Prediction of Site Effects and Ground Motion Parameters through Deep Neural Networks Using Horizontal-to-Vertical Spectral Ratios

(水平/上下スペクトル比を用いた深層ニューラルネットワークによる地盤特性と地震動強さの予測に関する研究)

DA PAN（潘 達）

Graduate School of Advanced Sciences and Engineering

Hiroshima University

Doctor of Philosophy

September 2024

Supervised by Prof Hiroyuki Miura

Abstract

Ground motion models (GMM) are key tools for predicting seismic ground motions in future earthquakes and are also critical for regional disaster planning and seismic-resistant building design. Seismic ground motion models are influenced by source characteristics, propagation path effects, and site effects (also known as site amplification factors, SAFs). SAFs play a significant role in controlling the amplitudes of seismic waves at the surface. One method to evaluate SAFs is the generalized spectral inversion technique (GIT), which derives source characteristics, propagation path effects, and SAFs from Fourier spectral data recorded at multiple sites. However, the GIT is limited to strong seismic records and cannot be applied to sites without such observations. Boreholes can be used to obtain information on subsurface soil stratification. After obtaining the one-dimensional site model, it is assumed that seismic waves propagate only in the vertical direction and pass through the horizontally stratified medium. This assumption is known as the 1D SH assumption. Subsequently, the propagation of S-waves through each soil layer is simulated, and the site effect is evaluated by multiple reflection theory. However, this approach requires detailed site subsurface soil modeling.

Recently, researchers have explored SAF estimation methods based on the horizontal-vertical spectral ratio (HVR). Since the proposal of the HVR method, the relationships between site effects and HVR have been discussed based on microtremor and earthquake observation data in various regions. The predominant frequency of HVR coincided with those of SAF at most sites, and the spectral shape of HVR and SAF were similar. However, significant discrepancies between the amplitude of MHVR and SAF were found in most cases. As a result, this study explores the development and application of deep neural network (DNN) models for predicting site effects and ground motion parameters using HVR data. The study is structured into seven chapters, each addressing a specific aspect of the research.

Chapter 1 provided a comprehensive background on earthquakes, emphasizing the

importance of accurate ground motion prediction and the critical role of site effects. It discusses the significance of horizontal-to-vertical spectral ratios (HVRs) in assessing site effects and outlines the objectives and structure of the dissertation. This chapter sets the stage by establishing the importance of predicting ground motion and understanding site effects using HVR.

Chapter 2 explored the first objective of the study, develop a robust deep neural network (DNN) model utilizing microtremor horizontal-to-vertical spectral ratios (MHVR) for assessing SAFs. It details the data collection, model construction, and validation processes. The DNN model demonstrated superior performance in estimating SAFs compared to traditional methods, showcasing the potential of using MHVR data for site effects assessment. The developed DNN model does not require any hard-to-get data such as seismic velocity structures and damping models thus providing a significant cost-benefit.

Chapter 3 expanded on this foundation by incorporating transfer learning techniques to adapt the pre-trained DNN model for new regions with varying geological conditions. It discusses the compilation of SAFs and MHVRs from various locations and the construction of the transfer learning model. Transfer learning significantly improved the performance in estimating SAFs in data-limited areas, demonstrating the feasibility of extending the applicability of the DNN model beyond its initial training region.

Chapter 4 introduced a model to predict pseudo-EHVR (pEHVR) from MHVR, addressing the challenge of obtaining reliable EHVR data in regions without direct seismic measurements. It discusses the comparison between MHVR and EHVR, model construction, and validation. The model successfully predicted pEHVR from MHVR with high accuracy. Future research should focus on improving high-frequency predictions and expanding the dataset to enhance generalizability.

Chapter 5 proposed the DNN model for predicting seismic ground motion parameters using EHVR. It integrates EHVR into existing ground motion prediction equations to improve prediction accuracy. The DNN model incorporating EHVR

outperformed traditional models in predicting ground motion parameters, particularly for spectral amplitude and shape. However, the availability of EHVR data and potential data leakage are limitations that need addressing in future research.

Chapter 6 summarized the findings of the previous chapters and discusses potential areas for future research. The research demonstrated the effectiveness of using HVR data for seismic risk assessment and ground motion prediction. Future work should focus on improving prediction accuracy, validating the models in diverse geological settings, and exploring the integration of additional seismic features to enhance model robustness.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

As the opening chapter of the entire paper, this section first provides some background on earthquakes and related seismic damage, including ground motion. It then discusses some key factors influencing the prediction of ground motion caused by earthquakes. Subsequently, it introduces the main focus of this paper, site effects, including their concept and methods of acquisition. Following this, the second focus of this thesis, the significance of horizontal-to-vertical spectral ratio, and its application are introduced. Finally, the paper briefly outlines its objectives and overall structure.

1.1. Background

An earthquake is a sudden vibration or shaking of the Earth's surface, usually caused by a sudden release of energy from underground rock formations. This release of energy results in the propagation of seismic waves, which cause the ground motion. Earthquake-related hazards can be categorized into several key areas:

1. Building damage: Strong ground motion by an earthquake can cause buildings to be damaged or even collapse. Older or inadequately designed buildings are more likely to be damaged. The main cause of the structural damage is the build-up of excess stress within structural components during ground motion, which eventually exceeds the strength of the materials and leads to their failure. Besides the failure of structural components, various other parts (such as, pipes, valves, wires, and decorative elements) can also be damaged during the motion process.
2. Infrastructure destruction: Ground motion can severely impact critical infrastructure, including roads, bridges, railways, airports, and ports, disrupting transportation and complicating rescue and recovery efforts. Damage to power plants and energy pipelines can lead to power outages or hazardous leaks, further intensifying the crisis.
3. Personal injury and death: Ground motion may result in injuries and deaths from collapsed houses, flying debris and so on. Confusion and panic during an earthquake may also lead to injuries during evacuation.
4. Liquefaction of ground motion: Liquefaction phenomenon refers to that the sudden increase in water pressure in saturated, loose soil due to strong ground motion in earthquakes reduces the contact forces between soil particles. Liquefaction can seriously affect the foundation bearing capacity of a building, which in turn can lead to tilting or even collapse of the building.
5. Secondary hazards: Earthquake can trigger other hazards such as fires, explosions, landslides, mudslides, and tsunamis.

Among the various types of hazards mentioned above, ground motion due to earthquakes is usually the main cause of damage. Therefore, studying ground motion prediction can help people in a wide range of fields to better understand and manage the risk of earthquakes, potentially reducing injuries and damage when earthquakes occur.

The ground motion prediction, is a model used to predict the strength and characteristics of ground motion when an earthquake occurs, also referred to as 'ground motion model'. Note that the model does not apply to predicting when and where earthquakes will occur.

To properly develop a ground motion model, it is first necessary to identify the different factors that are involved in the propagation of seismic waves from an earthquake source to sites, causing ground motion to occur at the sites. There are three main factors that impact ground motion: earthquake source, propagation path, and site effect, as shown in Figure 1-1.

Here earthquake source refers to the characteristics and the physical processes of the earthquake itself, including its magnitude, depth, faulting mechanism, and others. These factors determine the energy released at the source and the nature of the earthquake waves generated.

The propagation path refers to the movement of seismic waves within the earth. The energy generated by an earthquake source travels in the form of seismic waves through various media within the earth, reaching up to the surface. The different media in the propagation path affect the propagation velocity, intensity, period, and other important factors of seismic waves. Therefore, the propagation path is also considered in models predicting ground motion.

Site effects, sometimes referred to as site response, are the impacts of different geological and soil conditions on the propagation of seismic waves from bedrock to the surface. These conditions, such as soil type, soil thickness, water table elevation, and the type and structure of subsurface rock, can alter the propagation characteristics of seismic

waves. These effects can significantly influence the intensity, duration, frequency, and other important factors of vibrations propagating to the ground surface. Site effects have important implications for the structural safety and seismic design of buildings. Building engineers and seismologists need to consider these factors to better assess seismic risk and design safer buildings. For example, by conducting seismic micro zonation and analyzing seismic site effect [1], it is possible to more accurately predict the ground motion that different regions may experience during an earthquake. This thesis focuses on site effects.

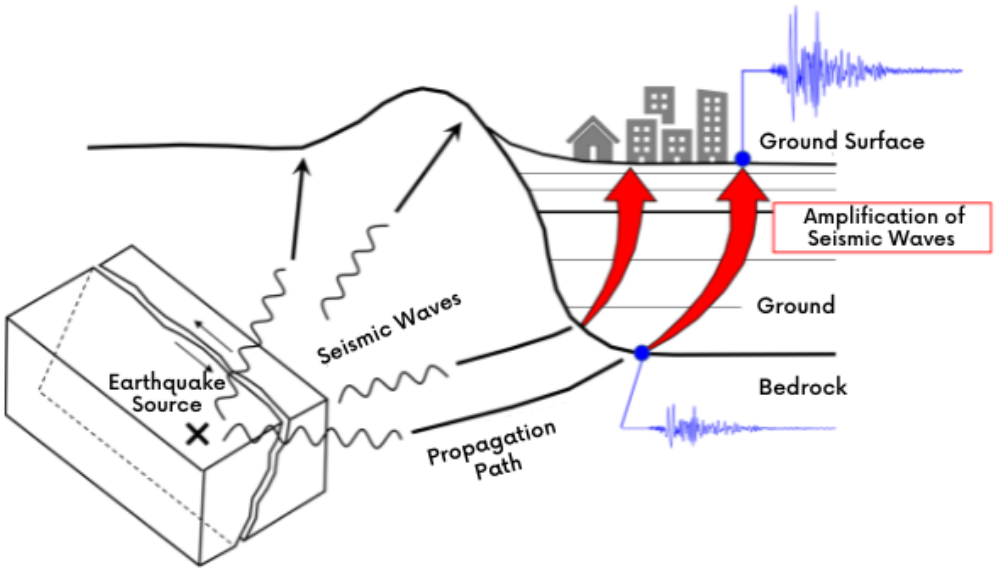

Figure 1-1. Three factors that impact the ground motion of earthquakes: source, path, and site effect. The red arrows indicate that seismic waves are impacted by site effect, which are the amplification of waves while travelling through the soil layer.

1.2. Literature Reviews

1.2.1. Site effects

The primary focus of this thesis is to study site effects—the third aspect of influencing seismic ground motion mentioned earlier (in Section 1.1). To better understand site effects, we first need to recognize the main components below the ground surface. As shown in Figure 1-1, in mountainous or plateau regions, the area below the surface is composed of bedrock, known as rock outcrops. In contrast, in plains or basin areas, the subsurface is made up of various types of topsoil and weathered rock layers. Depending on the properties of wave propagation, the amplitude is lower in hard soil layers and amplified in soft soils. Therefore, site effects often lead to the following phenomena: first and foremost, under certain specific geological conditions, the amplitude of seismic waves passing through the bedrock to the surface may be significantly amplified, a phenomenon known as site amplification effect (abbreviated as SAF in this thesis). This is especially true in basins or plains with soft soils or deep sedimentary layers. Secondly, as the waves pass through different geological structures, changes in the frequency components and duration of the seismic waves occur, affecting the characteristics of ground motion.

Currently, there are a variety of techniques for obtaining site effects. Boreholes can be used to obtain information on subsurface soil stratification. This information includes

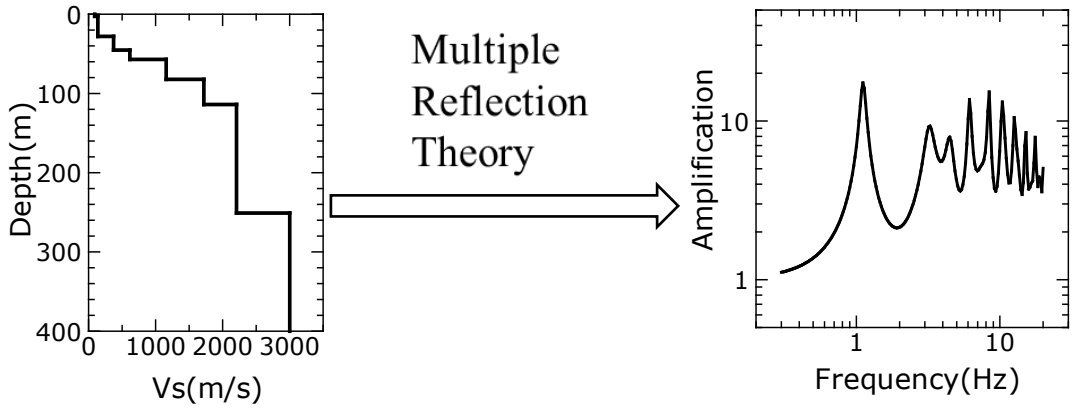

Figure 1-2. Site effect obtained used S-wave velocity structure model by multiple reflection theory.

the thickness, type, and physical properties of each soil layer. For each layer, parameters such as density, S-wave velocity (shear wave velocity), damping coefficient, and attenuation coefficient need to be determined. These parameters are essential for modeling the propagation of seismic waves through different media. S-wave velocity is a key parameter because it directly relates to the propagation speed of seismic waves and the magnitude of the amplification. After obtaining the one-dimensional site model, it is assumed that seismic waves propagate only in the vertical direction and pass through the horizontally stratified medium. This assumption is known as the 1DSH assumption [2]. Subsequently, the propagation of S-waves through each soil layer is simulated, and the site effect is evaluated by multiple reflection theory as shown in Figure 1-2. This approach requires detailed site subsurface soil modeling.

Another method involves recording seismic activity simultaneously at the surface and at the subsurface bedrock level [3]. By comparing the spectral ratios of the recordings from the surface and the bedrock, the site effects can be effectively quantified.

Generalized Spectral Inversion Technique (GIT) has been proposed as a method to separate seismic source, propagation, and amplification factors from seismic records [4,5]. This technique posits that the Fourier spectra of seismic records can be expressed as convolutions of the seismic source spectra, the damping characteristics of the propagation path, and the site amplification spectra. This approach assumes that the Fourier spectra of seismic records can be depicted as convolutions of the seismic source spectra, the damping characteristics of the propagation path, and site amplification spectra. By employing simultaneous equations derived from multiple seismic record spectra observed at various sites, GIT determines the source spectra, the Q-values of the seismic wave propagation path, and the amplification for each site. The amplification derived from multiple seismic records by GIT. As such, GIT enables the assessment of site amplification without the necessity for borehole data and site subsurface soil modeling. Recently, this technique has been utilized in several regions to evaluate the amplification factors at seismic observation sites, facilitated by the growing availability of seismic data

from the expansion of seismic observation networks [6–8]. However, it is important to note that it is impossible to obtain amplification factors at sites lacking seismic observation data.

1.2.2. Horizontal to Vertical spectral ratio

Horizontal to vertical spectral ratio (HVR) technique is a widely used technique in geophysical and seismological studies to estimate site effects, particularly site amplification, and to characterize subsurface geological conditions, by collecting and processing signals on the ground surface. This method proposed by Nakamura [9] in 1989, and overcomes the difficulties of the conventional use of seismic stations. It provides new ideas and methods for the assessment of site effects.

Nakamura [9] explained HVR technology based on two assumptions. The initial proposal by Nakamura [9] was that the HVR could to some extent replace the S-wave amplification factor for the horizontal component of the earthquake. The HVR derived from the data obtained at the surface can be expressed by Eq. (1-1):

$$HVR_S = \frac{H_S}{V_S} \tag{1-1}$$

Here H_S is the horizontal component of the record acquired on the ground surface and V_S is the corresponding vertical component, as shown in Figure 1-3. All are Fourier spectral records. Then it is assumed that the seismic waves propagate uniformly in the horizontal-vertical direction at the bedrock, as Eq. (1-2). The subscript B stands for bedrock.

$$HVR_B = \frac{H_B}{V_B} = 1 \tag{1-2}$$

Additionally, assuming that the vertical component of the seismic wave is not amplified,

it can be expressed as:

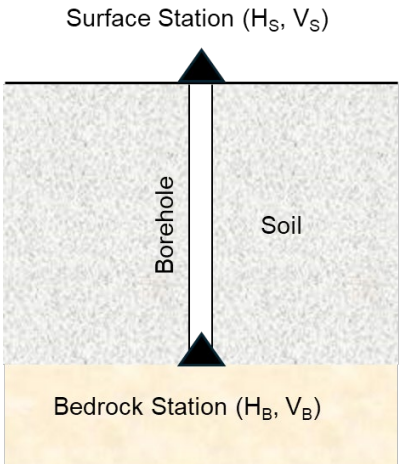

Figure 1-3. Concept of the HVR technique.

$$\frac{V_S}{V_B} = 1 \tag{1-3}$$

Then the amplification of seismic waves transmitted from bedrock to the ground surface can be expressed as the amplification of the horizontal component. Combining the above two equations, the following equation can be obtained. Based on these two assumptions, Nakamura [9] suggested that the HVRs can be equated to some extent with site effects.

$$\frac{H_S}{H_B} = \frac{H_S}{V_S} \times \frac{V_S}{V_B} \times \frac{V_B}{H_B} = HVR_S \tag{1-4}$$

The HVR method is currently being discussed and applied in some of the following areas.

1. Site Classification: HVR is widely used to classify sites based on their

predominant periods, crucial for assessing earthquake hazards. By comparing HVR curves with empirical curves, researchers can identify site categories and assess seismic risks[10,11]. Improvements and new methods, like spectral ratio curve matching and integration with local soil properties, enhance accuracy and applicability across different regions[12–14].

2. Site Effect Analysis: HVR helps in understanding site effects on ground motion during earthquakes. By analyzing HVR curves from ambient noise or earthquake recordings, researchers can estimate site amplification and resonance frequencies [15–17]. This method provides a cost-effective alternative to traditional spectral ratio methods, though it sometimes struggles with distinguishing source effects.

3. Inversion of Velocity Structure: HVR is employed to invert shear-wave velocity profiles, revealing subsurface conditions. Researchers use ambient noise or ground pulsation data to develop models of the shallow structure [18,19]. This application aids in site effect analysis by providing detailed subsurface information, essential for accurate seismic hazard assessments.

This paper will discuss the application of HVR to site effect analysis and ground motion prediction, with a more detailed discussion to follow.

1.3. Objectives

This paper aims to address several critical research topics to enhance site effect assessment and ground motion prediction accuracy.

The first objective is to develop a robust deep neural network (DNN) model utilizing microtremor horizontal-to-vertical spectral ratios (MHVR) for assessing site effects. Accurate assessment of site effects is essential for predicting earthquake impacts on specific locations, as traditional methods usually require a sufficiently large number of ground-shaking records or site subsurface formations. The MHVR-based DNN model does not require complex information and is available at any site.

Expanding applicability of the model is another significant goal. By incorporating data from diverse geographical regions and applying transfer learning techniques, the research aims to ensure that the model can effectively adapt to different geological settings. This approach enhances the generalizability of the model, making it a versatile tool for global site effect assessment by using MHVR.

Additionally, this study seeks to explore the relationship between microtremor HVR (MHVR) and earthquake HVR (EHVR) and develop a predictive model for evaluating EHVR from MHVR data. EHVR provides critical information about ground motion behavior during earthquakes, but obtaining EHVR data requires earthquake occurrences, which may be infrequent. MHVR, in contrast, can be measured continuously and non-invasively. By establishing a predictive relationship between MHVR and EHVR, the study aims to leverage the more readily available MHVR data to infer EHVR, thereby enhancing seismic hazard assessments in regions without recent earthquake records.

Finally, the research aims to improve the accuracy of existing ground motion prediction models by integrating EHVR data directly into ground motion prediction equations (GMPEs). Current GMPEs often rely on proxy parameters such as V_{S30} to account for site effects, which may not fully capture the complexity of local site conditions. Incorporating EHVR data provides a more accurate representation of site

effects, leading to better predictions of ground motion parameters such as peak ground acceleration (PGA), peak ground velocity (PGV), and spectral acceleration (SA).

1.4. Thesis structure

0: Introduction

This chapter introduces the background on earthquakes and the associated seismic damage, focusing on ground motion and site effects. It discusses key factors influencing ground motion prediction and highlights the significance of HVR in site effect analysis. The chapter also outlines the objectives and structure of the thesis. The introduction sets the stage for the research by establishing the importance of accurately predicting ground motion and understanding site effects using HVR.

Chapter 2: Predicting Site Amplification Factors (SAFs) from MHVRs using a DNN Model

This chapter proposes a methodology for estimating S-wave site amplification factors (SAFs) from MHVRs using a DNN model. It details the data collection, model construction, and validation processes. The DNN model demonstrated superior performance in estimating SAFs compared to traditional methods, showcasing the potential of using MHVR data for site effects assessment.

Chapter 3: Expanding MHVR-to-SAF Prediction to Data-Limited Areas Using Transfer Learning

This chapter explores the use of transfer learning to adapt a pre-trained DNN model to new regions and datasets. It discusses the compilation of SAFs and MHVRs from various locations and the construction of the transfer learning model. Transfer learning significantly improved the performance in estimating SAFs in data-limited areas, demonstrating the feasibility of extending the applicability of the DNN model beyond its initial training region.

