Deep Learning-based Prediction of Seismic Responses and Losses of Nonlinear Structural Systems

딥러닝 기반 비선형 구조시스템의
지진 응답 및 손실 예측

2021년 2월

서울대학교 대학원

건설환경공학부

김 태 용

Deep Learning-based Prediction of Seismic Responses and Losses of Nonlinear Structural Systems

딥러닝 기반 비선형 구조시스템의
지진 응답 및 손실 예측

지도교수 송 준 호

이 논문을 공학박사 학위논문으로 제출함
2021년 2월

서울대학교 대학원
건설환경공학부
김 태 용

김태용의 공학박사 학위논문을 인준함
2021년 2월

위 원 장	김 호 경	_(인)_
부 위 원 장	송 준 호	_(인)_
위 원	문 주 혁	_(인)_
위 원	권 오 성	_(인)_
위 원	윤 형 철	_(인)_

Abstract

Structural failures caused by a strong earthquake may induce a large number of casualties and huge economic losses. To enhance life safety and disaster-resilience of communities, the current approach aims to design a structure that can withstand a design level earthquake event. To achieve the critical design objective, it is necessary to estimate the nonlinear structural responses under strong earthquake ground motions. Although nonlinear time history analysis is the only possible way to precisely estimate the structural responses in many situations, high computational cost and modeling time may hamper the adoption of the approach in routine engineering practice. Thus, in modern seismic design codes, various simplified regression equations are introduced to replace the onerous and time-consuming nonlinear time history analysis, but the accuracy of the estimated responses is limited. Moreover, the existing methods cannot quantify the uncertain errors in the response estimation, mainly caused by the loss of information in representing input data by selected features, especially regarding ground motion characteristics. To effectively predict the seismic responses without performing dynamic analysis, this study introduces a deep neural network as a regression function and a structural system is considered as a single degree of freedom (SDOF) system.

First, a deep neural network (DNN) model that can predict seismic responses of structural systems is developed using a neural network architecture is motivated by the earthquake excitation mechanism. In the DNN model, a convolutional neural network (CNN) is introduced to extract the important features of structural systems from hysteretic behaviors. To train the proposed DNN model, a seismic demand database for three different idealized hysteretic behaviors is constructed by

performing a large number of nonlinear time history analysis. Numerical investigation confirms that the proposed DNN model provides superior performance compared to the existing nonlinear static procedures which are developed based on limited dataset and parameters.

Although the DNN model is a good replacement for an onerous and complicated nonlinear time history analysis, the response prediction is deterministic, which cannot quantify the variabilities stemming from the nonlinear behavior of the structural system, i.e., varying seismic demands given the same earthquake intensity measure values. In order to quantify such uncertainties and improve the prediction accuracy, a probabilistic deep neural network (P-DNN) model is proposed based on a Bayesian deep learning method. By introducing a loss function which is proportional to the negative log-likelihood of the Gaussian distribution function, the mean and variance of the structural responses can be obtained. This assessment is important especially for earthquake engineering applications because large randomness in the input ground motion details significantly impacts the structural responses. Thorough numerical examinations are carried out to demonstrate the performance of the proposed P-DNN model.

The DNN and P-DNN models show a superior level of accuracy compared to the existing simple regression-based methods, but the models were trained based on the idealized hysteresis. Given that the hysteresis of realistic structural systems generally incorporates various hysteretic characteristics such as stiffness/strength degradations and pinching effects, the developed DNN and P-DNN models cannot guarantee the prediction accuracy for the structural systems having sophisticated hysteretic characteristics. To this end, a modified Bouc-Wen-Baber-Noori (m-BWBN) model is first proposed and a parametric investigation is carried out to

define the feasible parameter domain to describe the structural hysteresis. A new seismic demand database is constructed, and a Bouc-Wen hysteresis-based deep neural network (BW-DNN) model is proposed and trained. It is found that the BW-DNN model can predict the seismic responses of a wide class of structural systems that are not covered by the DNN and P-DNN models.

Three different scales of earthquake engineering problems, i.e., seismic response prediction of structural elements, seismic fragility estimation of structural system, and seismic loss assessment of an urban community, are presented to demonstrate the effectiveness and applicability of the developed DNN models. In addition to the applications, a novel web-service (http://ERD2.snu.ac.kr) is developed to provide source codes, constructed databases, and interactive visualization of the proposed DNN predictions. The compelling results confirm the merits and potential of the outcomes of this study, which eventually enhance the current seismic design provisions and seismic risk assessment of structural systems.

Keyword: Earthquake engineering; Deep learning; Convolutional neural network; Bayesian deep learning; Single degree of freedom system; Bouc-Wen model; Regional seismic loss assessment; Fragility analysis; Uncertainties; Nonlinear time history analysis

Student Number: 2017-37449

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor, Professor Junho Song for his patience and thoughtful guidance. His teaching style, which did not tell me what to do but continued to make suggestions on things to think about, encourages me to grow not only as a researcher but also as an independent instructor. Over the past six years of my graduate studies, Professor Song has been a perfect role model as a mentor, teacher, advisor, and researcher. It has been my extreme honor to be his student.

I would also like to thank my research co-advisor, Professor Oh-Sung Kwon for his intellectual supports and considerate advices. His insightful comments with solid conviction always encourage me to attempt challenges and expand my knowledge. During my Ph.D. study working with Professor Kwon, I experienced and learned how to collaborate with other researchers as a junior member and a colleague. He is truly a great advisor and it was the utmost pleasure to have studied under his supervision.

I was honored to have Professor Ho-Kyung Kim, Professor Juhyuk Moon, Professor Hyungchul Yoon in the dissertation committee and grateful for their invaluable advice and warm encouragement.

I have privileged to have interaction with ingenious members in Structural System Reliability Group who will be my life long companion: Se Hyeok, Eujeong, Byeong-Seong, Sang-ri, Ji-Eun, Jihwan, Chulyoung, Jungho, Minkyu, Seungmin, Changuk, Dongkyu, Seonghyun, Youngjun, Woojun, Jaehwan, Sebin, and Sumin. All the discussions and advice that we have made enforced me to move forward and strive for a wide perspective. I would like to express my special thanks to Mr. Tarun

Singhal for giving me lots of special experiences such as hackathons and the comments for developing ERD^2 website. Without him, my research leave in Toronto would not be memorable.

Last but not least, I would like to thank my family for their help, support, and encouragement. I dedicate my work to them.

Chapter 1. Introduction

1.1 Motivation

In an earthquake event, a sudden rupture of a fault may induce a strong ground motion which propagates and eventually shakes structures and infrastructure in the area. The response of structures subjected to such a seismic excitation highly depends on the hysteretic force-displacement behavior of structural systems. Figure 1.1 illustrates how seismic waves propagating from a rupture induce the vibration of a structural system.

To accurately estimate the seismic responses of a structural system, which forms the basis for seismic design and performance assessment, a comprehensive understanding of both earthquake ground motion and structural systems is needed. Figure 1.2 summarizes inelastic seismic analysis procedures in terms of the level of incorporated details of structural model and characterization of ground motion. Note that although the relative uncertainties associated with each procedure is illustrated as shade in the figure, the actual uncertainties depend on the number of parameters considered in each analysis. Due to the significant nonlinearity and inelasticity of a structural system, a time history analysis using both refined numerical models and recorded ground motions is considered as the most accurate way to estimate the structural responses. This approach solves the dynamic equilibrium equation at every time step using a numerical integration scheme, then the corresponding responses of the structural system can be numerically estimated. It is, however, noted that the nonlinear time history analysis subjected to earthquake ground motions involves exceedingly high computational efforts, which makes the approach impractical in

most routine engineering processes. In addition, due to the lack of recorded ground motions that are relevant to the structure's specific location, and the uncertainty in the hysteretic behavior of a structure, adopting a highly precise yet computationally inefficient method cannot be fully justified.

To estimate seismic responses of nonlinear structural systems without performing nonlinear time history analysis, many research efforts have been made using single degree of freedom (SDOF) systems (ATC 40, 1997; FEMA 440, 2005; Nassar and Krawinkler, 1991). Note that although a three-dimensional structural model with detailed structural properties is desirable for a dynamic analysis, an idealized structural model, such as an equivalent SDOF system, is often employed, especially when the structure is in design phase (Ibarra et al., 2005; Ruiz-García and Miranda, 2003; Tothong and Cornell, 2006; Vamvatsikos and Cornell, 2006). Because the peak transient responses during the excitation are the most important features when designing and assessing structural system, most of such methods usually aim at predicting the peak transient value of structural responses (but for simplicity, the term *peak response* is used in this study). Note that the peak responses refer to maximum absolute values, i.e., without considering the sign or direction during the excitation. The existing prediction equations are developed based on regression analysis with a rather small number of earthquake events, which inevitably leads to a large variation in the prediction. Moreover, those are applicable to a limited number of structural systems with idealized hysteretic behaviors in order to employ the corresponding regression coefficients.

Meanwhile, some researchers have investigated various machine learning algorithms as an alternative to time-consuming analysis procedures used in the field of earthquake engineering (Adeli, 2001; Adeli and Panakkat, 2009; De Lautour and

Omenzetter, 2009; Giovanis et al., 2015; Lagaros and Fragiadakis, 2007; Mitropoulou and Papadrakakis, 2011; Möller et al., 2009; Nakamura et al., 1998; Xie et al., 2020 among many others). Since most of such machine learning models are trained for a specific type of structural system and hysteretic behavior, the applicability of the trained models to other structural systems is limited. Furthermore, relatively small number of nonlinear functions in the machine learning-based prediction models cannot fully identify the intricate relationships between structural responses and the corresponding information of structural system and earthquake events.

This study attempts to address the limitation of previous studies by improving the accuracy of response prediction and extending the applicability of the method to various types of structural system and ground motions. To this end, *deep* neural network (DNN), which is one of the machine learning algorithms, is employed as a tool to predict the structural responses. Since the DNN achieved unprecedented performance in terms of predicting the intricate relationship between input and output variables, it is possible to cover comprehensive information including both earthquake events and structural characteristics. Once the DNN model is developed, it is possible to accurately predict the seismic responses for various structural systems without performing time history analysis in various earthquake engineering problems.

1.2 Background and related research efforts

As previously discussed in Section 1.1, various methods have been developed as a replacement of onerous and complicated nonlinear time history analysis. This section

first introduces well-known nonlinear static procedures which are adopted in modern seismic design codes: (1) capacity spectrum method, (2) coefficient method, and (3) R-μ-T relationship. Then, the basic concepts of deep learning are introduced, which is followed by several instances of applying deep learning to earthquake engineering problems, especially for prediction of structural responses and assessment of seismic performance.

1.2.1 Current practices of seismic response prediction

1.2.1.1 Capacity spectrum method

Capacity spectrum method was developed based on the assumption that the peak or absolute maximum inelastic deformation of a nonlinear SDOF system can be approximately estimated from the peak deformation of a linear elastic SDOF system having a period and a damping ratio that are larger than the initial values of those for the nonlinear system (ATC 40, 1997; FEMA 440, 2005). The process begins with the generation of a force-deformation relationship from nonlinear pushover analysis in acceleration-displacement response spectrum (ADRS) format. By superimposing the pushover curve on an ADRS demand spectra, i.e., response spectrum of ground motion, the estimated force-deformation relationship becomes a capacity spectrum. Using simple regression equations, effective period and viscous damping for demand curves are recursively calculated, and then the intersection point between the radial effective period and the response spectrum for the effective viscous damping are estimated. The iterative procedure is carried out until the difference between the intersection points for the consecutive iterations lies within acceptable tolerance. By determining the point, it is possible to estimate the peak displacement or damage of the structural system subjected to a given earthquake. The schematic diagram of the

capacity spectrum method is depicted in Figure 1.3.

1.2.1.2 Coefficient method

Coefficient method (ASCE 41-13, 2013) calculates the target peak displacement u_{peak} of nonlinear structural system by modifying the peak displacement of a linear oscillator using several empirical coefficients as follows:

$$u_{peak} = C_0 \cdot C_1 \cdot C_2 \cdot \frac{S_a(T)}{\omega^2} \cdot \text{g} \qquad (1.1)$$

where C_0, C_1, and C_2 are the modification factors that represent different aspects of the building when characterizing as the simplified SDOF system, ω denotes the circular natural frequency of the structural system, $S_a(T_1)$ is the spectral acceleration at the first mode period, and g stands for the acceleration of gravity (= 9.81 m/s^2). The modification factors are estimated based on simple regression functions with predefined values which vary along with types structures and earthquake of interest.

1.2.1.3 R-μ-T relationship

Veletsos and Newmark (1960) found out that the peak displacements of elastic and inelastic structural systems having medium to large periods are identical to each other. Based on these empirical findings, the seismic responses of inelastic system can be readily estimated from the equivalent elastic system having the same period T. In other words, if the stiffness is independent of strength, the displacement ductility μ can be estimated from strength reduction factor R. Such an approximate prediction is referred to as R-μ-T relationship and the corresponding graphical representation

is shown in Figure 1.4. Numerous studies of R-μ-T relationship have been proposed by improving this approximation and extending to various types of structural systems (Chopra and Chintanapakdee, 2004; Miranda and Bertero, 1994; Nassar and Krawinkler, 1991).

1.2.2 Literature reviews: Deep learning in earthquake engineering

1.2.2.1 Introduction to deep learning

Machine learning is a field of study that provides systems or computers with the ability to automatically learn and improve from experiences without being explicitly programmed (Samuel, 1959). To properly figure out patterns of phenomenon, a large number of datasets is inevitably required. With the advent of high level of computing performance, machine learning has played a pivot role in many areas such as finance, medical, and engineering.

Deep learning is a part of machine learning algorithms based on artificial neural network (ANN) which is designed to mimic a human brain's structure and operation. The classic ANN typically consists of three layers: input, hidden and output layers whose configuration is depicted in Figure 1.5. Each node or unit represents a matrix operation including a nonlinear function whose parameters are adjusted through training by available input and output datasets. Deep neural network (DNN) is a variant of ANN which stacks multiple hidden layers widely and deeply, which yields a highly generalized regression equation (LeCun et al., 2015). Owing to the multiprocessing layers and nonlinear transformation, it is possible to identify critical features of input data and find hidden patterns of highly complex problems, thus to obtain superior predictions compared with other machine learning methods.

The objective function of a neural network for regression analysis, often termed

"loss function" is defined as follows for given input data x_i and its corresponding output y_i:

$$L = \frac{1}{N} \sum_{i=1}^{N} \|y_i - f(x_i)\|^2 \qquad (1.2)$$

where N is the number of input-output data points, f denotes the output estimated by the neural network model, and $\|\cdot\|$ represents the metric that computes the distance between the real and the estimated values of the output, e.g., Euclidian distance. Note that different loss functions are assigned to classification problems. The optimal DNN parameters for a given dataset is obtained by minimizing the loss function using an optimization algorithm with back-propagation scheme.

1.2.2.2 Deep learning methods for seismic response prediction

Owing to huge benefits of machine learning algorithms their application to earthquake engineering keeps increasing. Although there exist numerous applications for each subfield of earthquake engineering, this section focuses on seismic response prediction or performance assessment of structural systems, especially using ANN or DNN. Throughout the literature reviews, benefits from the deep learning method to each research effort are highlighted.

As discussed earlier, a large number of datasets are required to construct a DNN model to predict responses of structural system subjected to ground motions. Thus, datasets from numerical investigations are often employed instead of those from laboratory tests. In general, input of the model is the structural parameters and ground motions, while output is the structural responses. Starting from Conte et al. (1994) who used ANN as a surrogate model to predict linear elastic behavior of

multi-story buildings, ANN has been widely implemented to various buildings and infrastructures such as frame structure (Joghataie and Farrokh, 2008; Xu et al., 2004), concrete dam (Karimi et al., 2010), concrete bridge (Jeng and Mo, 2004), and steel moment frames (Akbas et al., 2011). Moreover, recently, using long short-term memory which is one of the DNN methods, Zhang et al. (2019) predicted the time series of seismic responses of a nonlinear hysteretic system. Once ANN or DNN models are developed, further dynamic analysis is not required to predict the responses of future earthquakes for the target structural system.

Structural fragility defined as the conditional probability of failure at a given value of seismic intensity is another important feature to evaluate the seismic performance of structural systems, which provides information to evaluate casualties, total economic losses, and repair costs. However, in order to estimate the structural fragilities considering various uncertainties lying on both seismic demand and capacity, a huge number of dynamic analysis is required. To reduce the computational costs, many research efforts employed ANN and DNN as a surrogate of the dynamic analysis (Lagaros et al., 2009; Liu and Zhang, 2018; Mitropoulou and Papadrakakis, 2011; Wang et al., 2018). Thereby, it is possible to estimate structural fragility for given earthquake hazard scenarios with less computational costs.

1.3 Objectives and scopes

The research described in this dissertation has two main objectives: (1) developing DNN models that predict seismic responses of structural systems having different hysteretic characteristics, and (2) demonstrating applications of the developed DNN

models to various earthquake engineering problems including prediction of seismic losses of an urban community.

The dissertation first focuses on developing a DNN model that can predict seismic responses of structural systems. The DNN model is developed using the seismic responses of three different idealized hysteresis. Next, to incorporate the latent uncertainties in the seismic features, a probabilistic deep neural network (P-DNN) model is developed by introducing a new loss function. Finally, an extension is made to predict the seismic responses of generalized hysteresis having stiffness and strength degradations, and pinching effect using a deep neural network. The extended model is referred to as BW-DNN model, because Bouc-Wen hysteresis model is employed to construct the datasets for training the deep neural network model. Compared with the existing nonlinear static procedures, the proposed models have the following benefits: (a) accuracy of response prediction is significantly improved with less computational efforts; (b) the models can be applied to various structural systems subjected to different kinds of ground motions; (c) once the models are developed and trained, no further training is required to predict seismic responses of structural systems that are not used in training; (d) it is possible to predict various structural responses including peak displacement, velocity and acceleration; (e) the P-DNN model provides the mean and variance of structural responses so that large randomness in the input ground motion details and their significant impact on the structural responses are quantitatively examined; (f) since important features of structural systems are automatically extracted from the DNN models, there is no need to parameterize the hysteresis of a structural system using a mathematical model such as Bouc-Wen model; (g) to facilitate the use of the proposed method in practice and research field, a novel web-service providing an

interactive visualization of the proposed method is developed and serviced at http://ERD2.snu.ac.kr. The advantage of the developed methods will be thoroughly investigated and presented throughout the dissertation.

Next, comprehensive applications of the developed models to three different scales of structural systems are presented. Since the seismic responses prediction of a structural system is the first step in many topics in earthquake engineering problems, the developed method can be implemented to various areas. Starting from the response prediction of structural elements, the method can be employed to evaluate the structural fragility and assess the seismic loss of an urban community in a probabilistic manner.

The intellectual merits provided by this research will lead to meaningful discussions and initiation of future research efforts in both neural network communities and earthquake engineering fields. Moreover, although the dissertation is focused on prediction of structural system subjected to earthquake ground motions, it is possible to extend the proposed concept to a wide array of engineering problems such as various hysteretic systems subject to stochastic excitations, which enriches the convergence between different academic fields.

1.4 Organization

The dissertation is organized into 6 chapters. The first objective described in the previous section is addressed in Chapters 2, 3, and 4, in which three different DNN models are developed with the corresponding databases and training methodologies. On the other hand, the second objective is handled in Chapter 5, which provides three different scales of earthquake engineering problems. More details on the specific

subjects covered in each chapter are presented below.

Chapter 2 describes the development of DNN model that can predict seismic responses of three idealized hysteresis which is widely used in civil engineering. To this end, a seismic demand database is developed using 54,090 different structural systems subjected to 1,499 ground motions. Inspired by the natural phenomenon of seismic excitations of structural systems mechanism, a new DNN architecture is proposed and trained using the database. Prediction accuracy of the DNN model that predicts the peak displacement is compared with the nonlinear static procedures presented in Section 1.2.1. This is followed by the development of DNN models that can predict other structural responses such as peak acceleration and velocity.

Chapter 3 begins with a discussion on latent uncertainties in the input of the DNN model, especially for ground motion information. Since the stochastic excitation is characterized by several seismic intensities, loss of information inevitably occurs. Due to this characterization, the DNN model always produces the same responses for the same DNN input, even though ground motions and the corresponding structural responses are different. To address this limitation, the probabilistic deep neural network (P-DNN) model is proposed by employing a new loss function which is proportional to the negative log-likelihood of the Gaussian distribution assumption. The predicted mean and variance from the P-DNN model are comprehensively investigated with the developed DNN model in Chapter 2 and Bayesian linear regression method, respectively.

Chapter 4 extends the DNN model in order to incorporate more complicated and sophisticated hysteretic behaviors. Since the DNN models in Chapters 2 and 3 are trained based on idealized hysteresis, there is a limitation to the prediction of seismic responses of structural systems having stiffness and strength degradations,

or pinching effects which are commonly shown in most of structural materials, especially in reinforced concrete (RC) structures. To this end, Bouc-Wen-Baber-Noori (BWBN) model which is versatile in describing various hysteresis is employed. However, since the current version of the BWBN model has a limitation to express the yield strength of structural systems, modified Bouc-Wen-Baber-Noori (m-BWBN) model is suggested and employed in this research. Moreover, the lower and upper bounds of the parameters of the m-BWBN model are determined from sensitivity analysis to cover every practical range of civil structural systems. Using the developed database and the new DNN architecture, the BW-DNN model is trained and its performance is verified.

Chapter 5 presents three different scales of earthquake engineering problems as an application of the developed DNN models. First, seismic responses of RC columns are predicted by using the BW-DNN and DNN models. Second, a new deep learning-based fragility estimation method is proposed using the predicted mean and variance from the P-DNN model. Third, probabilistic seismic loss assessment is carried out using the P-DNN model and the results are compared with the ones obtained by the DNN model and the coefficient method.

Finally, Chapter 6 provides a summary of the main results and conclusions of the study. Recommendations for future studies are also presented.
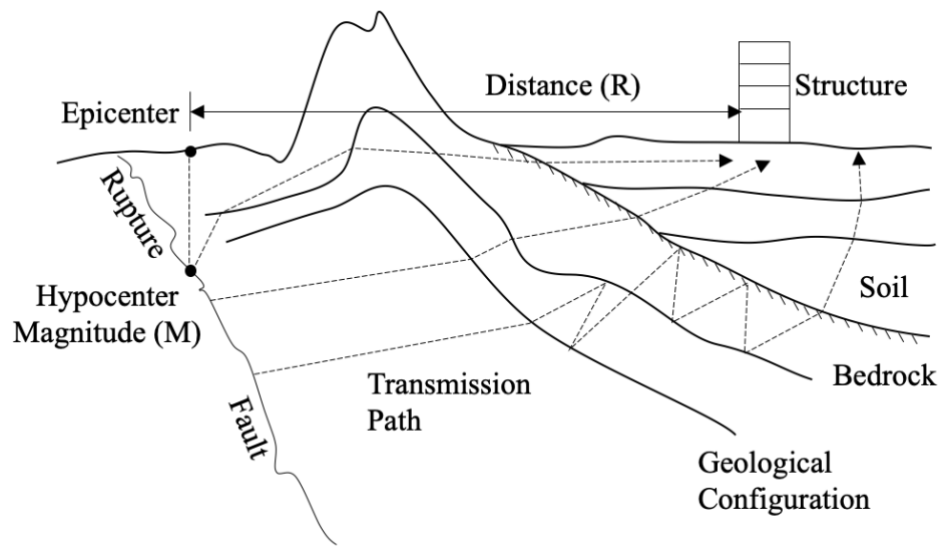
Figure 1.1 Illustration of structural vibration induced by earthquake event
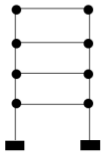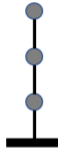
	S_a Response Spectra period	$\ddot{u}_g$ Ground Motions time
Uncertainty low high		
Detailed structural model		Nonlinear time history analysis
Equivalent MDOF	Multi-mode pushover analysis (MPA)	Simplified MDOF time history analysis
Equivalent SDOF	Nonlinear static procedure (NSP)	Simplified SDOF time history analysis

Figure 1.2 Inelastic seismic analysis procedure for various structural models and ground motion characterization along with the level of uncertainties
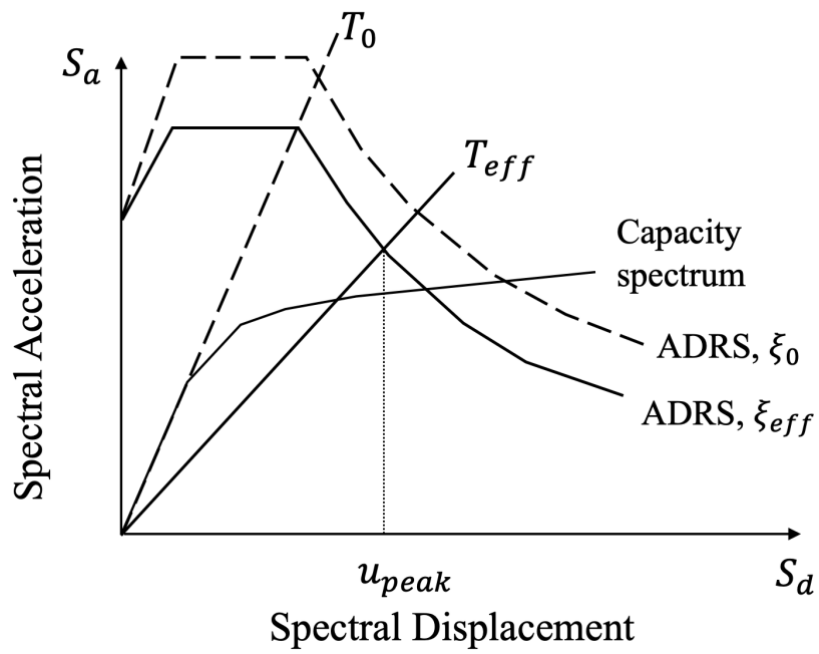
Figure 1.3 Schematic diagram of the capacity spectrum method
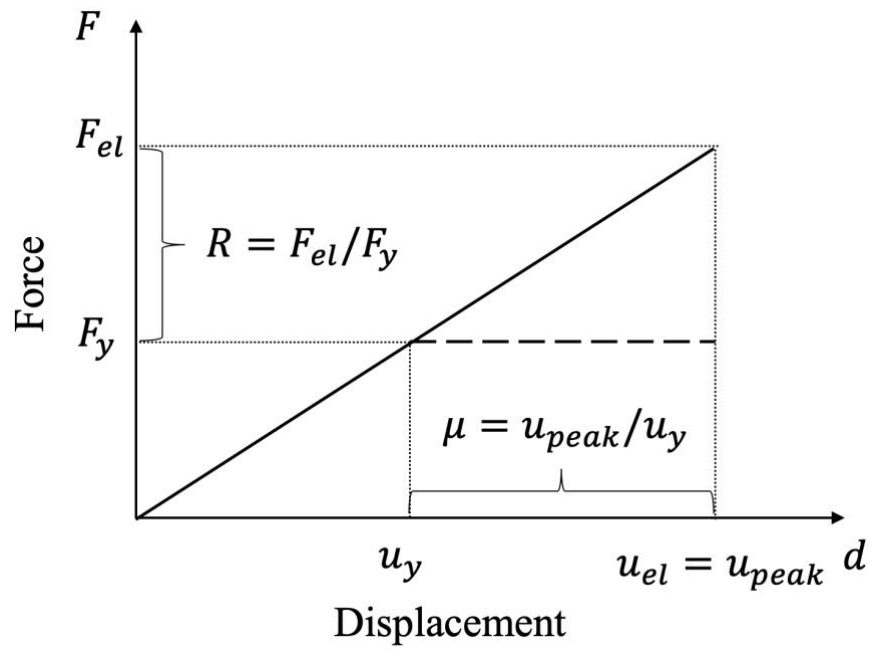
Figure 1.4 Schematic diagram of R-μ-T relationship
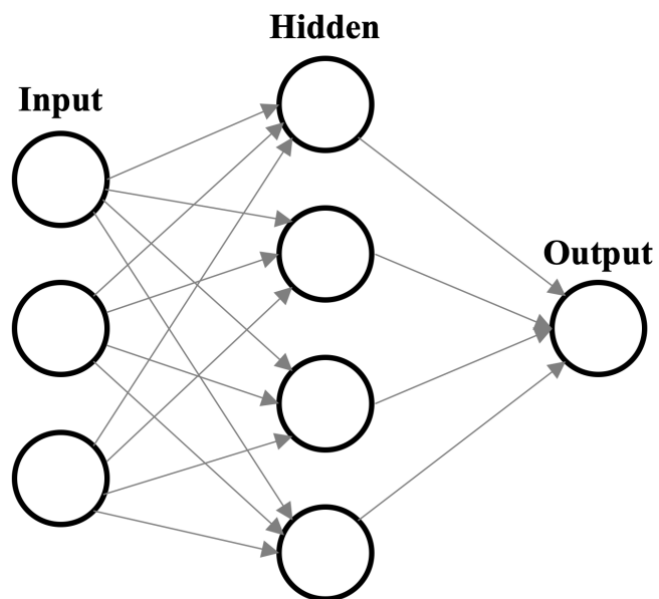
Figure 1.5 Illustration of ANN with 4 units/nodes in hidden layer

Chapter 2. Seismic Response Prediction of Idealized Hysteresis Using Deep Learning

2.1 Introduction

Deep neural network (DNN) has shown unprecedented performance in terms of predicting an intricate relationship between the input and output variables in diverse fields including business, medical, and engineering (LeCun et al., 2015; Schmidhuber, 2015). In particular, development of deep convolutional neural network (CNN) has achieved practical successes, especially in face recognition (Lawrence et al., 1997), image classification (Krizhevsky et al., 2012), and speech recognition (Sainath et al., 2015). Since the CNN shows a clear advantage in dealing with the data having a strong spatial correlation by using multiple overlapping filters, the CNN is employed to extract features of a nonlinear hysteretic system from its hysteresis loops. The features extracted from the hysteretic information are, then, merged with those representing stochastic excitation information to form a deep ANN that can predict the seismic responses of a structural system. Thereby, the trained DNN model can cover comprehensive information including both the stochastic excitations and characteristics of nonlinear hysteretic systems.

This chapter first presents the overview of seismic response assessment of structural systems and summarizes the structural and earthquake information which are employed as the inputs for the DNN model proposed in this study. Next, an architecture of the DNN model that can predict seismic responses of a single degree of freedom (SDOF) system is proposed by merging the extracted structural and earthquake information. After training the proposed DNN model using seismic

responses obtained from a huge amount of time history analyses, its efficiency and applicability are demonstrated through comparison with existing methods (Kim et al., 2019).

2.2 Seismic responses of structural systems and selected features

This section describes the overall procedure of nonlinear time history analysis and provides reasoning on the critical features that are selected as inputs for the DNN model. The structural system is represented as a hysteretic loop, i.e., a relationship between imposed displacement and normalized restoring force, whereas the earthquake ground motion is represented by a set of widely used features.

2.2.1 Characterizing single degree of freedom structural system

An equivalent SDOF system is often used as a replacement of a detailed numerical model, especially when the structure is being analyzed in the design phase (Ibarra et al., 2005; Ruiz-García and Miranda, 2003; Tothong and Cornell, 2006; Vamvatsikos and Cornell, 2006). The governing differential equation of an SDOF system subjected to earthquake ground motion is given by

$$m(\ddot{u} + \ddot{u}_g) + c\dot{u} + f_s = 0 \qquad (2.1)$$

where m is the mass, c is the damping coefficient, f_s represents the restoring force function which may depend on the history of structural responses, $\ddot{u}_g$ denotes the ground acceleration, and u, $\dot{u}$ and $\ddot{u}$ respectively stand for the displacement, velocity and acceleration of the mass relative to the ground. Eq. (2.1) can be rearranged after dividing both sides by the mass m:

$$\ddot{u} + 2\xi\omega\dot{u} + F_s = -\ddot{u}_g \tag{2.2}$$

where ξ represents the damping ratio (typically 5% is used in practice), ω is the circular natural frequency of the structural system, and F_s denotes the normalized restoring force, i.e., f_s/m. Given that ω depends on the relationship between u and F_s, F_s is the only term in the equation that affects the responses of a structure subjected to the specific ground acceleration history $\ddot{u}_g$. Figure 2.1 provides three different hysteretic models which are widely used in the field of civil engineering: linear elastic, bilinear kinematic hardening, and bilinear stiffness degrading systems. It should be noted that, in fact, the hysteretic behavior of a real structural system can be much more complicated, which will be later handled in Chapter 4. The three simplified hysteretic models in Figure 2.1 are used in this chapter to train the DNN model and demonstrate the effectiveness of the proposed method. The three idealized hysteretic models in Figures 2.1(a), (b) and (c) are, respectively, denoted as HM1, HM2, and HM3 in this study.

In order to identify the hysteretic behavior of a structural system, a quasi-static cyclic analysis (i.e., push and pull the nonlinear hysteretic system) of an SDOF system is performed by numerical simulations using predefined displacement history. The process to estimate a hysteretic behavior for the input of the DNN model is illustrated in Figure 2.2. Most of the existing methods directly employ the parameters of simplified mathematical models of a nonlinear hysteretic system such as stiffness and yield strength to an input of ANN or DNN models. However, in the proposed method, hysteretic behaviors (i.e., the displacement and force vectors in Figure 2.2) are employed as the input parameter to a CNN layer which is often used to extract complicated natural data to lower-level features. Thereby, the CNN would

automatically extract useful information for estimating the responses of the nonlinear hysteretic system by training the neural network model. Note that the displacement history is predefined based on the normalized yield force and stiffness of the structures, and the corresponding restoring force history is estimated from the quasi-static cyclic analysis.

