan, D. (2019). FEMA P-58-based seismic performance evaluation of reinforced concrete frame structures.
Yang, Q. Z., & Zhang, Y. (2006). Semantic interoperability in building design: Methods and tools. *Computer-Aided Design, 38*(10), 1099-1112. doi:https://doi.org/10.1016/j.cad.2006.06.003
Yang, T. Y., Moehle, J., Stojadinovic, B., & Kiureghian, A. D. (2009). Seismic Performance Evaluation of Facilities: Methodology and Implementation. *Journal of Structural Engineering, 135*(10), 1146-1154.
Yin, R. K. (2014). Case Study Research Designs and Methods. *Sage Publications*.

Zangeneh, P., & McCabe, B. (2020). Ontology-based knowledge representation for industrial megaprojects analytics using linked data and the semantic web. *Advanced Engineering Informatics, 46*, 101164. doi:https://doi.org/10.1016/j.aei.2020.101164

Zeng, X., Deng, K., Kurata, M., Duan, J., & Zhao, C. (2020). Seismic performance evaluation of damage-controlled composite steel frame with flexible-gel-covered studs. *Engineering Structures, 219*, 110855. doi:https://doi.org/10.1016/j.engstruct.2020.110855

Zeng, X., Lu, X., Yang, T. Y., & Xu, Z. (2016). Application of the FEMA-P58 methodology for regional earthquake loss prediction. *Natural Hazards, 83*(1), 177-192.

Zhang, N. (2015). *safety and economic evaluation of seismic wall structure and analysis of influencing factors.* Chongqing University,

Zhang, Y., & Rockett, P. I. (2011). A generic optimising feature extraction method using multiobjective genetic programming. *Applied Soft Computing, 11*(1), 1087-1097. doi:https://doi.org/10.1016/j.asoc.2010.02.008

Zheng, S., Xiang, Z., & Zheng, J. (2016). Research on China's building earthquake insurance system and the determination of insurance rates. *Journal of Catastrophology, 31*(003), 1-7,19.

Zhong, B., Gan, C., Luo, H., & Xing, X. (2018). Ontology-based framework for building environmental monitoring and compliance checking under BIM environment. *Building and Environment, 141*, 127-142. doi:https://doi.org/10.1016/j.buildenv.2018.05.046