Chapter 4: Prediction of Earthquake HVR (EHVR) from MHVR using a DNN Model

This chapter focuses on developing a DNN model to predict pseudo-EHVR (pEHVR) from MHVR. It discusses the comparison between MHVR and EHVR, model

construction, and validation. The model successfully predicted pEHVR from MHVR with high accuracy, providing a valuable tool for seismic risk assessment in regions without seismic records. Future research should focus on improving high-frequency predictions and expanding the dataset to enhance generalizability.

Chapter 5: Ground Motion Prediction by DNN Model using EHVR

This chapter proposes the DNN model for predicting seismic ground motion parameters using EHVR. It integrates EHVR into existing ground motion prediction equations to improve prediction accuracy. The DNN model incorporating EHVR outperformed traditional models in predicting ground motion parameters, particularly for spectral amplitude and shape. However, the availability of EHVR data and potential data leakage are limitations that need addressing in future research.

Chapter 6: Conclusions and Future Work

This chapter summarizes the findings of the previous chapters and discusses potential areas for future research. The research demonstrated the effectiveness of using HVR data for seismic risk assessment and ground motion prediction. Future work should focus on improving prediction accuracy, validating the models in diverse geological settings, and exploring the integration of additional seismic features to enhance model robustness.

Chapter 7: Appendix

In this chapter, we will provide the usage methods for the deep neural network (DNN) models developed in this thesis, primarily focusing on utilizing several publicly available pre-trained models.

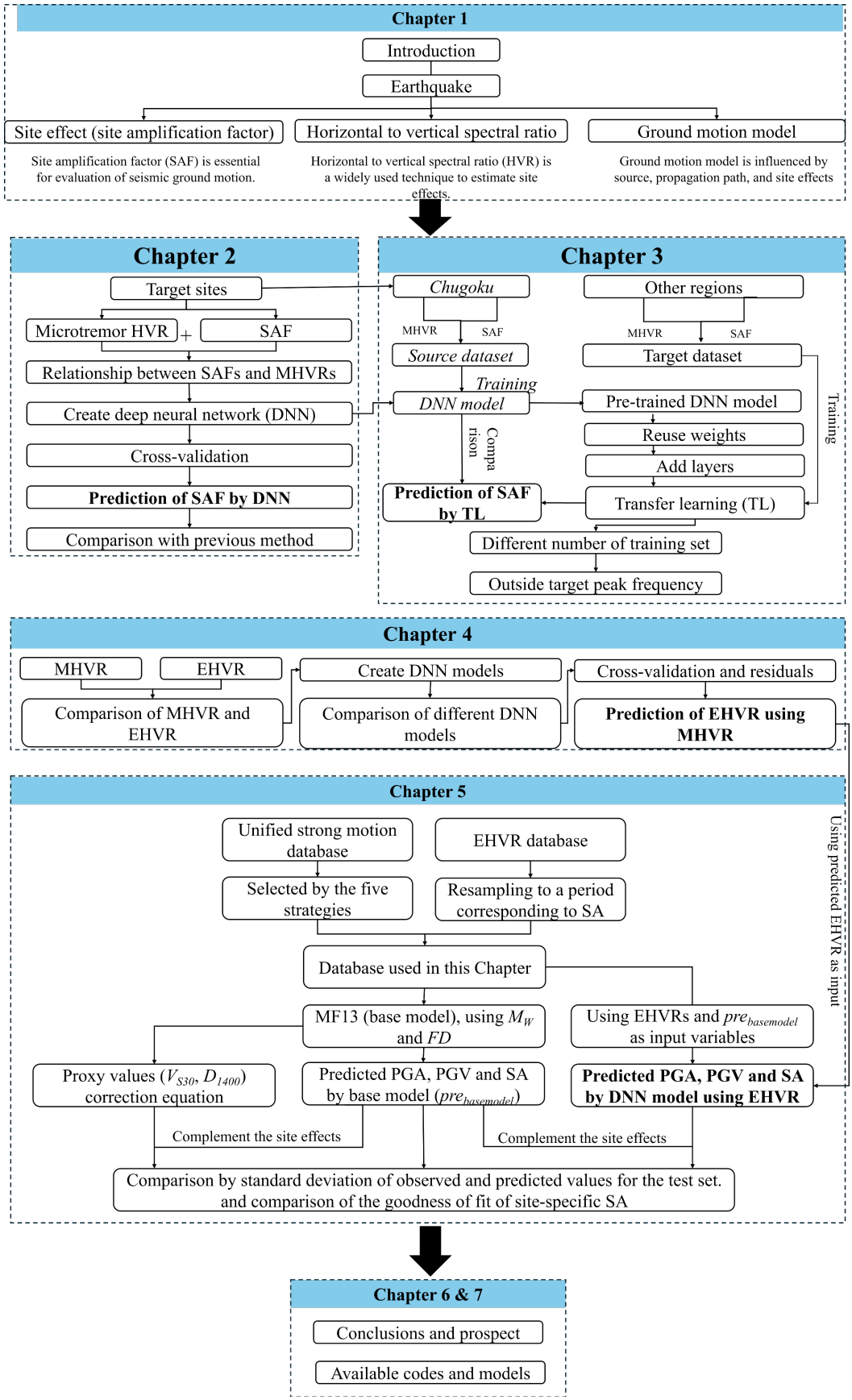

Figure 1-4. Outline of this dissertation.

Chapter 2. Predicting Site Amplification Factors (SAFs) from Microtremor H/V Spectral Ratios (MHVRs) by Using Deep Neural Network Model

This chapter proposed a novel methodology for directly estimating S-wave site amplification factors (SAF) from microtremor horizontal-to-vertical spectral ratio (MHVR) based on deep neural network (DNN) model. We analyzed site amplifications obtained from generalized spectral inversion technique and microtremor data observed at K-NET and KiK-net sites in Chugoku district, western Japan. The DNN model was developed using peak frequency and the frequency-dependent relationship between MHVRs and SAFs. The sites were divided into training set, validation set and test set. The training set and validation set were used in k-fold cross validation technique to evaluate and select optimal model. Once the optimal model had been determined, the model was employed on the test set that was completely independent of the training and validation set for evaluating the generalization performance. Residuals and root mean square error between the estimated and observed SAFs were evaluated to discuss the applicability of the proposed model. From the comparison of the results by the DNN model with those by existing double empirical correction method, we confirmed that the DNN model shows better performance in estimating SAFs.

2.1. Background

Seismic ground motions at surface can be represented as a product of seismic source, propagation, and site effects. Especially for soft soil sites, the site effect is essential for evaluation of seismic ground motion characteristics since the site effect strongly controls frequency-dependent amplitudes of ground motions. Site effect for S-wave of seismic ground motion is typically described as amplification of SH wave, which can be theoretically obtained by multiple reflection theory [2] when S-wave velocity structure model is available from PS logging and/or borehole data. It is, however, difficult to obtain amplification factors at sites where S-wave velocity structure models have not been obtained.

Generalized spectral inversion technique (GIT) has been proposed as a method for separating seismic source, propagation, and amplification factors from seismic records [4,5]. The approach assumes that Fourier spectra of seismic records at ground surface can be represented as products of seismic source spectra, damping characteristic of propagation path, and site amplification factors (SAFs). The source spectra, the Q-values of the seismic wave propagation path, and SAF are derived by using simultaneous equations from multiple spectra of seismic records observed at multiple sites. Thus, SAF can be evaluated without PS logging or borehole data by the GIT. Recently, this method has been applied in several regions to evaluate empirical amplification factors of seismic observation sites since more seismic data have been available with recent development of seismic observation networks [6–8]. It is, however, obviously impossible to obtain SAF without seismic observation data.

Another approach to obtain SAF is to use microtremor data. Microtremors have been considered as an effective tool to easily evaluate site effect since they can be observed everywhere and anytime once a measuring instrument is available. In particular, horizontal-to-vertical spectral ratio of microtremors (MHVR) typically presented by Nakamura [9,20] has been widely used for site effect. The relationships between SAF and MHVR have been discussed from various observation data in various regions [15,21–24].

These studies indicated that the predominant frequency of MHVR coincided with those of SAF at most sites, the spectral shape of MHVR and SAF were similar, but the significant discrepancy between amplitude of MHVR and SAF were found in most cases. These results suggested that MHVR should not be directly used as SAF despite the opinion by Nakamura [20].

Many studies have also been conducted to estimate the S-wave velocity structure via inverse analysis using MHVR. Arai and Tokimatsu [25] proposed a method to estimate the S-wave velocity structure based on the inverse analysis of MHVR and surface wave phase velocity data. Additionally, another method to infer S-wave velocity structure via diffuse field concept (DFC) has been proposed by Sánchez-Sesma et al. [26] and Kawase et al. [27]. Besides, a DFC-based approach was applied to estimate the S-wave velocity structure models from the MHVRs with help of existing geotechnical data as a constraint [28]. However, it cannot avoid the trade-off between S-wave velocity and layer thickness in the inverse analysis without a priori soil information. Therefore, it is still difficult to accurately estimate the S-wave velocity structure exclusively from MHVR [29].

Recently several studies have been performed to directly estimate SAFs from MHVRs. Sardina and Midorikawa [11] applied a pattern recognition approach for estimating soil response from MHVR, and suggested the possibility of MHVRs in direct estimation of SAFs. Such direct estimation approaches have been applied to horizontal-to-vertical spectral ratio not only of microtremors but also of earthquake motions (EHVRs). As introduced by Kawase et al. [30], EHVRs can be theoretically represented by transfer functions of P- and S-waves. Since EHVR can be described by simpler formula than MHVR in DFC, EHVRs have been used to directly estimate SAFs at seismic observation sites [3,31]. The EHVR-based approach cannot be applied to arbitrary sites where seismic observation records are not available. Kawase et al. [32] expanded the empirical method to estimate SAFs from MHVRs by developing double empirical corrections to firstly transform EHVRs from MHVRs, and to subsequently estimate SAFs from the transformed EHVRs using empirically derived spectral ratios of vertical motion

at surface to horizontal motion at bedrock. This suggested clear correlations between MHVR and EHVR (called EMR factor), and between EHVR and SAF (called VHbR factor). The double correlation model for estimating site amplification from MHVR through EHVR is shown in Eq. (2-1). Here, *pSAF* indicates pseudo site amplification factor, VH_bR indicates vertical-to-horizontal spectral ratio of bedrock motion. EHVR was estimated from MHVR using *EMR* that indicates earthquake-to-microtremor ratio in horizontal-to-vertical spectral ratio.

$$\begin{aligned} pSAF(f) &= EHVR(f) \times VH_bR(f) \\ &= \frac{EHVR(f)}{MHVR(f)} \times MHVR(f) \times VH_bR(f) \\ &= \{EMR(f) \times VH_bR(f)\} \times MHVR(f) \end{aligned} \tag{2-1}$$

However, includes two major uncertainties; uncertainty between MHVRs and EHVRs, and uncertainty between EHVRs and site amplification factors. These uncertainties would increase errors in estimating site amplifications.

Inspired by these previous studies, this study aims to propose an alternative empirical model to directly estimate SAFs from MHVRs. Recently, deep learning-based techniques, one of the artificial intelligence technologies, have been applied in many fields such as image classification [33], earthquake predictions [34,35] and other fields [36]. Deep learning refers to machine learning techniques which utilize multiple layers of artificial neural networks, such as deep neural networks (DNN) [37] which can repeatedly learn from a large amount of training data to efficiently extract eigenvalues.

In this study, we proposed a DNN model for directly estimating site amplification factors from MHVR. First, the relationships between S-wave site amplification factors (SAFs) and MHVRs were analyzed using observation data at K-NET and KiK-net in Chugoku district, western Japan. Second, we developed a DNN model to estimate the SAF based on a combination of multiple Affine and nonlinear activation layers Third, the

developed DNN model was validated through k-fold cross validation, residual, and root mean square error (*RMSE*). Then, the generalization performance of the developed model was tested by applying it to a completely unknown test set. Also, the results were compared with the double empirical correction method proposed by [32].

2.2. Data

2.2.1. Target sites

Nationwide seismic observation networks in Japan named as K-NET and KiK-net are operated by the National Institute for Earth Science and Disaster Resilience (NIED). We selected the target area of this study as the seismic observation sites in Chugoku district, western Japan. The locations of the K-NET and KiK-net sites in Chugoku district are shown in Figure 2-1. Chugoku district includes Hiroshima (HRS), Okayama (OKY), Shimane (SMN), Tottori (TTR) and Yamaguchi (YMG) prefectures. Totally 155 observation sites are located in the district.

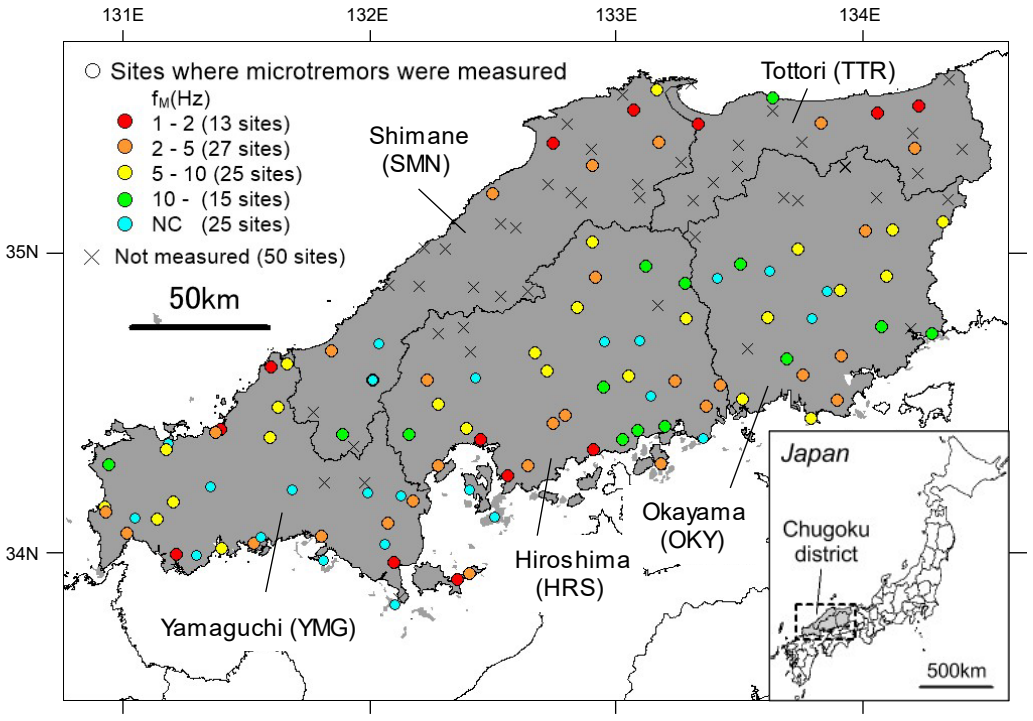

Figure 2-1. Locations of K-NET and KiK-net sites in Chugoku district, Japan.

Circles and diagonal crosses indicate sites where microtremor was measured and not measured, respectively. Color of circle represents predominant frequency of microtremor H/V spectral ratio (MHVR), f_M (Hz). NC means that peak was not confirmed in the observed MHVR.

2.2.2. Site Amplification Factors (SAFs) from Generalized Spectral Inversion Technique (GIT)

Takeda et al. [38,39] extracted the frequency-dependent SAFs of S-waves at the K-NET and KiK-net sites in Chugoku district by GIT. In the spectral inversion, the Fourier spectrum of the seismic observation record was assumed to be represented as the following Eq. (2-2).

$$O_{ij}(f) = S_i(f)\frac{1}{R_{ij}}exp\left(-\frac{\pi f R_{ij}}{Q(f)Vs}\right)G_j(f) \tag{2-2}$$

Here $O(f)$ is the observed spectrum at frequency f, the subscripts of i and j represent earthquake number and station number, respectively, and S is the source spectrum, Q is the Q-value for frequency-dependent attenuation of S-wave, Vs is S-wave velocity of the seismic bedrock, and G is the site amplification factors. R is the distance from the earthquake i to the station j, R^{-1} was adopted as the geometrical spreading of seismic wave since S-wave parts of the observation records were analyzed in this study. The exponential term represents the viscous damping factor of the propagation path. Using observation data recorded by multiple earthquakes at multiple sites, $S_i(f)$, $Q(f)$, and $G_j(f)$ were obtained for each frequency from Eq. (2-2) by solving the simultaneous equations.

Considering the criteria for data selection in the previous GIT studies [8,40], we preliminary assessed the observed ground motion records as follows. To exclude effects of surface waves mainly contained in later part of the seismic records, the horizontal

ground motions at surface with the duration of 10.24 s from the arrival of S-waves were analyzed. The records whose peak ground accelerations were larger than 200 cm/s^2 were excluded in the GIT to avoid nonlinear site response due to strong ground shakings. The Fourier spectra of NS and EW (north-south and east-west) components were calculated (H_{NS} and H_{EW}), and the horizontal spectra were obtained by root of the sum of squares from the following Eq. (2-3).

$$O(f) = \sqrt{{H_{NS}}^2(f) + {H_{EW}}^2(f)} \tag{2-3}$$

The seismic records observed from 2001 to 2008 in Chugoku district for 119 earthquakes with the magnitude of 4.0 or larger were analyzed in the GIT. The distribution of epicenters of the earthquakes used in the GIT is shown in Figure 2-2. The figure also

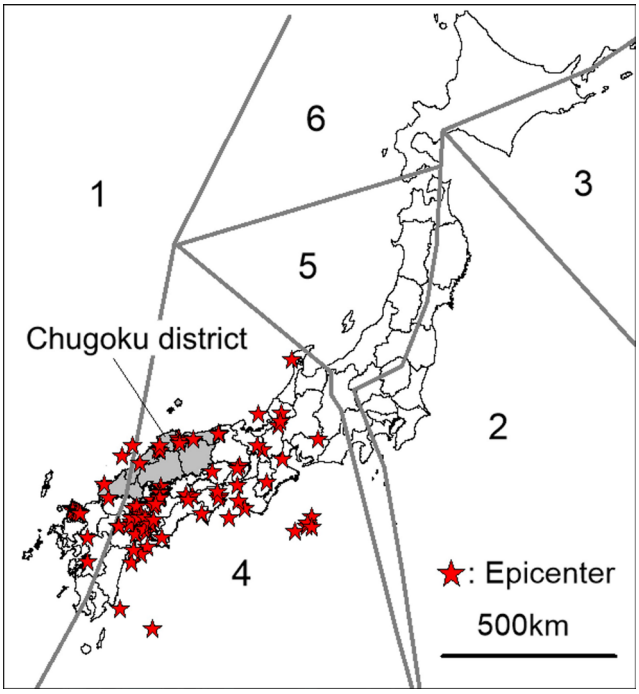

Figure 2-2. Epicentral locations of earthquakes used in our GIT and classification of regions (1-6) by Nakano et al. [8].

illustrates the classification of regions for evaluating different *Q*-value in each region by Nakano et al. [8]. Since the epicenters used in the study are mostly within the region four, single *Q*-value was estimated in the study. There were 68 crustal earthquakes, 49 intraplate earthquakes, and only two interplate earthquakes. The relationship between magnitude and depth of the earthquakes is shown in Figure 2-3. Different *Q*-value was considered for each earthquake type in most of the previous GIT studies in order to accurately evaluate the propagation path effects. However, Salazar et al. [41] pointed out that almost same SAFs were obtained from data observed in subduction and crustal earthquakes. Since the main purpose of the GIT was to extract reliable site amplification factors from larger number of observation records, the records from all the earthquake events were analyzed in the GIT.

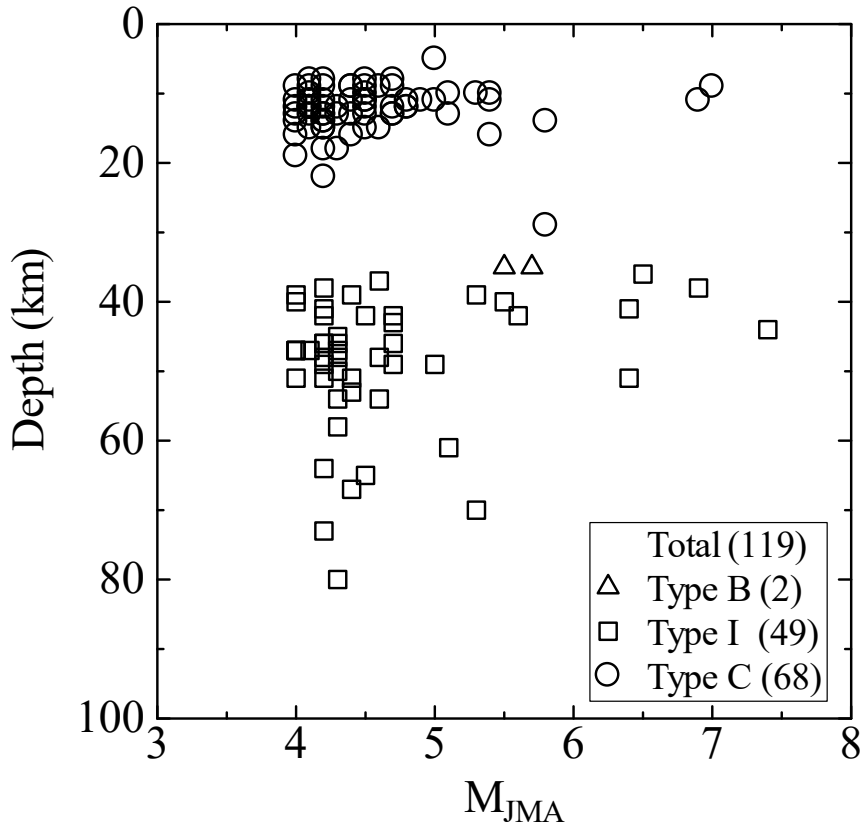

Figure 2-3. Relationship between magnitude and depth of earthquakes used in our GIT. Symbols for type B, I and C indicate interpolate, intraplate and crustal earthquake, respectively.