2.2.2 Characterizing earthquake ground motions

Ground motion records include many features of an earthquake event including soil condition of the site, fault mechanism, earthquake magnitude, epicentral distance, and attenuation relationship. These features influence the duration of the motion, frequency content, and the number of cycles, which eventually affect the dynamic responses of nonlinear structural systems. Several investigators have characterized such features of earthquake ground motion with a simple intensity measure (Housner and Jennings, 1982; Kim et al., Accepted; Marafi et al., 2016; Riddell, 2007 among many others). However, considering the described characteristics of stochastic excitation, it is desirable to use multiple features together to properly describe the earthquake ground motion. Thus, three different types of ground motion information are selected as the input of the DNN model, which have been demonstrated to correlate well with the responses of a nonlinear structural system: (1) earthquake characteristics, (2) peak values of ground motion time histories, and (3) response spectrum. The various intensity measures are employed instead of ground motion acceleration time history because a seismic demand of a site of interest is defined as a design spectrum according to modern seismic design codes (Eurocode 8, 2003; KMOLIT, 2018). Moreover, it is found that the DNN model can be more efficiently trained when the various intensity measures that show high correlation with the target

structural responses are employed as the input than the ground acceleration is used.

First, the features that are selected to represent earthquake characteristics are magnitude (M), epicentral distance from the site of the structure (R), and the soil class of the site. The soil class is important because it can affect the stability of structure by amplifying the ground motions based on its consolidation level (BSSC, 2003). For example, BSSC (2003) uses five different soil classes. Although the fault type information should be incorporated as another feature of an earthquake event, such information is not included in the strong motion database (Power et al., 2006) which is adopted in this study to carry out dynamic analysis. Thereby, it is decided not to use it as an input parameter. Second, the features that represent peak values of ground motion time history are the absolute maximum value of each recorded acceleration (peak ground acceleration; PGA), velocity (peak ground velocity; PGV), and displacement (peak ground displacement; PGD). Lastly, the features that represent frequency contents of ground motion acceleration are 5%-damped spectral accelerations in the period range of 0.005 sec to 10 sec (total 110 steps; $S_a(T)$) are used, i.e., response spectrum. Note that spectral acceleration $S_a(T)$ represents the peak acceleration of a linear elastic SDOF system with a specified natural period and given damping ratio subjected to the ground motion, usually given in a unit of gravity acceleration, g. Most of the previous studies described in Chapter 1 employ only a few spectral acceleration values of certain periods including the first mode of a target structural system as an input parameter of ANN models. By contrast, in this study, spectral acceleration of all possible periods of civil/architectural structural systems, i.e., response spectrum, is used as an input, so that the estimation of the target structure's natural frequency is not required. In other words, the neural network automatically identifies and employs the dominant spectral acceleration values to

predict seismic responses. Moreover, the effect of spectral shape is taken into account, which is found as an important factor affecting dynamic responses (Baker and Cornell, 2006; Haselton et al., 2011). The effect becomes significant, especially when the effective structural period changes due to its nonlinear behavior. In summary, three categorized features of earthquake ground motions are proposed by using a total of seven parameters (magnitude, distance, soil class, PGA, PGV, PGD, and response spectrum).

2.3 Architecture details

A new framework of the DNN model is proposed to incorporate information about both structural systems and earthquake ground motions. Motivated by the earthquake excitation mechanism of structural systems depicted in Figure 1.1, a new deep neural network architecture is constructed as shown in Figure 2.3. Each feature group of earthquake ground motion is first processed by the ANN (green boxes in Figure 2.3), then merged with the hysteretic features extracted from the CNN (orange box in Figure 2.3). This is to reflect the fact that each earthquake ground motion information produces distinct responses for a given hysteretic behavior. The merged layer is again processed by the ANN to provide 32 units (blue boxes in Figure 2.3) for each. The 32 units from each group of earthquake ground motion information are finally merged with the extracted hysteretic units again to form an ANN that can predict seismic responses.

Another architecture of the DNN model was also tried such that the three different kinds of earthquake ground motion information are processed by individual ANNs, and then directly merged with the hysteretic information extracted from the

CNN. In other words, the part of blue-gray boxes in Figure 2.3 was removed. It was, however, found that the testing results of the train and test datasets show large discrepancy and the prediction accuracy is worse than that from the proposed network structure in Figure 2.3.

A more detailed diagram of the proposed DNN architecture is provided in Figure 2.4. All convolutional and neural network layers are followed by a rectified linear operator (ReLU; Nair and Hinton, 2010) except for the first part of the convolutional layers for which tangent hyperbolic operator (Tanh) is used as the activation function. In addition, the batch normalization (Ioffe and Szegedy, 2015) is adopted for all layers after the ReLU operator is applied. Since the initial stiffness of the hysteresis shows large variability (the ratio of the largest to the smallest stiffness that covered in this study is about 40,000), the Tanh activation function is used for normalizing parameters describing the hysteretic behaviors. It was observed that, if the ReLU is adopted rather than the Tanh activation function at the first convolutional part, the hysteretic features extracted from the CNN was overfitted, especially for those having large initial stiffness.

Four types of filter sizes (2, 4, 8, and 16) are employed to capture the features having different scales. This is helpful, particularly when extracting features from the hysteretic behaviors because the input hysteretic behavior can have different scales (e.g., number of cycles). Moreover, 2×1 max pooling is applied to reduce the spatial size of the representation, i.e., training parameters, in the convolutional phase. The first fully connected layer that concatenates the output of the four convolutional layers containing 64 units is followed by the second fully connected layer with 48-dimensional outputs. This architecture is motivated by autoencoder (Li et al., 2013; Socher et al., 2013) which forces the layer to engage features having

different scales by intentionally reducing the units. Finally, the output of the network is a single unit that is fed to a linear activation function to estimate a continuous variable, i.e., regression problem. To the best of the author's knowledge, this study provides the first attempt to predict a variety of seismic responses using *deep* neural networks.

2.4 Training methodology

To train the proposed DNN model, a seismic demand database is first developed. To reduce the physical size of the database, a novel database architecture, i.e., schema, is proposed. Using the database, the DNN model is trained to predict the peak displacement of structural systems subject to ground motions. It was found that the performance of the DNN model varies along with the training methodologies even with the same network architecture and dataset. Thereby, optimal training methodology is also proposed based on a large number of numerical investigations. Since any type of structural response can be used as an output for the model, the trained results using other structural responses, such as the peak acceleration and velocity, will be also presented in the next section. Note that the work reported in this chapter was performed using TensorFlow (Abadi et al., 2016) and the Compute Canada's CPU and GPU clusters.

2.4.1 Database

A seismic demand database is constructed by performing a large number of nonlinear time history analyses using OpenSees (Mazzoni et al., 2006). The database is constructed using SQLite 3 whose schema is shown in Figure 2.5. As shown in the

figure, the database consists of 9 different tables each of which contains different information of structural systems (orange boxes), earthquake ground motions (blue boxes) and the results of dynamic analysis (yellow box). Note that arrows in the figure represent how the different tables are interrelated with each other using a foreign key. Since 85% of the dynamic analysis cases behave in linear range during earthquake excitation, the proposed architecture can dramatically reduce the size of the database. In other words, the seismic responses of structural systems having the same period are identical to each other when their normalized yield forces are greater than the corresponding spectral acceleration value. Thus, it is possible to replace a set of structural responses by a single index which indicates the location of the corresponding responses in the seismic demand table (yellow box).

The database contains responses of three different hysteretic models (HM1, HM2, and HM3 in Figure 2.1) subjected to 1,499 ground motions obtained from the NGA-West database (Power et al., 2006). To cover every practical range of structural characteristics, 90 steps of the structural period from 0.05 sec to 10 sec, 30 steps of the normalized yield strength from 0.05g to 1.5g, and 10 steps of the post-yield stiffness ratio from 0 to 0.5 are selected. The upper and lower limits of these values are determined based on the capacity curves in HAZUS-MH 2.1 (FEMA 2012) which can describe force-deformation relationships of a wide range of building structures and the intermediate values are uniformly distributed between the upper and lower limits. The total number of hysteretic behaviors is 54,090 (90 for HM1 and $90 \times 30 \times 10 = 27,000$ for each bilinear system). The seismic demand database is available at http://ERD2.snu.ac.kr.

Data preprocessing and augmentation technique enable a deep neural network to achieve improved accuracy with a relatively small amount of training time and to

produce a reasonable predicting accuracy for data points that are not used in training. Since it is widely reported that some of the input and output datasets follow a lognormal distribution (Ellingwood, 2001), the natural logarithm is applied to several input parameters, i.e., response spectrum, PGA, PGV, and PGD, as well as the output to resolve the skewness. Moreover, in order to allow the network to better generalize the structural hysteretic behaviors, the data augmentation technique (Krizhevsky et al., 2012) is applied, especially for the information of the structural system, by shifting and flipping the hysteresis during training.

2.4.2 Initialization

It is widely known that the performance of a deep neural network is highly affected by the initialization of weights (Glorot and Bengio, 2010; He et al., 2015; LeCun et al., 1998). In this study, it was found that using randomly initialized weights, the convolutional part of the network cannot extract the features properly, which, in turn, made the performance of the network much worse than those of existing nonlinear static procedures. To address such an issue, pre-training is carried out with a small number of samples of the hysteretic behaviors. Only HM1 and HM2 with zero post-yield stiffness ratio (i.e., elasto-perfectly plastic) are selected with randomly selected 894 ground motions for pre-training (the number of data set is $2{,}790 \times 894 = 2{,}494{,}260$). Adam optimizer (Kingma and Ba, 2014; Reddi et al., 2018) is used as the optimization algorithm to reduce the mean squared error (MSE) between the predicted structural responses and the ground truth values with 512 batch sizes and 54 epochs of training.

2.4.3 Training

It was found that the features extracted from the CNN were overfitted when the structural model was randomly selected along with earthquake ground motion. To address this issue, the same set of structural models is applied to each ground motion when training the model. In this regard, to check the over-fitting of the DNN model by monitoring the loss on the test set, ground motions are split into a train set of size 1,199 (80%) and a test set of size 300 (20%). However, note that even the entire set of ground motions used in this study is insufficient to generalize the ground motion information due to the nature of a random excitation based on the author's experience.

Although a total of 54,090 hysteretic behaviors need to be trained for the DNN model, a subset of hysteretic behaviors is selected to increase the efficiency of the training and overcome the memory constraints during training. This is based on the premise that DNN has an ability to estimate the intermediate variables of distinct input values through interpolation. Therefore, 15 steps of yield strength and 7 steps of post-yield stiffness ratio are considered rather than 30 and 10 steps of the original dataset, respectively (18,990 structural models are selected). When sparsely selecting the structural models, relatively smaller yield strengths are selected more often than bigger ones to incorporate enough number of cases in which the structural system behaves nonlinearly during earthquake excitation. That is, hysteretic behaviors used for training are prone to behave nonlinearly during earthquake excitation compared with those not used for training. This is because it was found that the nonlinear cases are more difficult to estimate (i.e., bigger MSE) than the linear cases. Moreover, 280 ground motions are selected among 1,199 ground motions for each epoch of training to overcome the limitation of the computational resources, i.e., 5,317,200 datasets for each epoch. Similar to the initialization, Adam

optimizer is adopted as the optimization algorithm to reduce the MSE of the output

with 512 batch sizes and 300 epochs of training.

2.5 Performance of deep neural network

This section focuses on investigating the performance of the trained DNN model by

comparing its accuracy with those by the three existing nonlinear static procedures.

Moreover, to give an insight about the proposed network, the features extracted from

the CNN are analyzed thoroughly by introducing the other hysteretic models which

are not used in training. It is numerically demonstrated that the proposed method has

superior performance in terms of its applicability and effectiveness.

2.5.1 Training results

To check whether the trained DNN model is overfitted or not, MSE ($1/N \cdot$

$\sum_{i=1}^{N}(y_i - f(x_i))^2$) and mean absolute error (MAE, $1/N \cdot \sum_{i=1}^{N}|y_i - f(x_i)|$) are

calculated for both train and test datasets. Note that the MSE and MAE are calculated

after the natural logarithm is applied to the structural responses. The results are

shown in Table 2.1.

The MSE and MAE of training results show that using the DNN model it is

possible to predict the structural responses that are fairly close to those by nonlinear

time history analysis. Although the MSE and MAE of the test dataset may be shown

as significantly higher than those of the train datasets, the discrepancy between the

train and test sets is not critical, in that the ratios of the predicted value to the correct

value in the original scale (i.e., not applying the natural logarithm) are around

$e^{0.05} \approx 1.05$ and $e^{0.12} \approx 1.13$, respectively. This also indicates that 1,119 ground

motions may not be enough for properly training the proposed network in terms of random earthquake ground motion information. The author believes that if the more ground motions are used for training, the better results can be obtained for the test dataset.

Since the hysteretic behavior is sparsely selected during the training process, the MSE is computed for both trained and non-trained hysteretic models in order to demonstrate the trained network's ability to interpolate over hysteretic behaviors. 200 ground motions are randomly selected from the entire set of ground motions to obtain the results shown in Table 2.2. In training the DNN model, the hysteretic models, which are prone to exhibit responses in nonlinear range, are more selected against the linear cases. Given that the DNN model predicts linear cases with less MSE than nonlinear cases, the MSE of nontrained cases in Table 2.2 is less than that of trained cases even though the hysteretic behaviors are not used for training. The results further confirm that the DNN model has an ability to interpolate regarding hysteretic behaviors. Therefore, it is expected that the trained DNN model will be able to predict the structural responses even if specific hysteretic behaviors were not used in training

2.5.2 Comparison with existing methods

To demonstrate the superior accuracy of the proposed method, the prediction results are compared with those by the three methods described in Chapter 1: (1) R-μ-T relationship developed by Nassar and Krawinkler (1991), (2) capacity spectrum method (FEMA 440, 2005), and (3) the coefficient method (ASCE 41-13, 2013). The elasto-perfectly plastic systems with randomly selected 200 ground motions are provided as inputs for the methods. Among 540,000 cases (= 2,700 × 200) of the

entire dataset, the cases showing nonlinearity during excitation (8.3% of the dataset, i.e., 44,842 cases) are investigated. The differences between estimation by nonlinear dynamic analyses and prediction by each method are shown in Figure 2.6 and the MSE are presented in Table 2.3. The results demonstrate that even though the structures handled in this example are a simple hysteresis, the existing methods cannot predict the peak displacement properly compared to the DNN model. In other words, the proposed method can predict the nonlinear responses with a smaller size of the uncertain error compared to the existing methods.

2.5.3 Validation of features extracted from CNN

Despite the great success of deep learning, there exist concerns and criticisms that the method provides black-box models (Alain and Bengio, 2016; Shwartz-Ziv and Tishby, 2017). In order to understand the input parameters passed through the CNN in the DNN model, features extracted from the CNN are plotted and investigated based on the knowledge of the structural dynamics. First of all, the output of the final FC which was obtained right after the application of the ReLU activation function, i.e., 64 vectors (orange box in Figure 2.3), is plotted in Figure 2.7 (i.e., 64 lines). Herein, hysteretic models of HM1 (90 hysteretic behaviors) are imposed as the input of the DNN model. The extracted values of the hysteretic behaviors are plotted such that natural frequency (period) is big (short) to small (long) as left to right. As shown in Figure 2.7, each element of the 64 units is activated along with the different natural frequency of a nonlinear hysteretic system (bell shape plot at specific structural systems), which resembles the modal analysis finding the various periods at which the system naturally resonates.

To give further insight of the DNN model, the features extracted from the CNN

are compared over the hysteretic types. The experiment is based on the assumption that if different hysteretic models share some characteristics such as the same initial stiffness (ω^2 in Figure 2.1 for HM1, HM2, and HM3), it is expected that some of the features extracted from the CNN of different hysteretic models may also be equivalent to each other and show meaningful trends.

Figure 2.8 shows the features extracted from the CNN when all hysteretic models of HM1 (90 hysteretic behaviors) and zero post-yield stiffness (i.e., $\alpha = 0$) cases of HM2 (2,700 hysteretic behaviors) and HM3 (2,700 hysteretic behaviors) are used. The figure is plotted as follows: When one of the hysteretic behaviors of HM1 is applied as an input of the CNN, 64×1 vectors can be obtained as an extracted feature vector. Then, each element of the vector is plotted in separate boxes. After the features extracted of every hysteretic behavior of HM1 are mapped out in Figure 2.8(a), the 64 separate boxes are obtained whose number of indices along the x-axis is 90. Using the same procedure, Figure 2.8(b) and Figure 2.8(c) represent the output vector when HM2 and HM3 are imposed as an input of the CNN, respectively. In addition, the features extracted of the hysteretic behaviors are plotted as the following order from left to right for each box: From hysteretic behaviors whose period is short to long, and the yield strength of each period is small to large, especially for HM2 and HM3, sequentially. In order to reduce the repeated plots in Figure 2.8, 5 indices among 64 are illustrated whose values can represent distinct information.

As shown in Figure 2.8, it is observed that overall shapes of the corresponding box of the first, second, and third rows of features extracted from the CNN are similar except for the relatively short period range (the 34th boxes) and smaller yield strength (the plots fluctuate along with yield strength) in which the structure is prone to

behave nonlinearly during vibration. Because bilinear models need additional information (i.e., nonlinearity of the hysteretic system) compared to a linear model, some of the indices in Figure 2.8 only activate for the input of HM2 and HM3 only. For instance, the 22nd index does not provide any information to differentiate hysteretic behaviors of different linear models (i.e., period), but it activates when the hysteretic behaviors of HM2 and HM3 are employed for the input of the CNN. Based on this investigation, it is concluded that the features extracted from the CNN may give some physical meanings of hysteresis even though it is hard to interpret the features as widely used mathematical parameters such as initial stiffness or normalized yield strength.

2.5.4 Prediction of other structural responses

In earthquake engineering, peak acceleration and velocity are also important structural responses to predict because they are closely related with the base shear capacity and the performance of non-structural components, respectively. The training is carried out in the same environments that are used for the peak displacement except for the replacement of the output datasets. To reduce the computational time for training, a transfer learning method (Pan and Yang, 2010; Yosinski et al., 2014) is employed to train each response using the DNN model that was trained to predict the peak displacement. Moreover, parameters of the DNN model are fixed except those located after the final merged layers, because the layers before the final merged layers extract the features from both earthquake information and structural information, which should not be different even if the different structural responses are predicted. After 30 epochs of training, the MSE and MAE are computed for the natural logarithms of the predicted peak acceleration (Table 2.4)

and velocity (Table 2.5). Moreover, Figure 2.9 and Table 2.6 show the MSE in the peak acceleration and velocity employing the input dataset used in Section 2.5.2. The errors are equivalent to those of the prediction of the peak displacement, which confirms that the proposed DNN model can provide highly accurate predictions regardless of the responses type.

2.6 Conclusions

To facilitate efficient prediction of the seismic responses of nonlinear structural systems without compromising accuracy, this study proposed a new deep learning-based model for response prediction of SDOF systems subjected to various ground motions. To this end, the CNN was introduced to extract the structural information which is described in terms of a displacement and force relationship. The extracted information of hysteresis was, then, merged with ground motion information to predict structural responses. To train the proposed DNN model, the database of seismic responses of structural systems was constructed through a large number of nonlinear dynamic analyses. A small set of data was applied to initialize weights of the DNN model and a proper training scheme was used to overcome the limitation of computational resources. The performance of the proposed method was tested through various numerical investigations. The examinations demonstrated that the proposed method successfully decomposes hysteretic behaviors into the lower-level features and is not overfitted to a certain data set. Moreover, it was found that the accuracy of the proposed method is superior to that of the three existing simple methods that are widely used in practice. It is expected that a variety of applications in earthquake engineering field can be developed using the proposed method such as

efficient and accurate regional seismic loss estimation.

Table 2.1 Comparison of MSE and MAE between train and test datasets

Ground motion set	Mean squared error (MSE)	Mean absolute error (MAE)
Train (1,119)	0.0044	0.0485
Test (300)	0.0253	0.1202

Table 2.2 Comparison of MSE between trained and non-trained hysteretic
behaviors for randomly selected 200 ground motions

Hysteretic behaviors	Mean squared error (MSE)
Trained	0.0090
Non-trained	0.0077

Table 2.3 Comparison of MSE between the proposed and existing methods

Methods	Mean squared error (MSE)
R-μ-T	0.9857
Capacity spectrum method	0.3526
Coefficient method	0.6110
Deep neural network	0.0407

Table 2.4 MSE and MAE of peak acceleration for train and test datasets

Ground motion set	Mean squared error (MSE)	Mean absolute error (MAE)
Train (1,199)	0.0131	0.0903
Test (300)	0.0267	0.1243

Table 2.5 MSE and MAE of peak velocity for train and test datasets

Ground motion set	Mean squared error (MSE)	Mean absolute error (MAE)
Train (1,199)	0.0278	0.1319
Test (300)	0.0433	0.1624

Table 2.6 MSE of the dataset for the elasto-perfectly plastic system subjected to randomly selected 200 ground motions used in Section 2.5.2

Structural responses	Mean squared error (MSE)
Peak acceleration	0.0101
Peak velocity	0. 0273

F_s

ω^2

u

(a)

F_s

F_y

$\alpha\omega^2$

ω^2

u

(b)

F_s

F_y

$\alpha\omega^2$

ω^2

u

(c)
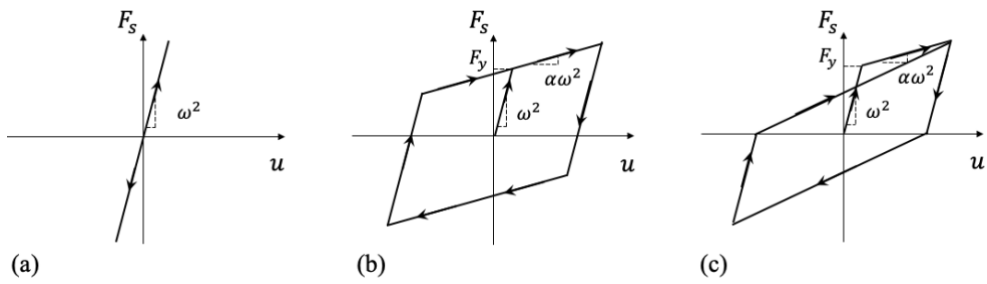
Figure 2.1 Hysteretic behaviors of structural systems: (a) linear (HM1), (b) bilinear kinematic hardening (HM2), and (c) bilinear stiffness degrading system (HM3)
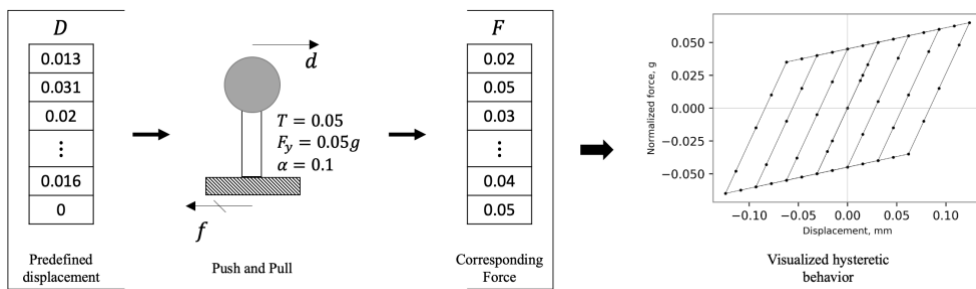
Figure 2.2 Procedure for generating a hysteretic behavior of structural systems
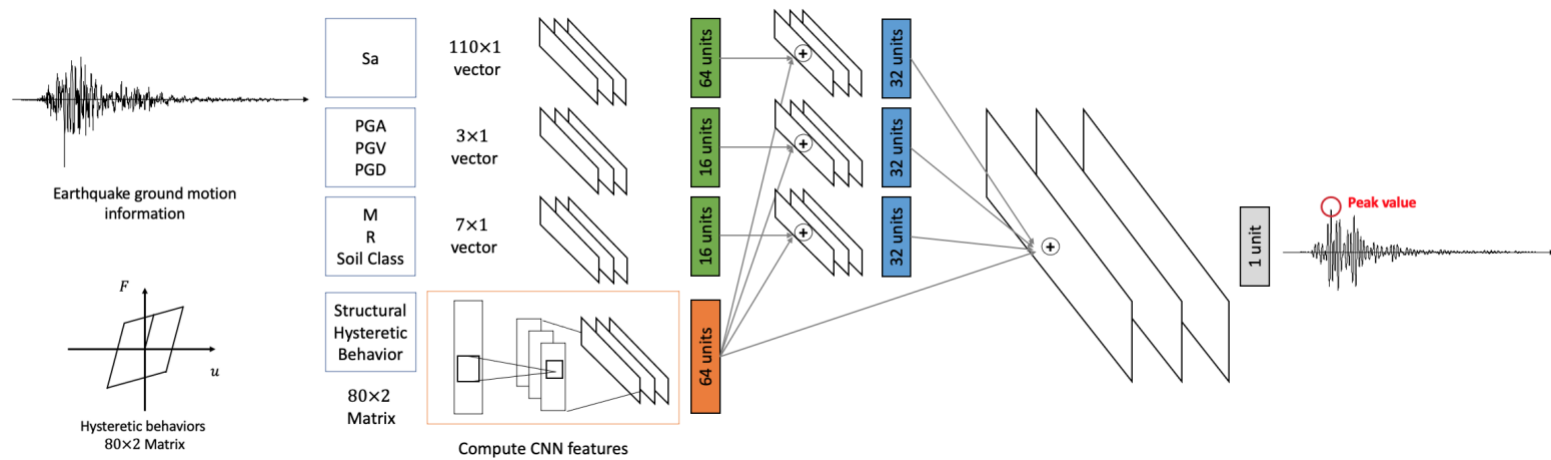
Figure 2.3 Schematic diagram of the DNN model
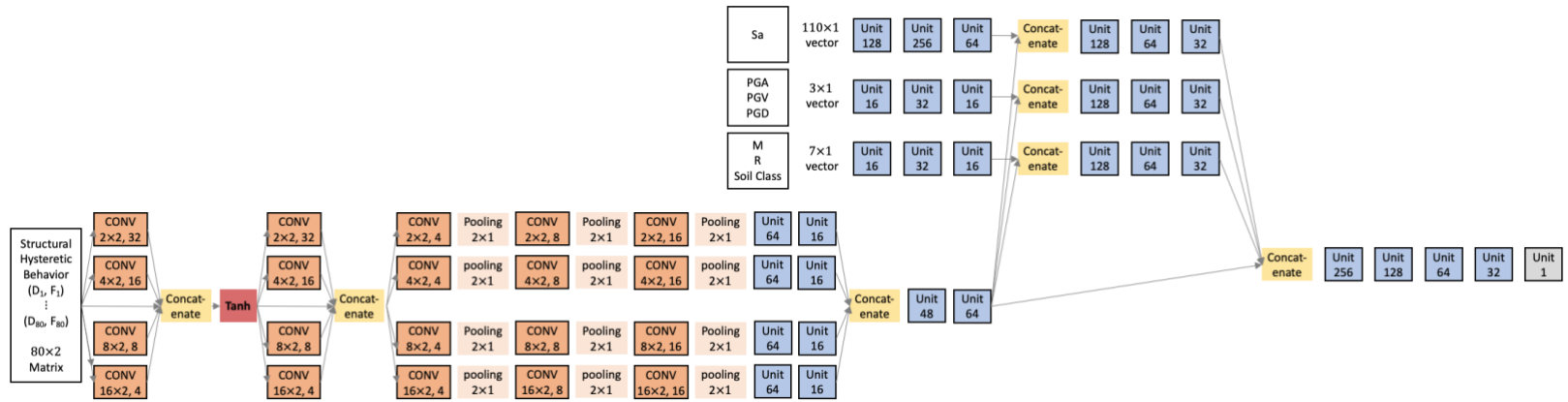
Figure 2.4 Detailed diagram of the proposed DNN architecture
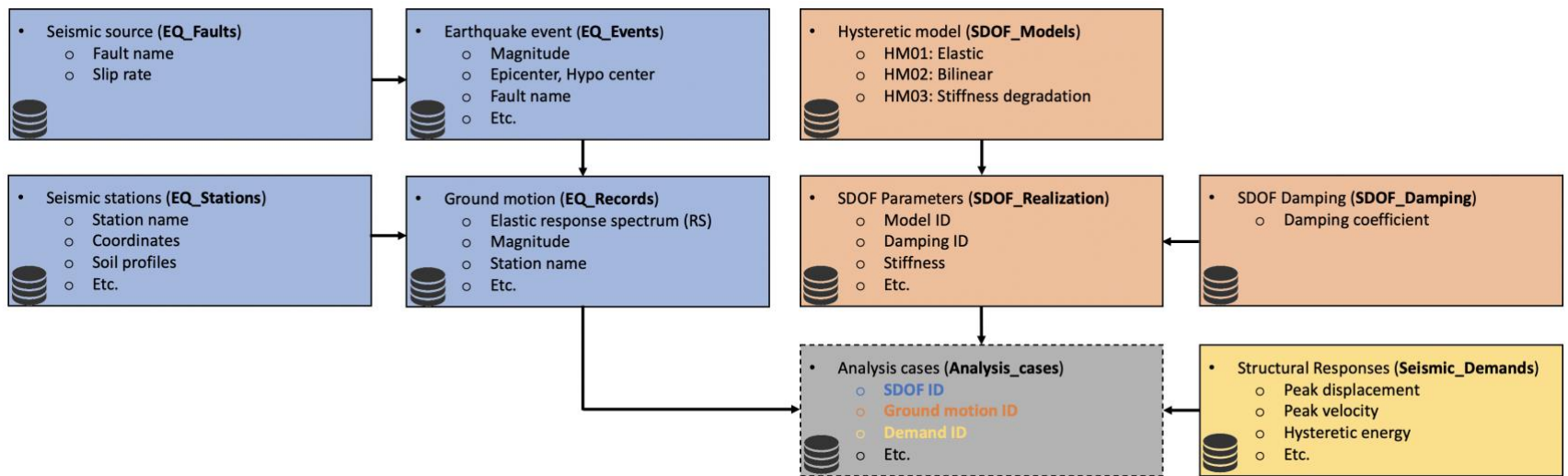
Figure 2.5 Schema of seismic demand database
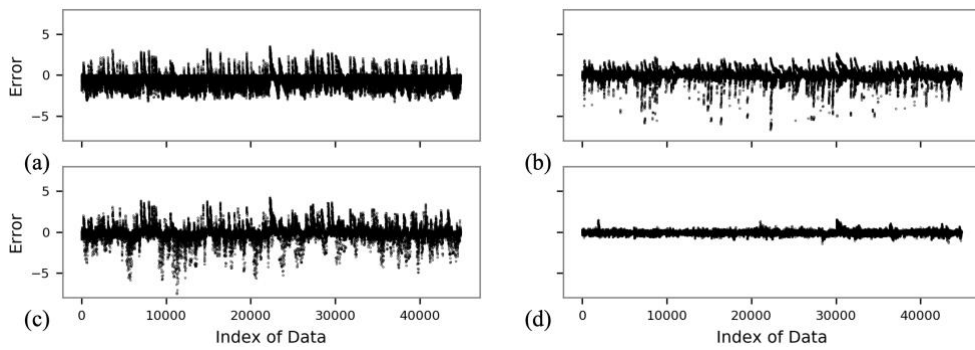
Figure 2.6 The difference between the natural logarithms of peak displacement by nonlinear time history analyses and those by (a) R-μ-T relationship by Nassar and Krawinkler (1991), (b) capacity spectrum method (FEMA 440, 2005), (c) the coefficient method (ASCE 41-13, 2013), and (d) the DNN model
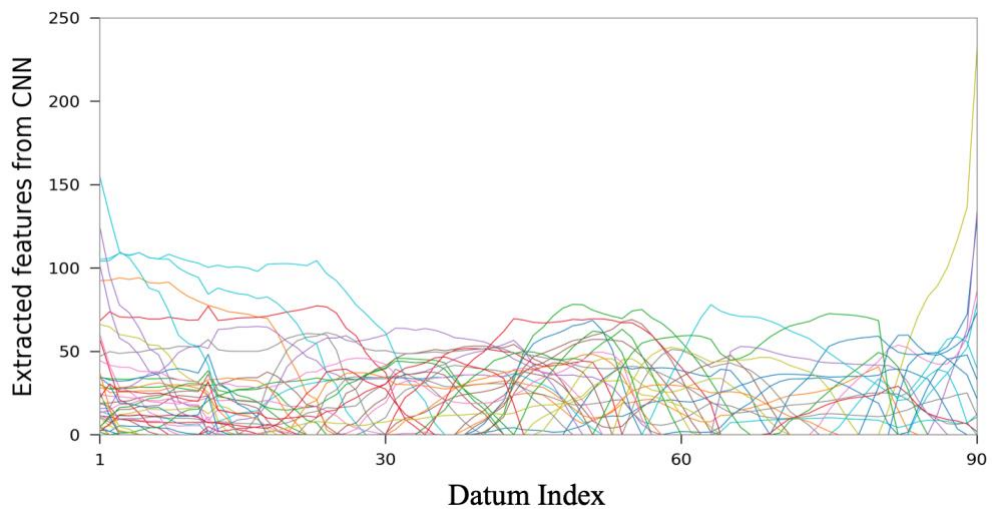
Figure 2.7 Features extracted from the CNN when HM1 is used as an input. *x*-axis represents each of hysteretic behavior and *y*-axis represents the values obtained from the CNN. To find the pattern clearly, the values of the same element in the 64 units are connected each other
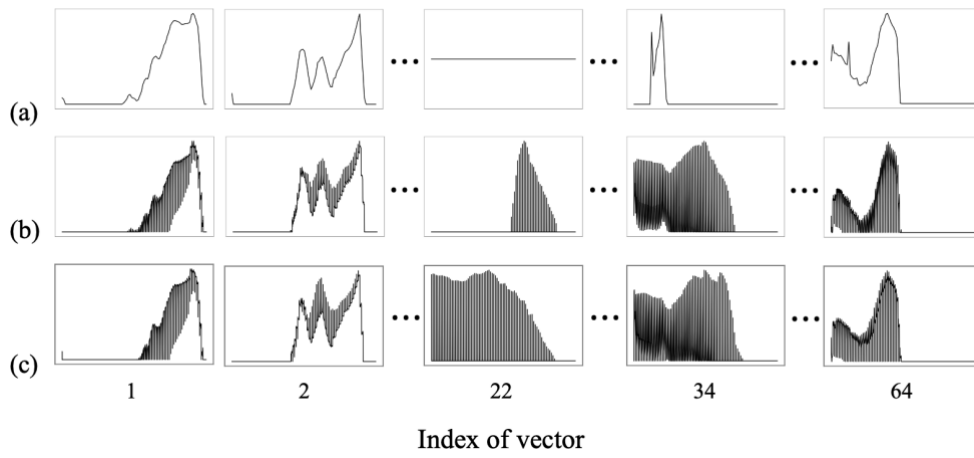
Figure 2.8 Features extracted from the CNN, i.e., the output of the final FC layer when (a) linear elastic (HM1), (b) elasto-perfectly plastic (HM2 with $\alpha = 0$), and (c) elasto-perfectly plastic with stiffness degradation (HM3 with $\alpha = 0$) are used as the input. In words, Figure 2.7 is the combination of Figure 2.8(a)
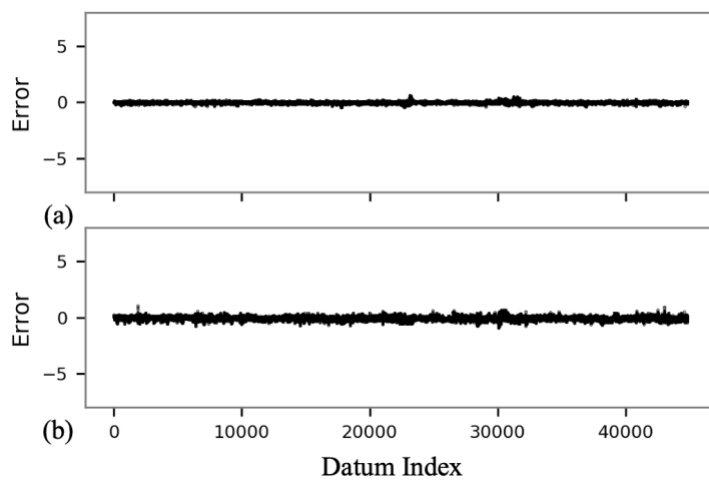
Figure 2.9 MSE of the natural logarithms of the predicted responses: (a) peak acceleration, and (b) velocity when applying the dataset used in Section 2.5.2

Chapter 3. Probabilistic Evaluation of Seismic Responses by Bayesian Deep Neural Network

3.1 Introduction

A proper estimation of structural responses under seismic excitation is one of the most essential tasks to establish an effective disaster risk management framework against seismic hazards. However, it is challenging to accurately assess the seismic response of a nonlinear structural system because of the variabilities, especially those in the ground motions lie in the ground motions (Behmanesh et al., 2017; Celik and Ellingwood, 2010; Kwon and Elnashai, 2006; Vamvatsikos and Fragiadakis, 2010; Yin and Li, 2010). Nonlinear time history analyses using a suite of ground motions is considered one of the most accurate ways to consider such variabilities, but the method entails large computational efforts. As an alternative of the time history analysis that can overcome such challenges, approximate nonlinear static procedures and the DNN model have been developed as respectively described in Chapters 1 and 2. However, their predictions are deterministic, thus it is difficult to quantify uncertain errors in the response estimations. Such an uncertainty is usually caused by the loss of information in representing input data by selected input features. The uncertainty needs to be quantified because the responses of a nonlinear structural system could vary widely even if ground motions are considered similar in terms of the selected measures of seismic intensity (Cornell et al., 2002; Deniz et al., 2018; Riddell, 2007).