The simultaneous equations were solved by giving the theoretical amplification factors at reference sites as constraint conditions in the GIT. Takeda et al. [38,39] selected five KiK-net sites as the references where PS-logging data up to almost V_S=3000 m/s (the seismic bedrock) were obtained and the seismic bedrock depth is comparatively shallow. Figure 2-4 shows the S-wave velocity structure models for each site and the theoretical amplification factors at surface from outcropping bedrock obtained from the S-wave velocity structure models.

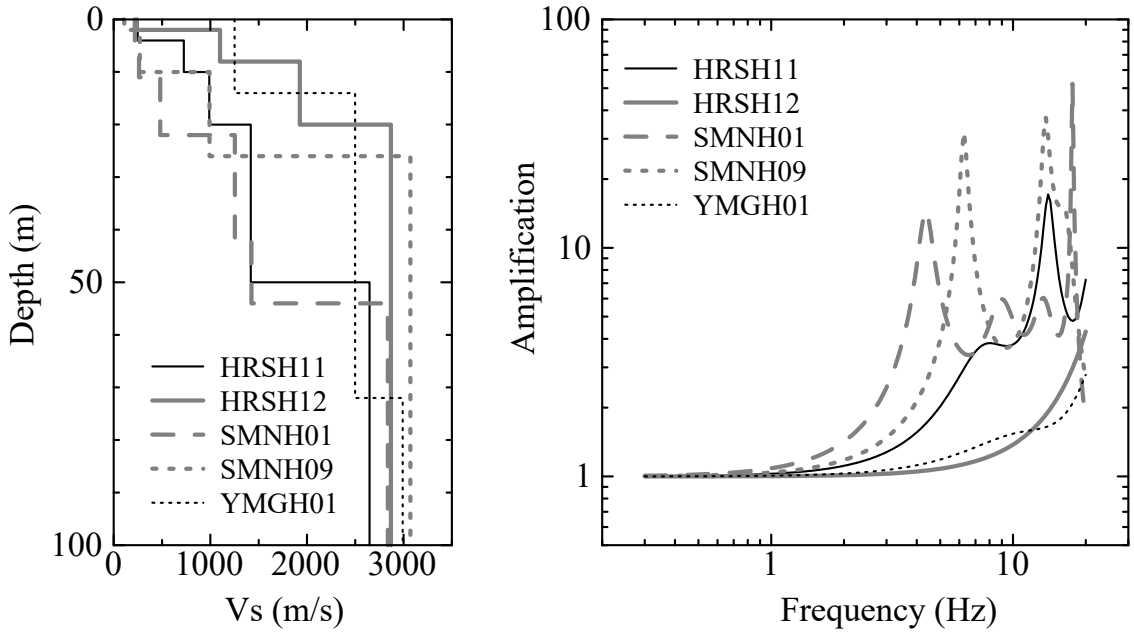

Figure 2-4. S-wave velocity structures and theoretical amplification factors at surface from outcropping bedrock used as constraint conditions in our GIT.

The target frequency range was from 0.3 to 20 Hz, and 200 frequency increments were given with equal intervals in the logarithmic scale by resampling the data. By applying the GIT, the source spectra, Q-value in the propagation path of the seismic bedrock and site amplification factors from the seismic bedrock to ground surface were obtained. The derived frequency-dependent Q-value in the target area is shown by Eq. (2-4).

$$Q(f) = 360f^{0.53} \tag{2-4}$$

Figure 2-5 shows the comparison of Q^{-1} in this study with the results in the region four (see Figure 2-2) by Nakano et al. [8] and Kawase and Matuo [40]. The Q^{-1} in this study was slightly smaller than the values by the previous studies probably because locally smaller Q^{-1} was expected in Chugoku district than in other western Japan region from the previous three-dimensional tomography analysis of seismic records.

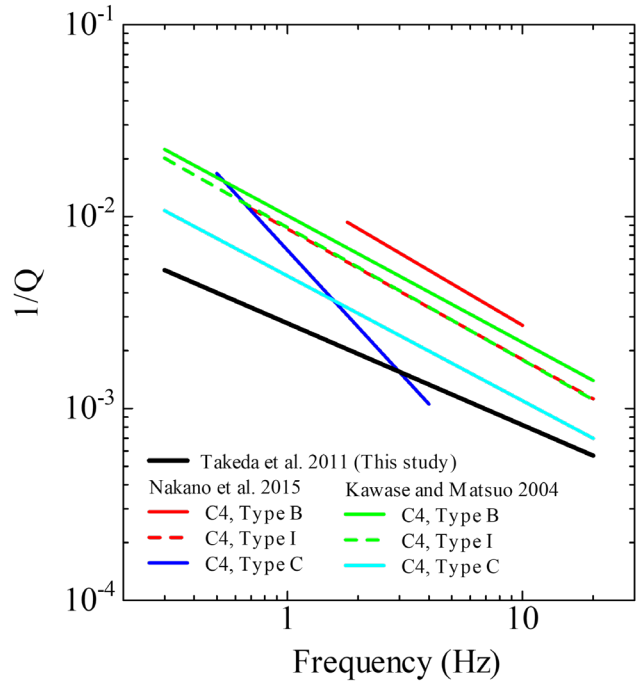

Figure 2-5. The comparison of the Q^{-1} estimated in this study with that of Nakano et al. [8] and Kawase et al. [40] for region four (C4).

The SAFs of each site obtained from the above analysis can be considered to represent the amplification of S-waves at ground surface from the outcropping seismic bedrock. Figure 2-6 shows the comparison of SAF at TTRH02 and OKY005 estimated in

our GIT and that estimated in Kawase and Matsuo (2004) and Nakano et al. (2019), respectively. As shown in Figure 2-6, we confirmed that our obtained SAFs show good agreement with those in the previous studies despite different datasets were analyzed in these GITs. The SAFs obtained in our GIT were used in the following analysis.

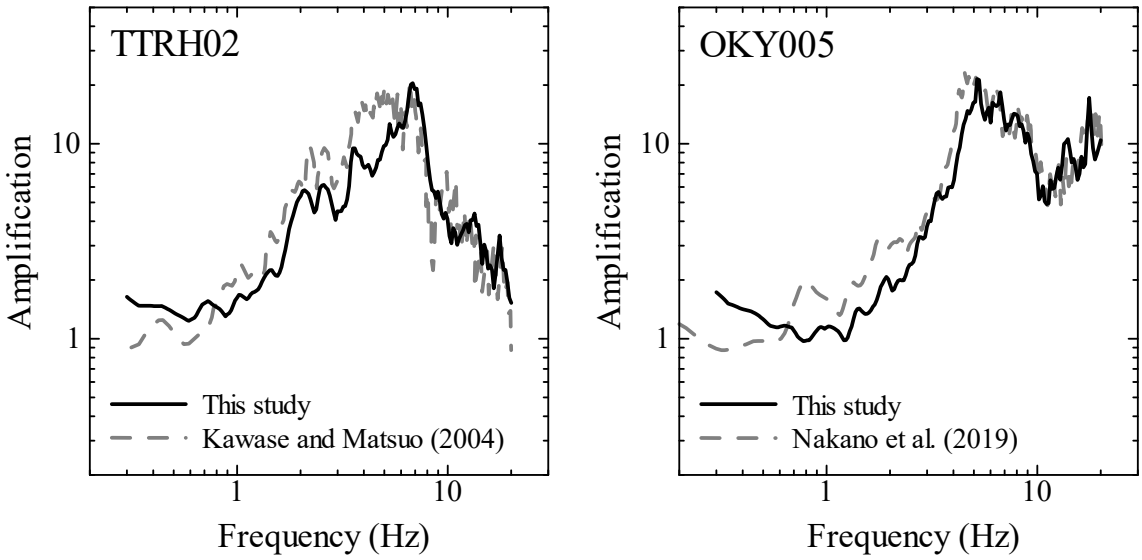

Figure 2-6. Comparison of site amplification factors at TTHR02 and OKY005 estimated in our GIT and those estimated by GIT.

2.2.3. Microtremor H/V Spectral Ratios (MHVRs) in Chugoku Region, Japan

The authors observed microtremors at the target sites from 2012 to 2020. The GEODAQS-2S3D (Figure 2-7) velocimeters developed by ANET Inc., Japan were used for the observations. The sensor covers the period of ambient motions of up to 2 s (frequency of 0.5 Hz). We observed the microtremors at the free grounds near K-NET and KiK-net stations by 100 Hz sampling.

For the microtremor records, first all-time history waveforms are visually inspected. Once it has been determined that the effect of local time-varying noises (e.g., traffic and artificial noises) are less, the observation recordings are split from the observed records

into 10 recording segments of 20.48 s. Then, MHVR is calculated from the horizontal and vertical amplitude Fourier spectra as Eq. (2-5).

$$MHVR(f) = \frac{\sqrt{FS_{NS}{}^2(f) + FS_{EW}{}^2(f)}}{FS_{UD}(f)} \tag{2-5}$$

Figure 2-7. The GEODAQS-2S3D velocimeters developed by ANET Inc.

Here, *FS* indicates the Fourier spectral amplitude, and the subscripts of *NS*, *EW* and *UD* indicate the horizontal (north-south and east-west) and vertical (up-down) components, respectively. The Fourier spectra were calculated by smoothing with Parzen window of 0.3 Hz bandwidth. The average and one standard deviation of MHVRs from 10 retained recording segments are calculated. Since peak frequency of MHVRs used for analysis are in 1.0 - 20.0 Hz range, the 10 recording segments of 20.48 s correspond to the recommended recording duration suggested in the guideline by Bard [42].

In this study, we used the microtremor observation data obtained at 105 sites in the Chugoku district. The observed MHVRs at six sites are illustrated in Figure 2-8, showing clear peaks identified in these MHVRs. However, the shapes of the MHVRs at 25 sites

out of 105 sites were flat and no significant peaks were observed within the target frequencies (0.3 – 20.0 Hz). Since peak frequency of MHVR (f_M) played a very pivotal role in estimating site amplifications, it would be difficult to accurately estimate amplification factors from such MHVRs. The MHVRs with peak values less than a factor of 2.0 were excluded in the following analysis. Peak frequency of MHVR at each site was defined as the frequency where the largest MHVR value within 0.3 to 20.0 Hz was observed. Figure 2-8 also shows the SAFs obtained at the sites in our GIT. As introduced in the previous studies [6–8], the peak frequencies of the MHVRs show good agreement with those of the SAFs.

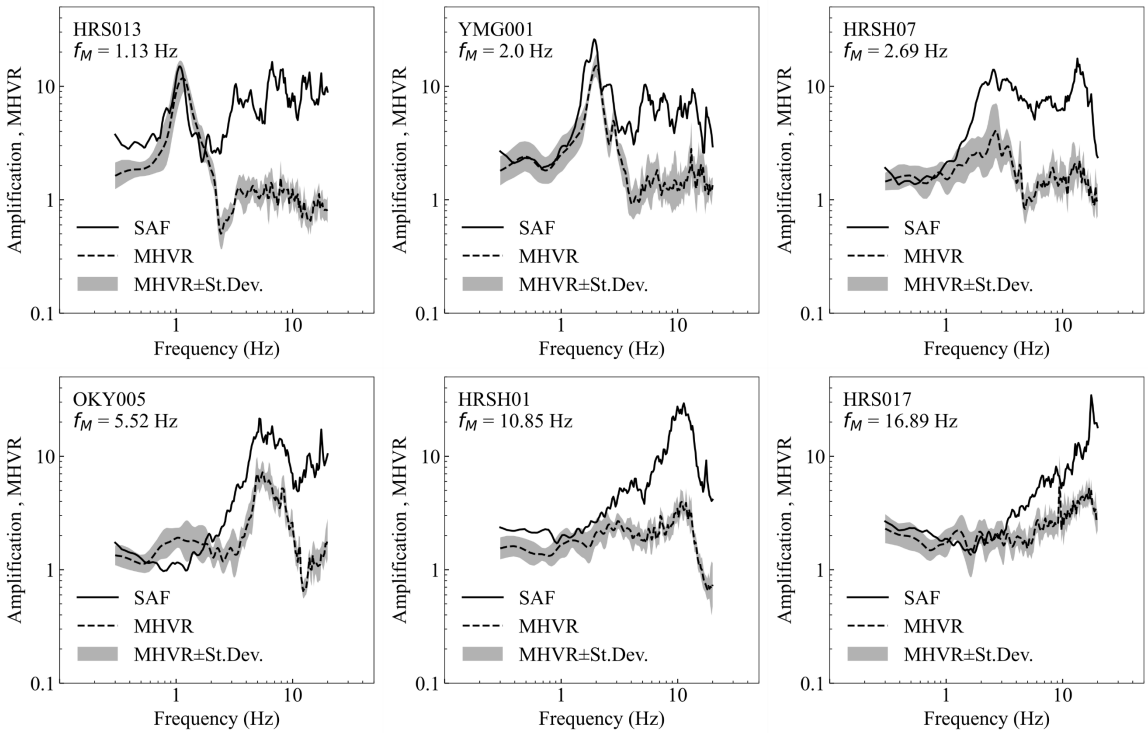

Figure 2-8. Comparison of observed MHVRs and SAFs by GIT at six sites. Average and one standard deviation of observed MHVRs from 10 retained recording segments are shown by dashed line and gray area, respectively.

Figure 2-9 shows the relationship of peak frequencies between SAFs and MHVRs at 80 sites. We found strong correlation between the two peak frequencies since the correlation coefficient is higher than 0.90. As shown in Figure 2-9, we need to pay attention to the fact that the coverage of f_M is 1.0 – 20.0 Hz in our dataset.

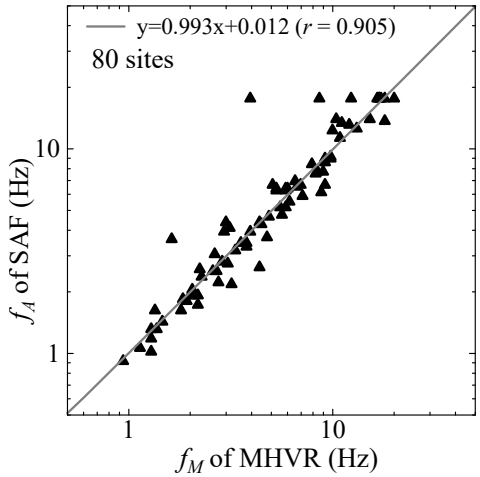

Figure 2-9. Comparison of predominant frequency between MHVRs (f_M) and site amplification factors (f_A).

The correlation coefficient (r) is 0.905, which shows the strong linearity of the relationship between the two peak frequencies. Using the least squares method, the linear equation between the two was calculated as shown in Eq. (2.5). Here x represents log(f_M) and y represents log(f_A).

$$y = 0.993x + 0.012 \tag{2-6}$$

2.2.4. Relationship between SAFs and MHVRs

In order to discuss the possibilities of H/V spectral ratio technique in estimating SAFs, the relationships between theoretical and observed values of SAF, EHVR and

MHVR are examined. The number of sites where S-wave velocity profiles reach the seismic bedrock is limited even in the KiK-net sites. Besides, one-dimensional theoretical transfer functions obtained from the S-wave velocity profiles of KiK-net did not match observed amplification factors at more than two-third of the sites due to uncertainties of the velocity models and multi-dimensional effects [3]. We selected HRSH18 as the target site since the S-wave velocity structure up to approximately the seismic bedrock (V_S=2,530 m/s) was obtained in the PS-logging. Theoretical one-dimensional transfer function for SH-wave was calculated from the S-wave velocity structure. The density of i-th layer was provided from the S-wave velocity shown in Eq. (2-7), and the Q-value of each layer was given by Eq. (2-8) [43].

$$\rho_i = 1400 + 670\sqrt{\frac{Vs_i}{1000}} \tag{2-7}$$

$$Q_i = 19.05f^{0.52} \tag{2-8}$$

Figure 2-10 (a) shows the comparison of the theoretical transfer function (1D-SH) and the observed SAFs by the GIT, indicating that the predominant frequencies are almost identical. However, the theoretical amplification factor is slightly smaller than the observed value probably due to uncertainties of damping factors in the layers.

Figure 2-10 (b) shows the comparison of the observed and theoretical earthquake H/V spectral ratio (EHVR) at the site. The observed EHVRs was obtained from the S-wave parts in the seismic records used in the GIT as is the analysis based on Eq. (2-6). Mean values ± the standard deviation are illustrated in the figure. The theoretical EHVR was calculated from the ratio of transfer function of S-wave to transfer function of P-wave with the S-wave and P-wave velocity at the bottom layer based on the DFC theory [30]. The theoretical EHVR shows good agreement with the observed value, and the

predominant frequencies of the EHVRs also correspond to those of the transfer function. The results confirm that EHVR could be a powerful tool to assess the site amplification factor and/or to estimate velocity structure model through inversion technique as discussed in the previous studies [3,27,31,35].

Figure 2-10 (c) shows the comparison of the observed and theoretical MHVR at the site. The theoretical MHVR was calculated from the velocity structure based on the DFC theory [26], by using the HV-Inv [44]. Although the spectral shape of the theoretical value is similar to the observed value, the theoretical value is much larger than the observed value. One of the reasons for the discrepancy would be that the HV-Inv neglects damping in site response. If the damping factors of the soil layers are properly provided in the theoretical calculation of MHVR, the theoretical value would be close to the observed MHVR. It suggests that there still remains uncertainty in applying the DFC theory to the inversion of MHVR, and it is still difficult to estimate the S-wave velocity structure accurately only by MHVR without a priori soil information.

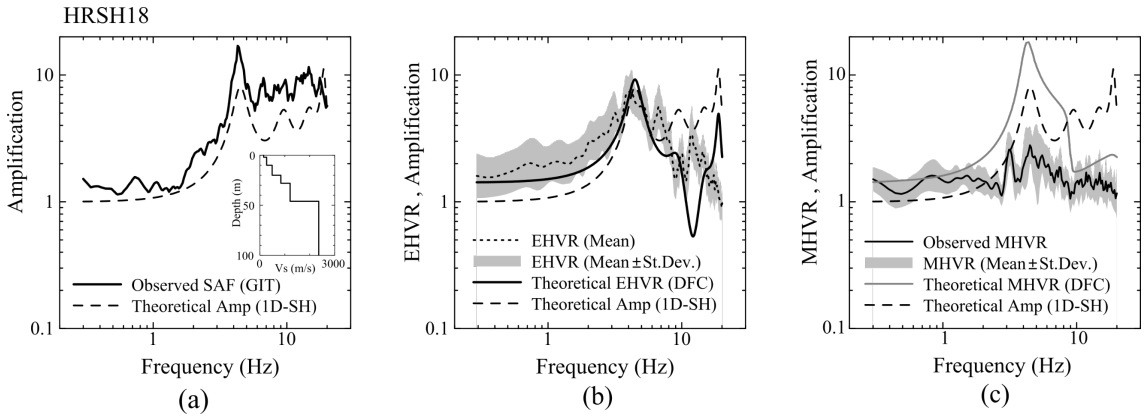

Figure 2-10. Comparison of theoretical and observed values for (a) SAF, (b) EHVR, and (c) MHVR at HRSH18. (a) Comparison between observed SAF derived by GIT and 1D theoretical amplification of SH-wave, and S-wave velocity profile, (b) Comparison of observed EHVRs of seismic recordings and theoretical EHVR calculated by DFC, (c) Comparison of observed MHVR and theoretical MHVR calculated by DFC.

2.3. Methods

2.3.1. Deep neural network model

The neural network, which is the basis of deep learning, is described.

In deep learning, a neuron is actually a computational unit. Figure 2-11 is the functional flow of a standard neuron.

a) A neuron is a computational unit that needs to receive several input values to start the computation. Here n values ($x_1, x_2 \ldots x_n$) are input.

b) Each input x_i (i=1~n) has a corresponding weight w_i. All the input values are then multiplied by the corresponding weights and then summed.

c) An activation function in the neuron converts the '$\sum_{i=1}^{n} xiwi + bias$' into an output value ($y_j$), and the output value (y_j) calculated with this function is used as input value to subsequent neurons.

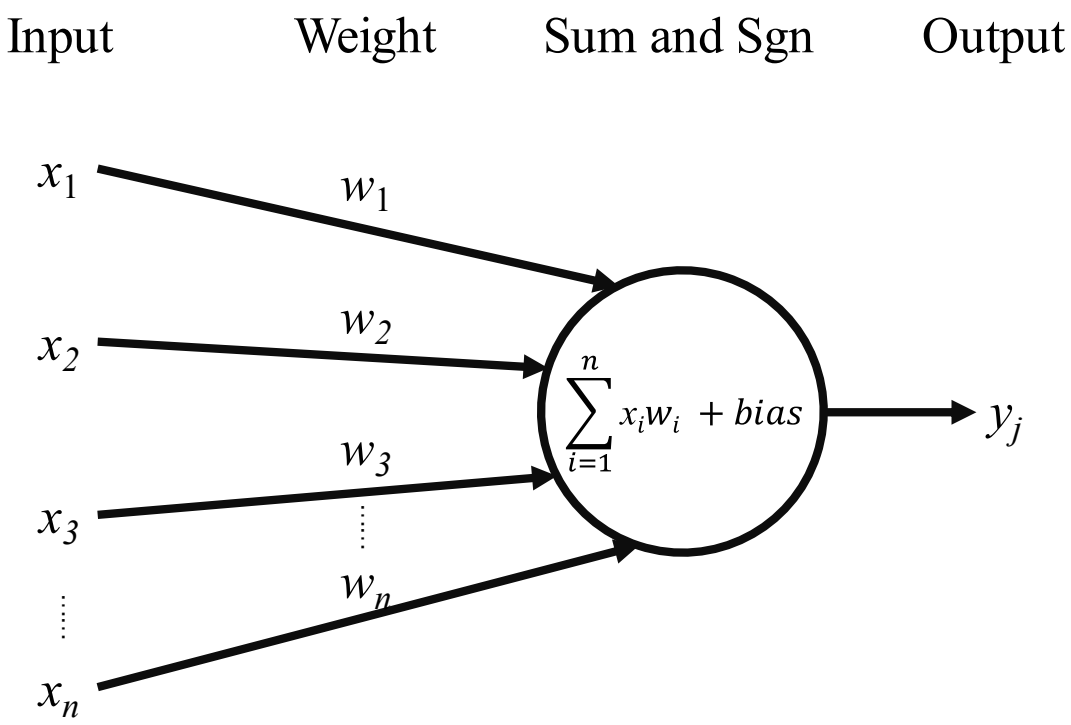

Figure 2-11. Schematic diagram of a neuron.

Neural networks are a conceptual imitation of human neuronal networks and are the basis for deep learning. In a neural network, multiple neurons are combined to form a single layer, consisting of input layer, hidden layer and output layer. Multiple inputs are passed through several layers of neurons and finally become multiple outputs. Through repeated training and learning of various input values, the neural network gradually grows to be able to return an appropriate output value for any input value.

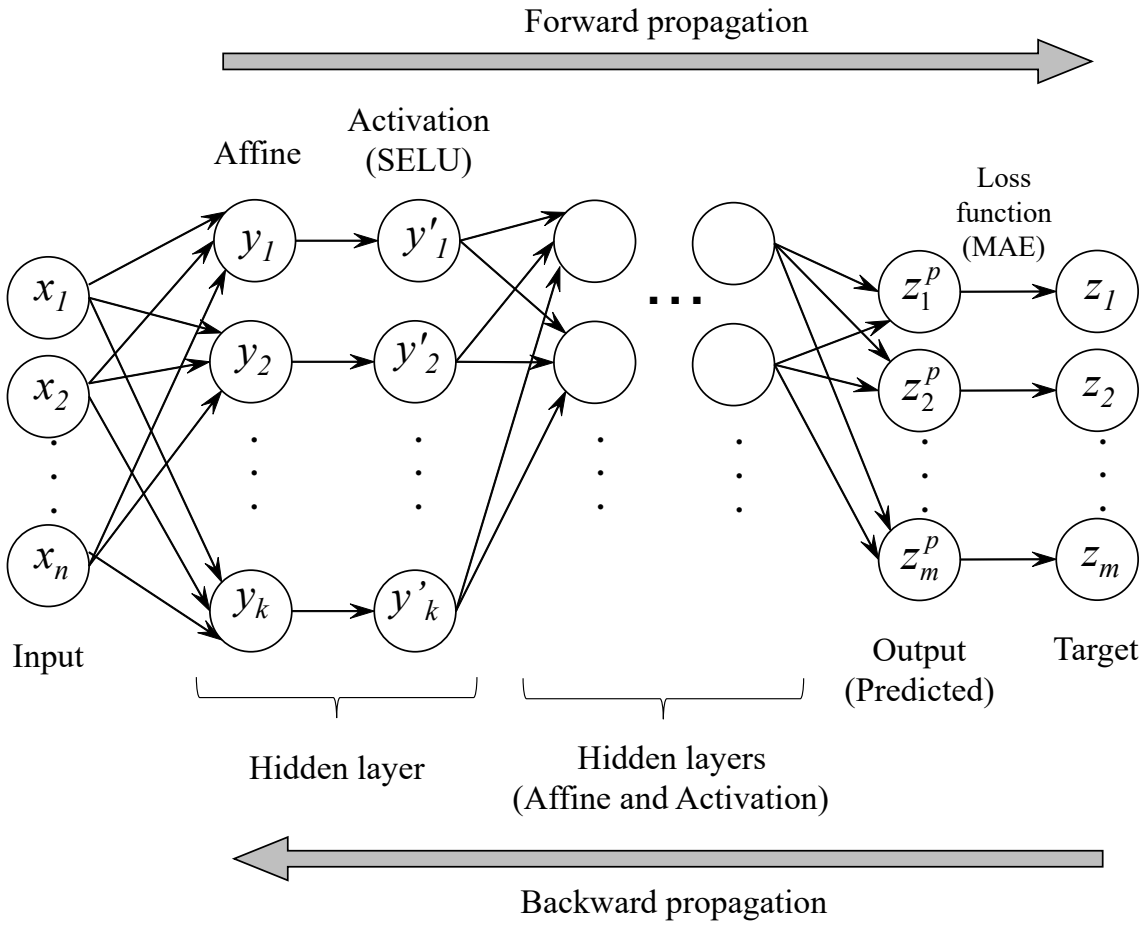

Figure 2-12. Schematic diagram of propagations of neurons in input, hidden and output layers for deep neural network. Here, *n*, *k* and *m* indicate the number of neurons in each layer.

Figure 2-12 shows the schematic diagram of forward and backward propagations of neurons in input, hidden and output layers of deep neural network (DNN) model. As

shown in Figure 2-12, DNN is a neural network that contains input layer, multiple hidden layers and output layer. Each layer of the neural network has several neurons. Neurons between layers are connected, while neurons within layers are not connected, and neurons in the next layer connect all neurons in the previous layer. This type of neural networks that have connections between all neurons in the adjacent layers are known as fully connected neural networks. The typical steps of the neural network are as follows: forward propagation to find the error using the multiple layers, then followed by backward propagation to update the weights of the hidden layer, and finally repeated given times (epoch) to find the minimum error. Since DNNs can fit almost any function, the nonlinear fitting ability has been considered to be generally excellent.

2.3.2. Evaluation metrics for goodness-of-fit

In general, goodness-of-fit metrics are to measure how well the observed data correspond to the assumed model. As in linear regression, in essence, goodness-of-fit metrics compares the observed values to the estimated values.

To numerically assess the estimation accuracies of the models in this study, goodness-of-fit metrics were calculated using the observed and estimated values. Table 2-1 shows the formulas and descriptions of the metrics analyzed in this study.

We calculated Pearson correlation coefficients, r, between the estimated and observed values for each site. The value range of r is from -1 to +1, the closer to 0 that represents the worse correlation, and +1 and -1 represent positive correlation and negative correlation, respectively. However, the Pearson correlation coefficient can only evaluate the spectral shape difference between the observed values and the estimated values. Thus, in addition to the spectral shape, the index of agreement (d) and the mean absolute error (MAE) between the estimated and observed values were also calculated to assess the proximities of the amplitudes in absolute and relative values, respectively. Here, d takes in the range from 0 to 1, and larger value represents smaller difference between the variables. On the contrary, smaller MAE represents the better accuracy. In the previous

study [3], it was suggested that for better evaluation accuracy, the double evaluation metrics were recommended to be used together (one for evaluating the spectral shape and one for evaluating the amplitude differences).

Table 2-1. Evaluation metrics used in this study.

Evaluation metrics	Formula	Range	Annotation
Pearson correlation coefficients, *R*	$\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$	(-1:1)	Closer to 0 that represents worse correlation, +1 and -1 represent positive and negative correlation.
Index of Agreement, *d*	$1-\frac{\sum_{i=1}^{n}(x_i-y_i)^2}{\sum_{i=1}^{n}(\lvert y_i-\bar{x}\rvert+\lvert x_i-\bar{x}\rvert)^2}$	(0:1)	Lager *d* represent smaller differences between the variables.
Mean absolute error, *MAE*	$\frac{\sum_{i=1}^{n}\lvert y_i-x_i\rvert}{n}$	(0:∞)	Smaller MAE represents better accuracy.

Note: *n*: sample size, *x*: observed site amplification factor, *y* : pseudo amplification factor, *i* : frequency increment.

2.4. Construction of DNN model for prediction

2.4.1. Architecture of DNN model

The DNN model used in this study was determined by comparing the different results obtained by changing different variables, layers or parameters in the neural network. This chapter describes in detail the comparison of the results of models employing different variables.

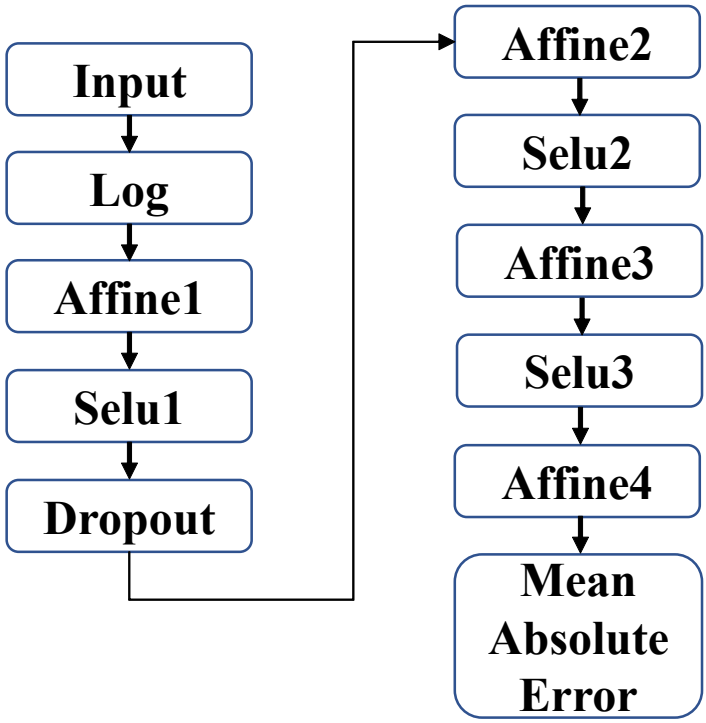

Figure 2-13. DNN architecture adopted in this study. Affine layers control weights of input data, and SELU layers indicate nonlinear activation functions.

Figure 2-13 shows the adopted DNN architecture in this study. In the fully connected neural network, the hidden layers basically consist of a combination of Affine and nonlinear activation layers. The log layer is used to process the input data (MHVR) in a logarithmic calculation to make the values uniformly distributed. Affine layer is the layer that performs the affine transformation used in the matrix product performed in the

forward propagation of the neural network. The input neurons are transferred by Affine layer represented as the Eq. (2-10).

$$y_j = \sum_{i=1}^{n} w_{ij} x_i + b_j \tag{2-9}$$

Here, x and y indicate the neurons in the current and following layer, and the subscript i and j mean the number of current neurons and out shape in the following layer, respectively. And w and b indicate the weight and constant term, respectively. The Affine layer controls the weights of the input training data. The weights and constant term were optimized by backpropagation of errors during training process.

Nonlinear activation layer represents the activation function to increase the weight in computing necessary data and to suppress the weights for unnecessary data. Since the logarithmic scale was used for data preprocessing as described later, some negative values appear in the hidden layer that may be favorable to the results. Therefore, Scaled Exponential Linear Unit (SELU) [45] was used as the activation function. The SELU activation function is expressed in Eq. (2-10) and enables to output a certain range of negative values while suppressing the appearance of very small negative values in the hidden layer, as shown in Figure 2-14.

$$y'_j = \begin{cases} \lambda y_j & , \ y_j > 0 \\ \lambda\alpha(e^{y_j} - 1) & , \ y_j \leq 0 \end{cases} \tag{2-10}$$

Here y and y' are the input data and output result, respectively. We used 1.67 and 1.05 for the parameters λ and α, respectively, as introduced by Klambauer et al. [45].

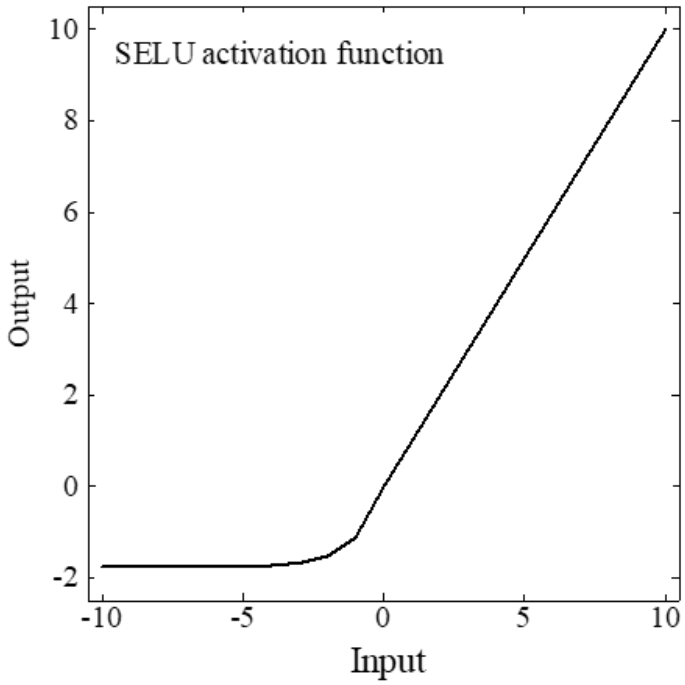

Figure 2-14. Scaled Exponential Linear Unit **(SELU)** activation function.

Dropout [46] was applied in the first hidden layer., which is a learning method that suppresses overlearning by randomly removing neurons. Figure 2-15 shows the schematic concept of dropout. Dropout refers to temporarily turn off a portion of the network nodes during the training round, and in principle the neurons removed are randomized. This has been commonly used to suppress the overfitting in deep learning.

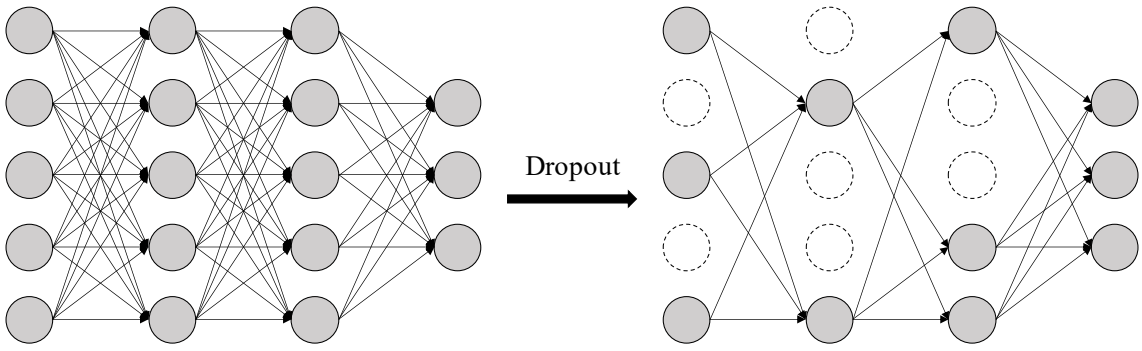

Figure 2-15. The schematic concept of Dropout

The neurons to be removed are chosen randomly in each training session. This makes it possible to train a different model each time. During evaluation, the signals of all neurons are transmitted, and the output of each neuron is multiplied by the percentage of neurons eliminated during training to obtain the average value of the model. In other words, dropout can suppress overlearning by training multiple models in a single pseudo-network.

In this model, Mean Absolute Error (MAE) shown by Eq.(2-11) was used as the loss function, which is the sum of the absolute values of the deviation between the target values, z and predicted values, z^p. Here, n_d means the total number of training or validation samples.

$$\mathrm{MAE} = \frac{1}{n_d}\sum_{l=1}^{n_d}|z_l - z_l^p| \tag{2-11}$$

2.4.2. Cross-validation techniques

In employing the DNN, the dataset was divided into the training set, validation set and test set. Training set is the set of data used to learn and fit the model. Validation set is the set of data used to evaluate while refining the model and selecting features. Test set (also called external test set) is the set of data used to evaluate generalization performance once the optimal model has been decided. A separate validation set sometimes does not provide a sufficiently clear evaluation especially when the data is limited, therefore a popular method is to use k-fold cross-validation [47] to tune the model hyperparameters rather than a separate validation set. Since it is generally difficult to know the best values of hyperparameters for a given issue, the authors searched optimal hyperparameters for this issue by trial and error through the cross-validation. We tuned the hyperparameters such as number of layers, neurons in hidden layers, learning rate, and number of epochs. In the method, the dataset (training set and validation set) is divided into k parts to test

the effect of the model. One of the parts is used as validation set, and the remaining *k*-1 parts are used as the training set during the training process, and it repeats *k* times by changing the validation and training sets. As shown in Figure 2-16, the model is tuned by *k*-fold cross-validation to get the optimal model, and then a separate external test set is used to evaluate the generalization performance of the optimal model. In this study, we split the data of 80 sites into training set, validation set and test set in 8:1:1 scale. Table 2-2 shows validation sets of the sites for 9-fold cross validation and test set. We manually selected eight sites in test set and each validation set to equally contain the sites with low and high peak frequencies. Meanwhile four K-net and four KiK-net were selected for use as test sets. Nine times trainings and validations were performed by changing the combination of the training and validation datasets for the 9-fold cross validation.

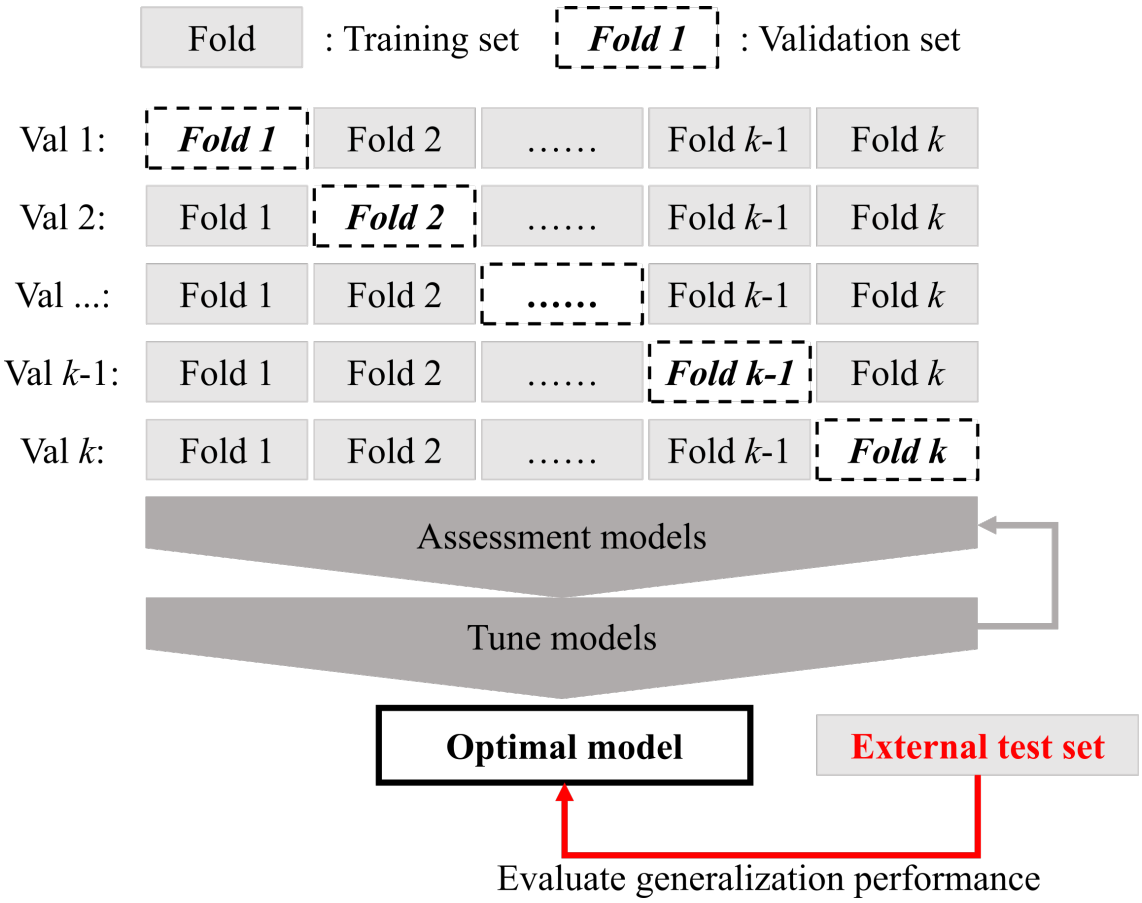

Figure 2-16. General flowchart used for *k*-fold cross validation on training sets and validation sets, and evaluation of generalization performance on test set.