To assess the variabilities of structural responses for given values of the selected

intensity measures, a new probabilistic deep neural network (P-DNN) model is proposed in this chapter. To this end, the deterministic DNN model from Chapter 2 is further developed by introducing a new loss function based on the Bayesian deep learning approach (Gal 2016; Kendall and Gal, 2017). The new loss function captures the variabilities of the output given the input dataset based on the assumption that the corresponding output follows a Gaussian probability density function (PDF). Using the Bayesian concept, the mean and variance of the structural response can be predicted for given hysteretic models and a set of ground motion information.

This chapter first provides a brief overview of Bayesian deep neural network along with an introduction of the new loss function. Next, a new architecture for the P-DNN model is proposed which is trained using the constructed database in Chapter 2. Thorough numerical investigations of the P-DNN model are carried out to understand the relationship between inherent uncertainties in the input and the corresponding predicted mean and variance. The P-DNN model enables us to quantify the impact of each ground motion feature on the uncertainty of structural responses, i.e., uncertainty quantification (Kim et al., 2020).

3.2 Bayesian neural network

The Bayesian deep learning method (Kendall and Gal, 2017) was originally developed to account for two types of uncertainties: the uncertainties in the estimated parameters of the deep learning model (i.e., model uncertainties) and the randomness inherent in the input data. The former type of uncertainties can be addressed by so-called Monte Carlo (MC) dropout assuming that the model parameters follow a

probabilistic distribution. The latter can be facilitated by using a new loss function which will be discussed in the following subsection. Unlike the latent uncertainties in the input data, the model uncertainties can be reduced by using enough amount of data in training. Thus, this study aims at quantifying the impact of the inherent uncertainties of the input data on the prediction of structural response using the new loss function. This uncertainty arises because several input features, e.g., M, R, soil type, PGA, PGV, PGD, and response spectrum, characterize a ground motion time history and used as the input of the DNN model. It is noted that this particular type of randomness of the outcome remains significant in the DNN model even if a good understanding about the input-output relationship and a sound initial condition for training a DNN model are accompanied. This section introduces background information of how such uncertainties can be captured through a neural network. Then, the influence of uncertainties in the input data, especially those in the ground motion, on the seismic response prediction is discussed in detail.

3.2.1 Quantification of input uncertainties

In the Bayesian deep learning method, the impact of latent uncertainties in the input data is quantified by obtaining the probability distribution of the output prediction from a deep neural network. In other words, the output of a deep neural network is presented as a Gaussian random variable, characterized by the conditional mean and variance for given input values. The conditional variance can be modeled following one of the two assumptions: homoscedasticity and heteroscedasticity (Gal and Ghahramani, 2015; Kendall and Gal, 2017). The former assumes that the conditional variance of the prediction is constant over the entire input data space. For example, the widely used probabilistic relationship between a seismic intensity measure (IM)

and a damage measure (DM), $DM = a \cdot IM^b \cdot \varepsilon$ (Cornell et al., 2002), adopts the homoscedasticity assumption because the conditional variance of the damage measure is assumed to be independent of IM in the equation. The latter, by contrast, assumes that the conditional variance varies over the input data space.

In order to take into account the fact that the prediction uncertainty actually depends on the input value, the heteroscedasticity assumption is adopted. To this end, the following loss function (Kendall and Gal, 2017) is used instead of Eq. (1.2):

$$L = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{2\sigma(x_i)^2} \|y_i - f(x_i)\|^2 + \frac{1}{2} \log \sigma(x_i)^2 \right) \tag{3.1}$$

where $f(x_i)$ and $\sigma(x_i)^2$ respectively represent the conditional mean and variance of the output prediction given the particular input x_i. The loss function in Eq. (3.1) is proportional to the negative log-likelihood based on the Gaussian distribution assumption. Therefore, the training of a deep neural network with the loss function in Eq. (3.1), i.e., finding model parameters minimizing the loss function, is equivalent to using the maximum a posteriori probability (MAP) approach employing a non-informative prior (Murphy, 2012).

When the discrepancy between the predicted mean value $f(x_i)$ and the train data y_i is large at x_i, the optimization algorithm will rely on reducing the value of the first term in Eq. (3.1), which tends to make the corresponding conditional variance $\sigma(x_i)^2$ of the optimal solution relatively large. However, the existence of the second term in Eq. (3.1) prevents the conditional variance from growing too large, which, in turn, helps a DNN model learn the input-output relationship rather than providing a good match at x_i in the particular train dataset only. In other words, the loss function of the P-DNN model helps figure out the structure of the input-output

relationship, whereas a DNN model training with Eq. (1.2) would require additional efforts, e.g., cross-validation to achieve the bias-variance tradeoff.

3.2.2 Role of conditional variance for seismic response prediction

Since the P-DNN model propagates the inherent randomness in the input to the output, the conditional variance of the structural responses, $\sigma(x_i)^2$ represents the uncertainties caused by feature-based description of both hysteresis loops and ground motions. However, in this study, the conditional variance mostly represents the uncertainties from the ground motion features. This is based on the fact that the hysteresis loop obtained from the quasi-static cyclic analysis in Section 2.2.1 includes most of the information needed to distinguish one another, whereas the severity and characteristics of the ground motions cannot be perfectly characterized even using multiple seismic intensity measures. In other words, ground motions having the same values of the selected features may lead to a significantly different level of structural responses if nonlinear time history analyses are used.

It should be also noted that the impact of the input randomness on the variance of structural responses is affected by the behaviors of the structural system during seismic excitations. With growing interest in properly describing the impact of a given earthquake shaking on structural response, the spectral acceleration at the first mode period is one of the most commonly used intensity measures in earthquake engineering because of its high correlation with the seismic responses of a linear structure. It is, however, widely known that there are no intensity measures that are fully correlated with the structural responses when the structure exhibits nonlinear behavior (Cornell et al., 2002; Deniz et al., 2018; Riddell, 2007). Thus, it is expected that the deep learning prediction for the linear structural behavior with a response

spectrum input shows smaller variance compared to the nonlinear structural behavior with the same input (or even with the additional features of intensity measure), which will be discussed later in Section 3.4.4.

3.3 Development of probabilistic deep neural network model

This section first presents the architecture of neural network model to predict two outputs, i.e., mean and variance of structural responses, by modifying the DNN model constructed in Chapter 2. The modified neural network architecture is followed by the training methodology that includes optimizer, number of epochs, and batch sizes. Note that, as similar to the DNN model, the P-DNN model was also constructed using TensorFlow (Abadi et al., 2016) and trained using the Compute Canada's GPU cluster.

3.3.1 Architecture details

The architecture proposed in Chapter 2 is enhanced by revising the final merged layers which are depicted as parallelograms in Figure 2.3. The diagram in Figure 3.1 shows details of the revised final layers developed through an extensive exploration of various types of network structures. Each yellow box denoted as "Batch norm." in the figure represents the batch normalization (Ioffe and Szegedy, 2015), which is applied after a rectified linear operator (ReLU; Nair and Hinton, 2010). The orange box, which represents the information extracted from the hysteretic behavior, is merged with each group of earthquake ground motion information (blue boxes). Then, three hidden layers are implemented to convolute different types of information. The last layer with 256 units is diverged into two ways with four hidden

layers each to estimate the mean and variance of structural responses, respectively. The architecture of the P-DNN model is developed by replacing the final output of the DNN model as two outputs following the approach by Kendall and Gal (2017). While Kendall and Gal (2017) used no hidden layers after diverging into two ways, it is found that the prediction performance of the P-DNN model can be enhanced by increasing the number of the hidden layer, especially for data points for which the stiffness of the hysteretic behavior is relatively large. On the other hand, a large number of hidden layers, say more than 4, may hamper proper convergence of the loss function.

3.3.2 Training

Using the database constructed in Section 2.4.1, the P-DNN model is trained to predict peak displacement. To resolve the skewness of the distribution of response spectrum, PGA, PGV, PGD, and structural responses, the natural logarithm is applied. Thereby, the structural responses predicted by the P-DNN model always take positive value, and are assumed to follow the lognormal distribution. To check whether the P-DNN model properly trains the relationship between the input and the output, only 80% of ground motions (1,199 ground motions) are selected and used for training. Moreover, in order to overcome the limitation of the computational resources during training, 300 ground motions among 1,199 ground motions are randomly selected for each epoch of training. The trained parameters of the DNN model are utilized as the initial weights of the P-DNN model, and Adam optimizer (Kingma and Ba, 2014) is adopted as the optimization algorithm to minimize the loss function in Eq. (3.1) with 512 batch sizes and 70 epochs during the training.

The MSEs of predicted responses randomly drawn from the estimated Gaussian

distributions are shown in Figure 3.2 as increasing the epochs. It is noted that 50 ground motions among 300 ground motions of test dataset are randomly selected for each epoch of training to calculate the MSE. Since the transfer learning technique is introduced, the initial value of MSE is relatively small for both train and test datasets compared to the general loss curve whose value of the first epoch is relatively large. As shown in the trend of MSE drop, one can infer that the more epochs are preformed, the better prediction accuracy can be obtained for the test dataset.

3.4 Performance of probabilistic deep neural network model

Five different numerical investigations are carried out to demonstrate the efficiency and effectiveness of the trained P-DNN model. First, the prediction accuracy of the P-DNN model is compared with other methods including the DNN model developed in Chapter 2. Next, the conditional variance predicted by the P-DNN model is examined along with different datasets of input. One can observe the variation of the conditional variance along with the nonlinear behaviors of structural systems. Third, to verify the variance estimated by the P-DNN model, Bayesian linear regression is introduced. The prediction results by both methods using the same dataset are compared with each other. Forth, the impact of ground motion features on the predicted uncertainties are comprehensively analyzed using the Bayesian deep neural network. Finally, trainability of the P-DNN model for other structural such as peak acceleration and velocity responses is demonstrated.

3.4.1 Prediction accuracy

The error measures, i.e., MAE, MSE and mean of the estimated conditional variance

(MV), are computed for the train and test datasets, respectively as shown in Table 3.1. Since the residual error and the conditional variance have a positive relationship with each other as explained in Section 3.2.1, it is expected that MV is also proportional to MSE and MAE (i.e., large MSE and MAE tend to result in large MV). Note that, in the same manner as in Section 2.5, all error measures are calculated in terms of the natural logarithms of the structural responses.

The low values of MAE and MSE for both train and test datasets in Table 3.1 confirm that the predictions by the P-DNN model are fairly close to those by the nonlinear dynamic analysis. However, it might seem that MAE of the test dataset is significantly larger than that of the train dataset, which could be considered as an evidence that the trained P-DNN model is overfitted. However, it is noted that, the average ratios of the simulated responses to the predictions from the P-DNN model in the original scale (i.e., before taking the natural logarithm) are estimated based on MAE values as $e^{0.05} \approx 1.05$ and $e^{0.13} \approx 1.13$, respectively, which are fairly close to one. In addition, MV of the test dataset is a little bit bigger than that of the train dataset, as expected.

In order to test and demonstrate the performance of the proposed method, its prediction accuracy is compared with the aforementioned three simplified methods as well as the DNN model in Chapter 2. For this test, nonlinear cases of the elasto-perfectly plastic systems (i.e., zero post-yield stiffness) subjected to randomly selected 200 ground motions are provided as inputs for the methods. This is 8.3% of the total dataset, i.e., 44,842 cases among 540,000 cases (=200×2,700). The MSE is calculated between the peak displacement obtained from dynamic analyses and that from each method, as shown in Table 3.2. The results confirm that the P-DNN model shows superior accuracy compared to three existing methods, but the MSE

calculated from the P-DNN model is slightly larger than that from the DNN model. Since several research efforts reported in the literature showed the performance of the Bayesian deep learning method exceeds that of the deterministic deep learning method with enough epochs of training (Gal and Ghahramani, 2015, Gal and Ghahramani, 2016), it is expected that better results could be obtained by additional training, especially by increasing the number of ground motions. The model was trained with 70 epochs to produce the results reported in this study while, based on rule of thumb, more than 10^3 epochs are needed.

3.4.2 Investigation of predicted responses

To confirm the performance of the P-DNN model in predicting the mean and variance of the peak displacement, the probabilistic predictions are plotted together with the structural responses obtained by the dynamic analysis. The predictions by the deterministic DNN model developed in Chapter 2 are plotted together to provide the comparison between the probabilistic and the deterministic DNN models. Since it is difficult to plot all data points in a single plot, the predictions for two different hysteretic models under 1,499 ground motions are shown in Figure 3.3. In each plot, the data points in the test results (i.e., structural responses from dynamic analysis and the DNN model) are rearranged in the increasing order of the mean values by the P-DNN model. The probabilistic predictions are represented by the mean curve (black solid line) and the orange shaded area showing mean ± 1 standard deviation interval, which covers approximately 70% of the probability distribution of the predicted structural response when the Gaussian distribution is assumed. The responses observed during the dynamic analysis and the predictions by the deterministic DNN model are shown by blue circles and red x-marks, respectively.

The horizontal dashed line in each figure represents the yield displacement of the corresponding structural system to distinguish cases showing linear and nonlinear behavior. To visualize the performance of the proposed network in detail, close-up images of two parts of the linear and nonlinear behaviors in each figure (dashed rectangles) are provided at the right-hand side of Figure 3.3. Note that the plots show the natural logarithms of the peak displacement in meter.

The mean curves of the probabilistic model successfully capture the central tendencies of the structural responses under given ground motions, which implies that the developed probabilistic model achieves unbiased predictions without overfitting. It is also found that the majority of the structural responses from the time history analysis (blue circle) and those predicted by the deterministic DNN model (red x) fall within the mean ± 1 standard deviation intervals in both plots. The variation of the conditional variance over the input data is visualized by the varying width of the standard deviation interval presented by the orange shade. It is also noted that the standard deviation interval is relatively small and almost constant where the peak displacement is small, but starts increasing at a certain point. This is because a structural system behaves nonlinearly under a ground motion with a relatively large intensity, but the features selected to train the probabilistic neural network cannot fully describe the ground motion characteristics which influence the nonlinear behavior of the structural system. The increases of the widths of the standard deviation interval around the yield displacement lines indicate that the P-DNN model can capture the uncertainties increased by the nonlinear behavior. Although not shown in this study, the standard deviation interval of the linear hysteretic model (HM1) is almost constant and small over all ground motions as seen in the linear behavior parts of Figure 3.3.

3.4.3 Verification of estimated variance

In order to verify the predicted conditional variances, a simple numerical investigation is conducted by using the Bayesian linear regression (MacKay, 1992). Since the estimated mean of the specific structural system with various ground motions shows almost linear curve when the data representing the x-axis are rearranged in increasing order of the mean of the predicted value (see Figure 3.3), it is possible to check the estimated variance by comparing the results between the Bayesian regression and the P-DNN model. For this purpose, the cases corresponding to the datum index from 601 to 1,200 in Figure 3.3(b) are selected and plotted in Figure 3.4(a). Bayesian linear regression is performed to predict the observation data (blue circle) given the datum index as shown in Figure 3.4(b). The average values of the standard deviations for the P-DNN model and the linear regression are 0.135 and 0.145, respectively, which shows similar results. Since the P-DNN model has more flexibility in capturing the mean curve, its standard deviation is slightly smaller than the one from the linear regression.

3.4.4 Impact of ground motion characteristics on uncertainties of responses

To investigate how individual features of earthquake ground motion (e.g., earthquake characteristics, peak values of ground motion time histories, and the response spectrum) affect the variabilities of the response prediction, various P-DNN models are trained using different subsets of the input parameters instead of using the complete feature set as done in Section 3.3. Table 3.3 lists seven different cases made

in terms of the type and the amount of ground motion information used in the training process as an input of the P-DNN models. It is noted that the results in Table 3.1 match Case 7 in Table 3.3. For example, if the earthquake characteristics are used as only ground motion information in training, i.e., Case 1 in Table 3.3, the layers related with peak values of ground motion time histories and the response spectrum are discarded while training and predicting the peak displacement. Table 3.3 also reports the error measures to compare the performance.

The trends of MAE, MSE and MV in Table 3.3 are similar to those in Table 3.1 (i.e., Case 7), except Case 1 in which the MSE is bigger than the MAE. This indicates that the neural network model in Case 1 cannot predict the structural responses, properly. As expected, the neural network models trained with the response spectrum show lower MAE and MSE values compared with other models. The smaller MSE value of Case 5 compared to Case 6 implies that the peak values of ground motions (PGA, PGV, and PGD) affect the structural response more than earthquake characteristics information (M, R, and soil class). It is also found that the MAE and MSE values for Cases 2-6 in Table 3.3 are bigger than those calculated from Case 7. Each group of information stands for different feature selections of the earthquake ground motion, and thus the results show that each feature tends to balance their limitations, i.e., insufficient information.

In order to further investigate the relationship between the input data and the structural response, the lognormal PDFs of the peak displacement predicted from the P-DNN are plotted in Figure 3.5. Table 3.4 lists the corresponding distribution parameters. The PDFs from Case 1 are not shown due to their poor accuracy. For the purpose of illustration, the bilinear stiffness degrading system (HM3) with period of 1 sec, normalized yield strength of 0.05g, and post yield stiffness ratio of 0.1 are

selected along with two ground motions for linear (1970 Lytle Creek earthquake) and nonlinear cases (1968 Borrego Mountain earthquake).

Table 3.4 shows that the coefficient of variation (c.o.v) of the linear case is much smaller than that of the nonlinear case. For the linear case, neural network models trained using the response spectrum, i.e., Case 3, 5, 6 and 7 show a good agreement with the peak displacement obtained from the time history analysis (0.0035 m) and also have smaller coefficient of variations. On the other hand, for the nonlinear case, Case 7 provides the median value close to the actual observation, and smaller coefficient of variation than that of Case 3, 5 and 6. This result confirms that the features representing earthquake characteristics (M, R, and soil class) and the peak values (PGA, PGV, and PGD) are closely related with predicting the nonlinear behavior of the structural system.

Following the procedure described above, one can quantify the contributions of a specific input data such as ground motion duration and fault mechanism to the uncertainties of responses, or find the best combination of the input data by checking the level of decrease in MSE and MV. However, it is worth noting that although a deep learning method has an ability to analyze the relationship between the input and output variables using raw datasets, a careful preprocessing of the input data is needed due to potential issues regarding convergence of loss functions.

3.4.5 Prediction of other structural responses

To demonstrate the trainability of the probabilistic deep neural network framework for other structural responses, peak acceleration or velocity is selected as the output of the neural network model. To utilize the knowledge gained during the aforementioned training, the parameters of the P-DNN model trained for the peak

displacement is used in developing P-DNN models for the peak velocity and acceleration. Since the proposed model first extracts the features from the input, and then merges those to predict the structural response, only final merged layers depicted in Figure 3.1 are newly trained. After 40 epochs of training, MAE, MSE, and MV are computed for the train and test datasets respectively, as shown in Table 3.5 and 3.6. Since the results show similar trends to those for the peak displacement, it is concluded that the proposed deep learning method can accurately predict the peak velocity and acceleration as well. Moreover, prediction of the peak acceleration shows better performance in terms of accuracy, i.e., lower MAE and MSE values, compared to other responses. It is speculated that this is because the response spectrum whose values are the maximum acceleration of the entire period (0.05 sec to 10 sec), i.e., response spectrum, is employed as one of the inputs for training the model.

To further investigate the performance of the P-DNN models, the estimated peak velocity and acceleration of the hysteretic models are plotted for 1,499 ground motions in Figure 3.6 and 3.7, respectively. The P-DNN models show a good performance while capturing larger uncertainties in these responses of the nonlinear systems, which is similar to the results for the peak displacement (Figure 3.3). In Figure 3.6 and 3.7, different markers are used to differentiate the linear and nonlinear cases for peak velocity and acceleration. This is because, in Figure 3.3, the nonlinearity of the system can be easily distinguished in terms of the pre-defined yield displacement, which is the linear elastic limit, but it is difficult to identify the nonlinear cases in terms of the other structural responses. There is no concrete discrimination point of the linear and nonlinear cases for the peak velocity because the structural system can be considered as a linear case even if the peak velocity is

bigger than the pseudo yield velocity, i.e., multiplying the yield displacement to the circular natural frequency. However, the red horizontal line in Figure 3.7 clearly distinguishes the linear and nonlinear cases just as the case of the peak displacement in Figure 3.3. This is because when the structural damping is relatively low, the peak acceleration can be directly estimated from the peak displacement based on the Duhamel's integral. The numerical example demonstrates that the proposed method, developed for the displacement first, is effective in predicting other types of structural responses after retraining neural network models with velocity or acceleration data.

3.5 Conclusions

In this chapter, the P-DNN model was proposed to evaluate the seismic responses of structural systems while quantifying the prediction uncertainties. The method predicts two outputs, the conditional mean and variance of the structural responses which are the main components of the loss function developed based on the Gaussian assumption. By introducing the mean and variance as the output of the neural network, the developed P-DNN model can capture the effects of randomness inherent in the input data, especially for the loss of information in the process of representing the earthquake ground information by selected features. The performance of the proposed method is comprehensively tested through five numerical experiments which confirm that the P-DNN model can successfully predict the relationship between the input and the output of the datasets. In particular, it produces a relatively small variance to the structural system behaving in the linear elastic range, while a relatively large variance is obtained for the nonlinear cases.

The characteristics of the developed model and framework are expected to facilitate various applications in earthquake engineering such as fragility assessment and regional seismic loss assessment of urban area.

Table 3.1 Comparison of MAE, MSE and MV for peak displacement between the
       train and test datasets

Ground motion set	MAE	MSE	MV
Train (1,199)	0.0531	0.0060	0.0028
Test (300)	0.1334	0.0308	0.0032

Table 3.2 Comparison of MSE between the proposed and existing methods

Methods	Mean squared error (MSE)
R-μ-T	0.9857
Capacity spectrum method	0.3526
Coefficient method	0.6110
DNN model	0.0407
P-DNN model	0.0518

Table 3.3 Comparison of MAE, MSE, and MV for the train and the test datasets
when different combinations of ground motion features are used as the
input of the P-DNN model during the training process

Case	Input	Ground motion Set	MAE	MSE	MV
Case 1	Earthquake characteristics (M, R, and soil class)	Train	0.5695	0.6699	0.0404
		Test	0.7227	0.8970	0.0367
Case 2	Peak values (PGA, PGV, and PGD)	Train	0.2162	0.0946	0.0580
		Test	0.2946	0.1454	0.0147
Case 3	Response spectrum (Sa)	Train	0.0650	0.0096	0.0034
		Test	0.1551	0.0422	0.0037
Case 4	Earthquake characteristics & Peak values	Train	0.1607	0.0555	0.0050
		Test	0.2773	0.1294	0.0052
Case 5	Peak values & Response spectrum	Train	0.0597	0.0070	0.0027
		Test	0.1354	0.0317	0.0029
Case 6	Response spectrum & Earthquake characteristics	Train	0.0592	0.0082	0.0028
		Test	0.1558	0.0442	0.0031
Case 7	Earthquake characteristics & Peak values & Response spectrum	Train	0.0531	0.0060	0.0028
		Test	0.1334	0.0308	0.0032

Table 3.4 Median and the coefficient of variation (c.o.v) of the peak displacement
from the P-DNN model for the linear and nonlinear cases. Note that the
median and c.o.v are readily calculated by the estimated lognormal
distribution

Case	Linear		Nonlinear	
	Median (m)	c.o.v	Median (m)	c.o.v
Case 2	0.0031	0.0706	0.0578	0.1983
Case 3	0.0036	0.0387	0.1150	0.2781
Case 4	0.0030	0.0451	0.0723	0.3476
Case 5	0.0036	0.0428	0.0839	0.2211
Case 6	0.0036	0.0292	0.0964	0.3318
Case 7	0.0037	0.0352	0.0877	0.2350

Table 3.5 Comparison of MAE, MSE, and MV for the peak velocity between the train and test datasets

Ground motion set	MAE	MSE	MV
Train (1,199)	0.0550	0.0066	0.0023
Test (300)	0.1342	0.0315	0.0026

Table 3.6 Comparison of MAE, MSE, and MV for the peak acceleration between the train and test datasets

Ground motion set	MAE	MSE	MV
Train (1,199)	0.0446	0.0039	0.0014
Test (300)	0.1162	0.0237	0.0015
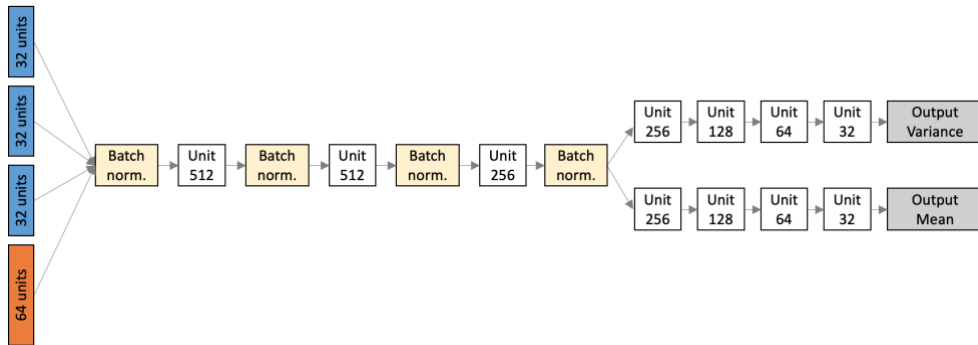
Figure 3.1 Detailed diagram of the final layers developed for P-DNN model
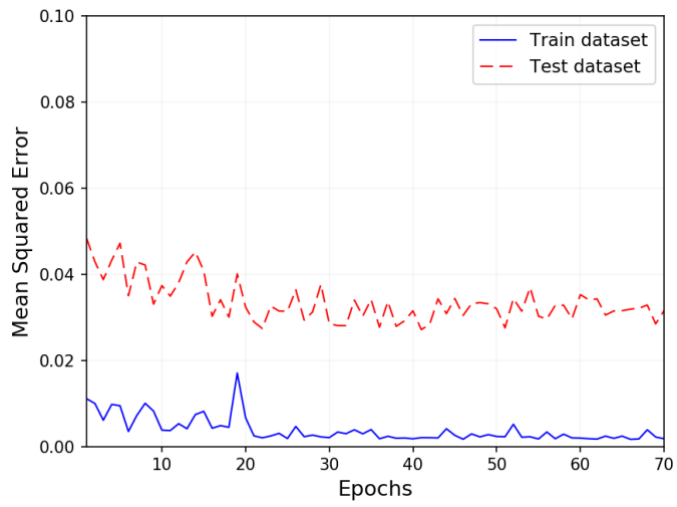
Figure 3.2 MSE of probabilistic predictions by the P-DNN model for the train (blue solid line) and test (red dashed line) datasets as an increase of epochs
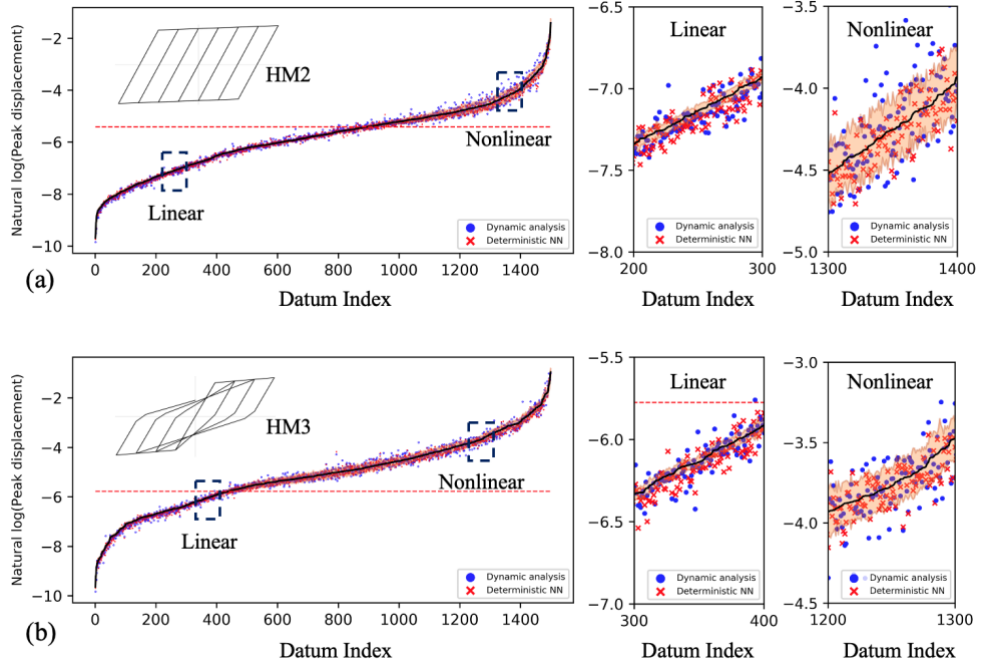
Figure 3.3 Performance of the P-DNN model trained to predict the peak displacement for (a) bilinear kinematic hardening (HM2) with period of 0.3 sec, normalized yield strength of 0.2g, and post yield stiffness ratio of 0.02, and (b) bilinear stiffness degrading system (HM3) with period of 0.5 sec, normalized yield strength of 0.05g, and post yield stiffness ratio of 0.05
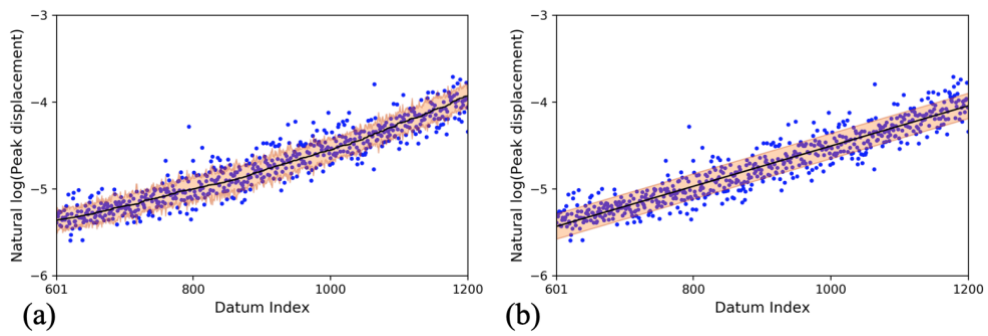
Figure 3.4 Comparison of the standard deviation estimated from (a) the P-DNN model and (b) the Bayesian linear regression

(a)

(b)
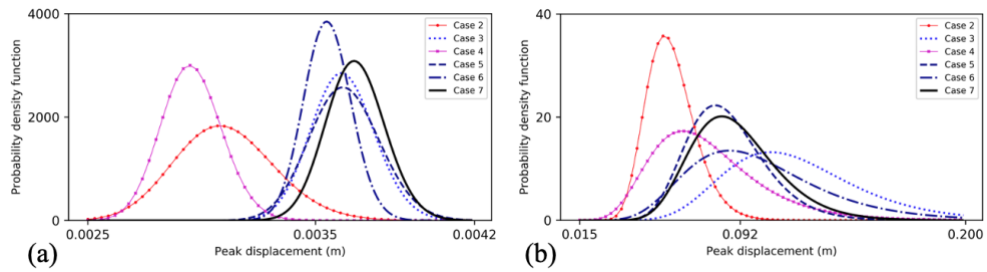
Figure 3.5 Lognormal PDF constructed using the conditional mean and variance from the P-DNN model for systems showing (a) the linear behavior, and (b) the nonlinear behavior during the earthquake excitement (i.e., peak displacement is bigger than the yield displacement 0.0124 m). The observed peak displacements by time history analysis for the linear and nonlinear are 0.0035 m and 0.092 m, respectively
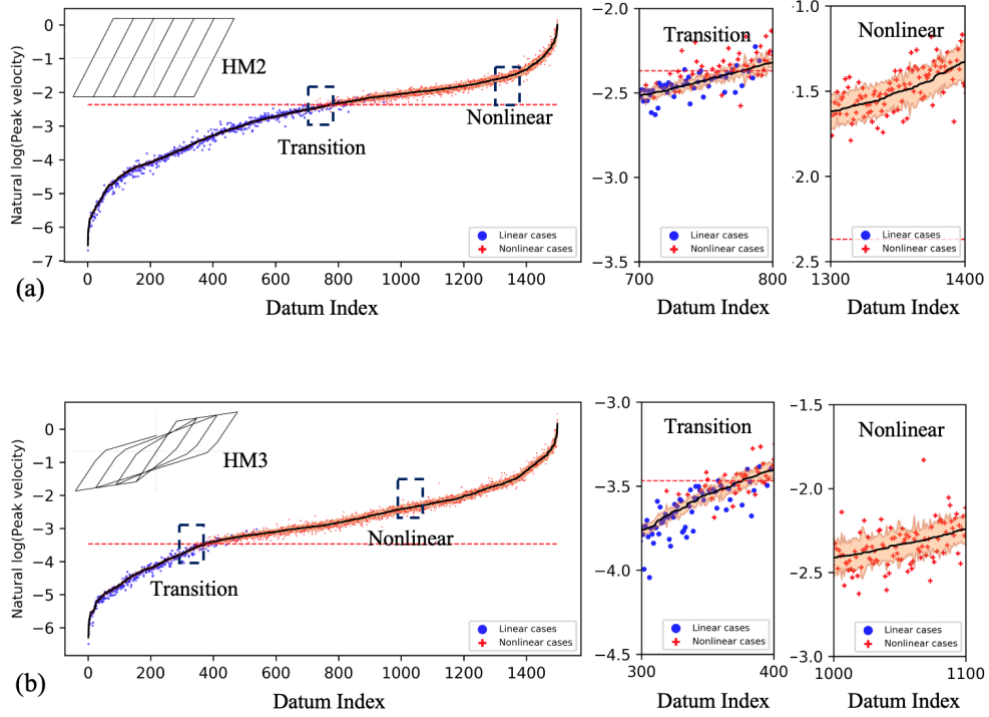
Figure 3.6 Performance of the P-DNN model trained for the peak velocity using (a) the bilinear kinematic hardening (HM2) with period 0.6 sec, yield strength 0.1g, and post yield stiffness ratio 0, and (b) the bilinear stiffness degrading system (HM3) with period 0.4 sec, yield strength 0.2g, and post yield stiffness ratio 0.1. The observed responses from the dynamic analysis for the linear and nonlinear cases are represented as the blue circle and the red plus mark, respectively. The red dotted horizontal line represents the pseudo yield velocity which is estimated by multiplying the yield displacement to the circular natural frequency
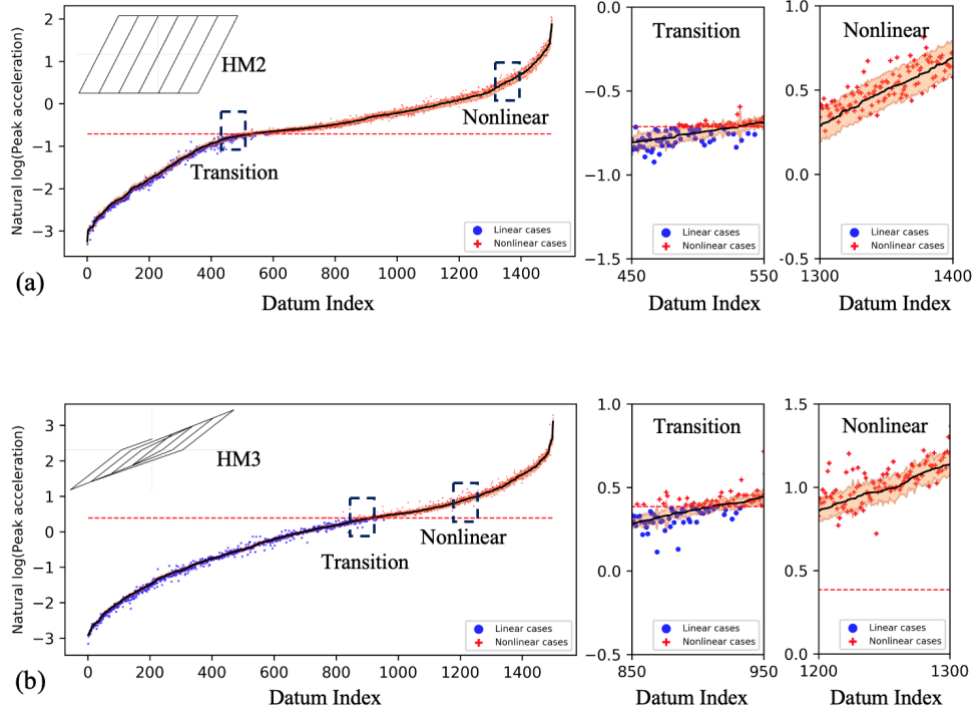
Figure 3.7 Performance of the P-DNN model trained for the peak acceleration using (a) the bilinear kinematic hardening (HM2) with period 0.05 sec, yield strength 0.05g, and post yield stiffness ratio 0, and (b) the bilinear stiffness degrading system (HM3) with period 0.08 sec, yield strength 0.15g, and post yield stiffness ratio 0.5. The observed responses from the dynamic analysis for the linear and nonlinear cases are represented as the blue circle and the red plus mark, respectively. The red dotted horizontal line represents the pseudo yield acceleration which is estimated by multiplying the yield displacement to square of the circular natural frequency

Chapter 4. Deep Learning-based Seismic Response Prediction of Generalized Hysteresis

4.1 Introduction

The DNN and P-DNN models, which were presented in Chapter 2 and Chapter 3, respectively, show superior prediction accuracy compared to the existing simple regression-based methods used in practice. Furthermore, it was demonstrated that the P-DNN model can properly quantify the inherent uncertainties of the stochastic excitation. However, since the models were trained based on the seismic responses of idealized hysteretic behaviors defined by a small number of parameters (stiffness, yield force, and post-yield stiffness ratio), it is difficult to guarantee the prediction accuracy of seismic responses for hysteretic behaviors showing significant stiffness/strength degradations and pinching effects. Such behaviors are usually shown in structural systems having plastic hinges or those constructed by deterioration materials such as reinforced concrete or wood, whose properties govern the seismic performance when the behavior of the structural system is significantly nonlinear (Deniz et al., 2017).

To construct a deep neural network model that can predict the seismic responses of realistic hysteresis, a new seismic demand database is required to incorporate more sophisticated hysteretic characteristics. In this regard, this study employs Bouc-Wen model extended by Baber and Noori (1985) which can describe the stiffness/strength degradations and pinching effects. However, since the model cannot cover every practical range of civil structural systems, a modified Bouc-Wen-Baber-Noori (m-BWBN) model is newly proposed. Moreover, to properly describe

hysteretic loops of various structural systems, experimental hysteretic data of reinforced concrete (RC) columns compiled by the Pacific Earthquake Engineering Research (PEER) center (Berry et al., 2004) is used to determine the feasible domain of the model parameters. After developing the seismic demand database based on the m-BWBN model, a deep neural network model is constructed and trained to predict seismic responses of general hysteresis. Note that the trained deep neural network model in this chapter is referred to as the BW-DNN model to differentiate from the DNN and P-DNN models that were already developed in Chapters 2 and 3, respectively.

This chapter first presents the mathematical formulation of the m-BWBN model and a procedure to identify the upper and lower limits of the model parameters. Next, an architecture of the BW-DNN model is proposed by increasing the number of network units, especially layers in the CNN part, and including additional seismic intensity measures that represent the cumulative characteristics of earthquake ground motions. To take into account the sophisticated hysteretic characteristics as an input of the BW-DNN model, a framework to estimate predefined displacement history for the quasi-static cyclic analysis of a structural system is also suggested. After training the BW-DNN model, its performance is investigated in terms of accuracy.

4.2 Modified Bouc-Wen-Baber-Noori model

Bouc-Wen model, which was originally proposed by Bouc (1967) and later generalized by Wen (1976), is one of the most widely used differential equations that can extensively describe the dynamic behavior of a structural system. Although a wide class of Bouc-Wen models has been developed to consider various hysteretic

characteristics (Baber and Wen, 1981; Baber and Noori, 1985; Foliente et al., 1996; Song and Der Kiureghian, 2006), this study adopts the Bouc-Wen-Baber-Noori model which can describe stiffness and strength degradations and pinching phenomena (Baber and Noori, 1985). However, because the model has a limitation in controlling the yield force of the hysteresis which is an important feature that characterizes the strength of a system, some modifications are made. The revised version of the hysteretic model is denoted as the modified Bouc-Wen-Baber-Noori (m-BWBN) model.

Just as other Bouc-Wen class models, the m-BWBN model could produce hysteresis loops that may not satisfy the physical phenomenon of structural systems, especially when the constraints of the parameters are not explicitly defined. Thus, using the experimental hysteresis data of RC columns, the upper and lower limits of the m-BWBN model parameters need to be determined. Moreover, sensitivity analysis is performed to identify the parameters that are not significant in modeling the hysteretic characteristics in order to reduce the computational costs for constructing the database and training the BW-DNN model. Based on the results, several model parameters are fixed to their representative values, which, in turn, addresses the computational challenges.

4.2.1 Mathematical formulation

The m-BWBN model describes the restoring force f_s in Eq. (2.1) as a combination of elastic and hysteretic components, i.e.,

$$f_s = f(u, z) = \alpha k_0 u + (1 - \alpha)F_y z \qquad (4.1)$$

where α is the ratio of initial to post-yield stiffness, k_0 is the initial stiffness, and

F_y is the yield force of a system. The auxiliary variable z satisfies the following non-linear differential equation:

$$\dot{z} = \frac{h(z,\varepsilon)}{\eta(\varepsilon)}[\dot{u} - \{\beta|\dot{u}|z|z|^{n-1} + \gamma\dot{u}|z|^n\}v(\varepsilon)]\frac{k_0}{F_y} \tag{4.2}$$

where $h(z,\varepsilon)$ controls the pinching of the hysteresis, $\eta(\varepsilon)$ and $v(\varepsilon)$ are the parameter affecting the stiffness and strength degradations, respectively, and β, γ, and n are the dimensionless quantities controlling the shape of hysteretic loops. The deterioration parameters can be expressed as $\eta(\varepsilon) = 1 + \delta_\eta\varepsilon$ and $v(\varepsilon) = 1 + \delta_v\varepsilon$, where δ_η and δ_v stand for stiffness and strength degradation rates, respectively, and ε is the normalized cumulative hysteretic energy which is defined by the following rate equation

$$\dot{\varepsilon} = (1 - \alpha)\frac{k_0}{F_y}z\dot{u} \tag{4.3}$$

In addition, the pinching function is written as (Foliente et al., 1996, Hossain et al., 2013, Pelliciari et al., 2018)

$$h(z,\varepsilon) = 1 - \zeta_1(\varepsilon)\exp(-(z\cdot\text{sgn}(\dot{u}) - qz_u)^2/\zeta_2^2(\varepsilon)) \tag{4.4}$$

where sgn is the signum function, q is a constant value that controls the pinching level, and z_u is the ultimate value of $z(\varepsilon)$ given by

$$z_u = \left(\frac{1}{v(\varepsilon)(\beta + \gamma)}\right)^{1/n} \tag{4.5}$$

The parameters $\zeta_1(\varepsilon)$ and $\zeta_2(\varepsilon)$, which respectively controls the pinching severity and determines the pinching region to spread, are defined as follows:

$$\zeta_1(\varepsilon) = \zeta_0(1 - exp(-p\varepsilon)) \tag{4.6}$$

$$\zeta_2(\varepsilon) = (\psi + \delta_\psi \varepsilon)(\lambda + \zeta_1(\varepsilon)) \qquad (4.7)$$

where ζ_0 is the total slip, p controls the rate of the initial drop in the slope, ψ contributes to the amount of pinching behavior, δ_ψ measures the dispersion rate of the pinching phenomenon, and λ controls the variation of the parameter $\zeta_1(\varepsilon)$ and $\zeta_2(\varepsilon)$.

Compared to the original Bouc-Wen-Baber-Noori model that does not include the yield strength as a model parameter (Baber and Noori, 1985), the m-BWBN model can properly describe the strength of structural systems as well as other complicated hysteretic behaviors such as degradation and pinching effects. However, since the m-BWBN model includes the strength deterioration term in the formulation, the yield strength of a structural system estimated from pushover analysis or quasi-static cyclic analysis can be smaller than the specified level of F_y in Eq. (4.1). This phenomenon stands out especially when the system has positive deterioration values. Moreover, the level of deterioration is significantly affected by given displacement steps, which can be inferred from Eq. (4.3) and (4.5). An incremental algorithm that estimates the restoring forces given displacement steps is presented in Appendix A based on that of other Bouc-Wen class model (Haukaas and Der Kiureghian, 2003; Hossain et al., 2013; Ning et al., 2016).

4.2.2 Parameter constraints and bounds

Several research efforts reported in the literature have proved that the parameters of the Bouc-Wen model are functionally redundant, that is there are multiple sets of parameters which provide an identical response for a given excitation (Foliente, 1995;

Ma et al., 2004). Moreover, based on the combination of the parameter, the Bouc-Wen model could produce unrealistic or non-physical hysteresis. Therefore, the constraints and bounds of the parameters need to be provided to properly describe the hysteresis of structural systems. To this end, experimental hysteresis datasets of RC columns are introduced (Berry et al., 2004). Using the genetic algorithm, the parameters of the m-BWBN model are identified so that the corresponding hysteresis matches each experimental dataset. By combining all sets of identified parameters, the feasible parameter domain can be determined.

4.2.2.1 Genetic algorithm

Genetic algorithm is one of the heuristic optimization algorithms which is widely employed for the purpose of parameter estimation. The algorithm mimics the fact that the population that is akin to an optimal design is more likely to survive for each generation. In general, the generation consists of the following four steps: (1) population, (2) selection, (3) crossover, and (4) mutation. The optimal generations which produce the minimum value of the objective function are identified by repeating the 4 steps multiple time. Details of the genetic algorithm can be found in Goldberg and Holland (1988). In this study, the standard genetic algorithm is employed, but a small modification has been made in the mutation step such that the level of randomness decreases as the generation proceeds. This is to apply the simulated annealing effects (Bertsimas and Tsitsiklis, 1993) on the genetic algorithm in order to explore broadly at an early stage, which eventually pursues to avoid premature convergence or the local optima.

4.2.2.2 Determination of bounds on model parameters

Before performing the genetic algorithm to identify the feasible domain of the m-BWBN model parameters, the bounds of several parameters are first determined from the structural engineering point of view. Similar to Section 2.4.1, the ranges of the stiffness k_0, yield strength F_y, and post-yield stiffness ratio α are determined according to HAZUS-MH 2.1 (FEMA, 2012). In addition, β and γ, which are known to control the shape of hysteretic loops especially hardening and softening, are respectively set to 0.5 to meet the following condition: $\beta + \gamma = 1$. This is because, as shown in Eq. (4.5), the ultimate value of z needs to be bounded over $(-1, 1)$ when $\delta_v = 0$ in order to ensure the yield strength of the system is equivalent to the specified value (Constantinou and Adnane, 1987). In other words, if $\beta + \gamma$ is smaller than 1, z becomes bigger than 1, which leads that the restoring force $F(t)$ is bigger than the yield force F_y at the yield point. Moreover, it is possible to find various combination of β and γ satisfying the condition However, it is found that 0.5 for each parameter gives the most reliable results. Lastly, due to the numerical convergence issues when estimating hysteretic parameter z from the differential equation, the upper and lower limits are introduced to some of the m-BWBN parameters as follows: $1 \leq n \leq 5$, $0 \leq \zeta_0 \leq 1$, $q \geq 0.01$, $\psi \geq 0.1$, and $\lambda \geq 0.01$. Note that these bounds are obtained from a large number of numerical investigations.

The most optimal way to determine the parameter bounds of the m-BWBN model for civil/architectural structural systems is to examine various combinations of the model parameters based on the hysteresis of different kinds of structural systems which are obtained by either simulations or experiments. However, since it

is difficult to obtain such a large volume of hysteresis loops for various structural systems, the experimental data of the RC columns under quasi-static cyclic analysis are introduced as a replacement (Berry et al., 2004). The experimental dataset is comprised of displacement and restoring force vectors for two different shapes of cross-section, i.e., rectangular and circular, whose number of datasets is 163 and 251, respectively. Note that the lengths and step sizes of hysteresis vary along with each experiment.

In performing the genetic algorithm, for each population, the restoring forces are obtained using the displacement vector from the dataset. By setting the mean squared error (MSE) between the experimental and the estimated restoring forces as the objective function, it is possible to identify the optimal parameters for each dataset. Several representations of the hysteresis from the m-BWBN model and the corresponding experimental datasets are shown in Figure 4.1. In the figure, the black solid line represents the experimental hysteresis, while the red dashed line stands for the hysteresis obtained by using the m-BWBN model with the parameters identified from the genetic algorithm. It is found that the m-BWBN has the capability to properly illustrate the hysteresis of RC columns. However, it should be noted that although the estimated hysteresis enables a proper illustration of the experimental hysteresis as shown in Figure 4.1, it is difficult to guarantee that the estimated parameters produce the equivalent restoring force when the history of the input displacement is changed. This is due to the inherent functional redundancy of the Bouc-Wen model as explained earlier.

To identify the upper and lower limits of the parameters, normalized frequency diagrams of the estimated deterioration and pinching parameters are presented in Figure 4.2. The upper limit of the input parameters is defined as the 95th percentile

of each normalized frequency diagram. This is due to the following two reasons. First, in performing the genetic algorithm, some parameters are overfitted to a specific dataset, which provides an extreme value. However, such values are only applicable to special hysteresis that is not suitable for general civil structural systems. Second, it is found that the hysteretic loop does not significantly transform its shape when the value that exceeds the 95th percentile of the distribution is assigned as the m-BWBN parameter. In other words, minor details of the hysteresis are ignored in order to reduce the computational cost for generating the seismic demand database.

4.2.3 Sensitivity analysis of model parameters

Sensitivity analysis aims at evaluating the degree of impact on model output for each input parameter, i.e., the level of change in the output of the model by varying input parameters. Although various methods have been developed to this end (Kim and Song, 2018; Sobol, 2001), an entropy-based sensitivity analysis method (Liu et al., 2006) is adopted in this study to consider probability distributions of each parameter. The method compares the Kullback-Leibler (KL) divergence of two probability density functions (PDFs), obtained before and after a random variable is fixed to a certain representative value. The mathematical expression of the KL divergence $D_{KL}(p_1|p_0)$ is written by

$$
\begin{aligned}
&D_{KL}(p_1|p_0) \\
&= \int_{-\infty}^{\infty} p_1\big(y(x_1, \dots, \bar{x}_i, \dots, x_n)\big) \log \frac{p_1\big(y(x_1, \dots, \bar{x}_i, \dots, x_n)\big)}{p_0\big(y(x_1, \dots, x_i, \dots, x_n)\big)} \, dy
\end{aligned} \tag{4.8}
$$

where p_0 stands for the PDF of y given input random variables $x_1, \dots, x_n$, and p_1 is the PDF of y when x_i is fixed to a certain value $\bar{x}_i$. In short, the KL divergence superficially represents the distance between two different PDFs. The

larger the distance between p_0 and p_1, the higher $D_{KL}(p_1|p_0)$ is calculated. If the variability of random variable x_i does not significantly affect the variability of y, the corresponding $D_{KL}(p_1|p_0)$ would be small. However, it should be noted that the KL divergence does not exactly stand for the distance between two PDFs, in that $D_{KL}(p_1|p_0)$ and $D_{KL}(p_0|p_1)$ are not equivalent.

The sensitivity analysis is performed for 9 parameters: n, δ_ν, δ_η, ζ_0, p, q, ψ, δ_ψ, and λ. The procedure consists of the following 4 steps: (1) randomly generate a set of input parameters by assuming that each parameter follows uniform distribution whose upper and lower limits are defined in the previous subsection, (2) perform a quasi-static cyclic analysis for each set of generated input parameters using the predefined displacement history (i.e., restoring forces are obtained for the given displacement history), (3) repeat the first and second steps while fixing a certain input parameter to a representative value, and (4) estimate the KL divergence between the fixed and unfixed distributions of restoring forces for each displacement step and average them. The displacement history is defined as four cycles in each of which the peak displacement is gradually increased, and scaled by the yield displacement of a target structural system to properly investigate the nonlinear range, i.e., detailed pinching and deterioration phenomenon (FEMA P695, 2009; FEMA P795, 2011). In the sensitivity analysis, the uniform distribution is selected as the prior distribution which means that no particular information is provided to each parameter, i.e., uninformative. If the distributions of the model parameters are determined from structural systems having various characteristics, one can use the identified PDFs as the prior distribution in the sensitivity analysis and obtain results that are slightly different from those in this study.

As a target structural system, an SDOF system is considered: period T is 0.5 s, normalized yield force F_y is 0.1g, and post-yield stiffness ratio α is 0.01. A total of 500 samples are randomly generated with the Latin hypercube sampling to efficiently cover the domains of the input parameters. The lowest value of the parameter domain is selected as the represented value, except n in which the median values are chosen. This is because, except n, the parameters of the lowest value do not take into account strength and stiffness degradations nor pinching effects in the system. In other words, it is possible to quantify the level of these hysteresis phenomena with respect to parameters when the lowest values are employed for a representative value. The results of the sensitivity analysis are enlisted in Table 4.1 with their descriptions and the feasible domain defined in the previous subsection. Note that the value inside the parenthesis in the 'sensitivity rank' column of the table indicates the calculated D_{KL} which is normalized by their maximum value.

It is observed from Table 4.1 that δ_η is the most sensitive parameter which is followed by δ_v, p, and ζ_0. On the other hand, q, ψ, and δ_ψ are the relatively insensitive parameters whose contributions are below 20% compared to the effect of δ_η. Moreover, although, in the table, n which controls the level of sharpness from linear to nonlinear behavior has ranked as 5th from the quasi-static cyclic analysis, it is found that n does not critically affect the seismic responses during dynamic analysis. Therefore, to address some computational challenges when developing the seismic demand database, three pinching related parameters and n are fixed to their median value, i.e., $n = 3$, $q = 0.027$, $\psi = 0.117$, and $\delta_\psi = 0.003$. However, it should be noted that since no correlation between parameters is considered during

the sensitivity analysis, the results of the investigation may be changed along with the additional information.

To check the level of compromising the details of hysteretic characteristics, the genetic algorithm is reperformed by fixing 4 parameters to their representative values and applying the identified parameter bounds, and then the mean squared errors (MSE) are compared with the one that does not consider any constraints. The MSE between estimated and experimental hysteretic behaviors with and without considering the constraint is 0.0012 and 0.0011, respectively, and the relative error between the two MSEs is calculated as 7.96%. Based on these findings, it is concluded that there exists a trade-off between about 8% of hysteretic details and the computational efforts.

4.3 Training methodologies

The database is first constructed for various structural systems represented by the m-BWBN model with a large number of ground motions. Runge-Kutta integration method of order two and three is employed for the time history analysis. Next, a newly proposed deep neural network architecture is described by modifying the previously developed DNN model in Chapter 2 to properly investigate the onerous relationship between the input and output. Moreover, the data retrieval process from the databases is presented with details of hyperparameters for training. Note that, the database and the BW-DNN model were constructed using SQLite3 and TensorFlow (Abadi et al., 2016), respectively through the Compute Canada's CPU and GPU clusters.

4.3.1 Database

In order to construct the seismic demand database, parameters of the m-BWBN model are discretized based on the estimated feasible domain as shown in Table 4.1. First, the basic hysteretic parameters such as structural period, normalized yield force, and post-yield stiffness ratio are, respectively, discretized as 40, 10, and 4 steps, instead of 90, 30, and 10 steps in Section 2.4.1. This is due to the computational cost for a large amount of nonlinear time history analysis. Next, the two deterioration parameters are respectively discretized as 3 steps. Since the pinching effects are described by interwinding 6 different parameters, 9 possible combinations of pinching parameters are employed based on an extensive exploration of different hysteretic behaviors. The set of pinching parameters and their shapes are shown in Figure 4.3 as the red dashed line with the hysteresis of no pinching effects which is illustrated by the black solid line in each subplot. In words, the total number of structural systems considered in this database is $129{,}600 \ (= 40 \times 10 \times 4 \times 3 \times 3 \times 9)$.

Among various methods available for solving the system of nonlinear equations of motion, Bogacki–Shampine method, which is often termed as the Runge-Kutta method of order two and three, is selected for this purpose (Bogacki and Shampine, 1999). Since 1,499 ground motions from the NGA-West database (Power et al., 2006) is employed, a required number of time history analysis is $129{,}600 \times 1{,}499 = 194{,}270{,}400$. Although it is widely known that the accuracy of the dynamic analysis by the Runge–Kutta–Fehlberg method, i.e., Runge-Kutta method of order four and five is better than that of the Bogacki–Shampine method, this study adopts the relatively lower-order method to reduce the computational challenge (Bogacki and Shampine, 1989; Cheney and Kincaid, 1999). Moreover, it is found that the level of

accuracy of the two integration schemes is equivalent to each other in this problem.

4.3.2 Characterization of generalized hysteresis

Since a substantial amount of information is contained in the hysteresis of the m-BWBN model than the idealized one in Chapters 2 and 3, refined displacement and force steps are required for the input of the BW-DNN model. The hysteresis loop obtained from the quasi-static cyclic analysis should satisfy the following two requirements: (1) all hysteretic characteristics are incorporated in the hysteresis loop to distinguish one another, and (2) hysteresis loops should produce a reasonable level of performance in terms of accuracy and applicability when they are used as the input of the BW-DNN model. Inspired by literature (FEMA 440, 2005; FEMA P695, 2009; FEMA P795, 2011), a guideline to estimate predefined displacement history for the quasi-static cyclic analysis is proposed as the following 3 steps.

Step 1: A pushover analysis is performed by subjecting a structure to a monotonically increasing pattern of lateral forces.

Step 2: By superimposing a linear line whose slope is equal to 90% of the initial stiffness of the pushover curve, an intersection point is obtained. The intersection point is considered as the reference point whose role is the yield point of the structure.

Step 3: Multiply the yield displacement estimated from the reference point to the basic displacement history. The basic displacement history is defined by 241 steps with 5 cycles whose peak displacement is gradually increased as the cycle proceeds. In particular, the step size of the basic displacement history is 0.25 and its amplitude of each cycle is 0.5, 1.0, 1.5, 2.0, and 2.5. The

graphical representation of the basic displacement history is presented in Figure 4.4.

The predictions from a neural network are not reliable for regions of the variable space in which no train dataset is available, i.e., no extrapolation. When the normalized yield strength of a structural system exceeds 1.5g is employed in the developed framework, the yield displacement is estimated by dividing the stiffness to 1.5g instead of the actual yield strength of the structural system. Note that 1.5g is the maximum yield strength handled in this study. Moreover, the yield displacement and the corresponding displacement history for the linear structural systems are estimated in the same manner.

Using the estimated displacement history, the restoring forces are obtained from the incremental response equation described in Appendix A. As shown in Figure 4.3, various hysteretic behaviors are readily distinguished from each other. Moreover, it was found from a large number of numerical examinations that good prediction accuracy of the BW-DNN model is observed when such hysteresis loops are used as the input compared to other hysteresis obtained from different displacement histories.

4.3.3 Architecture details

The architecture of the BW-DNN model is illustrated in Figure 4.5 by improving the network model developed in Section 2.3. Compared to the previous version illustrated in Figure 2.4, five modifications have been made to properly predict the seismic demands of generalized hysteresis which are indicated in bold and underline in Figure 4.5. First, compared with the DNN and P-DNN models, the number of units in the BW-DNN model is increased, especially the CNN part that extracts

important features of structural systems. Since more sophisticated hysteretic characteristics are contained in the m-BWBN hysteresis than the idealized one, a sufficient number of units are required. From the extensive exploration of various deep neural network architectures, it was found that the number of units shown in Figure 4.5 is optimal for predicting the seismic responses. In other words, a smaller number of units than the proposed one gives less accuracy, whereas, it may hamper the convergence of the BW-DNN model in the opposite case. Second, since the normalized cumulative hysteretic energy ε which controls the level of deterioration and pinching effects is highly related to the cumulative features of ground motions, three new seismic intensity measures are added as the input of the BW-DNN model: strong motion duration t_{strg}, zero-crossing of ground acceleration during strong motion duration f_{strg}, and time integral of squared ground motion acceleration during strong motion duration in the unit of g CSA_{strg} (Kim et al., 2021). Note that the strong motion duration is defined as the time interval between 5% and 95% of the time integral of squared ground motion acceleration. Third, the number of units in layers is decreased and L2 regularizers are introduced to address the overfitting issue, especially where the ground motion features are extracted. Due to insufficient numbers of ground motions, it was found that the deep neural network model tends to overfit the train dataset. Even though the regularization technique can mitigate the overfitting problem, an increase of the number of ground motions is the only and most efficient way to extract the inherent patterns between input and output. Fourth, the features of earthquake information are connected to those of ground motions and frequency contents after processing by several layers of neurons to represent a more realistic structural excitation mechanism subjected to an earthquake ground motion.

An extensive numerical investigation demonstrates that this adjustment does not critically affect the performance of the BW-DNN model, but it is expected that the accuracy might be improved, especially when the seismic demands that are highly related to the source information such as hysteretic energy are selected as the output. Finally, the rectified linear operator (ReLU; Nair and Hinton, 2010) is changed to a Leaky version of a Rectified Linear Unit (LeakyReLU), which is known to produce better performance than using the ReLU in the neural network.