Table 2-2. Validation set of 9-fold cross validation and test set.

Fold 1		Fold 2		Fold 3		Fold 4		Fold 5	
site	f_M	site	f_M	site	f_M	site	f_M	site	f_M
SMN005	0.94	HRS013	1.13	HRS018	1.29	TTR004	1.29	TTR008	1.29
YMG012	1.84	OKY012	1.92	YMG001	2.00	YMG016	2.00	YMG011	2.05
OKY011	2.58	YMG014	2.64	HRSH07	2.69	YMG015	2.75	YMG013	2.87
SMN006	3.19	HRSH06	3.54	TTR003	3.78	TTR005	3.78	HRS012	3.94
YMG003	5.07	OKY014	5.29	YMGH08	5.29	OKY005	5.52	HRS010	5.64
HRS009	6.13	YMG006	6.54	HRSH13	6.67	YMG007	6.96	YMGH13	7.11
OKYH11	9.16	YMGH02	9.16	YMGH09	9.16	OKYH12	9.76	HRSH02	9.97
OKYH02	12.31	HRS021	13.11	OKYH04	15.20	HRSH08	16.54	HRS017	16.89
Fold 6		Fold 7		Fold 8		Fold 9		***Test set***	
site	f_M	site	f_M	site	f_M	site	f_M	***site***	***f_M***
YMG018	1.34	TTR002	1.46	HRS015	1.62	YMG017	1.80	***YMG002***	***1.37***
HRS019	2.18	SMN002	2.18	HRS014	2.23	YMGH17	2.28	***HRSH05***	***3.33***
YMGH14	2.93	HRS020	2.99	SMN004	3.06	SMN013	3.12	***OKYH01***	***4.38***
OKY013	3.94	HRS011	4.38	SMN015	4.76	OKY006	4.86	***HRSH18***	***4.47***
HRS001	5.88	HRS005	5.88	HRSH16	5.88	YMG008	6.01	***OKY008***	***7.90***
OKY007	8.24	YMGH10	8.42	HRS008	8.60	SMN001	8.78	***HRS003***	***8.97***
SMNH14	10.40	HRSH01	10.85	OKY004	11.08	TTR006	12.05	***OKY009***	***9.76***
HRS016	16.89	HRS002	17.25	YMG005	18.00	SMN011	20.00	***OKYH13***	***18.00***

2.4.3. Input and output data

The input and output combinations were determined by trial and error approach. When only $MHVR(f_i)$ was used as input data, strongly smoothed spectra were obtained. As shown in Figure 2-17, the pSAF when using five-values of $i\pm2$ steps as the input data is significantly better than single-value (i step) especially at high frequencies. Here, the subscript i indicates the target frequency number (from 1 to 200). To match the frequency number of the site amplifications shown above, the number of frequencies of the MHVRs was set to 200 by resampling the frequency interval from 0.3 to 20 Hz.

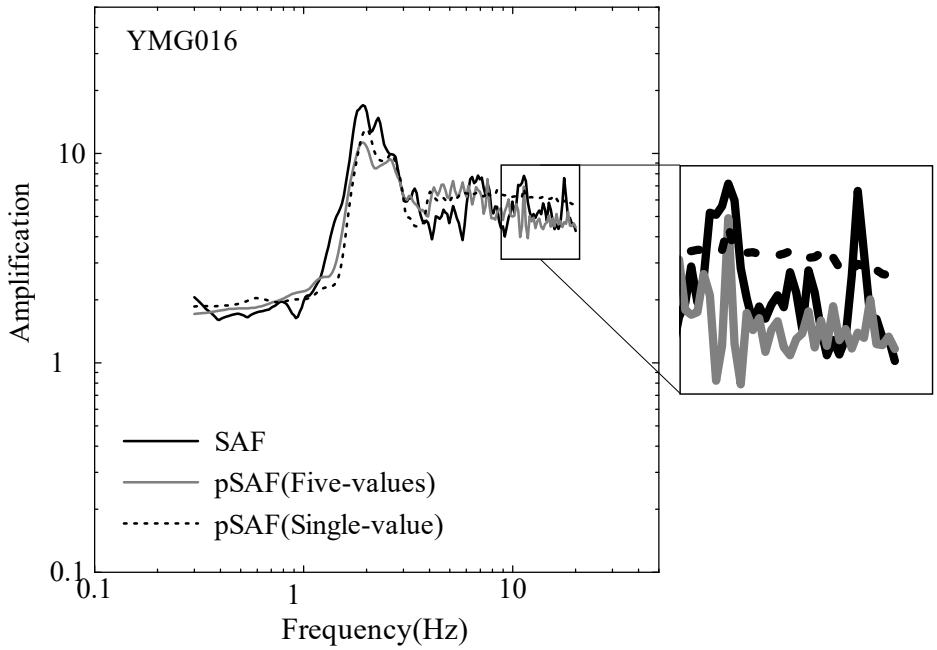

Figure 2-17. Comparison of observed and pseudo SAFs estimated from five-values ($i\pm2$ steps) and single-value (i step) as the input data at YMG016.

To accurately reproduce the shapes of the amplification factors, different combinations of input and output dataset ($i\pm1$, $i\pm2$, $i\pm3$, $i\pm5$, $i\pm7$, $i\pm9$ steps) were analyzed in this study.

The $i\pm2$ step means that we used the target frequency f_i, the peak frequency of

MHVR f_M, $MHVR(f_{i-2})$, $MHVR(f_{i-1})$, $MHVR(f_i)$, $MHVR(f_{i+1})$, and $MHVR(f_{i+2})$ as the input data for the training. Here, the subscript i indicates the target frequency number (from 1 to 200). To match the frequency number of the site amplifications shown above, the number of frequencies of the MHVRs was set to 200 by resampling the frequency interval from 0.3 to 20 Hz. We also used the ratios of SAF to MHVR (as AMR), $AMR(f_{i-2})$, $AMR(f_{i-1})$, $AMR(f_i)$, $AMR(f_{i+1})$ and $AMR(f_{i+2})$ as the output data to be trained. When $i\pm2$ is smaller than 0 or larger than 201, the $MHVR$ and AMR values at step i were given in the analysis. Table 2-3 summarizes the combination of the input and output dataset with some examples. In the same way, the $i\pm1$, $i\pm3$, $i\pm5$, $i\pm7$, $i\pm9$ steps also represent the same input and output dataset handling as above.

Table 2-3. Example of input and output dataset for training.

Input data (x)							Output data (y)				
f_i	f_M	$MHVR$ (f_{i-2})	$MHVR$ (f_{i-1})	$MHVR$ (f_i)	$MHVR$ (f_{i+1})	$MHVR$ (f_{i+2})	AMR (f_{i-2})	AMR (f_{i-1})	AMR (f_i)	AMR (f_{i+1})	AMR (f_{i+2})
0.300	1.06	2.80	2.80	2.80	2.83	2.86	1.34	1.34	1.34	1.28	1.23
0.306	1.06	2.80	2.80	2.83	2.86	2.89	1.34	1.34	1.28	1.23	1.17
0.312	1.06	2.80	2.83	2.86	2.89	2.93	1.34	1.28	1.23	1.17	1.12
0.319	1.06	2.83	2.86	2.89	2.93	2.96	1.28	1.23	1.17	1.12	1.07
0.326	1.06	2.86	2.89	2.93	2.96	2.99	1.23	1.17	1.12	1.07	1.02
0.333	1.06	2.89	2.93	2.96	1.06	0.33	1.17	1.12	1.07	1.02	1.01
0.340	1.06	2.93	2.96	2.99	1.06	0.34	1.12	1.07	1.02	1.01	1.00
0.348	1.06	2.96	2.99	3.00	1.06	0.35	1.07	1.02	1.01	1.00	0.98

For the different datasets mentioned above, the same DNN model was used for

training. The model used is shown in Figure 2-13. Batch size, which refers to the number of training examples utilized in one iteration, was set to 1,600. Number of epochs was set to 1,000, and learning rate was set to 0.001 since we confirmed in our preliminary analysis that the loss functions of the training and validation samples were not significantly reduced for larger than 1,000 epochs.

The output data of this DNN model is the correction factor from MHVR to amplification factor (*AMR*) for observed MHVR that corresponds to the $EMR \times VH_bR$ in Eq. (**2-1**). The pSAF can be estimated using *AMR* and observed MHVR at each frequency f_i as shown in Eq. (2-12).

$$pSAF(f_i) = AMR(f_i) \times MHVR(f_i) \tag{2-12}$$

Equation (2-12) means that the *AMR* can be derived from the proposing model by simply giving observed MHVR, and consequently SAF can be estimated only from the MHVR even at sites where seismic velocity structure is unknown and seismic observation records are not available.

For different input and output datasets, outshapes of the layers are set differently in the above DNN model due to the different number of input and output layers. Outshapes of the layers represents the number of neurons. Table 2-4 shows the hyperparameters and outshape of the layers when used the different dataset. As the number of input layers increases, the number of hidden layers also increases. The reason for this is that fewer layers cannot accurately train a larger number of input data.

Table 2-4. Hyperparameters and outshape of the layers when used the different dataset.

Layer	Hyperparameters	$i\pm1$	$i\pm2$	$i\pm3$	$i\pm5$	$i\pm7$	$i\pm9$
		Outshape	Outshape	Outshape	Outshape	Outshape	Outshape
Input		5	7	9	13	17	21
Log		5	7	9	13	17	21
Affine1		10	16	16	32	32	48
Selu1	α=1.67, λ=1.05	10	16	16	32	32	48
Dropout	0.5	10	16	16	32	32	48
Affine2		10	16	16	32	32	48
Selu2	α=1.67, λ=1.05	10	16	16	32	32	48
Affine3		10	16	16	32	32	48
Selu3	α=1.67, λ=1.05	10	16	16	32	32	48
Affine4		3	5	7	11	15	19
Mean Absolute error		3	5	7	11	15	19

With confirming that no overfitting had occurred, the pseudo site amplification factor (*pSAF*) was calculated using Eq. (2-12). Since cross-validation was used, all pseudo site amplification factor are the results of the validation dataset that was not involved in the training.

Several evaluation metrics used to validate the accuracy of the pseudo site amplification factors for the above models were described in detail in Section 2.3.2. To numerically assess the estimation accuracies of the different models, goodness-of-fit metrics were calculated using the observed and pseudo site amplification factor. Following the definition of good match, the accuracies of the pseudo site amplification factors were compared using two goodness-of-fit metrics with varying thresholds. And *r* was adopted for evaluating the spectral shape. Both *d* and MAE were calculated for evaluating the amplitude.

Following the definition of good match, the accuracies of the models were compared using two evaluation metrics with varying thresholds. And *r* was adopted for evaluating the spectral shape. Both *d* and MAE were calculated for evaluating the amplitude. Table 2-5 lists the percentages of good match sites when used the different datasets under different thresholds of the evaluation metrics. The accuracy of the DNN model used the *i*±2 dataset was highest than those of the other datasets especially at the threshold of $r > 0.65$ and MAE < 0.20. Therefore, we confirmed that the *i*±2 dataset used was quite successful in the model for estimating the site amplification factors from MHVRs.

Table 2-5. Success rates under different definitions of evaluation metrics.

dataset	r>0.60	r>0.60& d>0.60	r>0.65& d>0.65	r>0.60& MAE<0.25	r>0.65& MAE<0.20
***i*±1**	0.875	0.875	0.850	0.813	0.650
***i*±2**	0.875	0.875	0.850	0.838	0.700
***i*±3**	0.888	0.888	0.825	0.825	0.638
***i*±5**	0.875	0.875	0.825	0.825	0.650
***i*±7**	0.863	0.863	0.825	0.813	0.663
***i*±9**	0.875	0.875	0.825	0.813	0.600

2.5. Results

2.5.1. Result of validation sets

Through the above analysis, we adopted total amount of the input training samples was 12,800 (200 frequencies/site × 8 sites/fold × 8 folds). Batch size, which refers to the number of training examples utilized in one iteration, was set to 1,600. Number of epochs was set to 1,000, and learning rate was set to 0.001 since we confirmed in our preliminary analysis that the loss functions of the training and validation samples were not significantly reduced for larger than 1,000 epochs. The numbers of data in the input and output layers were set to 7 and 5, respectively, as mentioned above. Since we have divided each site data into 200 arrays in terms of frequency, the process needs to be calculated 200 times (i =1 - 200) in order to obtain the complete pSAF spectra. The number of the input and output data, and hyperparameters of the developed DNN architecture are summarized in Table 2-6.

Table 2-6. Hyperparameters and outshape of the layers in the DNN model.

Layer	Hyper-parameters	Outshape (Number of neurons)	Layer	Hyper-parameters	Outshape (Number of neurons)
Input		7	Selu2	α=1.67, λ=1.05	16
Log		7	Affine3		16
Affine1		16	Selu3	α=1.67, λ=1.05	16
Selu1	α=1.67, λ=1.05	16	Affine4		5
Dropout	0.5	16	Mean Absolute Error		5
Affine2		16			

Figure 2-18 shows the loss functions (MAE) for the training and validation data

during the training process. Each line indicates the result for each fold in the cross validations. Since the validation set was randomly selected according to f_M, and the model may reach a local minimum of the error at one or several sites, likewise one or several sites may not fit as well. Therefore, the validation error is more dispersed compared to the training error. However, both the training error and the validation error were converged to a certain value and stabilized, and the average values were 1.03 for the training error and 1.10 for the validation error, respectively. We also confirmed that all the loss functions for the training and validation data were well converged without overfitting. Since the model output was *AMR*, the final pSAF was calculated based on the Eq. (2-12).

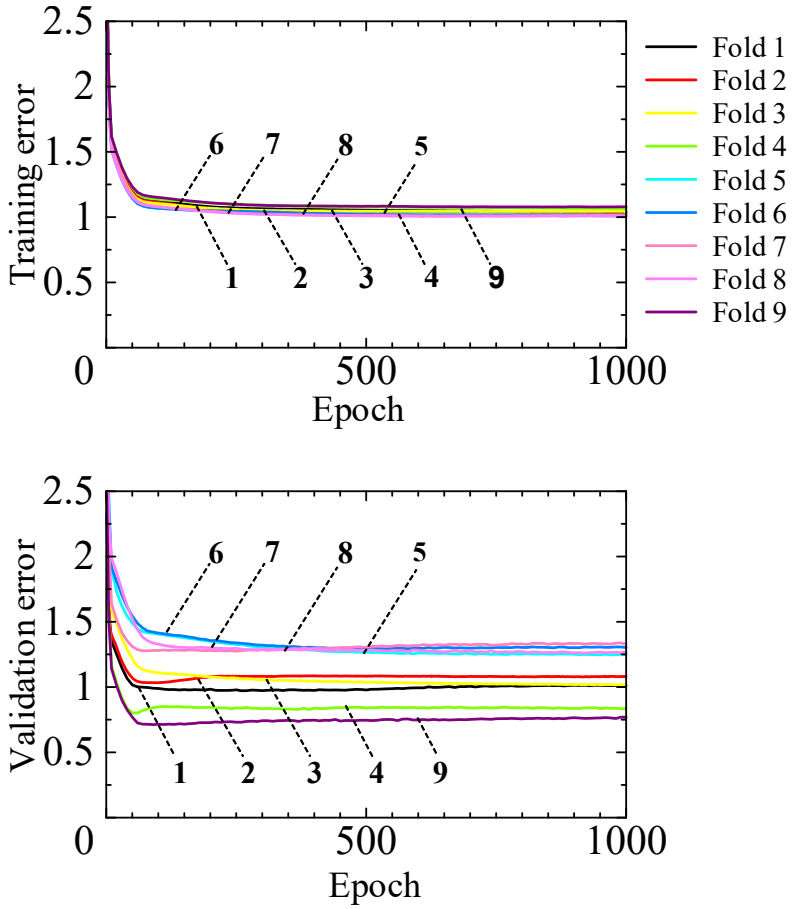

Figure 2-18. Mean absolute error (MAE) for training (upper) and validation (lower) samples obtained from the 9-fold cross validation. Number in the figure indicates the fold number in each validation dataset.

The pSAF estimated by the DNN model in the validation result for the fold 3 are illustrated in Figure 2-19 with the comparison of the observed SAFs by the GIT. Even though the SAFs in Figure 2-19 was not used in the training process, the pSAF estimated by the DNN model shows good agreement with the observation at most sites, which suggests the usability of the DNN model. With the 9-fold cross validation, the estimates for all 72 sites were obtained (All the validation results including fold 1 to fold 9 are shown in the Section 2.7, Appendix). To assess the accuracies of the estimations, we calculated residuals in common logarithmic scale between the observed and estimated SAFs at each frequency as shown by Eq. (2-13).

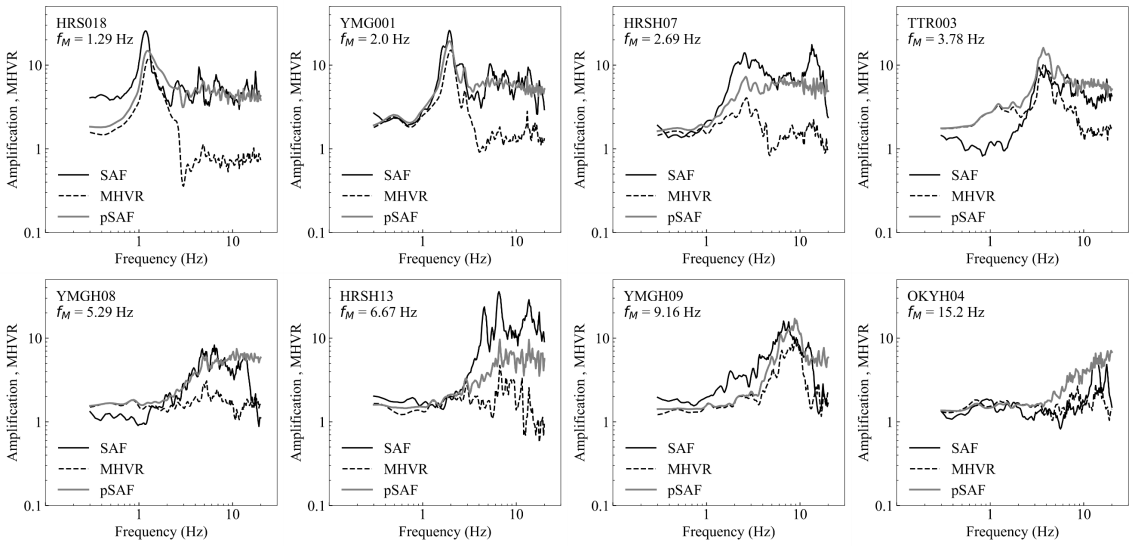

Figure 2-19. Comparison of observed SAFs by GIT and pSAFs estimated by the optimal DNN model and the double empirical correction (DEC) method based on *EMR*s and empirical *VHbR*s by Kawase et al. [32] for the test sites.

$$\text{Residual}(f_i) = log(SAF(f_i)) - log(pSAF(f_i)) \tag{2-13}$$

The residuals were calculated at each site in each fold. Figure 2-20 shows the mean

and one standard deviations of the residuals obtained from the validation results of the 9 folds. Whereas the residuals in the frequencies higher than 10 Hz almost reach 0.4 which illustrates a negative bias at high frequencies, the residuals in other frequencies are smaller than 0.2.

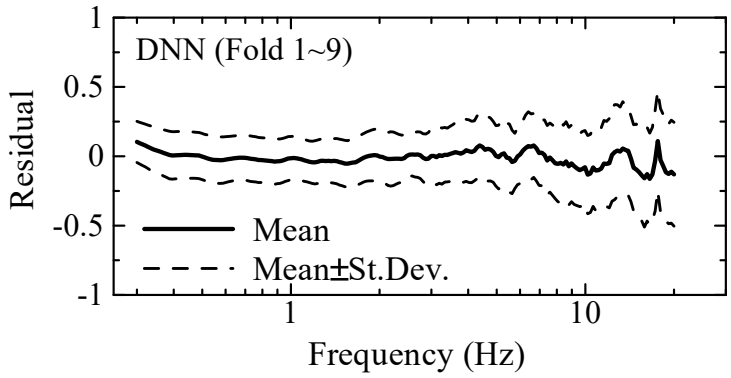

Figure 2-20. Residuals in logarithmic scale between observed and pseudo SAFs. Mean value and standard deviation in each frequency are calculated for all validation data at 72 sites on 9-fold cross validation.