4.3.4 Training methodologies

The BW-DNN model is trained using the seismic demand databases of idealized and generalized hysteresis to predict peak displacement during seismic excitation. After pretraining the BW-DNN model using only linear and bilinear perfectly plastic hysteresis, the model is trained for the entire structural systems. Since the number of datasets from both databases (idealized hysteresis models and m-BWBN model) is 202,499,910 (= 8,229,510 + 194,270,400), it is difficult to use every dataset for training at once due to memory issues. To address this issue, 45 ground motions are randomly selected with the corresponding 129,600 m-BWBN hysteresis, 90 hysteresis of HM1, and sparsely selected 7,560 hysteresis of HM2 and HM3 respectively for every three epochs, i.e., a total of 144,810 (= 129,600 + 90 + 7,560 + 7,560) for each ground motion. When selecting ground motions for each epoch, they are selected in terms of probability density functions (PDF) of PGA to cover a broad range of ground motion intensity. To check whether the BW-DNN model is overfitted or not, 80% of ground motions are used for training and 20% are used for testing. Adam optimization algorithm is adopted to minimize the mean squared error (MSE) between predicted and ground truth seismic responses.

4.4 Performance of deep neural network

To check the performance of the trained BW-DNN model, mean squared error (MSE) and mean absolute error (MAE) are estimated for train and test datasets as shown in Table 4.2 after 30 epochs of training. Because it was found that 30 epochs are not enough for a given dataset and network model, further investigation is required to obtain better performance of the BW-DNN model. Because the L2 regularization technique is employed and small amounts of noise are added to the ground motion information on purpose, the MSE and MAE of the train and test datasets are equivalent to each other compared to those of DNN and P-DNN models shown in Table 2.1 and Table 3.1, respectively. However, the level of error is slightly higher than the previous models because of an insufficient number of epochs as previously mentioned. Thus, it is expected that better performance would be achieved if more epochs of training are performed. Moreover, to examine the trainability and sensitivity of the BW-DNN model to other structural responses, training results of the BW-DNN model need to be provided by setting the output of the model as peak acceleration and velocity.

4.5 Conclusions

To properly predict the seismic responses of structural systems showing more sophisticated hysteretic behaviors, the BW-DNN model was developed. First, the modified Bouc-Wen-Baber-Noori (m-BWBN) model was proposed to describe the yield strength of structural systems as well as other sophisticated hysteretic behaviors such as degradation and pinching effects. A comprehensive numerical investigation

was performed using hysteresis loops of experimental data of RC column to estimate the feasible domains of the m-BWBN model parameters. The Runge-Kutta integration method was adopted to construct a seismic demand database of the m-BWBN model. Using the database, the BW-DNN model was trained to predict the seismic response of structural systems having sophisticated hysteretic behaviors. The performance of the trained BW-DNN model was tested in terms of accuracy by comparing the mean squared error between the predicted peak displacement and the one from dynamic analysis. Because the BW-DNN model is versatile to predict various seismic responses of structural systems, future investigation is needed to develop DNN models that can predict peak velocity and acceleration.

Table 4.1 Summary and description of parameters for m-BWBN model

No.	Parameter	Description	Bounds	Sensitivity rank
1	α	Post-yield stiffness ratio	$0 \leq \alpha \leq 0.5$	-
2	k_0	Stiffness	$0.05\ s \leq T \leq 10\ s$	-
3	F_y	Yield force	$0.05g \leq F_y \leq 1.5g$	-
4	β	Basic hysteretic shape control	$\beta = 0.5$	-
5	γ	Basic hysteretic shape control	$\gamma = 0.5$	-
6	n	Sharpness of yield	$1 \leq n \leq 5$	5 (0.1443)
7	δ_v	Strength degradation rate	$0 \leq \delta_v \leq 0.36$	2 (0.2552)
8	δ_η	Stiffness degradation rate	$0 \leq \delta_\eta \leq 0.39$	1 (1.0)
9	ζ_0	Measure of total slip	$0 \leq \zeta_0 < 1$	4 (0.2067)
10	p	Pinching slope	$0 \leq p \leq 1.38$	3 (0.2374)
11	q	Pinching initiation	$0.01 < q \leq 0.43$	7 (0.0732)
12	ψ	Pinching magnitude	$0.1 \leq \psi \leq 0.85$	8 (0.0595)
13	δ_ψ	Pinching rate	$0 \leq \delta_\psi \leq 0.09$	9 (0.0515)
14	λ	Pinching severity	$0.01 \leq \lambda \leq 0.8$	6 (0.0741)

Table 4.2 Comparison of MSE between train and test datasets

Ground motion set	Mean squared error (MSE)	Mean absolute error (MAE)
Train	0.0489	0.1694
Test	0.0586	0.1821
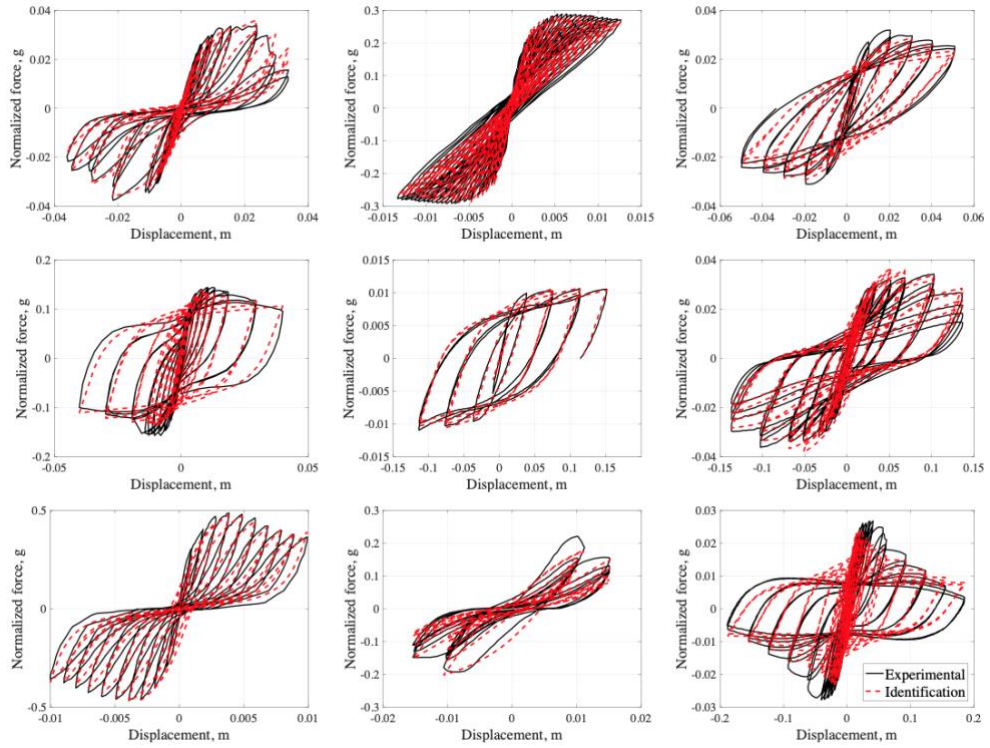
Figure 4.1 Comparison of the predicted hysteresis loop using genetic algorithm with experimental data of RC columns
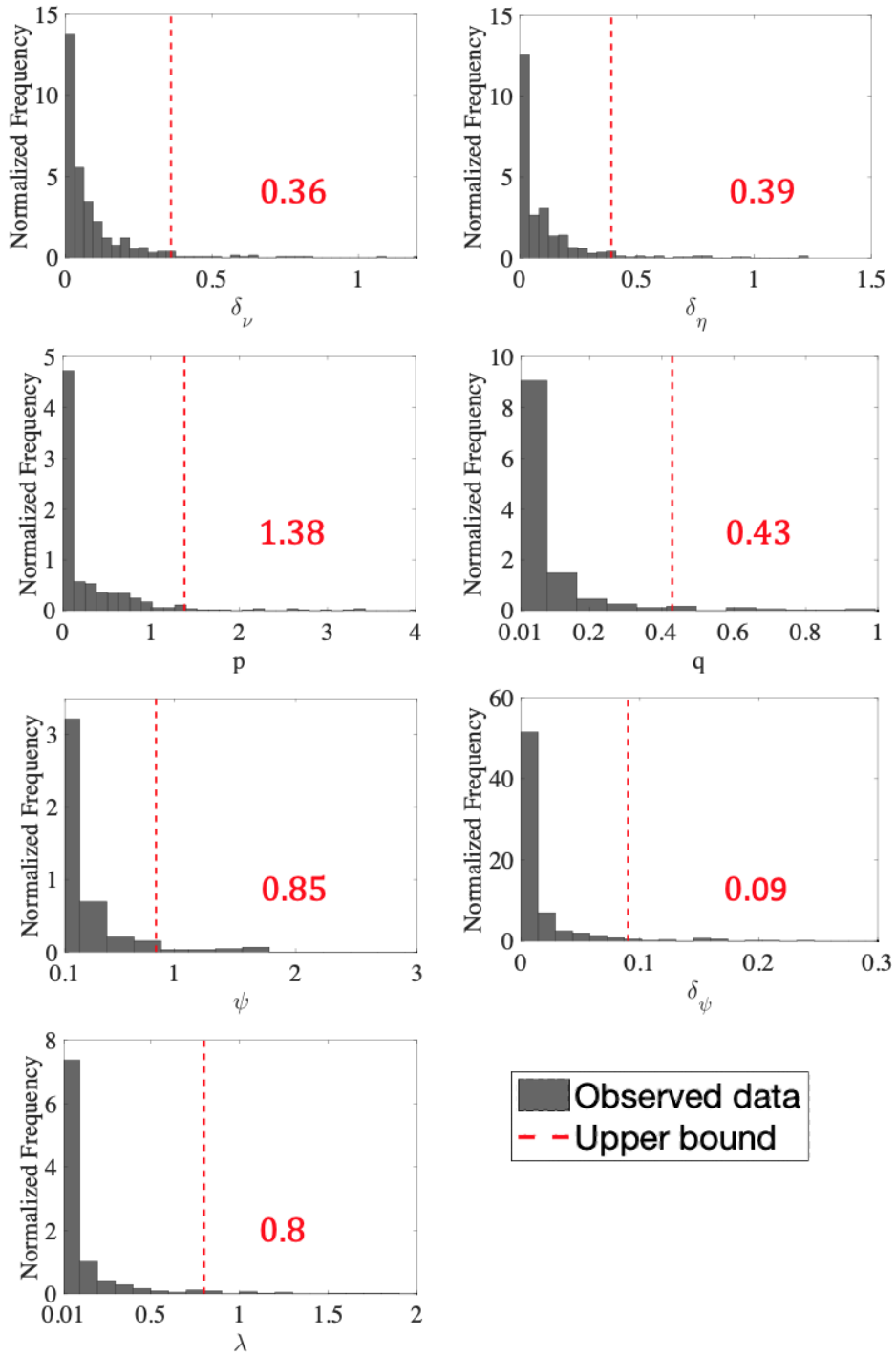
Figure 4.2 Normalized frequency diagrams of m-BWBN model parameters and
identified upper bounds
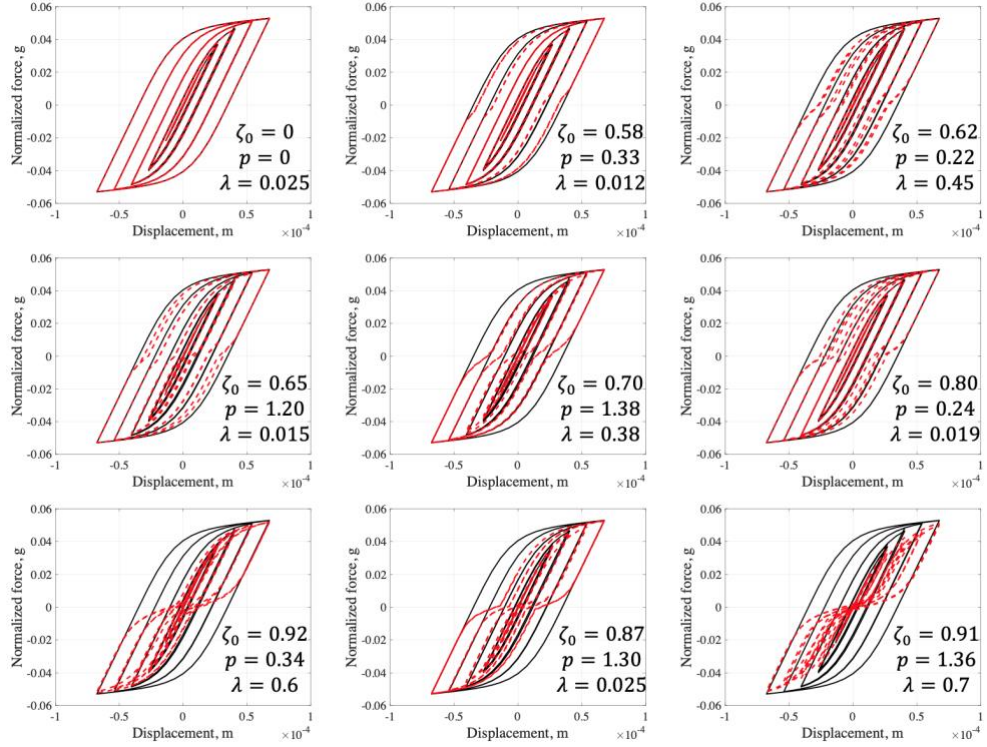
Figure 4.3 Nine different pinching effects considered in this study and their parameters for m-BWBN model
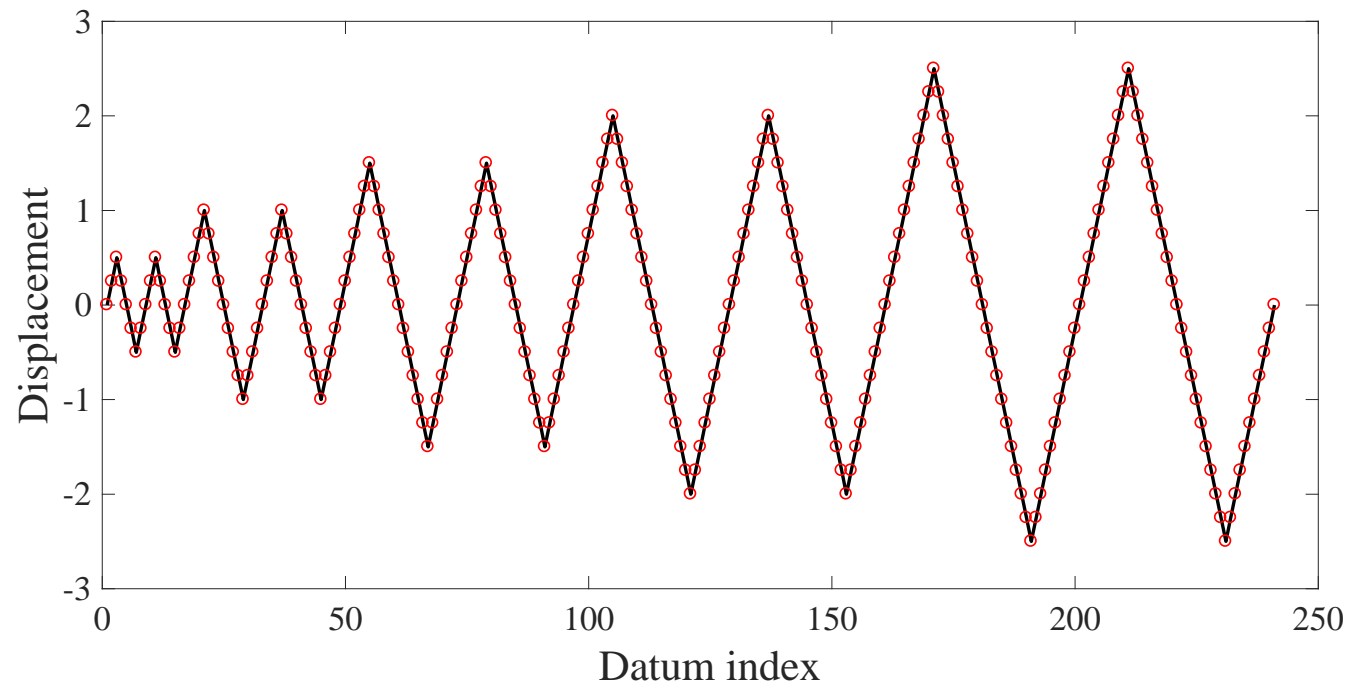
Figure 4.4 Illustration of basic displacement history for generating hysteretic loop
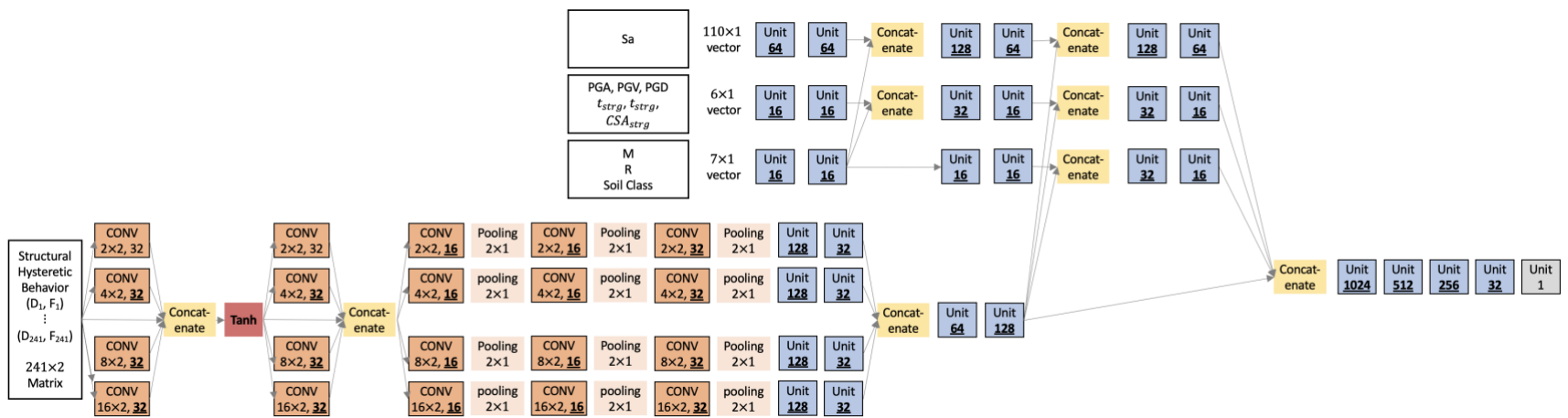
Figure 4.5 Detailed diagram of the BW-DNN model architecture

Chapter 5. Application of Trained Deep Neural Network Models

5.1 Introduction

In Chapters 2 – 4, the performance of three DNN models was comprehensively demonstrated through various numerical examples. To demonstrate the applicability of the developed methods, three application examples having different scales of target structural systems are provided in this chapter. First, seismic responses of structural elements, i.e., reinforced concrete (RC) columns are predicted using the DNN and BW-DNN models. In this regard, different periods of RC columns are introduced by changing their mass. Second, using the capabilities of the P-DNN model to predict the mean and variance of seismic responses, a new method that estimates a structural fragility is proposed. Owing to the merits of the P-DNN model, it is possible to estimate structural fragility without assuming the relationship between seismic demand and capacity. Third, risk assessment of an urban community subjected to seismic excitation is carried out using the DNN and P-DNN models. Given that an urban community is a complex system characterized by a large number of components and their correlation, the DNN and P-DNN models can efficiently predict the responses of a set of structures in the target region. Moreover, because the P-DNN model can quantify the uncertainties of the nonlinear responses of structural systems given the seismic intensity measure, such variabilities can be incorporated during the loss assessment process and the results are compared with the one using the DNN model. Apart from the applications, a user-friendly web-service (Earthquake Responses using Deep learning and Database; ERD[2]) is

developed to provide an interactive visualization of the proposed DNN predictions. The merits and role of the web platform in the earthquake engineering community are comprehensively described.

The chapter is organized as follows. The application of the trained BW-DNN model as a seismic response prediction tool for structural elements is first presented. Next, a deep learning-based fragility assessment framework is proposed and its application to reinforced concrete (RC) frame structure is illustrated. Third, a deep learning-based regional seismic loss assessment framework is developed and applied to an urban community which mimics the downtown Vancouver area. Finally, the ERD^2 is introduced with a detailed explanation of how other researchers can employ the outcomes and findings of this study to their own problems.

5.2 Application to structural element: Response prediction of RC columns

The BW-DNN model can predict the seismic responses of general hysteresis. To demonstrate its effectiveness and applicability, seismic responses of RC columns are predicted using the BW-DNN model. To this end, a deep learning-based seismic prediction framework is introduced. Using the proposed procedure, three RC columns having different first mode periods are introduced with 135 ground motions retrieved from the NGA-West database based on the following criteria: M > 6, 15 < R < 35, and soil classes A, B, C, and D. Note that the soil class is defined by Building Seismic Safety Council (BSSC, 2003).

5.2.1 Deep learning-based seismic response prediction

In order to use the BW-DNN model as a prediction tool for seismic responses of structural systems, two different types of information are required: (1) structural information, and (2) earthquake ground motion information. First, the structural information is represented by the hysteresis loop, i.e., force and displacement steps (241×2), which is extracted from the generalized hysteretic generation procedure proposed in Section 4.3.2. Since it was found that the prediction performance of the BW-DNN model is highly affected by the shape of hysteresis, the displacement history that is used to estimate the structural hysteresis is of importance. Regarding earthquake ground motion information, three different features are employed: Earthquake features (M, R, Soil class), ground motion features (PGA, PGV, PGD, t_{strg}, f_{strg}, CSA_{strg}), and frequency features (spectral acceleration). It should be noted that the natural logarithm needs to be applied when the features are employed as the input of the BW-DNN model except the soil class which is represented as a categorical data formulation. When adopting the DNN model in seismic design framework, a design spectrum of the site of interest is employed for the frequency features instead of the spectral acceleration of ground motions, and other intensity measure values are approximately obtained from a ground motion prediction equation or estimated based on random vibration theories. Since the BW-DNN model predicts the natural logarithm of the peak responses, the exponential function is required to convert the predicted responses to the original scale.

5.2.2 Application to RC columns

Concrete and reinforcement are one of the widely used structural materials in civil

and architectural engineering. Due to the inherent material properties and their interaction, stiffness/strength degradations and pinching effects are usually observed in the hysteresis of RC elements.

RC column is one of the most important structural elements in RC structural systems. A cross-sectional area of the RC column and its material properties used in this study are shown in Figure 5.1. By varying the mass of the structural element, three different first mode periods of RC column are obtained and their hysteresis are shown in Figure 5.2. Since the mass is the only parameter that is changed along with the structural elements, the yield strength and period are varied, whereas the shape of hysteresis is equivalent to each other. As shown in Figure 5.2, one can readily identify the stiffness/strength degradations and pinching effects in the hysteresis. Using 135 ground motions from the NGA-West database (Power et al., 2006), time history analyses are carried out. Figure 5.3 shows the results as a form of scatter plot between spectral acceleration and peak displacement. The spectral acceleration value at which the response variability occurs is equivalent to the normalized yield strength of structural elements obtained from the quasi-static cyclic analysis in Figure 5.2.

To investigate the prediction performance of the BW-DNN model, the mean squared error (MSE) and relative error (RE) of the natural logarithm of the predicted responses and the ones from the time history analysis are estimated and presented in Table 5.1. The MSE and RE obtained using the DNN model and the coefficients method are also presented in the same table. Note that when predicting the structural responses using the DNN model, the equivalent HM3 model is introduced. As shown in the table, relative errors of predicted responses from the DNN model and the coefficient method increase as more pinching and strength/stiffness degradations effects are incorporated in the hysteresis, while the BW-DNN model shows steady

prediction error regardless of the hysteretic characteristics. This finding indicates that the DNN model and the coefficient method have a limitation in predicting seismic responses of generalized hysteresis, while the BW-DNN model can cover a broad range of structural systems having sophisticated hysteretic characteristics. However, it is found that the prediction error by using the BW-DNN model is relatively high even though the system behaves in a linear range, i.e., RC column whose period is 0.115 sec. This is because the m-BWBN hysteresis models are intentionally more selected than the idealized hysteresis that behaves in the linear range when training the BW-DNN model. To increase the prediction accuracy of the linear seismic responses of the BW-DNN model, a new training methodology is required to address the issues.

5.3 Application to structural systems: Seismic fragility estimation

In order to consider the record-to-record variability of ground motions and the uncertainties of structural behaviors under excitations, modern seismic design codes adopt the probabilistic assessment framework which interwinds the uncertainties in the ground motions and the corresponding structural responses by using the total probability theorem (Deierlein et al., 2003). Such a framework is often referred to as performance-based earthquake engineering (PBEE) framework. In PBEE, the seismic performance of a structural system is often described by the conditional failure probability given earthquake intensity measure, which is termed as *structural fragility*. Although various methods have been developed to estimate structural fragility (Cornell et al., 2002; Deniz et al., 2018; Jalayer et al., 2014; Kim et al., Accepted; Shome et al., 1998; Vamvatsikos and Cornell, 2002), a large

computational cost is usually required. However, when using the conditional mean and variance of the structural responses from the P-DNN model, structural fragility can be directly computed without any assumed relationship between seismic capacity and demand nor a large number of dynamic analysis. Thus, this section provides the procedure to estimate structural fragility using the P-DNN model which is followed by the validation of the procedure with a 3-story RC structural system.

5.3.1 Deep learning-based structural fragility assessment

Since the P-DNN model was developed based on the idealized hysteretic behaviors of an SDOF system, hysteresis loops of the equivalent idealized SDOF systems need to be employed as the input of the P-DNN model. To this end, the hysteresis of an MDOF system is transformed into an equivalent SDOF system. After obtaining the usable hysteresis, a deep learning-based fragility estimation framework is proposed as follows:

Step 1: Identify the relationship between the base shear force and the rooftop displacement (i.e., hysteretic curve) through a pushover analysis (Fajfar, 2000).

Step 2: Determine the equivalent SDOF system using the hysteretic curve from Step 1 (ASCE, 2000; FEMA 440, 2005).

Step 3: Using the P-DNN model with inputs representing the hysteretic curve of the equivalent SDOF system and the features of the selected ground motions, estimate the conditional mean and variance of the displacement given each ground motion.

Step 4: Using a modal participation factor Γ, transform the estimated response of

the equivalent SDOF system to the top displacement of the original MDOF system (Fajfar, 2000).

Step 5: Calculate the failure probability, i.e., probability to exceed a displacement threshold representing the limit-state of interest for each ground motion, based on the assumption that the natural logarithm of the displacement follows a Gaussian distribution with the estimated mean and variance (i.e., the displacement in the original scale follows the lognormal distribution). As a result, the pair of intensity measure and the corresponding failure probability is obtained for each ground motion.

Step 6: A cumulative distribution function (CDF) model, e.g., lognormal CDF is fitted to the pairs of intensity measure and corresponding failure probability obtained from Step 5 to construct a smooth fragility curve.

5.3.2 Application to 3-story RC frame structure

The proposed framework is demonstrated using a 3-story RC structural system whose configuration is shown in Figure 5.4. Kwon and Elnashai (2006) defined the limit states 'Serviceability', 'Damage Control', and 'Collapse Prevention' in terms of the maximum inter-story drift ratio – 0.57%, 1.2%, and 2.3% respectively. By assuming that the roof drift ratio is proportional to 70% of the inter-story drift ratio, the failure criteria are converted in terms of the drift as 0.043 m, 0.090 m, and 0.173 m. This assumption is reasonable, in that the yield displacement estimated from the idealized SDOF system, 0.052 m (estimated by 0.042 m/1.24 as discussed below) lies between 'Serviceability' and 'Damage Control' limit-states.

The quasi-static cyclic analysis of the structural system is performed with the 1st mode force distribution. Figure 5.5 shows the relationship between the

normalized 1st mode base shear and equivalent 1st mode roof displacement of the RC model (black solid line) and that of the idealized SDOF system (red dashed line) estimated from the iterative procedure (FEMA 440, 2005). A bilinear model having a stiffness degradation mechanism is selected to describe the pinching effects and the strength degradation of the structure. Note that if the P-DNN model is trained with the database developed in Section 4, the hysteretic behaviors obtained from the quasi-static cyclic analysis can be directly used as the input of the deep learning model. The ground motions whose magnitude ranges from 6.5 to 8.0 and epicenter distance from 20 km to 30 km are selected for the fragility analysis (total 58 ground motions). Using the P-DNN model and the procedure introduced in the previous subsection, fragility curves are obtained for each limit-state as shown in Figure 5.6. Moreover, to compare the estimated fragility functions with those obtained from an existing method, the procedure proposed by Cornell et al. (2002) is employed, whose details are described in Appendix B. The fragilities are computed using the structural responses of the idealized SDOF system estimated from the nonlinear time history analysis, and depicted as the dotted lines in the same plot.

As shown in Figure 5.6, the variability of both types of fragility curves, visualized by their flatness, increases as more severe failure criteria are considered. It is noted that the variability of the curves is also affected by the method used for fragility estimation. The fragility curves obtained from the P-DNN model are stiffer than the conventional method for serviceability limit-state, while less stiff for damage control and collapse prevention limit-states. Since the conventional method relies on a simple regression model with a constant standard deviation (i.e., homoscedasticity) for both the linear and nonlinear range of the structural system when identifying the relationship between the spectral acceleration and drift, the

conventional method shows significant variabilities in the linear elastic range. For the same reason, the variability in the nonlinear range is underestimated due to the influence of the data in the linear elastic range on the estimate of the constant variance. In reality, however, the fragility curve of the serviceability limit state should be a vertical line at the performance point 0.076g (= stiffness×performance limit/Γ = $2.183 \times 0.043/1.2495$) in Figure 5.6, which matches the fragility curve developed by the proposed method.

This numerical example confirms that it is convenient and intuitive to develop a fragility function using the P-DNN model. Moreover, since the deep learning-based fragility does not use an assumption about the relationship between intensity measure and damage measure, its actual relationship can be properly incorporated into the fragility for each limit-state. This framework can be further extended to develop a ground motion selection algorithm for the purpose of assessing the structural system because one can readily identify the ground motions that significantly affect the variability of the fragility. For instance, Figure 5.6 implies that the ground motions with spectral acceleration values between 0.2g and 0.6g are needed to assess the structural system for collapse prevention limit-state.

5.4 Application to urban community: Regional seismic loss assessment

Community-level seismic risk assessments are indispensable to assure resilient urban communities against earthquake hazards. To this end, a loss curve is usually employed which is defined as the exceedance probability of certain performance or monetary loss (Bommer et al., 2002; Goda and Hong, 2008; Shiraki et al., 2007). Note that the mathematical definition of the loss curve and details of its estimation are described in Appendix C. The loss curve is estimated based on the seismic responses, e.g., usually peak displacement, of a set of structural systems in the target urban area. In estimating the seismic responses, some researchers have adopted structural fragility models (Bai et al., 2009; Diaz Gomez et al., 2015; Miller and Baker, 2015), whereas others have employed a nonlinear static procedure (NSP) method (ASCE 41-13, 2013; FEMA 440, 2005; Nassar and Krawinkler, 1991) based on assumed probabilistic distributions of structural parameters (Goda and Hong, 2008). Although both approaches can incorporate uncertainties in the structural behavior into predictions, some limitations still exist. The former method often demands a huge computational cost or effort for developing the fragility curve for each type of structural system under the seismic hazard at the site of interest. The latter approach is computationally efficient but may suffer from a large estimation error as confirmed in Chapters 2 and 3. This may result in significant under- or overestimation of the regional loss.

In order to overcome this challenge, the DNN and P-DNN models are introduced for accurate and efficient prediction of seismic responses of individual structures. It was confirmed that the accuracy of the DNN and P-DNN models is superior to that of existing nonlinear static procedure methods. Furthermore, the

uncertainties of the structural responses given intensity measure which stem from its nonlinear behavior can be quantified by the P-DNN model. This evaluation is important, in that the uncertainty in the damage of the structural system is propagated to that of the regional monetary loss of the urban community. Thus, this section first proposes a regional seismic loss assessment framework using the P-DNN model. Next, the proposed frameworks are demonstrated with the V-city example to demonstrate its applicability and effectiveness (Kim et al., 2020).

5.4.1 Deep learning-based regional loss assessment framework

Inspired by the simulation-based seismic loss assessment by Choi (2017), the following regional seismic loss assessment procedure using the P-DNN model is proposed. The buildings distributed in a target region are represented as idealized SDOF systems, because the regional loss assessment usually focuses on a group of structures. Furthermore, detailed computational simulations would demand exceedingly large efforts with regard to structural modeling and computational simulations. Regarding the earthquake ground motion, it was assumed that the earthquake rupture plane is on the fault surface, and a rupture has a maximum area up to that of the active fault where it is located. It should be noted that the proposed framework is still applicable even when the rupture plane (or epicenter) is defined using different approaches.

Step 1: Randomly choose an active fault according to the relative annual occurrence rates λ_s/λ_E, $s = 1, ..., S$ in Eq. (C.4).

Step 2: Generate earthquake magnitude M based on its cumulative distribution function (CDF) given the fault selected in Step 1.

Step 3: Determine a rupture area based on the magnitude-scaling relationship derived by Shaw et al. (2009), which is a function of M from Step 2.

Step 4: Introduce two random variables following the uniform distribution which, respectively, represent the relative vertical and horizontal coordinates of the center of the rupture on the fault surface. Using the estimated area in Step 3 and the location of the source, the active rupture is determined.

Step 5: Calculate the distance between the earthquake rupture determined from Step 4 and each property location in the urban area, i.e., R_{ij}, in Eq. (C.1).

Step 6: Using the GMPE in Eq. (C.1), simulate the intensity measures of each property location. The intra- and inter-event residuals are randomly generated. In particular, the intra-event residuals need to be generated based on the selected spatial correlation model and the distance between the property locations for each pair.

Step 7: Predict the seismic response of each building using the P-DNN model given the generated magnitude M (Step 3), the distance R_{ij} (Step 5), and the estimated intensity measures (Step 6). Since the seismic responses are predicted in a probabilistic manner, the responses of each building are randomly drawn from a Gaussian distribution with the estimated mean and standard deviation. It should be noted that the exponential function is applied to each sample because the P-DNN model estimates the natural logarithm of the peak structural response.

Step 8: Evaluate the aggregated monetary loss, i.e., the net losses of the structural systems in the region, based on the estimated structural responses determined in Step 7.

Step 9: Repeat Step 1 to Step 8 until reaching the target coefficient of variation (c.o.v)

to obtain the loss curve using Eq. (C.5). Note that the required number of samples N_{MCS} in MCS is determined by $N_{MCS} = (1-p)/(\delta_t^2 \cdot p)$ to achieve the target c.o.v δ_t for the probability level p. As Eq. (C.5) indicates, the total occurrence rate λ_E must be multiplied by the average of the indicator function values computed for the generated samples.

Using the probabilistic predictions from the P-DNN model, the proposed method is expected to facilitate accurate pre-earthquake loss assessments based on proper quantifications of uncertainties and prediction in the structural responses given the intensity measures.

5.4.2 Application to hypothetical urban areas

To demonstrate the proposed deep learning-based regional seismic loss assessment, the hypothetical region termed as "V-city" is introduced with details of structural characteristics and regional information. Then, the applicability and effectiveness of the regional loss assessments are demonstrated.

5.4.2.1 Damage and monetary loss

In order to estimate the regional monetary loss due to a potential earthquake event, it is necessary to evaluate the damage of each structural system in the urban area. Such damage evaluation is usually performed based on the seismic capacity and demand of the structural systems. In this example, the capacity and demand are described in terms of peak displacement. In particular, the capacity is represented by the ultimate displacement of the structural system, whereas the demand is defined as the displacement of the structure caused by the scenario earthquake. To introduce

unit-less measures, both capacity and demand are normalized by the yield displacement of the structural system, Δ_y. These normalized capacity and demand, which are termed as ductility capacity and ductility demand, are denoted respectively by μ_R and μ. Based on the assumption that the seismic loss initiates as the displacement demand exceeds the yield displacement of the structural system, the damage factor δ is defined as (Goda and Hong 2008)

$$\delta = \max\left(\min\left(\frac{\mu - 1}{\mu_R - 1}, 1\right), 0\right) \tag{5.1}$$

If the seismic demand is less than yield displacement, δ becomes zero to indicate no damage. On the other hand, the total collapse, i.e., $\delta = 1$, is assumed to occur when seismic demand exceeds the ultimate capacity.

The seismic losses of buildings can be described using the damage factor δ. For example, Goda and Hong (2008) categorized the monetary seismic losses into three types in terms of the damage factor, i.e., building-related loss $L_{BL}(\delta)$, contents-related loss $L_{CO}(\delta)$, and business-interruption related loss $L_{BI}(\delta)$ as follows:

$$L_{BL}(\delta) = \delta^{\beta_{BL}} L_{BL}(1), \quad L_{CO}(\delta) = \delta^{\beta_{CO}} L_{CO}(1), \quad L_{BI}(\delta) = \delta^{\beta_{BI}} L_{BI}(1) \tag{5.2}$$

where $L_{BL}(1)$, $L_{CO}(1)$, and $L_{BI}(1)$ represent the three types of losses (in the Canadian dollar, CAD) at the complete damage state, and β_{BL}, β_{CO}, and β_{BI} are the loss-damage model parameters of the building of interest. By summing up the losses of all buildings in the urban area, the regional seismic loss can be described as

$$L = \sum_{k=1}^{n_R} \left(L_{BL}(\delta_k) + L_{CO}(\delta_k) + L_{BI}(\delta_k) \right) \tag{5.3}$$

where n_R is the number of building in the area.

5.4.2.2 Description of V-city

To demonstrate the proposed regional seismic loss assessment framework, 18 different building types in Goda and Hong (2008) are considered in this V-city example. These buildings are associated with the HAZUS-Earthquake classifications (FEMA and NIBS, 2003) and are idealized as bilinear SDOF systems. The adopted bilinear model is associated approximately with a seismic design code (NRCC, 2005) whose minimum required design base shear V_d is given by

$$V_d = C_s \cdot W \tag{5.4}$$

where C_s is the design base shear coefficient defined for each structural type, and W denotes the total weight of the structural system. Note that the strength reduction factor, usually denoted as R in "R-μ-T relationship" (Nassar and Krawinkler, 1991), is used to calculate C_s (Goda and Hong, 2008). Then, the yield displacement of a building, Δ_y is defined as

$$\Delta_y = R_N V_d / k = R_N C_s W / k \tag{5.5}$$

where R_N indicates the ratio of the actual yield strength of a designed structure to the design base shear V_d, which takes into account the fact that the actual yield strength is greater than its design base shear, and k denotes the stiffness of the structural system. The values of these parameters used in the hypothetical example are summarized in Table 5.2. To consider the modeling uncertainty of the structural

system, in this example, R_N and μ_R are assumed to follow lognormal distributions with the mean values provided in Table 5.2 and the coefficients of variation (c.o.v) of 0.3 and 0.15, respectively (Ellingwood et al., 1980; Ibarra, 2003).

Using the 18 different types of buildings, a set of 200 hypothetical structural systems including 80 residential and 120 commercial buildings are introduced. The hypothetical region roughly mimics the existing building stock in the downtown Vancouver area which has 400 property lots with an area of $25 \times 50 \ m^2$. The number of buildings from each building type is also shown in Table 5.2. To examine the effect of spatial distributions of buildings on regional losses, two different examples of building distributions shown in Figure 5.7 are investigated. In the first example (Figure 5.7(a)), 200 buildings are rather uniformly distributed over a square area of 2.5 km by 2.5 km whose center is located at 49.2°N and 123.2°W. On the other hand, the second example in Figure 5.7(b) is an asymmetrically distributed case in which more than three-quarters of the buildings are allocated in the western part, especially the southwestern area of the urban community. Figure 5.7(a) and Figure 5.7(b) are referred to as Case V1 and Case V2, respectively, in this study. It is noted that, based on the parameters in Table 5.2 and Eq. (5.3), the maximum possible seismic losses of both Case V1 and Case V2 are calculated as $ 1,678,939,376 (CAD). This is the direct loss caused by the complete damage of all structural systems, i.e., $\delta_k = 1$ for all buildings.

5.4.2.3 Regional seismic loss assessment of V-city using P-DNN model

In order to represent the realistic seismic hazard environment, it is assumed that nine active faults are in the vicinity of V-city, as illustrated in Figure 5.8. The geometric and seismic properties of the active faults, i.e., the distribution of the magnitudes of

each rupture, are adopted from Field et al. (2013) and are modified to facilitate the numerical investigation. Table 5.3 provides a summary of the properties of the active faults. Note that the dip angle represents the angle between the fault and the horizontal ground surface while the upper and lower bounds are correspondingly the upper and lower distances of the rupture from the ground surface.

Using the GMPE proposed by Campbell and Bozorgnia (2008) along with the correlation model developed by Loth and Baker (2011), the annual exceedance probability of the seismic loss of V-city is evaluated. Note that for simplicity, the soil properties in both cases are defined in terms of the shear velocity $Vs_{30} = 760 \ m/s$. After performing 100,000 instances of MCS, the loss curves for Cases V1 and V2 are obtained using the P-DNN model which is depicted as the blue solid line in Figure 5.9. For comparison, the DNN model (red dashed line) and the coefficient method (ASCE 41-13, 2013; green dash-dot line) are also employed instead of the P-DNN model in Step 7 of Section 5.4.1. Moreover, to check the effect of the spatial correlation on the regional seismic loss estimation, the loss curve without considering spatial correlation is evaluated and illustrated as the yellow dash-dot-dot line in the same plot.

Since λ_E in Eq. (C.5) for the nine active faults is calculated as 0.151 from the last column in Table 5.3, the value of the exceedance probability is 0.151 when no damage/loss occurs. It is found that, in both building-distribution cases, the DNN and the coefficient method overestimate the regional monetary loss compared to that from the P-DNN model under relatively mild earthquake scenarios, i.e., small exceedance probability. This result can be explained in terms of the relationship between the seismic responses and the monetary loss, in particular, the power function in Eq. (5.2) and its model parameters β, which are less than 1. The P-DNN

model predicts the mean and variance of the natural logarithm of the structural responses, whose mean value is equivalent to the estimate by the DNN model. However, after calculating monetary losses from the structural responses using Eq. (5.2), it is found that the mean value of the monetary loss by the P-DNN model is smaller than that by the DNN model. In other words, the distribution of the monetary loss is negatively skewed. For instance, if δ_0 follows a normal distribution with a mean of 0.5 and standard deviation of 0.1, then using 10,000 samples, the skewness and the mean of L_0, which is defined as δ_0 to the power of 0.41, i.e., $L_0 = \delta_0^{0.41}$, are estimated as -0.40 and 0.749, respectively, while $0.5^{0.41} = 0.753$. Therefore, in order to assess the regional seismic loss accurately, it is important to consider the probabilistic distribution of the structural responses given the intensity measures. It should be noted that different results can be obtained according to the definition between the seismic response and the corresponding loss.

Figure 5.9 also indicates that the DNN model shows a large overestimation of the regional loss for the entire exceedance probability per occurrence, while the overestimation by the coefficient method tends to decrease as the earthquake scenario becomes more severe. Due to the large prediction errors by the coefficient method, for an earthquake scenario causing most of the buildings in the region to collapse (i.e., μ exceeds μ_R in Eq. (5.3)), only the underestimated structural responses affect the bias of the monetary loss. Thus, one can find that, past a certain value of the exceedance probability, the loss curve obtained from the coefficient method is similar to or smaller than that by the P-DNN model. Since the capacity spectrum method shows a similar level of estimation error as compared to the coefficient method confirmed in Chapters 2 and 3, just as the coefficient method, a similar amount of error is expected in the regional loss assessment based on HAZUS

(FEMA, 2012), which adopted the capacity spectrum method for estimating structural responses given seismic intensities.

Next, the effect of spatial correlation on the regional assessment is examined by comparing the blue solid and yellow dash-dot-dot line in Figure 5.9. These results confirm that the effect of the spatial correlation of the intra-event residuals greatly influences the loss estimation. In general, without considering this correlation, one may underestimate the probability of occurrence of the worst-case scenario and overestimate the frequent events. Finally, because the active ruptures are located at the southwestern part of the region, the estimated loss of Case V2 is larger than that of Case V1 for a given exceedance probability as shown in Figure 5.10.

5.5 Earthquake responses using deep learning and database (ERD^2)

To disseminate the latest research achievement and extend the impacts on industrial and academic communities, a web-service (**E**arthquake **R**esponses using **D**eep learning and **D**atabase; http://ERD2.snu.ac.kr) was developed in this study to provide an interactive visualization of the proposed DNN predictions. Based on the input information provided by users, the website can (1) estimate the mean and variance of the structural responses using the method proposed in this study, and (2) fetch seismic demand values from the database with similar hysteretic behavior and ground motion characteristics as shown in Figure 5.11. The constructed databases and relevant Python codes used to train the DNN models are also shared on the website. The novel website shows a good example for researchers on how sustainable academic environments can be developed. This is because, in most cases, research outcome is not shared nor opened to the public, thus it is difficult to

reproduce researchers' contributions. However, the research outcomes (e.g., database, source codes, and user interface) already provided in the web service can reduce the computational costs for emulating the framework and give guidance for further study.

The web-service was developed as a platform where state-of-the-art DNN models for structural engineering can be shared, such as DNN models for damage identification based on images after significant natural hazards, DNN models for regional loss prediction, and DNN models for site response prediction. The portal will be used by researchers and engineers (1) to study basic knowledge about deep learning and earthquake engineering, (2) download a large volume of datasets and relevant codes, (3) share the latest trained DNN models, and (4) discuss research outcomes with international researchers, which will eventually foster internationally sustainable academic environments.

5.6. Conclusions

In order to demonstrate the applicability of the proposed DNN models, three different scales of earthquake engineering problems were introduced. First, the merits and potential of the BW-DNN model were confirmed by predicting the seismic responses of reinforced concrete (RC) columns with varying periods. Although this study only presented the prediction results of structural elements under seismic excitation, it is possible to predict the seismic responses of multi degree of freedom (MDOF) systems whose first mode period governs the seismic behavior. Such an investigation is required on various types of structural systems. Second, the new deep learning-based fragility estimation method was developed using the P-

DNN model. Owing to the heteroscedasticity assumption in training the P-DNN model, it was confirmed that the conditional failure probability of the structural system is properly estimated using the developed framework. Third, the application of the DNN and P-DNN models to assess the seismic loss of an urban community was presented. It was demonstrated that the importance of considering the variabilities of nonlinear behaviors in regional seismic loss assessment. In particular, such variabilities with large prediction errors of structural responses result in over/underestimation of regional seismic losses. Finally, the developed user-friendly website (ERD^2) provides the supporting source codes and databases, which disseminates the finding and outcomes to other researchers and facilitate sustainable research environments.

Table 5.1 Comparison between the natural logarithm of seismic responses of RC columns from dynamic analysis and predicted ones by using the BW-DNN model, the DNN model, and the coefficient method

Methods	RC column ($T = 0.115$)		RC column ($T = 0.162$)		RC column ($T = 0.257$)	
	MSE	RE (%)	MSE	RE (%)	MSE	RE (%)
BW-DNN model	0.3507	8.199	0.2023	7.327	0.0979	7.580
DNN model	0.0480	2.873	0.1089	5.744	0.1685	11.130
Coefficient method	0.0431	2.174	0.3505	9.681	1.3124	23.401

MSE: Mean squared error $(1/N \cdot \sum_{i=1}^{N}(y_i - f(x_i))^2)$

RE: Relative error $(1/N \cdot \sum_{i=1}^{N}|y_i - f(x_i)|/|y_i|)$

Table 5.2 Structural parameters and damage-loss information of hypothetical buildings (Goda and Hong, 2008)

I_{BT}[a]	# of bldgs.	# of stories	Size (m)	Structural & occupancy types	$L_{BL}(1),\ L_{CO}(1),$ $L_{BI}(1)$ (CAD/ft^2)	$\beta_{BL},\ \beta_{CO},\ \beta_{BI}$	$T_n(s)$[b]	Mean R_N	Mean μ_R	Target C_s[c]
1	8	2	10×12	W1-RES1	87.6, 21.9, 19.9	0.75, 0.68, 0.57	0.4	2	6	0.12
2	8	1	8×12	W1-RES1	87.6, 21.9, 19.9	0.75, 0.68, 0.57	0.4	2	6	0.12
3	17	2	15×30	W2-RES3	111.4, 27.9, 26.3	0.81, 0.68, 0.62	0.4	2	6	0.12
4	12	2	15×30	W1-COM1	47.8, 26.5, 23.9	0.81, 0.68, 0.43	0.4	2	6	0.12
5	1	5	18×36	S4M-RES3	111.4, 27.9, 26.3	0.69, 0.58, 0.53	0.7	2.25	4	0.1
6	2	5	18×36	S4M-COM4	103.5, 51.7, 163.9	0.70, 0.58, 0.57	0.7	2.25	4	0.1
7	1	13	18×36	S4H-RES3	111.4, 27.9, 26.3	0.69, 0.59, 0.53	1.4	2.25	3	0.075
8	1	13	18×36	S4H-COM4	103.5, 51.7, 163.9	0.70, 0.59, 0.57	1.4	2.25	3	0.075
9	7	2	15×30	C2L-RES3	111.4, 27.9, 26.3	0.76, 0.64, 0.58	0.4	2.5	6	0.12
10	10	2	15×30	C2L-COM1	47.8, 26.5, 23.9	0.75, 0.64, 0.41	0.4	2.5	6	0.12
11	18	5	18×36	C2M-RES3	111.4, 27.9, 26.3	0.75, 0.64, 0.58	0.6	2.5	5	0.12
12	27	5	18×36	C2M-COM4	103.5, 51.7, 163.9	0.77, 0.64, 0.62	0.6	2.5	5	0.12
13	13	15	18×36	C2H-RES3	111.4, 27.9, 26.3	0.76, 0.64, 0.58	1.65	3	3	0.05
14	25	15	18×36	C2H-COM4	103.5, 51.7, 163.9	0.77, 0.64, 0.62	1.65	3	3	0.05
15	4	2	15×30	URMLR-RES3	111.4, 27.9, 26.3	0.81, 0.69, 0.62	0.35	2	5	0.08

16	34	2	15×30	URMLR-COM1	47.8, 27.9, 26.3	0.81, 0.69, 0.43	0.35	2	5	0.08
17	4	3	20×40	URMMR-RES3	111.4, 27.9, 26.3	0.81, 0.69, 0.63	0.5	2	3.3	0.08
18	8	3	20×40	URMMR-COM2	61.0, 33.4, 19.5	0.80, 0.69, 0.49	0.5	2	3.3	0.08

[a] I_{BT} is the building type index which is related to the structural and occupancy types defined in the HAZUS-Earthquake classifications (FEMA and NIBS, 2003)

[b] T_n is the fundamental period of the building.

[c] C_s is used to represent the seismic design level for existing buildings (see Eq. (9)).

Note that 5% post-yield stiffness ratio is assumed for structural systems.

Table 5.3 Properties of 9 active faults

Index	Average dip angle	Upper bound (km)	Lower bound (km)	Active fault length (km)	Annual occurrence rate
1	50	0	13	47.2168	0.0316
2	90	0	12.3	20.1137	0.0019
3	67	0	15.9	21.1298	0.0056
4	20	9	14	19.1865	0.0023
5	25	0	6	19.3854	0.0019
6	65	0	17.7	31.3851	0.0082
7	76	0	18.2	106.011	0.0327
8	90	0	11	170.693	0.0268
9	90	0	15.9	42.8961	0.0398
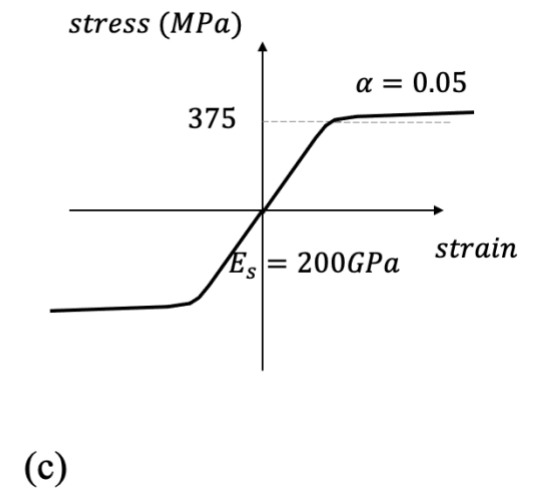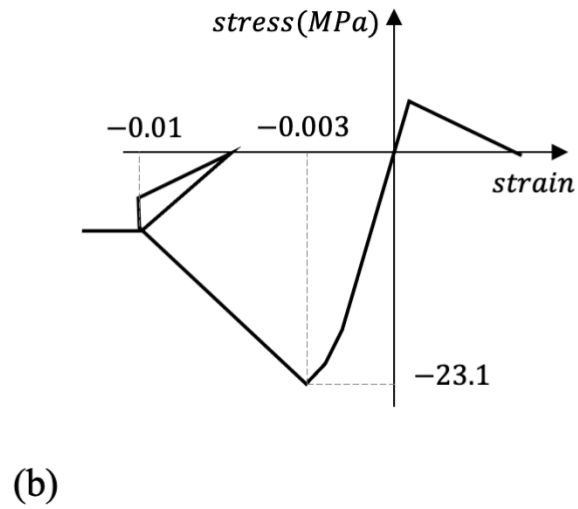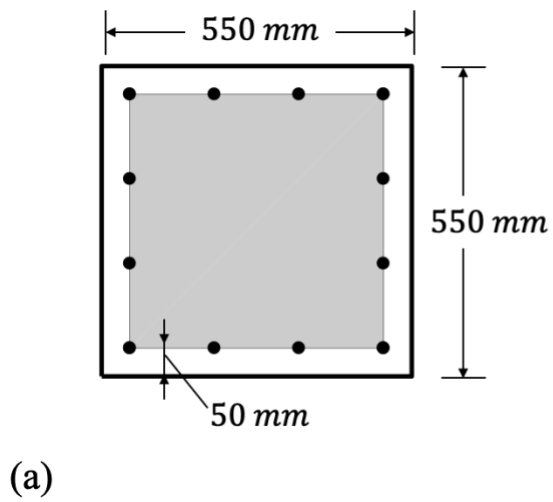
Figure 5.1 Sectional and material properties of RC columns
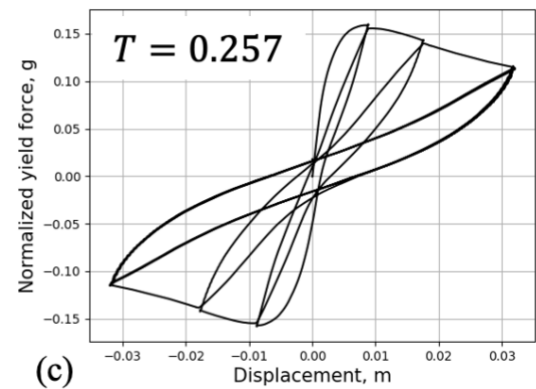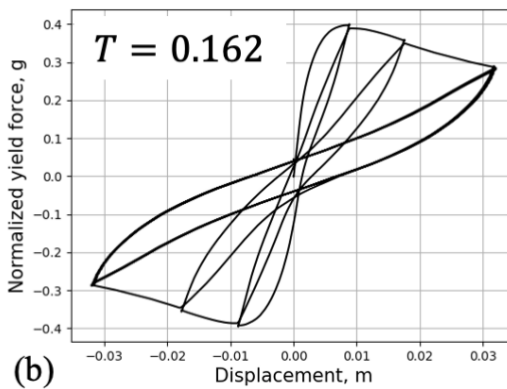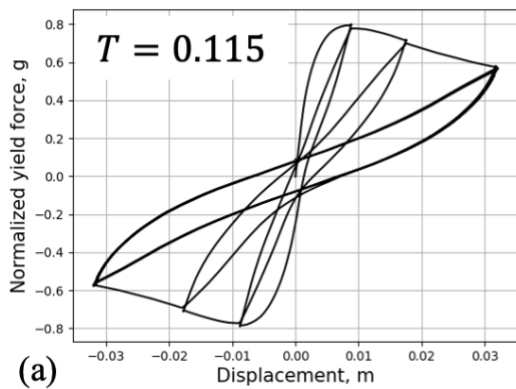
Figure 5.2 Hysteresis of three RC columns with different first mode periods
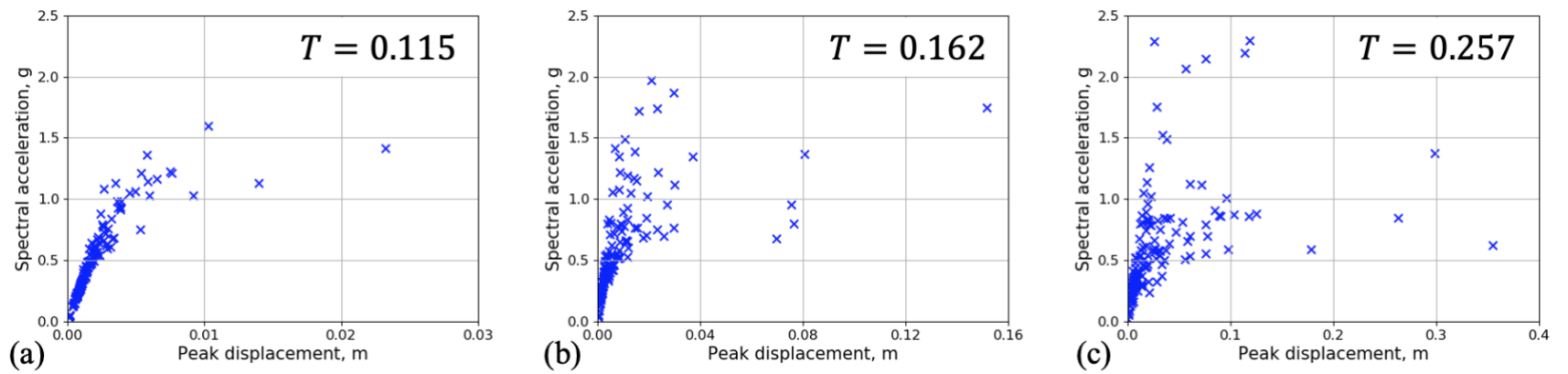
Figure 5.3 Relationship between spectral acceleration and peak displacement under 135 ground motions for the three RC columns