2.5.2. Result of test set

The above results validate that the predictive ability of the selected optimal model can be considered well. However, the model validation by the internal cross validation techniques is not enough, as although the validation set is not used in training, it is incorporated into the process of tuning and selecting the model in *k*-fold cross validation session. Therefore, a completely training-independent external test set for evaluating the generalization performance of the final model is considered necessary [48].

The final model was developed using data from 72 sites on training set and validation set (i.e. 14,400 input data) which exclude the sites of test set, and hereafter we called it as the optimal model. Then the pSAFs for the test set at the eight sites were estimated by applying the developed optimal model. As shown in Figure 2-21, the generalization

performance of the DNN model on a completely unknown external test set was also excellent. However, for the SAF with very large peak values, such as HRSH18, the pSAF was not fully fitted with the observation. This could be explained by the fact that there were less similar site data with large peak values in the training set. Since the number of training samples that include large peaks in SAF was not enough, it would be difficult to extrapolate such large peak values from the current optimal model, which could be an existing drawback of the DNN model. We also calculated logarithmic residuals between the observed and pseudo SAFs at each frequency as shown in Figure 2-22. The result for the test set seems to be superior to the validation set especially on the low frequency range. One of the reasons for smaller residuals on the test set would be that the number of training data in the optimal model (14,400 input data at 72 sites) was increased compared to that in the previous model developed in the cross-validation process (12,800 input data at 64 sites).

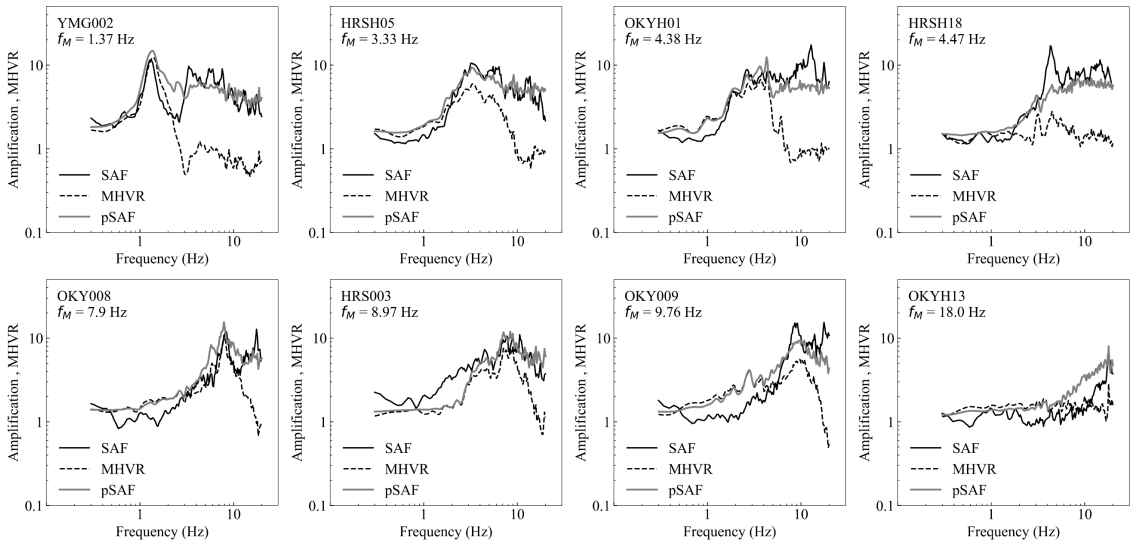

Figure 2-21. Comparison of observed SAFs by GIT and pSAFs estimated by the optimal DNN model on training-independent external test set for evaluating generalization performance.

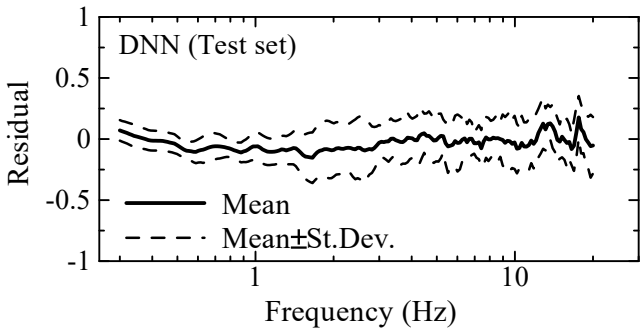

Figure 2-22. Residuals in logarithmic scale between observed and pseudo SAFs for test set.

2.5.3. Comparison with previous method

To validate the performance of the DNN model even more, we compared the estimated pSAF by the DNN model with those by the double empirical correction (DEC) method by cc. By delineating the *EMR*s and the empirical VH_bRs shown in Eq. (2-1) from the results in Kawase et al. [32], the DEC method was applied to the MHVRs at the test sites to calculate pSAFs. Figure 2-23 shows the comparison of the two pSAFs with the observed SAFs, indicating that the DNN model shows better agreement with the observed SAFs, especially at the lower frequency than 4.0 Hz.

In order to discuss the applicability of the DNN model for other districts in Japan, we applied the DNN model to the MHVRs observed outside Chugoku district. Since Kawase et al. [32] introduced the MHVRs and SAFs at KGW003, EHM011, FKOH01 and FKOH03, the SAFs were estimated by applying the DNN model to the delineated MHVRs. These sites were selected because the peak frequencies of the MHVRs are within the coverage of f_M in our dataset (1.0 – 20.0 Hz). Figure 2-24 shows the comparison of the observed SAFs in Kawase et al. [32] and pSAFs estimated by the DNN model and the DEC method. Surprisingly the pSAFs by the DNN model also show better agreement with the SAFs than the those by the DEC despite the observed MHVRs and SAFs were fully derived in Kawase et al. [32].

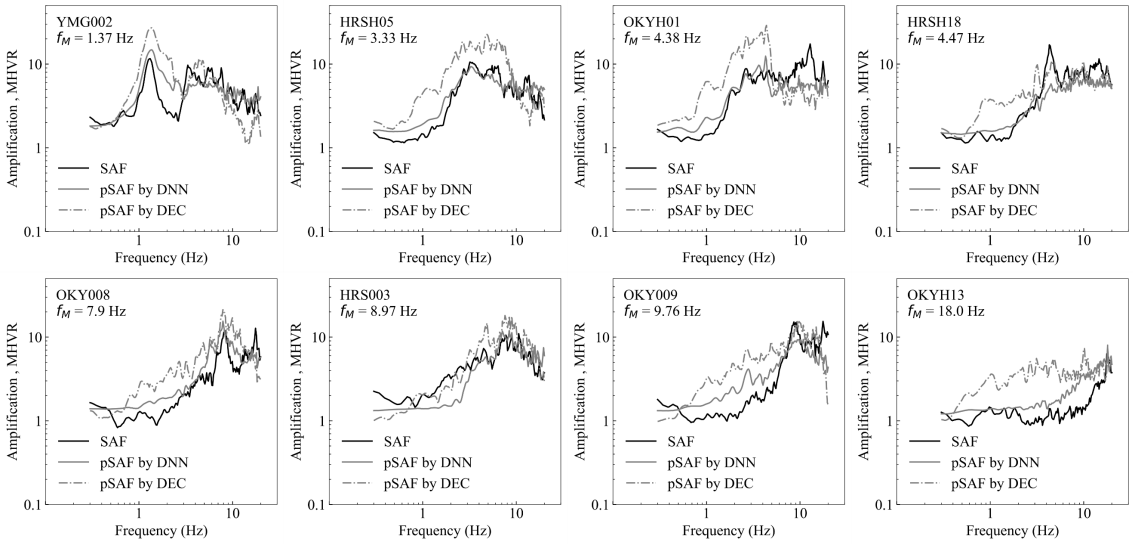

Figure 2-23. Comparison of observed SAFs by GIT and pSAFs estimated by the optimal DNN model and the double empirical correction (DEC) method based on *EMR*s and empirical *VHbR*s by Kawase et al. [32] for the test sites.

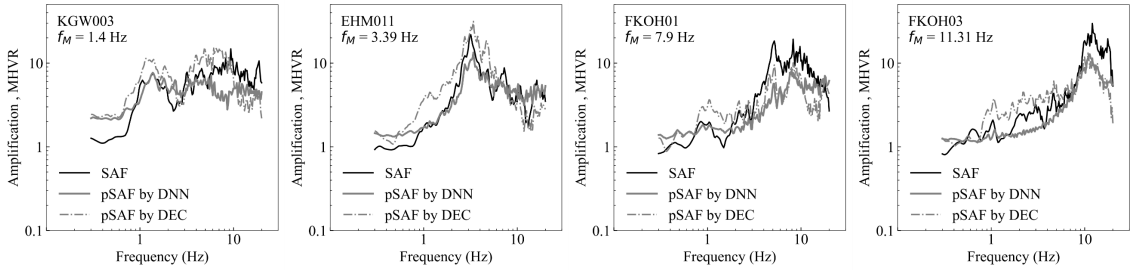

Figure 2-24. Comparison of observed SAFs by GIT and pSAFs estimated by the optimal DNN model and the DEC method at four sites (KGW003, EHM011, FKOH01 and FKOH03) located in non-Chugoku district introduced in Kawase et al. [32].

We assessed the residuals between SAFs and two pSAFs for the test eight sites in Chugoku district and the four sites in non-Chugoku district. As shown in Figure 2-25, the average residual of the DNN model is significantly smaller than the DEC when using the test sites in Chugoku district. In contrast, when using the four sites from Kawase et al.

[32], the difference between the two average residuals is relatively small. However, the residual of the DNN model seems smaller than that of the DEC.

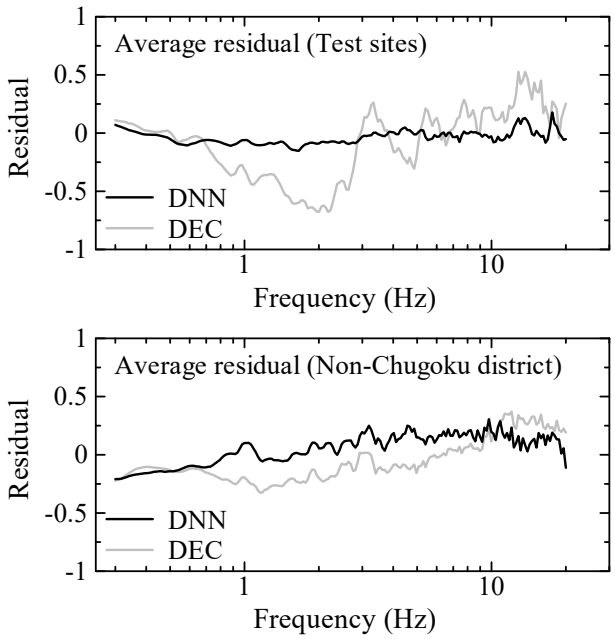

Figure 2-25. Comparison of the average residual of SAF and pSAF between the DNN model and the DEC method. The upper figure shows the average residuals for the test sites, and the lower figure shows the average residuals for the sites in non-Chugoku district.

In order to comprehensively quantify the estimation accuracies, root mean square errors (RSMEs) between the observed SAFs and estimated pSAFs by the DNN model and the DEC were calculated for the validation and test datasets in Chugoku district and data at four sites in non-Chugoku district shown above based on the equation below.

$$RMSE = \sqrt{\frac{1}{n}\left\{\sum_{i=1}^{n}(log(SAF(f_i)) - log(pSAF(f_i)))^2\right\}} \tag{2-14}$$

Figure 2-26 shows the distribution of the *RMSE* obtained at each site for (a) the validation sites, (b) the test eight sites and (c) the four sites in non-Chugoku district. In our results, the percentage of the number of sites in the validation set with *RMSE* below 0.2 was 50 %, the percentage with *RMSE* below 0.3 was 92 %, and the percentage with *RMSE* below 0.4 was 99 %, there was only one site with the *RMSE* of 0.42. All the *RMSE* in the test sites and non-Chugoku district sites are smaller than 0.25. We confirmed that smaller *RMSE*s are found in the DNN not only for test sites but also for the sites in non-Chugoku district. These results suggest that our proposed DNN model could be applicable not only for Chugoku district but also for other districts in Japan.

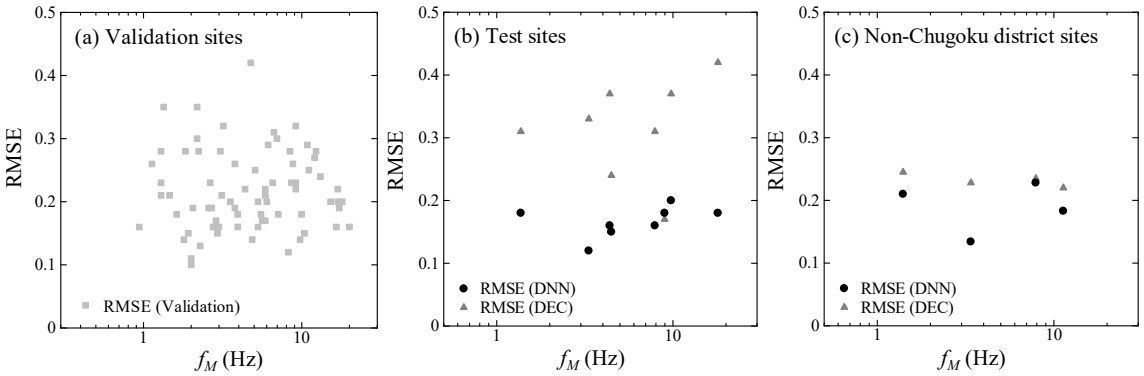

Figure 2-26. Distribution of the *RMSE* obtained at each site for (a) the validation sites, (b) the test eight sites and (c) the four sites in non-Chugoku district.

We have to pay attention to that the number of available samples in the current DNN model is still limited, indicating the limitation of the current DNN model in expanding to estimations of nationwide or global site amplification factors. As described before, increase of training samples and expansion of variation of samples especially with lower peak frequencies (i.e. < 1 Hz) and large amplification factors (i.e. > 10) would be key for more accurate and robust estimation of site amplifications in future studies.

2.6. Conclusions

In this study, we developed the deep neural network (DNN)-based method for estimating S-wave site amplification factors (SAFs) from microtremor H/V spectral ratios (MHVRs). We analyzed the site amplification factor derived by the generalized spectral inversion technique (GIT) and observed MHVRs at K-NET and KiK-net sites in Chugoku district, Japan. Excluding 25 site data where the observed MHVRs were flat or showed very small peaks, we confirmed that the peak frequencies of the MHVRs well agree with those of the site amplification factors at the 80 sites.

The DNN model developed in this study consists of multiple Affine and nonlinear activation layers with dropout. After the observation data was divided to training set, validation set and test set, the supervised learnings which requires input data and desired output data were performed to the training set. The most appropriate input and output data were selected for cross-validation. The input and output data for $i\pm2$ steps were used to accurately reproduce the spectral shape of the observed SAF. In this study, 9-fold cross validation technique was applied to validate and select the developed DNN architecture. The results of the cross validation showed that the loss functions were well converged in the learning process without overfitting not only in the training samples but also in the validation samples. Then the pseudo SAF was computed with the trained optimal model for eight sites in the test set that was completely independent of the training and validation set. The generalization performance of the DNN model on a completely unknown external site was also excellent. Finally, we compared the model with the existing method and concluded that the DNN model showed better performance in either Chugoku district or non-Chugoku district.

The developed DNN model does not require any hard-to-get data such as seismic velocity structures and damping models thus providing a significant cost-benefit. The DNN model can be easily developed as more records are collected and further improve accuracy compared to the conventional model such as regression-based approach. However, the number of the training samples analyzed in this study (80 sites) is still

limited, indicating the limitation of the current DNN model in expanding to estimations of nationwide or global site amplification factors. We expect that the model should be optimized with a larger dataset to get better performance in future studies.

2.7. Appendix

Result of all validation sets including fold 1 to fold 9. (Black line: Site amplification factor by generalized spectral inversion technique (GIT), Dash line: Observed microtremor H/V spectral ratio (MHVR), Grey line: Pseudo site amplification factor estimated from MHVR by the developed DNN model)

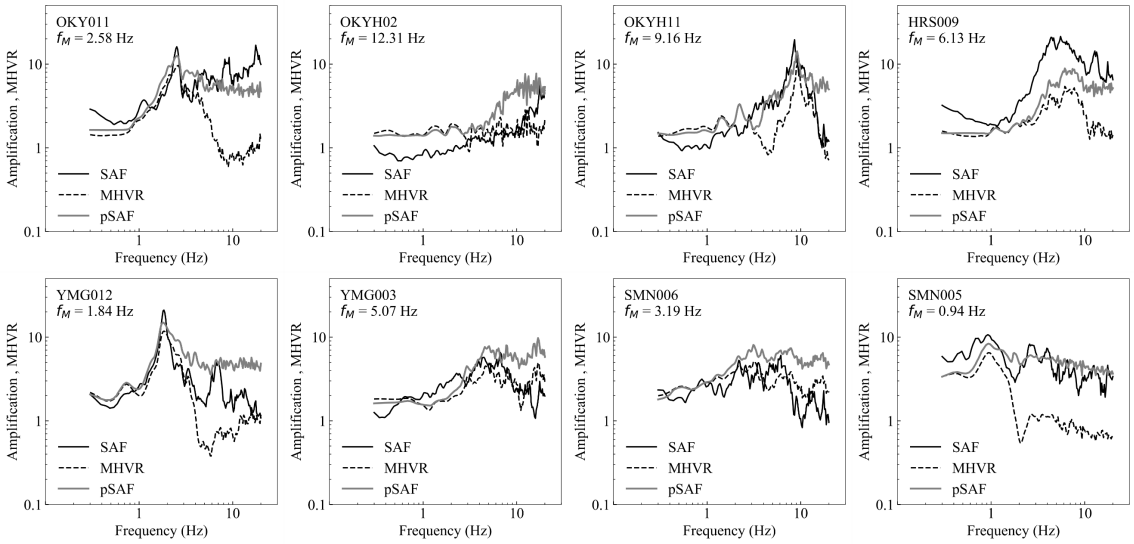

Figure S2-27. Result of validation for fold 1.

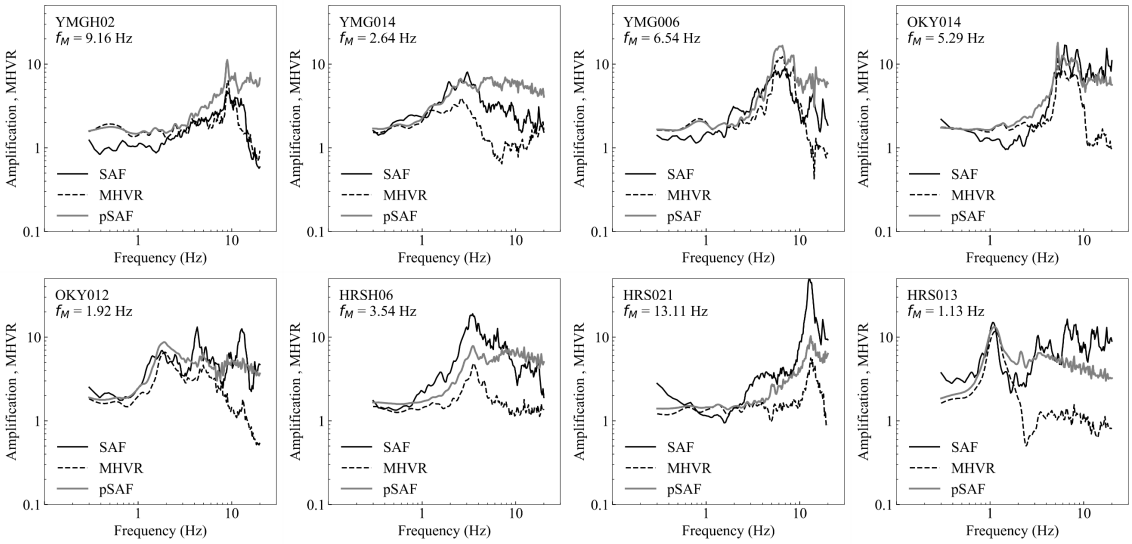

Figure S2-28. Result of validation for fold 2.

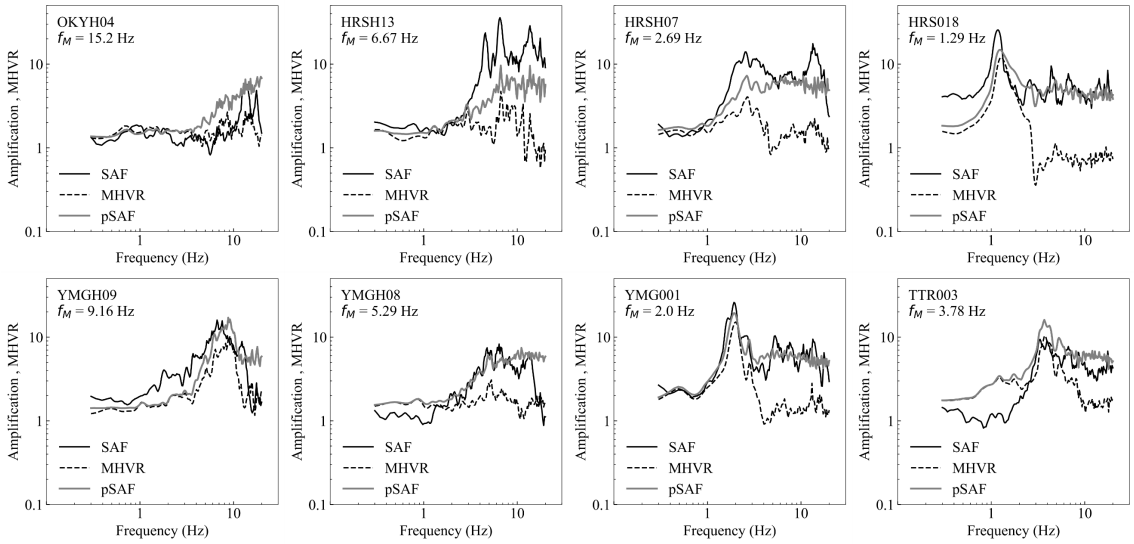

Figure S2-29. Result of validation for fold 3.

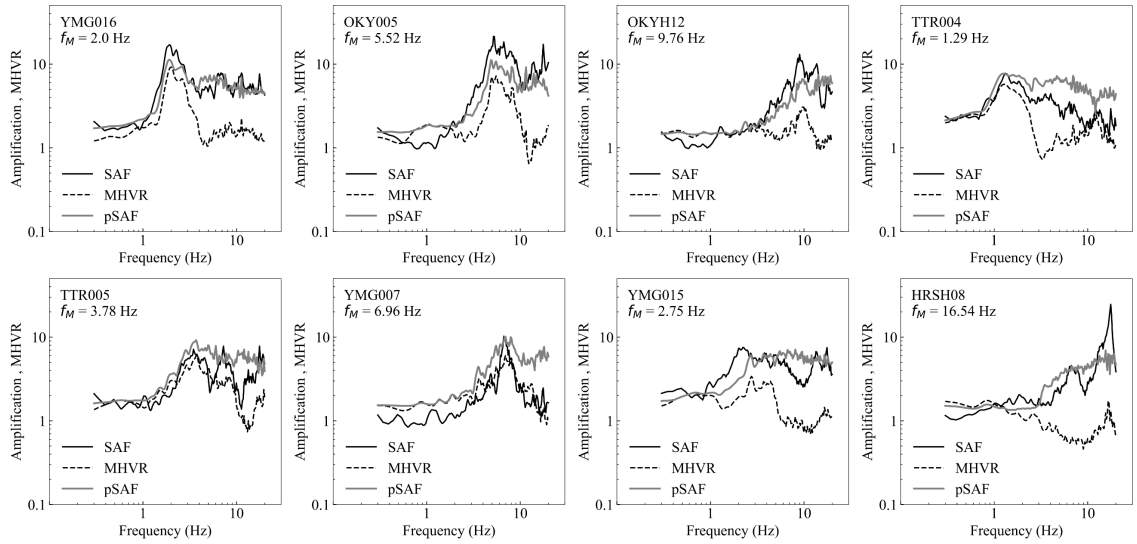

Figure S2-30. Result of validation for fold 4.

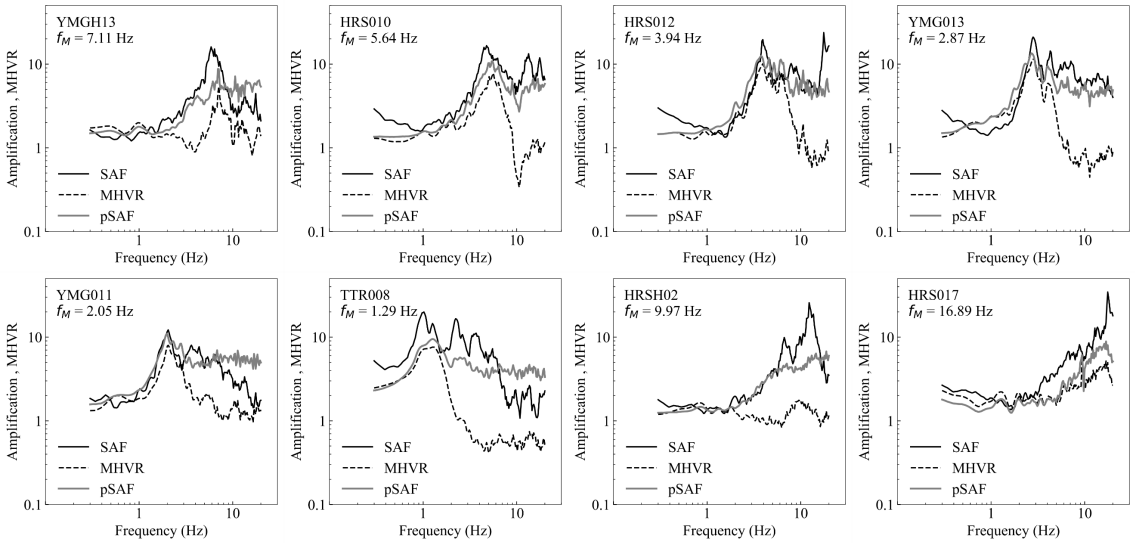

Figure S2-31. Result of validation for fold 5.

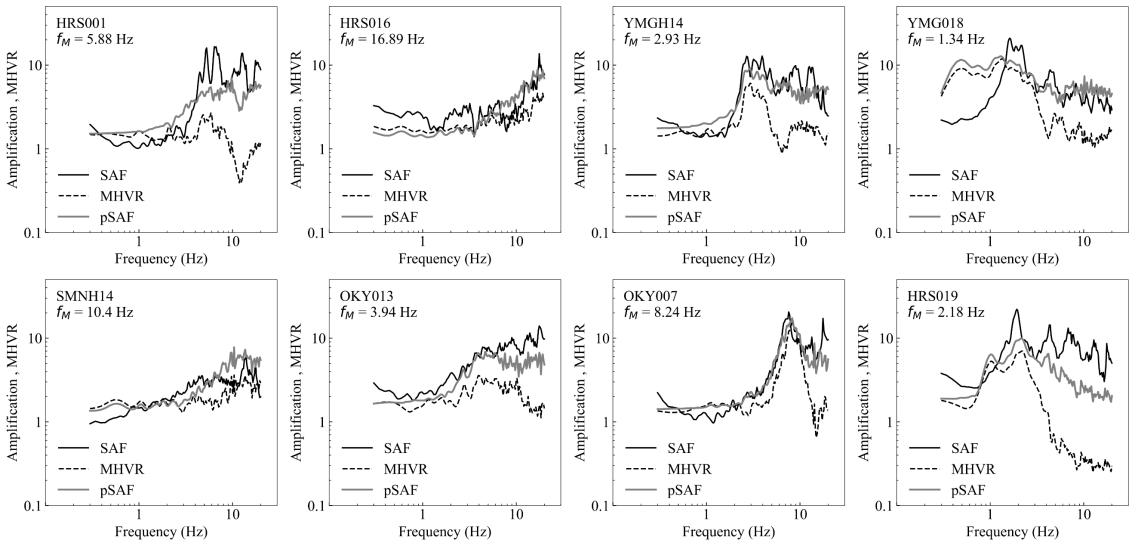

Figure S2-32. Result of validation for fold 6.

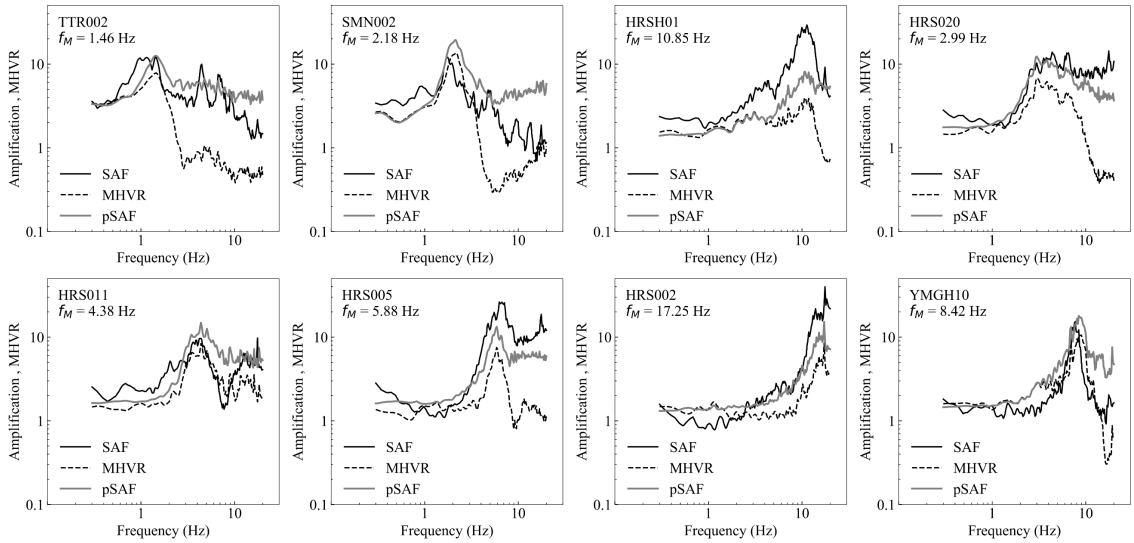

Figure S2-33. Result of validation for fold 7.

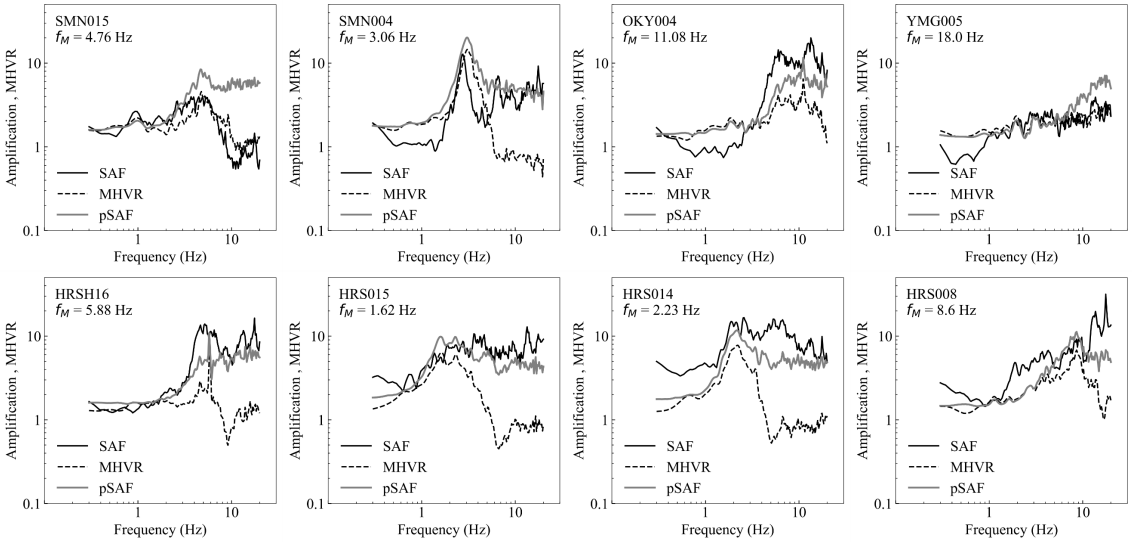

Figure S2-34. Result of validation for fold 8.

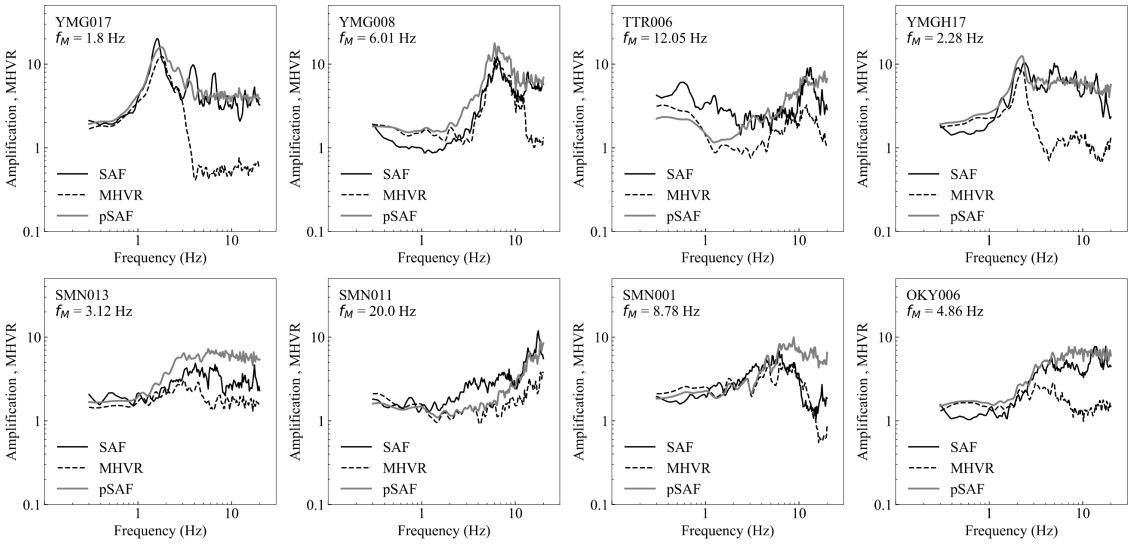

Figure S2-35. Result of validation for fold 9.

Chapter 3. Expanding MHVR-to-SAF Prediction to Data-Limited Areas Using Transfer Learning

In the previous chapter, we introduced a simple and cost-effective method for estimating site amplification factors (SAFs) using a deep neural network (DNN) model and microtremor horizontal-to-vertical spectral ratios (MHVR). Since the previous DNN model was based on the observed SAFs and MHVRs within a limited district in Japan, the applicability of the previous model to non-source regions with different site conditions was limited. In this chapter, we explore the application of a transfer learning (TL) technique to adapt an existing (pre-trained) DNN model to new regions and a different database. We compiled SAFs obtained through the generalized spectral inversion technique (GIT) at seismic observation stations (K-NET and KiK-net) in Japan as the ground truth for site effects. MHVRs recorded at these stations across several districts in Japan were collected to construct a dataset for developing the TL model. We then constructed a TL model by leveraging the neural network layers and their weights from the pre-trained model while incorporating additional layers to enhance performance. Comparisons between the various models are also discussed to evaluate the improvements achieved through the TL approach.

3.1. Background

Seismic ground motions are generally determined by source characteristics, propagation path characteristics, and site effects. Site effects, also known as site amplification factors (SAFs), significantly influence and modulate the amplitudes of seismic ground motions at the surface. This is because the subsurface ground conditions at a site can amplify incoming seismic waves. Theoretical and/or empirical approaches have been examined for evaluating SAFs. In the theoretical approaches, SAF at a target site can be obtained by multiple reflection theory based on subsurface seismic velocity structure model and damping factors [49,50]. The obtained theoretical SAF for S-wave is generally represented as a one-dimensional transfer function. However, this approach requires detailed seismic velocity structure down to a bedrock, and is limited in their applicability. Besides, accuracies of the transfer functions depend on the quantity of available soil information. Several studies have reported inconsistencies between the SAFs derived from theoretical approaches and the observed values at certain sites. These discrepancies are likely attributed to uncertainties in the velocity models and the complex nature of seismic wave propagation [51–54].

As presented in the above sections (see Section 1.2.1, Section 2.2.2), in empirical approaches, SAFs have been evaluated from seismic ground motion records by generalized spectral inversion technique (GIT). Source characteristics, propagation path effects and SAFs can be derived by the GIT from Fourier spectral data of seismic motion records at multiple sites [4,5]. The ability of GIT to operate without requiring detailed subsurface information at the observation sites has facilitated its application in a wide range of countries and regions [7,8,40,55–57]. However, the GIT based on strong seismic records does have the drawback that it is only a point-by-point estimation and cannot be applied to sites where strong seismic observations have not been observed.

Recently, another approach based on horizontal-vertical spectral ratio (HVR) of observed ground motion has been discussed in estimating SAF [9,20,24,58]. The HVR is further distinguished into HVR of earthquake motions (EHVR) and HVR of microtremor

data (MHVR). According to the previous studies [11,25,30,31], the HVRs have the potential to reliably identify the predominant frequency of seismic ground motion at a site and can be used to infer the S-wave velocity structure of the subsurface. However, HVR itself tends to underestimate the actual SAF in terms of amplitude [59]. In order to develop an additional correction factor, Kawase et al. [32] proposed a double empirical correction method in estimating SAF from MHVR.

On the other hand, machine learning (ML) techniques have been widely applied to seismological tasks in recent years; such as ground motion prediction [60,61], rapid seismic wave discrimination [62], etc. According to the review for applications of ML techniques in the seismic problems introduced by Mousavi and Beroza [63,64], the ML techniques, which can handle large amounts of high-dimensional data, can be a powerful tool for seismic analysis as the volume of seismic data continues to expand. In the task for evaluating SAF, Zhu et al. [65]compared the ML-based model for EHVR-estimated SAF with numerical simulations, and indicated that the ML model was superior to the conventional approach. In Chapter 2, we proposed a deep neural network model (DNN) for directly estimating SAF from MHVR using the spectral ratio between MHVR and SAF (amplification-to-microtremor ratio: AMR) as an intermediate medium [66]. These HVR-based approaches do not require large-scale subsurface structural profiles or large amounts of seismic motion data for inversion. Thus, the difficulty of estimating SAF at a target site can be greatly reduced. From an implementation point of view, MHVR is much easier and more cost-effective because microtremor data can be collected as part of a field survey, whereas EHVR still requires long-term observations and limited to seismically active regions.

The limitation of the previous HVR-based approaches is that the results obtained from the methods are limited to the domain of the training data, and the applicability of the methods to other domains is not certificated. As Zhu et al. [65] pointed out, the ML-based model based on a Japanese database did not yield satisfactory results in regions outside of Japan. Thus, in short, the task of estimating SAF from HVR still has the

following three impediments.

1) How to transfer the already developed model from the used region to other new regions.

2) How to extend the model to different databases collected. (Such as the difference in the GIT method used to obtain SAF).

3) How to improve the accuracy of the model for different regions with limited SAF data.

Transfer learning, one of the techniques in machine learning, has garnered significant attention and discussion due to its adaptability and effectiveness in diverse domains [67,68]. TL aims to adapt learned knowledge through domain similarity to new domains and tasks using the knowledge learned from source domains. The technique can be explained by transferring or fine tuning the parameters of a pre-trained model to a new domain to help improving the performance and generalization of the model. It makes the model to have better generalization ability and requires less training data [50,69–71]. Particularly, the TL is expected to be extremely valuable in some seismological applications for which the labeled data are limited.

MHVR-based approaches would be simple and effective in evaluating SAFs particularly for regions with sparse seismic observation stations because MHVRs can be readily acquired at any location without earthquake records. Moreover, MHVR-derived SAFs could potentially facilitate simplified and more accurate ground motion evaluations and predictions. However, as mentioned earlier, the applicability of our previous DNN model was limited when applied to new regions. Additionally, reconstructing new models for these new regions requires a substantial amount of site data. In this study, in order to address these tasks, we delved into the utilization of the TL approach to construct estimation models applicable to new regions. We developed a new TL model based on the weight parameters trained in the previous DNN model by Chapter 2 (Pan et al. [66]). The site data for the previous model originated from the Chugoku district in western Japan,

while the new TL model utilized site data from other districts across Japan. A small portion of the new dataset was used for training and validation, and most of the remaining site data was used as the external test set. We employed k-fold cross validation cross-validation strategy and residual analysis to evaluate the performance of the trained TL model on the test set. Additionally, we conducted further experiments to investigate the effects of different elements on the performance of the proposing TL model, including the comparison with the DNN models with extended training sets.

3.2. Data

3.2.1. Location of MHVRs

The sites located in regions completely different from Chapter 2 were used as the target in this study. As shown in Figure 3-1, we conducted microtremor observation at ground surfaces near K-NET and KiK-net sites, the nationwide seismic observation networks in Japan, which are operated by the National Institute for Earth Science and Disaster Resilience (NIED). The microtremor data in this study are distributed in the following four districts of central and eastern Japan: Chubu, Kanto, Tohoku, and Hokkaido (148 sites in total) whereas the sites by Chapter 2 were located only in the Chugoku district.

After collecting the microtremor data, we followed the same processing as Chapter 2. The Fourier spectra were calculated using a Parzen window smoothing with a 0.3 Hz bandwidth. We calculated the MHVR of the site using Eq. (**2-5**). Following Chapter 2, the average value and one standard deviation of the MHVR of the 10 recording segments of 20.48 s with less noise were calculated as the MHVR of the site.