17.8 N. sec^2/mm per each mass   49.849 N

3658 mm

3658 mm

3429 mm

5486 mm   5486 mm   5486 mm
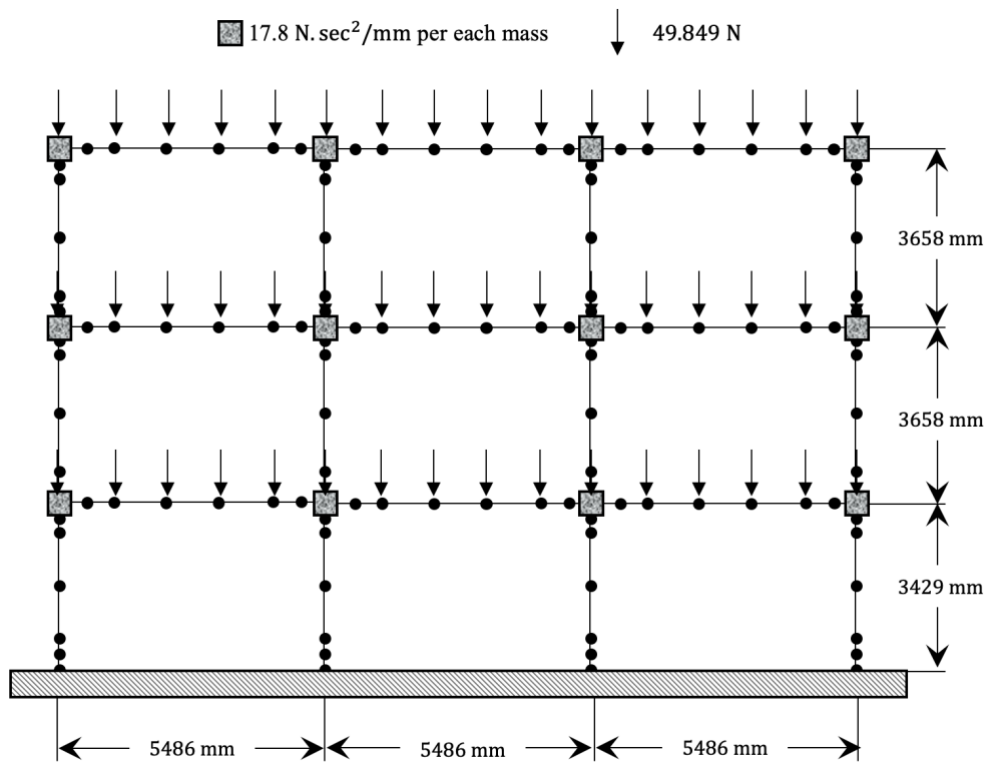
Figure 5.4 Analytical model of the RC frame structure
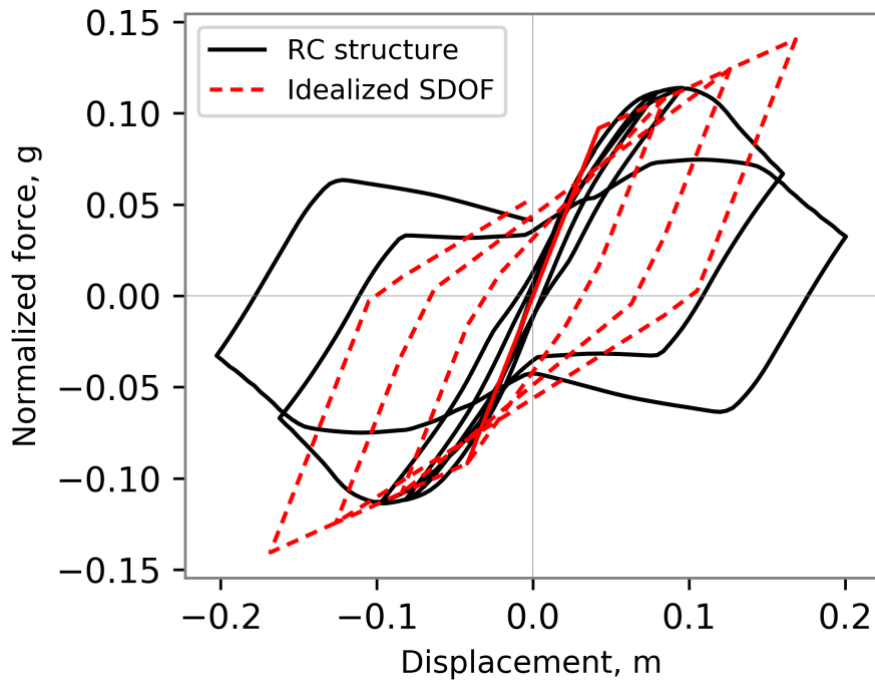
Figure 5.5 Hysteretic behavior generated from the RC structure (black solid line) and that of the idealized SDOF system (red dashed line). The idealized SDOF system is bilinear stiffness degradation (HM3) with period 1.36 sec, stiffness 2.183 g/m, yield strength 0.092g, and post yield stiffness ratio 0.177. The period of the idealized SDOF system is longer than the first mode period of RC building, because, when the structure is applied to lateral load, cracks are generated in the RC structure, which elongates its first mode period
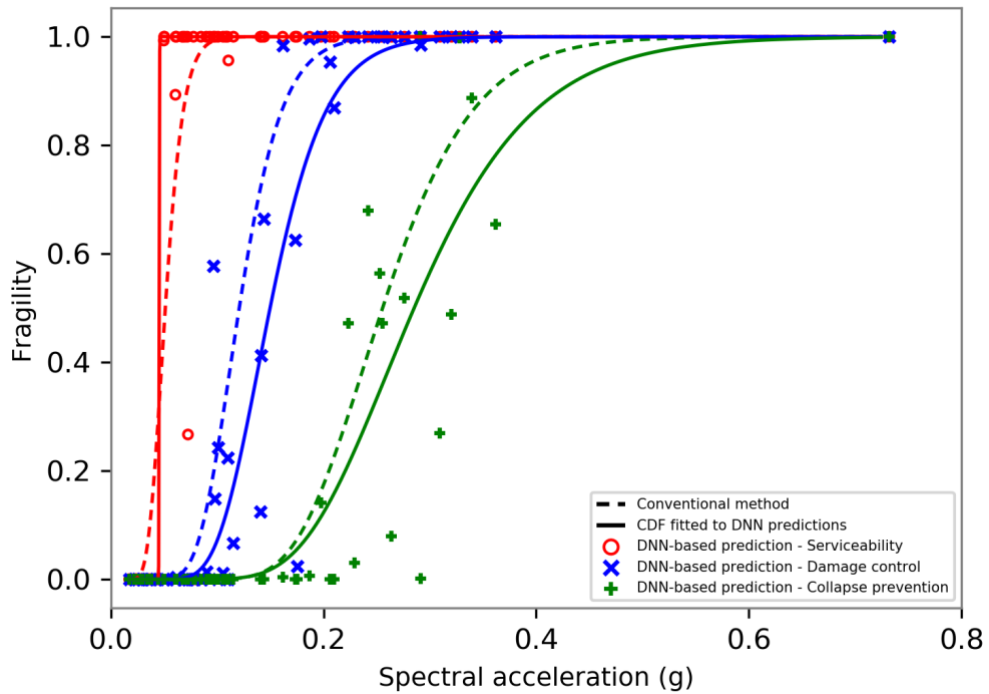
Figure 5.6 Fragility functions of the frame structure for three limit states, obtained by the proposed DNN model (solid line) and an existing regression-based method (dashed line)
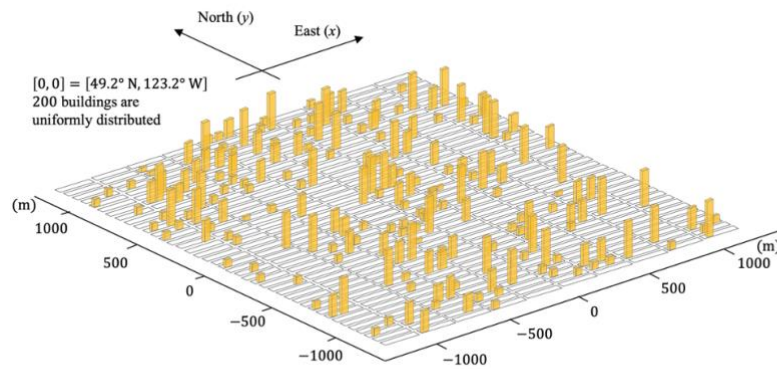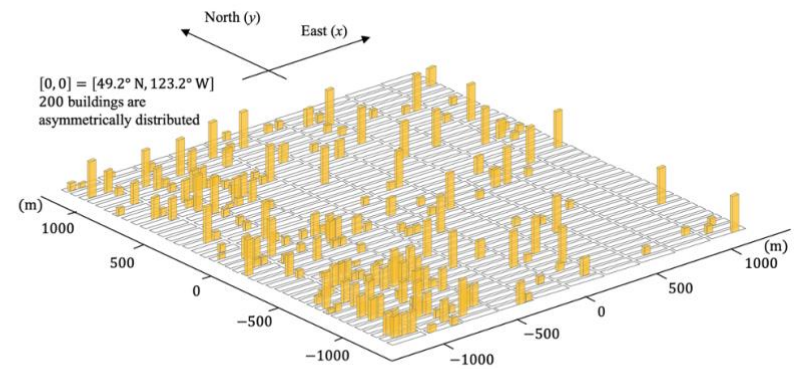
Figure 5.7 200 hypothetical buildings allocated on the 400 property lots (40 × 10): (a) Case V1: the set of buildings uniformly distributed in the region, and (b) Case V2: the set of the buildings asymmetrically distributed in the region
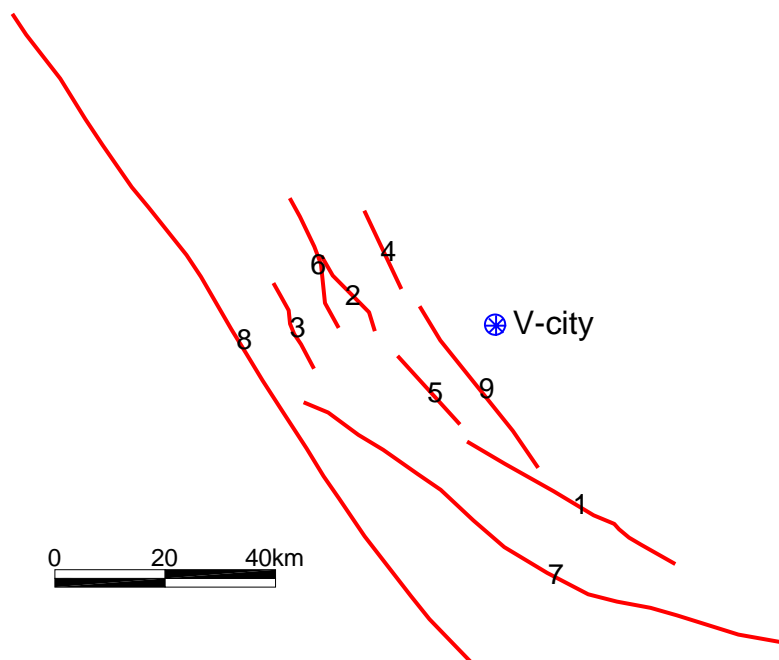
Figure 5.8 Illustration of V-city surrounded by 9 active faults
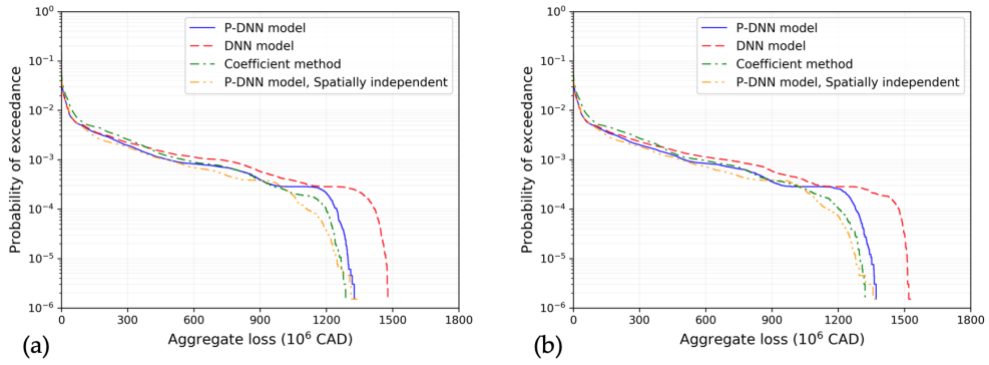
Figure 5.9 The aggregated losses of V-city surrounded by 9 active faults for (a) Case V1, and (b) Case V2, estimated by use of the P-DNN model (blue solid line), the DNN model (red dashed line), the coefficient method (green dash-dot line), and the P-DNN model without considering spatial correlation (yellow dash-dot-dot line)
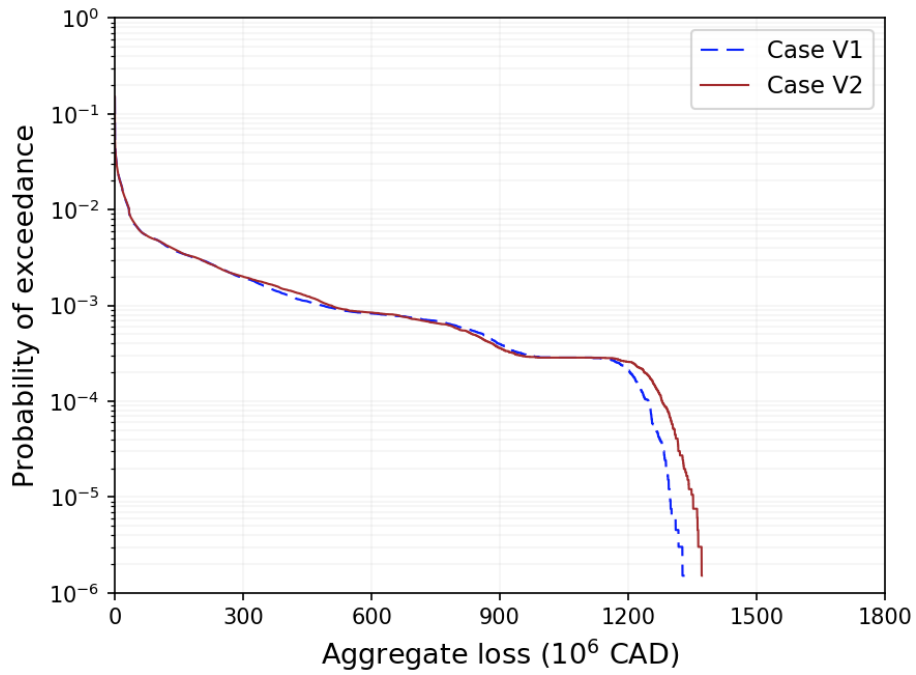
Figure 5.10 Comparison of aggregated loss for Case V1 (blue dashed line) and Case V2 (brown solid line) when the V-city is surrounded by 9 active faults
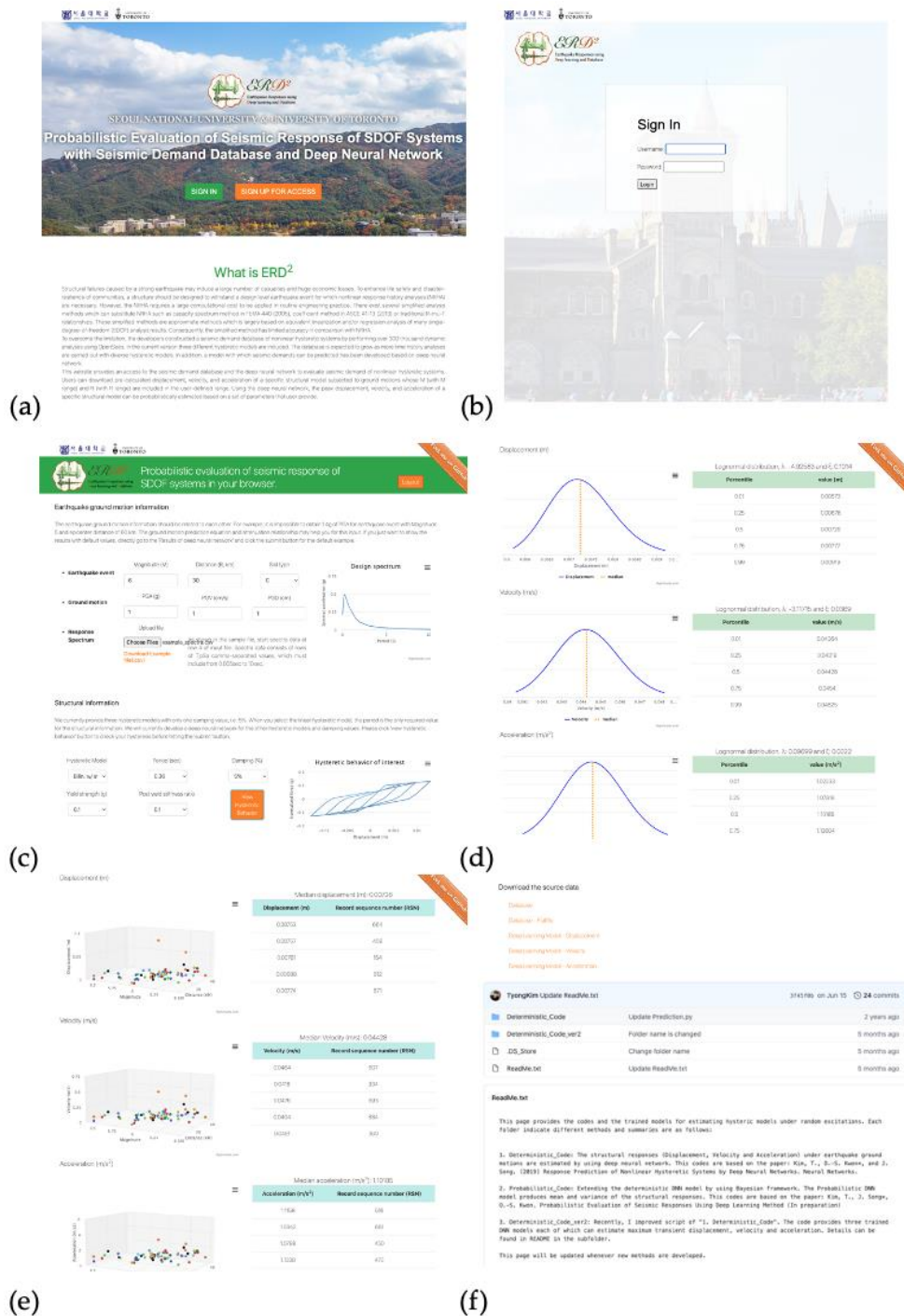
Figure 5.11 Earthquake Responses using Deep learning and Database (ERD^2) web-service: (a) the main page of the website, (b) sign-in page, (c) user-defined earthquake and structural information, (d) P-DNN model predictions, (e) retrieve seismic demands from the database, and (f) download source codes

Chapter 6. Summary, Conclusions, and Further Studies

6.1. Major developments and findings

This dissertation focuses on developing deep neural network (DNN) models that can predict seismic responses of a wide class of structural systems having various hysteretic behaviors. The two main objectives proposed in Chapter 1 were fulfilled: (1) three different DNN models that predict seismic responses of structural systems, i.e., DNN, P-DNN, and BW-DNN models, were developed. The DNN models deal with idealized hysteresis, latent uncertainties of ground motion information, and generalized hysteresis having stiffness/strength degradations and pinching effects. (2) applications of the developed DNN models were provided for three different scales of engineering examples, i.e., structural element, structural system, and urban community, which demonstrated that the developed methods are applicable to a broad range of earthquake engineering problems and used as a seismic risk mitigation strategy. The major developments and findings of this study are graphically presented in Figure 6.1 and summarized as follows:

- Two seismic demand databases considering idealized and generalized hysteretic systems were developed. To this end, a large number of dynamic analyses were carried out. Seven structural responses including peak displacement, peak velocity, peak acceleration, hysteretic energy, elastic energy, peak base reaction force, and residual displacement, were stored. Note that one can download both databases from the developed ERD[2] website (http://ERD2.snu.ac.kr).

- A deep neural network (DNN) model was developed to predict the seismic responses of idealized hysteretic systems. Three different hysteretic systems, i.e., linear (HM1), bilinear kinematic hardening (HM2), and bilinear stiffness degrading systems (HM3), were used as the structural information, while 1,499 ground motions from the NGA-West database (Power et al., 2006) were introduced for earthquake information. The architecture of the DNN model was inspired by the natural phenomenon of seismic excitation to structural systems. It was demonstrated that the DNN model shows superior accuracy compared to other simple regression-based nonlinear static procedures such as the capacity spectrum method, coefficient method, and R-μ-T relationship.

- In developing the DNN model, a convolutional neural network (CNN) which shows a clear advantage in dealing with the data having a strong spatial correlation was introduced to extract important features of hysteretic behaviors. Owing to this idea, the DNN model was able to predict the responses of a wide class of hysteretic systems, which can overcome the limitation of existing studies that have focused on the specific type of structural systems and hysteretic behaviors.

- A probabilistic deep neural network (P-DNN) model was constructed to take into account the inherent uncertainties of random stochastic excitation into response prediction. Using a loss function that is proportional to the negative log-likelihood based on the Gaussian distribution assumption, the P-DNN model was trained to predict the conditional mean and variance of the seismic responses. The performance of the proposed method was tested through various numerical experiments which confirmed that the P-DNN model can successfully quantify the prediction uncertainties. In particular, it produced a relatively small variance

to the structural system behaving in the linear elastic range, while a relatively large variance was obtained for the nonlinear case.

- A modified Bouc-Wen-Baber-Noori (m-BWBN) model was developed to represent a wide class of hysteretic behavior having stiffness/strength degradations and pinching effects. Using the genetic algorithm with the compiled hysteretic loops of reinforced concrete (RC) column database (Berry et al., 2004), feasible domains of the m-BWBN model parameters are defined to properly illustrate the structural hysteresis. Moreover, a sensitivity analysis was performed to distinguish the parameters that do not significantly affect the shape of hysteresis, which results in reducing the computational costs in developing the seismic demand database.

- Using the m-BWBN model and the corresponding seismic demand database, a Bouc-Wen hysteresis-based deep neural network (BW-DNN) model was trained which can predict seismic responses of structural systems having various hysteretic characteristics. In order to properly extract the structural information and improve the applicability of the trained BW-DNN model, a new framework of quasi-static cyclic analysis was proposed. The numerical investigation confirmed that the predictions by the proposed BW-DNN model have superior accuracy compared to existing methods that cannot properly consider the strength/stiffness degradations and pinching effects.

- Three different scales of earthquake engineering problems, i.e., structural component, structural system, and regional levels, were presented. In particular, the application consisted of seismic response evaluation of structural elements using the DNN and BW-DNN model, seismic fragility estimation of a structural system using the P-DNN model, and seismic loss assessment of urban

community using the DNN and P-DNN models. The numerical investigations demonstrated the applicability and effectiveness of the developed DNN models for a wide class of earthquake engineering problems.

- In the practice adopting the P-DNN model in structural fragility evaluation, the conditional probability given seismic intensity measure was directly estimated without any assumption between intensity measure and damage measure. The framework can be further extended to develop a ground motion selection algorithm for the purpose of assessing the structural system because it is possible to identify the ground motions that significantly affect the variability of structural fragility.

- In regional seismic loss assessment, it was found that the uncertainties in the responses of the structural system given the seismic input features may have a significant impact on the loss assessment. This is because most of the critical damage of the structural system occurs when the system behaves in nonlinear ranges, which amplifies the level of uncertainties in the response.

- According to the modern seismic design codes (Eurocode 8, 2003; KMOLIT, 2018), the responses of structural systems subjected to a design spectrum are estimated by nonlinear static procedures or linear/nonlinear time history analysis. Note that the design spectrum is defined by soil class, importance of the structural system, and return period of earthquake events. The former estimates the responses by a simple regression function, while the latter performs time history analysis using ground motion records that are spectrum-matched or spectrum-compatible to the design spectrum of the site of interest. The trained DNN models can be a good replacement for both methods, in that the method can improve the prediction accuracy compared to the nonlinear static procedures,

and reduce the computational costs in the time history analysis. By adopting the DNN models in the routine engineering practice, it is possible to establish an effective disaster risk management framework against seismic hazards.

- A website (Earthquake Responses using Deep learning and Database, ERD^2) was developed to provide the interactive visualization of the proposed P-DNN predictions and share the latest research outcomes (URL: http://ERD2.snu.ac.kr). The trained model, constructed database, and source codes can be downloaded from the website and will be updated when a new deep learning-based method is proposed. In near future, after the trained DNN models are fully demonstrated and validated their performance by international researchers, they can be adopted in OpenSeesPy (Zhu et al., 2018) or as one of the options in numerous structural analysis programs similar to the complete quadratic combination (CQC) or the square root of the sum of the squares (SRSS) that predicts the seismic responses before performing a nonlinear time history analysis.

6.2. Recommendations for further studies

In order to propose a new deep learning-based seismic response prediction method that covers more realistic hysteretic behaviors and advance the frontiers of applications relevant to this study, the following topics are recommended for future research:

- Although this study focuses on developing DNN models that predict responses of structural systems caused by earthquake excitations, the concept and framework can be applied to other engineering fields in which it is important to

address the hysteretic behavior of a structural system under stochastic excitations, e.g., high-rise buildings subject to wind load, or wire cable isolator subject to mechanical vibration.

- Introduction of the CNN to extract important features from hysteretic behaviors is a new attempt in the deep learning fields, in that the traditional deep learning problems usually deal with images or voice records. Therefore, as an analogy of this study, investigation of datasets having significant spatial/temporal correlation using the CNN can be a promising future research topic in various engineering fields.

- The P-DNN model only covers inherent uncertainties from the input dataset, i.e., loss of information in the process of representing earthquake ground information by the selected features. However, there exist uncertainties in the model parameters of the deep neural network, i.e., model uncertainties. By adopting Monte Carlo dropout with the loss function written in Eq. (3.1), two different uncertainties can be quantified simultaneously, and its impact on response prediction can be examined.

- Due to the functional redundancy of the Bouc-Wen model, it is difficult to identify the optimal parameters that produce the equivalent hysteresis from finite element methods or experiments. In other words, parameters that produce similar hysteresis for a certain displacement history do not guarantee the hysteretic behaviors when different displacement histories are assigned, e.g., varying number of cycles, intensity, and displacement steps. Therefore, a general procedure is required to identify the optimal m-BWBN parameters that produce similar hysteresis regardless of the displacement history.

- When the seismic behavior of a structural system is dominated by its first mode

period, it is possible to predict the responses by using the BW-DNN model with reasonable accuracy. Once this numerical investigation is properly carried out, the effectiveness and application of the BW-DNN model are further expanded.

- It is demonstrated the merits and potential of probabilistic prediction of seismic responses of structural systems as shown in Chapter 3. Because the seismic prediction from the P-DNN model is confined to the idealized hysteresis, it is difficult to directly employ the hysteresis of general structural systems to the input of the P-DNN model. Note that equivalent idealized hysteresis is employed in the DNN and P-DNN models. To address the issue, the probabilistic deep neural network model for Bouc-Wen hysteresis (P-BW-DNN model) needs to be developed, which probabilistically predicts the seismic responses of realistic structural systems. Once the P-BW-DNN model is developed, the seismic responses of structural systems having various hysteretic characteristics can be predicted in a probabilistic manner.

- The developed DNN models can predict the important structural responses that are represented by a scalar value, e.g., peak responses of the nonlinear hysteretic system. To this end, various features of earthquake ground motion that show a high correlation with the target structural responses are adopted for the input of the models. Because of such high correlation, it is possible to efficiently train the DNN models and predict the responses with relatively high accuracy rather than the classic nonlinear static procedures. However, since a deep neural network can extract hidden features or find unknown patterns, it is possible to increase the prediction accuracy by introducing a whole ground motion acceleration history as an input of the DNN model. Moreover, using a recurrent neural network (RNN), a full-time history of structural responses can be

predicted instead of a scalar response value of structural systems, e.g., peak displacement.

- Although the proposed deep learning-based response prediction of nonlinear hysteretic system framework is able to predict peak displacement, velocity, and acceleration, it is difficult to predict the imbalanced responses such as residual displacement or hysteretic energy whose values are zero when the hysteretic system behaves in a linear range. Therefore, further study is desirable to develop a framework that predicts the imbalanced structural responses using a deep neural network.

- In this study, four hysteretic models, i.e., linear, bilinear, bilinear stiffness degradation, and m-BWBN model, were introduced and a large number of dynamic analysis, i.e., simulation, were carried out with 1,499 ground motions from the NGA-West database. Although the trained DNN models are verified using the test dataset, their validation to experiment datasets is desirable to extend the applicability of the DNN models to real world engineering problems. In addition, a model refined framework needs to be proposed when the DNN models which are trained using the simulation dataset show poor prediction accuracy for the experimental dataset.

- This study presented the application of the trained DNN models to precisely estimate the seismic loss of an urban community which can enhance the pre-earthquake regional loss assessment. On the other hand, to facilitate near-real-time post-earthquake loss assessment, a deep neural network can be used as a surrogate to estimate area-wide structural damage given the spatial distribution of the seismic intensity levels. Since the surrogate model can be extended by combining an infrastructural system such as a lifeline or transport network,

stakeholders can readily determine the evacuation route or recovery plan from the predicted results.

- In this study, even though a large number of ground motions are introduced compared to other studies, most of them are low-intensity, which cannot make structural systems behave in nonlinear ranges. Moreover, since only 5% damping ratio is employed in dynamic analysis, it is difficult to investigate the effect of damping on structural systems. Therefore, the seismic demand database is needed to extend for various scaling factors for ground motions and different damping ratio values. When the database is extended, it is possible to develop an improved version of the DNN models that would be employed to various earthquake engineering problems, e.g., incremental dynamic analysis (IDA).

- Since the DNN models were developed based on the seismic demand database of various single degree of freedom systems, its application to multi degree of freedom (MDOF) systems, which show higher mode effects during excitation, is limited. To fully employ the merits of the deep learning-based framework, an extension of the developed DNN models to the MDOF systems is required.
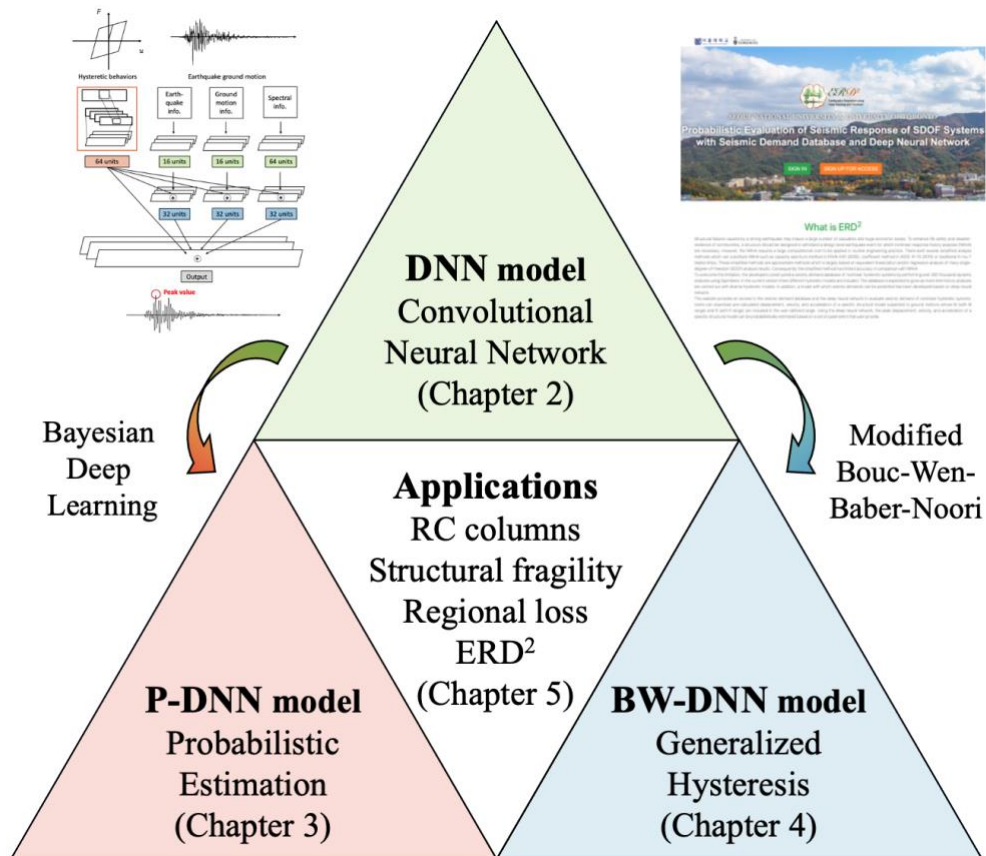
Figure 6.1 Diagram of the main contributions and findings of the dissertation

References

Abadi M., Agarwal A., Barham P., Brevdo E., Chen Z., Citro C., Corrado G.S., Davis A., Dean J., Devin M., Ghemawat S., Goodfellow I., Harp A., Irving G., Isard M., Jia Y., Jozefowicz R., Kaiser L., Kudlur M., Levenberg J., Mane D., Monga R., Moore S., Murray D., Olah C., Schuster M., Shlens J., Steiner B., Sutskever I., Talwar K., Tucker P., Vanhoucke V., Vasudevan V., Viegas F., Vinyals O., Warden P., Wattenberg M., Wicke M., Yu Y., Zheng X. (2016) TensorFlow: Large-scale machine learning on heterogeneous distributed systems. ArXiv160304467 Cs.

Adeli H. (2001) Neural networks in civil engineering: 1989–2000. *Comput.-Aided Civ. Infrastruct. Eng.* 16: 126–142.

Adeli H., Panakkat A. (2009) A probabilistic neural network for earthquake magnitude prediction. *Neural Netw.* 22: 1018–1024.

Akbas B., Shen J., Sabol T.A. (2011) Estimation of seismic-induced demands on column splices with a neural network model. *Applied Soft Computing* 11(8): 4820–4829.

Alain G., Bengio Y. (2016) Understanding intermediate layers using linear classifier probes. ArXiv161001644 Cs Stat.

ASCE. (2000) FEMA-356: Prestandard and commentary for the seismic rehabilitation of buildings. Developed by ASCE for FEMA, Washington, DC

ASCE (2013) Seismic Evaluation and Rehabilitation of Existing Buildings, ASCE/SEI 41-13 (Public Comment Draft), American Society of Civil Engineers, Reston, VA.

ATC (1997) Seismic evaluation and retrofit of concrete buildings, Report No. ATC-40, Applied Technology Council, Redwood City, CA.

Baker J.W., Cornell C.A. (2006) Spectral shape, epsilon and record selection. *Earthquake Eng. Struct. Dyn.* 35(9): 1077–1095.

Baber, T. T. Noori, M. N. (1985) Random vibration of degrading pinching systems. *Journal of Engineering Mechanics* 111: 1010–1026.

Baber T.T., Wen Y.K. (1981) Random vibrations of hysteretic degrading systems. *Journal of Engineering Mechanics* 107(6): 1069–1089.

Behmanesh I., Yousefianmoghadam S., Nozari A., Moaveni B., Stavridis A., (2018) Uncertainty quantification and propagation in dynamic models using ambient vibration measurements, application to a 10-story building. *Mech. Syst. Signal Process.* 107: 502–514.

Berry M., Parrish M., Eberhard, M. (2004) PEER structural performance database user's manual, Pacific Earthquake Engineering Research Center, University of California, Berkeley.

Bertsimas D., Tsitsiklis J. (1993). Simulated annealing. *Statist. Sci.* 8(1): 10-15.

Bogacki, P., Shampine L.F. (1989). A 3(2) pair of Runge–Kutta formulas. *Applied Mathematics Letters* 2(4): 321–325.

Bommer J., Spence R., Erdik M., Tabuchi S., Aydinoglu N., Booth E., del Re D,, Peterken O., (2002) Development of an earthquake loss model for Turkish catastrophe insurance. *Journal of Seismology* 6(3): 431–446.

Bouc R. (1967) Forced vibration of mechanical systems with hysteresis. Proc. of 4th Conference on Nonlinear Oscillation, Academia, Publishing House of Czechoslovak Academy of Science, Prague, Czechoslovakia.

BSSC. (2003) NEHRP recommended provisions for seismic regulations for new buildings and other structures, Part1: Provisions, FEMA 368, Federal Emergency Management Agency, Washington, D.C.

Campbell K., Bozorgnia Y. (2008). NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. *Earthquake Spectra* 24(1): 139-171.

Celik O.C., Ellingwood B.R., (2010). Seismic fragilities for non-ductile reinforced concrete frames - Role of aleatoric and epistemic uncertainties. *Struct. Saf.* 32: 1–12.

Cheney W., Kincaid D. (1999). Numerical mathematics and computing, 4th edition. Brooks/Cole, Pacific Grove, CA.

Choi B. (2017) Efficient simulation-based approaches for community-level probabilistic seismic risk assessment. MS. thesis, Seoul National University, Seoul, Korea.

Chopra A.K., Chintanapakdee C. (2004) Inelastic deformation ratios for design and evaluation of structures: single degree-of-freedom bilinear systems. *Journal of Structural Engineering* 130:1309-1319.

Constantinou M.C., Adnane M.A. (1987) Dynamics of soil-base-isolated structure systems: evaluation of two models for yielding systems. Report to NSAF. Philadelphia, PA: Department of Civil Engineering, Drexel University.

Conte J.P., Durrani A.J., Shelton R.O. (1994) Seismic response modeling of multi-story buildings using neural networks. *Journal of Intelligent Material Systems and Structures* 5(3): 392–402.

Cornell C.A., Jalayer A.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.

De Lautour O.R., Omenzetter P. (2009) Prediction of seismic-induced structural damage using artificial neural networks. *J. Eng. Struct.* 31(2): 600-606.

Deierlein G.G., Krawinkler H., Cornell C.A. (2003) A framework for performance based earthquake engineering. In: Pacific Conference on Earthquake Engineering 2003.

Deniz D., Song J., Hajjar J.F. (2017) Energy-based seismic collapse criterion for ductile structural frames. *Engineering Structures* 141: 1-13.

Deniz D., Song J., Hajjar J.F. (2018) Energy-based sideway collapse fragilities for ductile structural frames under earthquake loadings. *Engineering Structures* 174: 282-294.

Diaz Gomez L., Kwon O.S., Dabirvaziri M. (2015) Seismic fragility of steel MRFs in Vancouver and Montreal designed in the 1960s, 1980s, and 2010. *Canadian Journal of Civil Engineering* 42(11): 919-929.

Ellingwood B.R. (2001) Earthquake risk assessment of building structures. *Reliability Eng. and System Safety* 74(3): 251-262.

Ellingwood B.R., Galambos T.V., MacGregor J.G., Cornell C.A. (1980) Development of a probability based load criterion for American national standard A58. Washington,

DC: National Bureau of Standards. National Bureau of Standards Special Publication No. 577.

European Committee for Standardization. (2003) Eurocode 8: design of structures for earthquake resistance—Part 1: general rules, seismic actions and rules for buildings. Brussels, Belgium

Fajfar P. (2000) A nonlinear analysis method for performance based seismic design. *Earthquake Spectra* 16(3): 573-592.

FEMA. (2005) Improvement of nonlinear static seismic analysis procedures. Report No. FEMA-440, Washington, DC.

FEMA. (2009) Quantification of building seismic performance factors, FEMA P-695, prepared by Applied Technology Council for the Federal Emergency Management Agency, Washington, D.C.

FEMA. (2011) Quantification of building seismic performance factors: Component equivalency methodology, FEMA P-795, prepared by Applied Technology Council for the Federal Emergency Management Agency, Washington, D.C.

FEMA. (2012) Multi-hazard loss estimation methodology. Earthquake model – HAZUS MH 2.1 Technical Manual, Washington, USA.

Federal Emergency Management Agency (FEMA) and the National Institute of Building Sciences (NIBS). (2003). HAZUS-earthquake: technical manual. Washington, DC: FEMA and NIBS.

Field E.H., Biasi G.P., Bird P., Dawson T.E., Felzer K.R., Jackson D.D., Johnson K.M., Jordan T.H., Madden C., Michael A.J., Milner K.R., Page M.T., Parsons T., Powers P.M., Shaw B.E., Thatcher W.R., Weldon R.J. II, Zeng Y. (2013) Uniform California earthquake rupture forecast, version 3 (UCERF3)—The time-independent model: U.S. Geological Survey Open-File Report 2013–1165", 97 p.", California Geological Survey Special Report 228, and Southern California Earthquake Center Publication 1792, http://pubs.usgs.gov/of/2013/1165/.

Foliente G.C., Singh M.P., Noori M.N. (1996) Equivalent linearization of generally pinching hysteretic, degrading systems. *Earthquake Eng Struct Dyn*. 25(6): 611–29.

Gal Y., (2016) Uncertainty in deep learning. PhD dissertation, University of Cambridge.

Gal Y., Ghahramani Z. (2015) A theoretically grounded application of dropout in recurrent neural networks. arXiv preprint arXiv:1512.05287.

Gal Y., Ghahramani Z. (2016) Bayesian convolutional neural networks with bernoulli approximation variational inference, arXiv:1506.02158.

Giovanis D., Fragiadakis M., Papadopoulos V. (2015) Epistemic uncertainty assessment using incremental dynamic analysis and neural networks. *Bull. Earthq. Eng*. 14.

Glorot X., Bengio Y. (2010) Understanding the difficulty of training deep feedforward neural networks, in: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. Presented at the Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, pp. 249–256.

Goda K., Hong H.P. (2008) Estimation of seismic loss for spatially distributed buildings." *Earthq. Spectra*. 24: 889–910.

Goldberg D.E., Holland J.H. (1988) Genetic algorithms and machine learning. *Machine Learning* 3(2): 95–99.

Haselton C.B., Baker J.W., Liel A.B., Deierlein G.G. (2011) Accounting for ground-motion spectral shape characteristics in structural collapse assessment through an adjustment for epsilon. *J. Struct. Eng*. 137: 332–344.

Haukaas T., Der Kiureghian A. (2004). Finite element reliability and sensitivity methods for performance-based earthquake engineering. Berkeley: Pacific Earthquake Engineering Research Center, College of Engineering, University of California.

He K., Zhang X., Ren S., Sun J., (2015) Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. ArXiv:1502.01852 Cs.

Hossain M.R., Ashraf M., Padgett J.E. (2013) Risk-based seismic performance assessment of yielding shear panel device. *Eng. Struct.* 56: 1570–1579

Housner G., Jennings P. (1982) Earthquake design criteria. EERI Monograph Series, Earthquake Engineering Research Institute, Oakland, CA.

Ibarra L.F. (2003). Global collapses of frame structures under seismic excitations. Ph.D. dissertation, Stanford University, Stanford, CA.

Ibarra L.F., Medina R.A., Krawinkler H., (2005) Hysteretic models that incorporate strength and stiffness deterioration. *Earthq. Eng. Struct. Dyn*. 34: 1489–1511.

Ioffe S., Szegedy C. (2015) Batch normalization: Accelerating deep network training by reducing internal covariate shift. ArXiv150203167 Cs.

Jalayer F., Elefante L., De Risi R., Manfredi G. (2014) Cloud analysis revisited: efficient fragility calculation and uncertainty propagation using simple linear regression. Proceedings of the 10th National Conference in Earthquake Engineering, Earthquake Engineering Research Institute, Anchorage, AK, 2014.

Jeng C.H., Mo Y.L. (2004) Quick seismic response estimation of prestressed concrete bridges using artificial neural networks. *Journal of Computing in Civil Engineering* 18(4): 360–372.

Joghataie A., Farrokh M. (2008) Dynamic analysis of nonlinear frames by Prandtl neural networks. *Journal of Engineering Mechanics* 134(11): 961–969.

Karimi I., Khaji N., Ahmadi M.T., Mirzayee M. (2010) System identification of concrete gravity dams using artificial neural networks based on a hybrid finite element-boundary element approach. *Engineering Structures* 32(11): 3583–3591.

Kendall A., Gal Y. (2017) What uncertainties do we need in Bayesian deep learning for computer vision?, arXiv:1703.04977.

Kim T., Kwon O.S., Song J. (2019) Response prediction of nonlinear hysteretic systems by deep neural networks. *Neural Networks* 111: 1-10.

Kim T., Kwon O.S., Song J. (2021) Clustering-based adaptive ground motion selection algorithm for efficient estimation of structural fragilities, *Earthquake Engineering and Structural Dynamics*.

Kim T., Song J. (2018) Generalized reliability importance measure by Gaussian mixture. *Reliability Engineering and System Safety* 173: 105-115.

Kim T., Song J., Kwon O.S. (2020) Probabilistic evaluation of seismic responses using deep learning method. *Structural Safety* 84: 101913.

Kim T., Song J., Kwon O.S. (2020) Pre- and post-earthquake regional loss assessment using deep learning. *Earthquake Engineering and Structural Dynamics* 49(7): 657-678.

Kingma D.P., Ba J. (2014) Adam: A method for stochastic optimization. ArXiv14126980 Cs.

KMOLIT (Korea Ministry of Land, Infrastructure and Transport). (2018) General seismic design code. KMOLIT, Sejong-si, Korea

Krizhevsky A., Sutskever I., Hinton G.E. (2012) ImageNet classification with deep convolutional neural networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 1097–1105.

Kwon O.S., Elnashai A. (2006) The effect of material and ground motion uncertainty on the seismic vulnerability curves of RC structure. *Eng. Struct.* 28: 289–303.

Lagaros N.D., Fragiadakis M. (2007) Fragility assessment of steel frames using neural networks. *Earthquake Spectra* 23(4): 735–752.

Lagaros N.D., Tsompanakis Y., Psarropoulos P.N., Georgopoulos E.C. (2009) Computationally efficient seismic fragility analysis of geostructures. *Computers and Structures* 87(19–20): 1195–1203.

Lawrence S., Giles C.L., Tsoi A.C., Back A.D., (1997) Face recognition: A convolutional neural-network approach. *IEEE Trans. Neural Netw.* 8: 98–113.

LeCun Y, Bengio Y., Hinton G. (2015) Deep learning. *Nature* 521: 436–444.

LeCun Y., Bottou L., Bengio Y., Haffner P., (1998) Gradient-based learning applied to document recognition. Proc. IEEE 86: 2278–2324.

Li P., Liu Y., Sun M. (2013) Recursive autoencoders for ITG-based translation, in: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Seattle, Washington, USA, pp. 567–577.

Liu H.B., Chen W., Sudjianto A. (2006) Relative entropy based method for probabilistic sensitivity analysis in engineering design. *J. Mech. Des.* 128(3): 326–33.

Liu Z., Zhang Z. (2018) Artificial neural network based method for seismic fragility analysis of steel frames. *KSCE Journal of Civil Engineering* 22(2): 708–717.

Loth C., Baker J.W. (2011) A spatial cross-correlation model of spectral accelerations at multiple periods. *Earthq. Eng. Struct. Dyn.* 42: 397–417.

Ma F., Zhang H., Bockstedte A., Foliente G.C., Paevere P. (2004) Parameter analysis of the differential model of hysteresis. *Transactions of ASME* 71(3): 342–349.

McKay M.D., Beckman R.J., Conover W.J. (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21: 239–245.

Marafi N.A., Berman J.W., Eberhard M.O. (2016) Ductility-dependent intensity measure that accounts for ground-motion spectral shape and duration. *Earthq. Eng. Struct. Dyn.* 45: 653–672.

Mazzoni S., Mckenna F., Scott M.H., Fenves G.L. (2006) The Open System for Earthquake Engineering Simulation (OpenSEES) user command-language manual.

Miller M., Baker J.W. (2015) Ground-motion intensity and damage map selection for probabilistic infrastructure network risk assessment using optimization. *Earthquake Engineering and Structural Dynamics* 44(7): 1139–1156.

Miranda E., Bertero V.V. (1994) Evaluation of strength reduction factors for earthquake-resistant design. *Earthq. Spectra* 10: 357–379.

Mitropoulou C.C., Papadrakakis M. (2011) Developing fragility curves based on neural network IDA predictions. *Eng. Struct.* 33: 3409–3421.

Möller O., Foschi R., Quiroz L., Rubinstein M. (2009) Structural Optimization for Performance-Based Design in Earthquake Engineering: Applications of Neural Networks. *Struct. Saf.* 31: 490–499.

Murphy K.P. (2012) Machine learning: A probabilistic perspective. MIT Press.

Nair, V., Hinton, G.E., n.d. Rectified Linear Units Improve Restricted Boltzmann Machines 8.

Nakamura, M., Masri, S.F., Chassiakos, A.G., Caughey, T.K., 1998. A Method for Non-Parametric Damage Detection through the Use of Neural Networks. *Earthq. Eng. and Struct. Dyn.* 27, 997–1010.

Nassar, A., Krawinkler, H., 1991. Seismic Demands for SDOF and MDOF Systems, John A. Blume Earthquake Engineering Center Report No. 95, Stanford University, CA.

Ning C.L., Yu B., Li B. (2016) Beam-column joint model for nonlinear analysis of non-seismically detailed reinforced concrete frame. *Journal of Earthquake Engineering* 20(3): 476–502.

NRCC. (2005) National Building Code of Canada 2005. Ottawa, Canada: NRCC.

Pan S.J., Yang Q., (2010) A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* 22: 1345–1359.

Pelliciari M., Marano G.C., Cuoghi T., Briseghella B., Lavorato D., Tarantino A.M. (2018) Parameter identification of degrading and pinched hysteretic systems using a modified Bouc-Wen model. *Structure and Infrastructure Engineering* 14(12): 1573–1585.

Power M., Chiou B., Abrahamson N., Bozorgnia Y., Shantz T., Roblee C. (2008) An overview of the NGA project. *Earthq. Spectra* 24: 3–21.

Reddi S.J., Kale S., Kumar S. (2018) On the convergence of Adam and beyond. ICLR 2018 Conference Blind Submission.

Riddell R. (2007) On ground motion intensity indices. *Earthq. Spectra* 23: 147–173.

Ruiz-Garcia J., Miranda E. (2007) Probabilistic estimation of maximum inelastic displacement demands for performance-based design. *Earthquake Eng. and Struct. Dyn.* 36: 1235–1254.

Sainath T.N., Kingsbury B., Saon G., Soltau H., Mohamed A.R., Dahl G., Ramabhadran B. (2015) Deep convolutional neural networks for large-scale speech tasks. *Neural Netw.* 64: 39–48.

Samuel A.L. (1959) Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development* 3(3): 535–554.

Schmidhuber J. (2015). Deep learning in neural networks: An overview. *Neural Netw.* 61: 85–117.

Shaw B.E. (2009). Constant stress drop from small to great earthquakes in magnitude-area scaling. *Bulletin of the Seismological Society of America* 99(2A): 871–875.

Shiraki N., Shinozuka M., Moore J.E., Chang S.E., Kameda H., Tanaka S. (2007) System risk curves: Probabilistic performance scenarios for highway networks subject to earthquake damage. *Journal of Infrastructure Systems* 13(1): 4354.

Shome N., Cornell C.A., Bazzurro P., Carballo J.E. (1988) Earthquakes, records, and nonlinear responses. *Earthquake Spectra* 14(3): 469-500.

Shwartz-Ziv R., Tishby N. (2017) Opening the black box of deep neural networks via information. ArXiv170300810 Cs.

Sobol I.M. (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* 55: 271-280.

Socher R., Perelygin A., Wu J.Y., Chuang J., Manning C.D., Ng A.Y., Potts C., (2013) Recursive deep models for semantic compositionality over a sentiment treebank. Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing: 1631–1642.

Song J., Der Kiureghian A. (2006) Generalized Bouc–Wen model for highly asymmetric hysteresis. *J. Eng. Mech.* 132: 610–618.

Tothong P., Cornell C.A. (2006) An empirical ground-motion attenuation relation for inelastic spectral displacement. *Bull. Seismol. Soc. Am.* 96: 2146–2164.

Vamvatsikos D., Cornell C.A. (2002) Incremental dynamic analysis. *Earthquake Engineering and Structural Dynamics* 31(3): 491–514.

Vamvatsikos, D., Cornell C.A. (2006) Direct estimation of the seismic demand and capacity of oscillators with multi-linear static pushovers through IDA. *Earthq. Eng. and Struct. Dyn.* 35: 1097–1117.

Vamvatsikos, D., Fragiadakis M. (2010) Incremental dynamic analysis for estimating seismic performance sensitivity and uncertainty. *Earthq. Eng. and Struct. Dyn.* 39: 141–163.

Veletsos A.S., Newmark N.M. (1960) Effect of inelastic behaviour on the response of simple systems to earthquake motions. Proceedings of the 2nd World Conference on Earthquake Engineering, Tokyo, Japan, 895–912.

Wang Z., Pedroni N., Zentner I., Zio E. (2018) Seismic fragility analysis with artificial neural networks: Application to nuclear power plant equipment. *Engineering Structures* 162: 213–225.

Wen Y.K. (1976) Method for random vibration of hysteretic systems. *J. Eng. Mech. Div.* 102: 249–263.

Xie Y., Ebad Sichani M., Padgett J., Desroches R. (2020) The promise of implementing machine learning in earthquake engineering: A state-of-the-art review. *Earthquake Spectra:* 1-33.

Xu B., Wu Z., Chen G., Yokoyama K. (2004) Direct identification of structural parameters from dynamic responses with neural networks. *Engineering Applications of Artificial Intelligence* 17(8): 931–943.

Yin Y.J., Li Y. (2010) Seismic collapse risk of light-frame wood construction considering aleatoric and epistemic uncertainties. *Struct. Saf.* 32: 250–261.

Yosinski J., Clune J., Bengio Y., Lipson H. (2014) How transferable are features in deep neural networks? ArXiv14111792 Cs.

Zhang T., Ramakrishnan R., Livny M. (1996) "BIRCH: An efficient data clustering method for very large databases." In: Proc.1996 ACM SIGMOD Internat. Conf. on Management of data, 25(2): 103–114.

Zhu M., McKenna F., Scott M.H. (2018) OpenSeesPy: Python library for the OpenSees finite element framework. SoftwareX 7: 6–11.

Appendix

Appendix A. Incremental response equation

To compute the stress for a given displacement, incremental response equations need to be first identified. According to Eq. (4.1), the restoring force at time $t_{(i+1)}$ is obtained as

$$f(u, z)_{(i+1)} = \alpha k_0 u_{(i+1)} + (1 - \alpha) F_y z_{(i+1)} \tag{A.1}$$

The rate equation for z in Eq. (A.1) is discretized by a backward Euler solution scheme as follows:

$$z_{(i+1)} = z_{(i)} + \Delta t \frac{h_{(i+1)}}{\eta_{(i+1)}} \times$$

$$\left[1 - \left| z_{(i+1)} \right|^n \left\{ \beta \cdot \text{sgn}\left(\frac{(u_{(i+1)} - u_{(i)})}{\Delta t} z_{(i+1)} \right) + \gamma \right\} v_{(i+1)} \right] \tag{A.2}$$

$$\times \frac{(u_{(i+1)} - u_{(i)})}{\Delta t} \frac{k_0}{F_y}$$

where $\varepsilon_{(i+1)}$ is estimated by discretization of the Eq. (4.3) using the backward Euler scheme again:

$$\varepsilon_{(i+1)} = \varepsilon_{(i)} + \Delta t (1 - \alpha) \frac{k_0}{F_y} z \frac{(u_{(i+1)} - u_{(i)})}{\Delta t} \tag{A.3}$$

The incremental responses equations are estimated by Newton-Raphson algorithm for incremental displacement $\left(u_{(i+1)} - u_{(i)} \right)$. The details of the procedure are summarized as follows:

Step 1. Evaluate function $\Gamma(z_{(i+1)})$.

$\Gamma(z_{(i+1)})$ is given as

$$\Gamma(z_{(i+1)}) = z_{(i+1)} - z_{(i)} - h_{(i+1)}a_2(u_{(i+1)} - u_{(i)})\frac{k_0}{F_y} \qquad (A.4)$$

in which a_2 and $h_{(i+1)}$ are defined as

$$a_2 = \frac{1 - |z_{(i+1)}|^n a_1 v_{(i+1)}}{\eta_{(i+1)}}$$

$$a_1 = \beta \cdot \text{sgn}\left(\frac{(u_{(i+1)} - u_{(i)})}{\Delta t}z_{(i+1)}\right) + \gamma$$

$$h_{(i+1)} = 1 - \zeta_{1,(i+1)}$$

$$\cdot \exp\left(-\left(z_{(i+1)} \cdot \text{sgn}\left(\frac{(u_{(i+1)} - u_{(i)})}{\Delta t}\right) - qz_{u,(i+1)}\right)^2 / \zeta_{2,(i+1)}^2\right)$$

where,

$$v_{(i+1)} = 1 + \delta_v \varepsilon_{(i+1)}$$

$$\eta_{(i+1)} = 1 + \delta_\eta \varepsilon_{(i+1)}$$

$$\zeta_{1,(i+1)} = \zeta_0\left(1 - \exp(-p\varepsilon_{(i+1)})\right) \qquad (A.6)$$

$$\zeta_{2,(i+1)} = \left(\psi + \delta_\psi \varepsilon_{(i+1)}\right)\left(\lambda + \zeta_{1,(i+1)}\right)$$

$$z_{u,(i+1)} = \left(v_{(i+1)}(\beta + \gamma)\right)^{-1/n}$$

$\varepsilon_{(i+1)}$ can be obtained by using Eq. (A.3).

Step 2. Estimate the derivative of $\Gamma\left(z_{(i+1)}\right)$ with respect to $z_{(i+1)}$

The derivative of $\Gamma\left(z_{(i+1)}\right)$ with respect to $z_{(i+1)}$ is written as

$$\Gamma'\left(z_{(i+1)}\right) = 1 - \left(h'_{(i+1)}a_2 + h_{(i+1)}a'_2\right)\left(u_{(i+1)} - u_{(i)}\right)\frac{k_0}{F_y} \tag{A.7}$$

where

$$h'_{(i+1)} = a_3\left(\zeta'_{1,(i+1)} - a_4 + \zeta'_{2,(i+1)}a_5\right)$$

$$a'_2 = \frac{1}{\eta^2_{(i+1)}}\cdot\left(-\eta'_{(i+1)}\left(1 - \left|z_{(i+1)}\right|^n a_1 v_{(i+1)}\right)\right.$$

$$ - \eta_{(i+1)}\left(n\left|z_{(i+1)}\right|^{n-1} a_1 v_{(i+1)}\cdot sgn\left(z_{(i+1)}\right)\right) \tag{A.8}$$

$$ \left. + \left|z_{(i+1)}\right|^n a_1 v'_{(i+1)}\right)\right)$$

in which

$$a_3 = -\,exp\left(-\frac{\left(z_{(i+1)}\cdot sgn\left(\frac{\left(u_{(i+1)} - u_{(i)}\right)}{\Delta t}\right) - qz_{u,(i+1)}\right)^2}{\zeta^2_{2,(i+1)}}\right)$$

$$a_4 = 2\zeta_{1,(i+1)}\left(z_{(i+1)}\cdot sgn\left(\frac{\left(u_{(i+1)} - u_{(i)}\right)}{\Delta t}\right) - qz_{u,(i+1)}\right)\times \tag{A.9}$$

$$\left(sgn\left(\frac{\left(u_{(i+1)} - u_{(i)}\right)}{\Delta t}\right) - qz'_{u,(i+1)}\right)/\zeta^2_{2,(i+1)}$$

$$a_5 = 2\zeta_{1,(i+1)}\left(z_{(i+1)}\cdot sgn\left(\frac{\left(u_{(i+1)} - u_{(i)}\right)}{\Delta t}\right) - qz_{u,(i+1)}\right)^2/\zeta^3_{2,(i+1)}$$

and

$$z'_{u,(i+1)} = -\frac{v'_{(i+1)}(\beta + \gamma)}{n}\left(v_{(i+1)}(\beta + \gamma)\right)^{-\frac{n+1}{n}}$$

$$v'_{(i+1)} = \delta_v \varepsilon'_{(i+1)}; \quad \eta'_{(i+1)} = \delta_\eta \varepsilon'_{(i+1)}$$

$$\zeta'_{1,(i+1)} = \zeta_0 p\, exp\left(-p\varepsilon_{(i+1)}\right)\varepsilon'_{(i+1)}; \tag{A.10}$$

$$\zeta'_{2,(i+1)} = \psi\zeta'_{1,(i+1)} + \lambda\delta_\psi\varepsilon'_{(i+1)} + \delta_\psi\varepsilon'_{(i+1)}\zeta_{1,(i+1)} + \delta_\psi\varepsilon_{(i+1)}\zeta'_{1,(i+1)}$$

$$\varepsilon'_{(i+1)} = (1-\alpha)\frac{k_0}{F_y}\left(u_{(i+1)} - u_{(i)}\right)$$

Step 3. Update $z_{(i+1)}$

The trial value in the Newton Raphson scheme is obtained as follows:

$$z^{new}_{(i+1)} = z_{(i+1)} - \frac{\Gamma\left(z_{(i+1)}\right)}{\Gamma'\left(z_{(i+1)}\right)} \tag{A.11}$$

Then, update $z_{(i+1)}$ and store the old value for the convergence check

$$z^{old}_{(i+1)} = z_{(i+1)}$$

$$z_{(i+1)} = z^{new}_{(i+1)} \tag{A.12}$$

Step 4: Iterate

Iterate from Step 1 to Step 4 until $\left|z^{old}_{(i+1)} - z^{new}_{(i+1)}\right|$ is smaller than the prescribed tolerance.

Appendix B. Regression-based fragility estimation

Various methods have been developed to estimate structural fragilities. Strong earthquake ground motions are required to properly understand the behavior of structural system at the collapse level, but in general it is hard to obtain such records. To address this issue, some methods intentionally scale up relatively weak ground motions, while others relies on the relationship between intensity measure (IM), e.g., PGA, and engineering demand parameter (EDP) estimated by the dataset obtained using relatively weak ground motions. The latter approach often uses the following linear regression model of $\ln EDP$ on $\ln IM$ with homoscedasticity assumption (Cornell et al., 2002):

$$\ln EDP = a \cdot \ln IM + b + \sigma \cdot \varepsilon \tag{B.1}$$

where a and b are the coefficients obtained from the regression analysis, σ represents the conditional standard deviation, and ε is the standard Gaussian random variable. Following the regression function, the conditional mean $\mu_{\ln EDP}$ and variance $\sigma_{\ln EDP}^2$ of $\ln EDP$ given IM are obtained as $a \cdot \ln IM + b$ and σ^2, respectively. As a result, the probability that structural demand exceeds the structural capacity C at given IM value x is given as:

$$P(C - \ln EDP \leq 0 | \ln IM = \ln x)$$

$$= \Phi\left(-\frac{\mu_c - \mu_{\ln EDP}(x)}{\sqrt{\sigma_c^2 + \sigma_{\ln EDP}^2 - 2\rho\sigma_c\sigma_{\ln EDP}}}\right) \tag{B.2}$$

where μ_c and σ_c denotes the mean and standard deviation of the structural capacity, and ρ denotes the correlation coefficient between C and $\ln EDP$. When the uncertainty in the capacity is ignored, the fragility in Eq. (B.2) is simplified as

$$P(x) = \Phi\left(-\frac{\mu_c - \mu_{\ln EDP}(x)}{\sigma_{\ln EDP}}\right) \tag{B.3}$$

Appendix C. Probabilistic regional loss assessment of urban community under earthquake hazard

The natural logarithm of the intensity measure (IM) of the ground motion at site i caused by an earthquake event j can be described as

$$ln\, y_{ij} = f\left(M_j, R_{ij}, \theta_i\right) + \sigma_{ij}\varepsilon_{ij} + \tau_j\eta_j \qquad (C.1)$$

where $f\left(M_j, R_{ij}, \theta_i\right)$ denotes the attenuation law given as a function of magnitude M_j, seismological distance from the epicenter, R_{ij} and a set of other explanatory parameters θ_i. The variables, ε_{ij} and η_j, respectively represent the intra- and inter-event residuals, which are modelled as the Gaussian random variables with zero mean and standard deviation of one. The magnitudes of the uncertainties are represented by σ_{ij} and τ_j for intra- and inter-event, respectively. For a given earthquake event j, η_j is constant at all site, whereas ε_{ij} varies from site to site while showing spatial correlation. Such dependency is often modeled as an auto- and cross-correlation models which are given as functions of the distance between sites. Eq. (C.1) is often termed as a ground motion prediction equation (GMPE). In this study, four random variables $X = \{M_j, R_{ij}, \varepsilon_{ij}, \eta_j\}$ are introduced to represent the uncertainties in the spatial distribution of IM in the target region.

The annual probability that the regional seismic loss L exceeds the threshold l is calculated as the sum of the annual exceedance probabilities for seismic sources or faults $s = 1, \dots, S$, i.e.,

$$P(L \geq l) = \sum_{s=1}^{S} \lambda_s \cdot P(L \geq l | s) \qquad (C.2)$$

where λ_s denotes the annual occurrence rate of an earthquake event by source s,

and $P(L \geq l|s)$ is the conditional probability of exceedance given the earthquake event by source s. Eq. (C.2) can be rewritten in terms of the four random variables $\boldsymbol{X}$ as:

$$
\begin{aligned}
P(L \geq l) &= \sum_{s=1}^{S} \left[\lambda_s \int_{L(\boldsymbol{x}) \geq l} f_{\boldsymbol{X}|s}(\boldsymbol{x}) \, d\boldsymbol{x} \right] \\
&= \sum_{s=1}^{S} \left[\lambda_s \int_{L(\boldsymbol{x}) \geq l} f_{\boldsymbol{X}|s}(M, R, \eta, \varepsilon) \, dMdRd\eta d\varepsilon \right]
\end{aligned}
\tag{C.3}
$$

where $f_{\boldsymbol{X}|s}(\cdot)$ is the conditional probability density function (PDF) of the random variables $\boldsymbol{X} = \{M_j, R_{ij}, \varepsilon_{ij}, \eta_j\}$ given source s, and $L(\boldsymbol{x})$ denotes the functional relationship between the regional seismic loss L and the random variable $\boldsymbol{x}$ (see Section 5.4 for details). Introducing the annual occurrence rate of an earthquake event, i.e., $\lambda_E = \sum_{s=1}^{S} \lambda_s$, the exceedance probability can be described in terms of the conditional PDF given an earthquake event, $f_{\boldsymbol{X}|E}(\boldsymbol{x}) = \sum_{s=1}^{S} (\frac{\lambda_s}{\lambda_E}) f_{\boldsymbol{X}|s}(\boldsymbol{x})$, i.e.,

$$
P(L \geq l) = \lambda_E \int_{L(\boldsymbol{x}) \geq l} \left[\sum_{s=1}^{S} \frac{\lambda_s}{\lambda_E} f_{\boldsymbol{X}|s}(\boldsymbol{x}) \right] d\boldsymbol{x} = \lambda_E \int_{L(\boldsymbol{x}) \geq l} f_{\boldsymbol{X}|E}(\boldsymbol{x}) \, d\boldsymbol{x}
\tag{C.4}
$$

Based on Eq. (C.4), one can estimate the annual probability of exceedance by a Monte Carlo simulation (MCS), i.e.,

$$
\begin{aligned}
P(L \geq l) &= \lambda_E \cdot \int_{\boldsymbol{x}} I(L(\boldsymbol{x}) \geq l) \cdot f_{\boldsymbol{X}|E}(\boldsymbol{x}) d\boldsymbol{x} \\
&= \lambda_E \cdot E_{\boldsymbol{X}|E}[I(L(\boldsymbol{x}) \geq l)] \\
&\approx \lambda_E \cdot \frac{1}{N} \sum_{n=1}^{N} I(L(\boldsymbol{x}_n) \geq l)
\end{aligned}
\tag{C.5}
$$

where $I(L(\boldsymbol{x}) \geq l)$ denotes an indicator function which gives "1" if the given

realization of random variables makes the loss exceed the threshold l, otherwise "0",

and x_n is n^{th} sample generated from the conditional PDF $f_{X|E}(x)$, $n = 1, ..., N$.

Eq. (C.5) provides the loss curve, i.e., the complementary cumulative distribution

function (CCDF) of the total loss of the urban community.

초 록

김태용

건설환경공학부

서울대학교 대학원

도시 인프라 시스템의 복잡성이 증가함에 따라 지진에 대한 사회 경제적 취약도 역시 나날이 증가하고 있다. 따라서 향후 발생할 지진 재해에 의한 피해를 경감하고 도시 인프라 시스템의 복원력을 증대하기 위해서는 구조물 내진성능의 정확한 평가가 필수적이다. 이론적으로 구조물의 붕괴 및 파괴 지지력(Structural Capacity against Collapse and Failures)의 평가를 위해서는 여러 지진동을 이용한 매우 정밀한 비선형 구조물의 거동 이력을 필요로 한다. 그러나 비선형 시간이력 해석은 높은 계산 비용과 모델링 시간을 필요로 하기에 현재의 내진설계기준에서는 간단 회귀식 기반 방법론들이 이를 대체하여 주로 사용된다. 회귀식 기반 방법론들은 상대적으로 적은 수의 변수와 지진 응답 이력을 바탕으로 개발되었기에 응답 예측의 정확도가 높지 않다. 더불어 기존 방법론들은 지진동이 다르더라도 선택한 지진강도척도(Intensity Measures)의 값이 같다면 항상 같은 값을 추정하게 되어, 구조물 지진 응답의 변동성 정량화에 한계를 보인다. 본 학위논문에서는 비선형 시간이력을 수행하지 않고 효과적으로 구조물의 지진 응답 값 및 그 불확실성을 추정하기 위해 심층 인공 신경망, 즉 딥러닝(Deep Learning)을 도입하고자 한다.

지진에 의해 구조물이 진동하는 자연현상을 바탕으로 지진 및 구조물의 정보가 주어졌을 때 구조물의 지진 응답을 예측할 수 있는 심층 인공 신경망 모델(Deep Neural Network; DNN model)의 구조를 제안하였다. DNN model에는 Convolutional Neural Network (CNN)가 사용되어 구조물의 이력곡선으로부터 응답 예측에 필요한 주요 변수들을 추출하게 된다. 세 가지 이상화된 이력곡선과 미국 NGA-West

데이터베이스의 지진동을 활용해 비선형 동적해석을 수행하여 데이터베이스를 구축하였고 이를 이용해 제안한 DNN model을 학습시켰다. 수치 예제를 통해 학습된 DNN model은 기존 회귀모델기반 방법론에 비해 구조물의 지진 응답 예측 시 매우 높은 수준의 정확도를 보임을 확인하였다.

앞서 개발된 DNN model은 매우 높은 정확도를 보임을 확인하였으나 구조 시스템의 비선형 거동에 따른 예측 불확실성을 산정하지 못한다는 한계를 가진다. 이를 해결하기 위하여 베이지안 딥러닝 기법을 기반으로 구조물의 지진 응답을 확률론적으로 예측할 수 있는 확률론적 심층 인공 신경망 모델(Probabilistic Deep Neural Network; P-DNN model)을 제안하였다. 자연로그를 취한 가우스 분포 함수의 음의 값을 인공 신경망 모델의 손실함수로 사용하여 앞서 구축한 데이터베이스를 바탕으로 P-DNN model을 학습시켰다. 새로운 손실함수를 도입함으로써 구조물의 지진 응답에 대한 평균과 표준편차를 예측할 수 있으며, 입력변수의 불확실성 정도에 따라 표준편차의 크기가 달라짐을 확인하였다. 여러 수치 예제를 통해 개발된 모델이 응답의 불확실성을 효율적으로 정량화함을 증명하였고, 구조물의 불확실성은 비선형 거동의 정도가 커질수록 그 정도가 심해지기에 개발된 방법론은 구조물의 파괴나 붕괴를 다루는 내진공학 측면에서 매우 중요한 역할을 할 것이다.

DNN model과 P-DNN model은 기존 단순 회귀식 기반 모델에 비하여 뛰어난 예측 정확도를 보였지만, 이상화된 이력곡선을 기반으로 학습되었다. 다시 말해, 강도/강성 감소 및 핀칭 효과를 보이는 실제 구조 시스템의 지진 응답을 학습된 심층 인공 신경망을 통해 정확히 예측하기는 쉽지 않다. 이러한 문제를 해결하기 위하여, 먼저 구조물 이력곡선에 강도/강성 감소 및 핀칭 효과를 표현할 수 있는 modified Bouc-Wen-Baber-Noori (m-BWBN) 모델을 제안하였다. m-BWBN 모델과 미국 NGA-West 데이터베이스의 다양한 지진동을 사용하여 비선형 동적해석을 수행을 통해 데이터베이스를 구축 하였으며, 상대적으로 많은 정보를 담고 있는 이력곡선의 응답을 예측할 수 있는 BW-DNN

model을 제안 및 학습시켰다. 학습된 BW-DNN model은 기존 DNN model과 P-DNN모델에서 다루지 못한 광범위한 구조 시스템의 지진재해에 대한 응답을 예측할 수 있음을 확인하였다.

개발된 방법론의 효율성과 적용가능성을 입증하기 위해 구조 부재의 지진 응답 예측, 구조 시스템의 지진 취약도 추정, 도시의 지진 손실 평가 등 세 가지 서로 다른 규모의 지진 공학 문제를 도입하였다. 더불어 파이썬(Python)으로 개발된 심층 인공 신경망 소스 코드, 학습된 모델, 구축된 데이터베이스, 그리고 제안된 딥러닝 기반 예측 플랫폼을 제공하는 새로운 웹 서비스(http://ERD2.snu.ac.kr)를 개발하여 사용자의 편의성 및 방법론의 확장성을 확보하였다. 본 학위논문에서 개발된 딥러닝 기반 비선형 응답 방법론은 향후 일어날 지진에 따른 관심 구조물의 응답을 효율적으로 예측함으로써 지진 재해에 대한 피해 예측 및 복구를 위한 합리적인 대비에 기여할 것이다.

주요어: 지진 공학; 딥러닝; 합성곱 신경망; 베이지안 딥러닝; 단자유도 시스템; 북웬 모델; 도시의 지진 손실 평가; 취약도 해석; 불확실성; 비선형 시간 이력 해석.

학번: 2017−37449