We set the criteria for the observed MHVRs in developing a TL model. The first criterion is that the maximum amplitude of MHVR is larger than a factor of two in order to identify predominant frequency [42,66]. The second criterion is that the predominant frequency of MHVR is found between 0.3 to 10 Hz. The frequency range for this study was determined based on the range of SAFs covered, as detailed in the following section. We excluded the data at 36 sites because the MHVRs at the sites did not satisfy the criteria. Totally, we used a total of 112 sites as target sites in the following analysis. Figure 3-2 shows the average MHVR with one standard deviation and the MHVRs of 10 segments at AOM012 as a typical example. We can clearly identify the predominant frequency of around 1 Hz at the site.

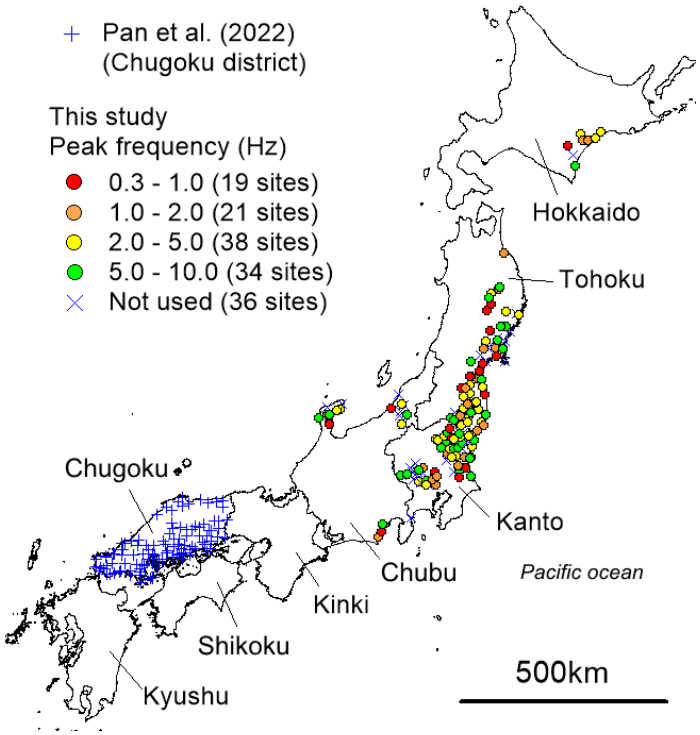

Figure 3-1. Locations of the seismic observatories for the MHVR and SAF sites in this study (circles) and not used sites (diagonal crosses); and locations of the training set sites used in the DNN model constructed by Pan et al. (2022) (crosses). The colors represent the peak frequencies in Hz.

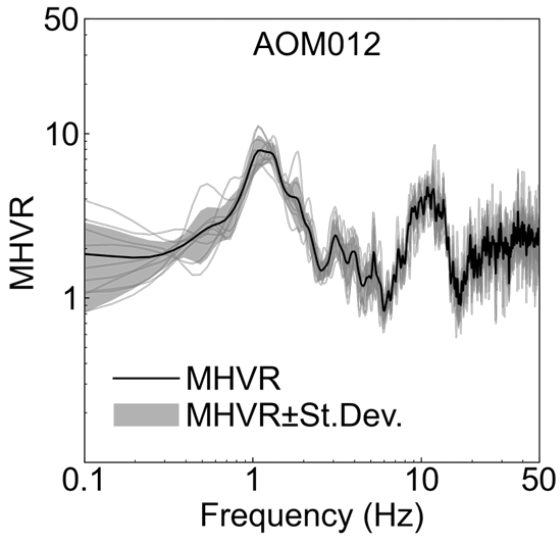

Figure 3-2. Average (Black line) and standard deviation (Gray area) of observed MHVRs at AOM012. Gray thin lines indicate the MHVRs of 10 segments.

3.2.2. SAFs

In this study, we used the SAFs obtained by the GIT in Nozu et al. [55] because they estimated the SAFs at all the K-NET and KiK-net sites in Japan whereas the SAFs in Chapter 2 were limited in the Chugoku district. Additionally, Nozu et al. [55] extracted the SAFs by analyzing long duration of the seismic motions including not only the S-wave parts but also the subsequent phases. On the other hand, the SAFs in Chapter 2 limited the amplifications of S-waves in the seismic motions. We aimed to extend the model to predict SAFs at different regions obtained by different GIT.

The conditions and criteria in the assessment of SAF by GIT in Nozu et al. (2007) are described below. The Fourier spectra of the seismic motion observation record can be represented as Eq. (3-1) below.

$$O_{ij}(f) = S_i(f)P_{ij}(f)G_j(f) \tag{3-1}$$

where O represents the observed spectra of seismic motion, S, P, and G represent the source characteristics, propagation path characteristics, and site effect, respectively; i and j are used to denote seismic events and observation sites.

Considering the geometric attenuation (1/r) of the spherical diffusion of body waves from the epicenter, the propagation path properties can be expressed as Eq. (3-2).

$$P_{ij}(f) = \frac{1}{R_{ij}} exp\left(-\frac{\pi f R_{ij}}{QV_s}\right) \tag{3-2}$$

where R is the source distance, V_s is the S-wave velocity in the propagation path, and Q is the attenuation coefficient. A factor of one in SAFs at rock sites was used as given conditions in solving the simultaneous equations shown in Eq. (3-1). The time window for computing the Fourier spectra of the observed records was selected at 160 s from

arrivals of P-waves. The target frequency range was 0.3-10 Hz. The obtained SAFs represent site amplifications of seismic waves at the ground surface from the outcropping seismic bedrock with the S-wave velocity of approximately 3,000 m/s.

3.2.3. Comparison of two SAFs

As described above, different time windows for analyzing Fourier spectra were applied in the GITs of Chapter 2 and Nozu et al. [55]. The time window of 10.24 s was applied in Chapter 2 in order to extract the SAFs of S-waves whereas much longer time window of 160 s was applied in Nozu et al. [55] to evaluate the SAF of S-waves and subsequent phases of ground motions. Figure 3-3 shows the comparison of two SAFs at four representative sites in the Chugoku district. Since the SAFs by Nozu et al. [55] contain the amplification of the subsequent phases, the SAFs by Nozu et al. [55] are larger than those in Chapter 2. The residuals of the two SAFs were calculated for each site using Eq. (3-3).

$$Residual(f_i) = \log\left(NewSAF(f_i)\right) - \log\left(PreSAF(f_i)\right) \tag{3-3}$$

Here, *NewSAF* refers to the SAFs by Nozu et al. [55], and *PreSAF* refers to the SAFs in Chapter 2. Figure 3-4 shows the average of the residuals and its standard deviation for the 80 sites in the Chugoku district. Positive residuals indicate that the SAFs by Nozu et al. [55] were larger than those in Chapter 2. Assessing the amplifications of later phases could be crucial for evaluating site effects in sedimentary basins, as basin-induced surface waves propagate within these basins and have been observed in the latter part of seismic records [54,72,73]. The use of the SAFs by Nozu et al. [55] indicates that the TL model developing in this study would expand the region for application. Additionally, the coverage of the previous DNN model was limited to sites with predominant frequencies of microtremors (f_M) higher than 1.0 Hz whereas the sites with f_M of around 0.3 Hz are

included in the study area as shown in Figure 3-1. It means that the proposed TL model can expand the coverage of predominant frequencies of microtremors.

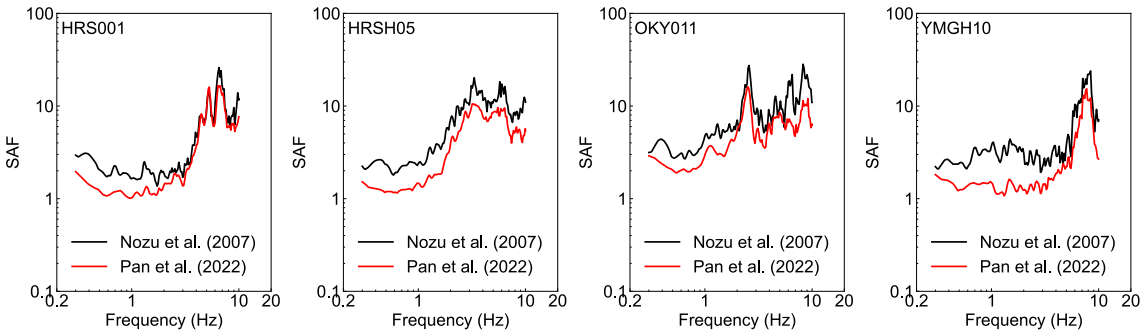

Figure 3-3. Comparison between the two SAFs located in the Chugoku region. Black and red lines represent the SAFs by Nozu et al. [55] and Chapter 2 (Pan et al. [66]), respectively.

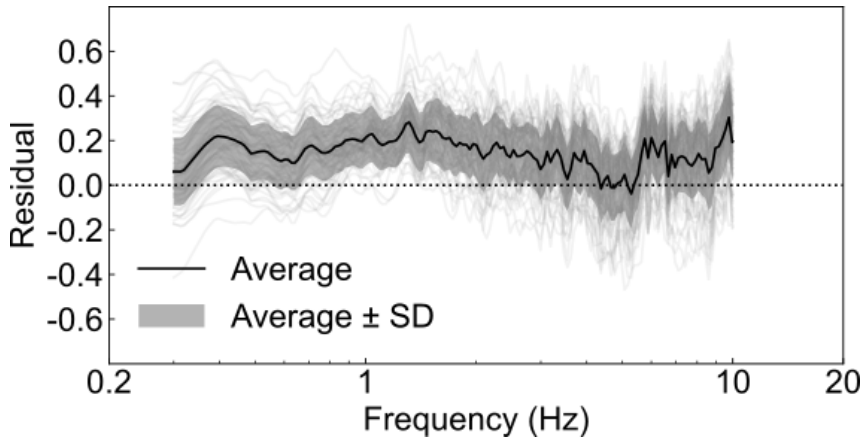

Figure 3-4. Residuals of the two SAFs (gray lines) at 80 sites in the Chugoku district. Black line and gray hatch indicate the average and that with the standard deviation, respectively.

3.3. Transfer learning model

3.3.1. Transfer learning technique

Transfer learning (TL) is a machine learning technique that involves applying knowledge gained from a source dataset to a different but related target dataset. Specifically, it refers to the process of leveraging or fine tuning a pre-trained model, which has learned representations of features from a large dataset, to a new problem or task with a smaller dataset. This approach aims to improve or expand the performance of the model on the new dataset by transferring the weights in different layers in the pre-trained model. It is important to note that the term "source" in transfer learning does not refer to the source of an earthquake location in seismology, but rather to the source dataset from which knowledge is transferred.

Based on the consistency between the feature space of the source and target dataset, TL can be categorized into homogeneous TL and heterogeneous TL [68]. Homogenous TL refers to the situation where the source and target dataset have the same input and output label space, but have different data distributions. In homogeneous TL, a common method is that the knowledge learned from the source dataset can be reused to the target dataset. As shown in Figure 3-5, the model constructed using the source dataset is called a pre-trained model. Then the knowledge is extended to a new target dataset, i.e. reuse the weights of layers. For the new model, most of these hidden layers are reused; a new fully connected layer is added, and only the weights of the newly added layer are updated during training. By leveraging the knowledge gained from the source dataset, TL can improve or expand the learning performance of the target dataset and reduce the need for additional data or computational resources. It is also because the number and the weights of layers to be trained in the new model are fewer than in the model of the source dataset. In this study, the DNN model proposed by Chapter 2 (Pan et al. [66]) was used as a pre-trained model that was trained on the source dataset (Chugoku district shown in Figure 1). The MHVRs and SAFs from other regions (Chubu, Kanto, Tohoku, and Hokkaido) were used as the target dataset. A new TL model was constructed by employing the above

strategy and training by the target dataset. Then the performance of the TL model was compared with the pre-trained model.

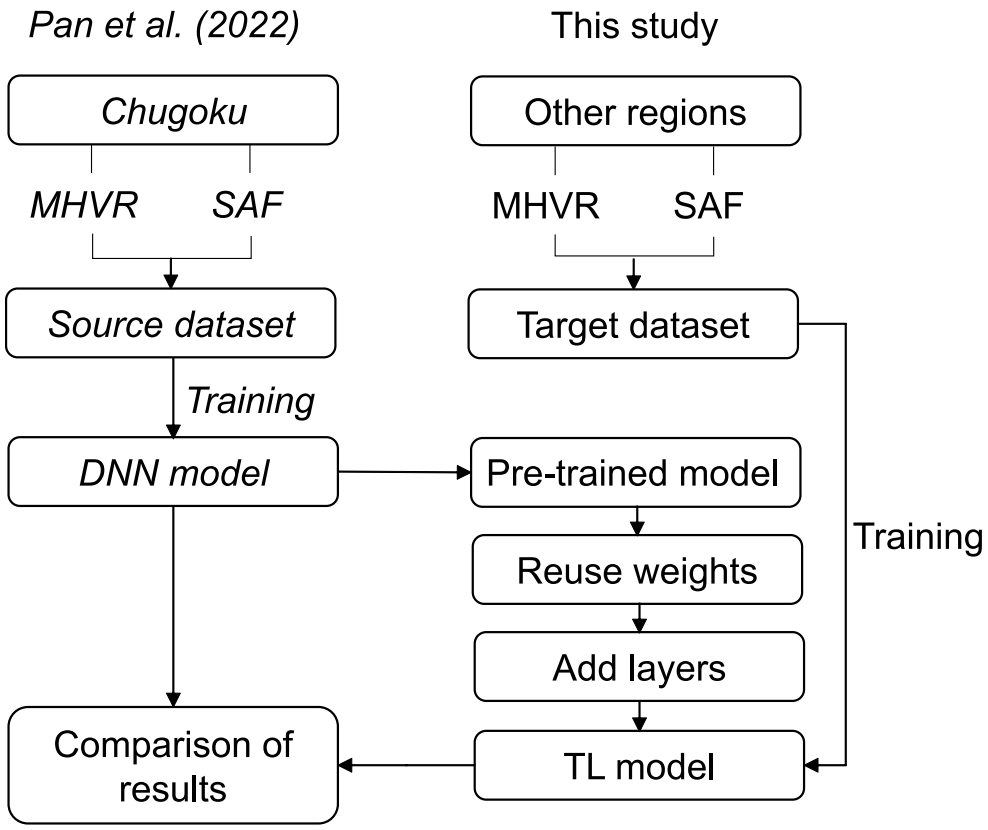

Figure 3-5. Flow of homogeneous TL model adopted in this study. Italics represent the work of Chapter 2 (Pan et al. [66]). This method achieves transfer of the models by reusing the weights of the neural network layers from the pre-trained model to the TL model.

3.3.2. Construction of TL models for prediction

The pre-trained DNN model by Chapter 2 was a deep neural network that contains input layer, multiple hidden layers, and output layer as shown in Figure 3-6. The hidden layers basically consisted of a combination of affine (linear) and nonlinear activation layers. The weights of the hidden layer were optimized by backpropagation of errors

during the training process. The activation function used in the hidden layers was Scaled Exponential Linear Unit (SELU), which is a type of activation function that allows for negative output values [45]. The loss function used in the training process was mean absolute error (MAE), which is a metric that measures the average absolute difference between the target values and predicted values.

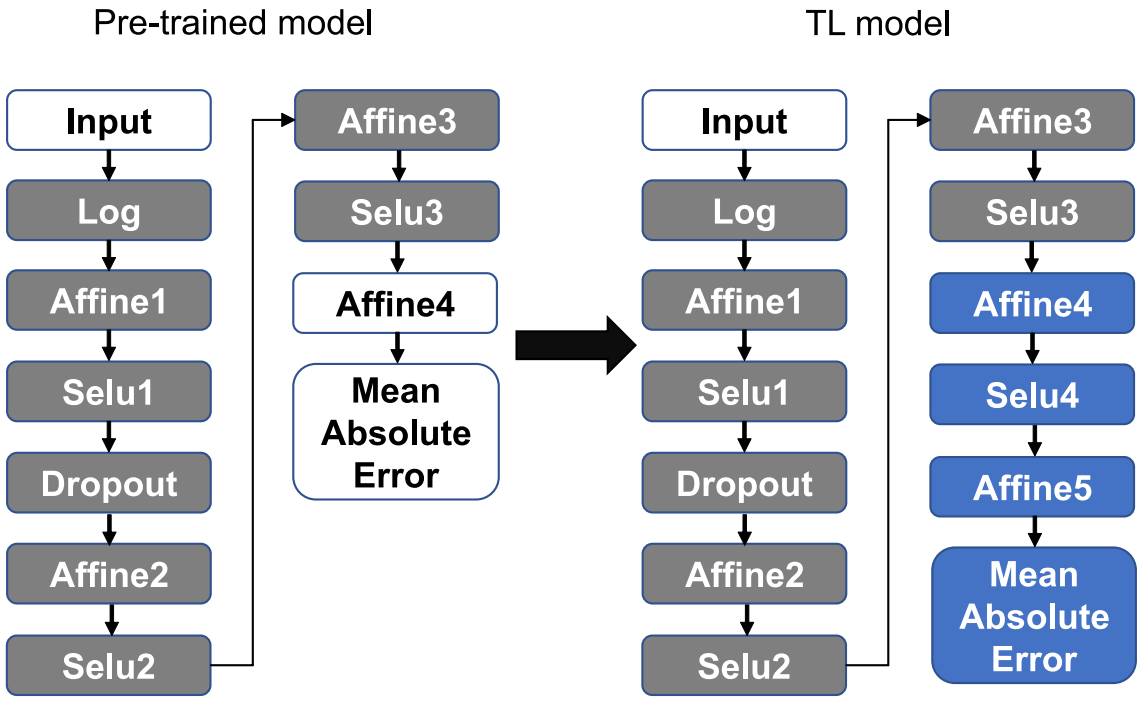

Figure 3-6. Comparison of the TL model with the pre-trained DNN model. Affine denotes a linear transformation consisting of a matrix multiplication and addition of a bias term in a neural network layer; Selu represents an activation function; DropOut serves as a regularization technique, involving the random deactivation of neurons during training to mitigate overfitting. The mathematical representations for Affine, Selu and Mean absolute error (MAE) can be referred in Section 2.4.1.

Considering the fact that the complementary correction between MHVR and SAF is generally frequency dependent. At the peak frequency and higher, MHVR tends to

underestimate SAF, and this difference is more pronounced than below the peak frequency. Therefore, the input of this pre-trained DNN model was frequency-dependent segmented values. In Chapter 2, we used the target frequency f_i, the peak frequency of MHVR f_M, $MHVR(f_{i-2})$, $MHVR(f_{i-1})$, $MHVR(f_i)$, $MHVR(f_{i+1})$, and $MHVR(f_{i+2})$ as the input, and the ratios of SAF to MHVR (as AMR), $AMR(f_{i-2})$, $AMR(f_{i-1})$, $AMR(f_i)$, $AMR(f_{i+1})$ and $AMR(f_{i+2})$ as the output data to be trained. Here, the subscript i refers to the target frequency number. In Chapter 2, we also pointed out that this type of input and output better reflected the complement of MHVR and SAF in terms of shape. Finally, the number of training datasets required to build the pre-trained DNN model was a total of 14,400 sets from the Chugoku district as shown in Figure 3-1. These corresponded to the source dataset in Figure 3-5.

In this study, the location of the data for our target dataset is shown in Figure 3-1 for the four districts, Chubu, Kanto, Tohoku, and Hokkaido. The target dataset has the same feature space as the source dataset, meaning that we used the same input and output strategies. This is consistent with the definition of homogeneous TL. Therefore, we construct a new TL model using the same strategy as in Figure 3-5. The comparison of the TL model with the pre-trained DNN model is shown in Figure 3-6. The hidden layers from the DNN model were reused and the hyperparameters of these hidden layers are frozen and not updated during the training process. Then we added a fully connected layer to the final layer and update the hyperparameters using the target dataset. One reason for these procedures is to leverage the additional correction effect of the pre-trained DNN model from the previous source dataset. It is expected that this correction would be effective to some extent and can be beneficial for constructing the new model. Then, to accommodate the difference between the source and target dataset, a new fully connected layer is added to the TL model, allowing the model to learn new representations specific to the target dataset.

We used the data from 16 sites as the training set for the TL model. To ensure consistency, the same resampling strategy was employed to resample the MHVRs and

SAFs to the frequency range of 0.3-10 Hz (168 frequencies/site). Table 3-1 provides a detailed summary of the shape and hyperparameter settings for the developed TL model. The table includes information about the number of layers, the number of neurons per layer, the activation function used for each layer, and whether the weight used during the training of the TL model was frozen. Additionally, during the training process for the developed model, a learning rate of 0.001 was used along with a batch size of 2,688 and a maximum epoch of 1,000. It is critical to note these settings for replicating the TL model.

Table 3-1. Hyperparameters and outshape of the layers in the TL model.

Layer	Hyper-parameters and weight setting	Outshape (Number of neurons)
Input		7
Log		7
Affine1	Weight freezing	16
Selu1	α=1.67, λ=1.05	16
Dropout	0.5	16
Affine2	Weight freezing	16
Selu2	α=1.67, λ=1.05	16
Affine3	Weight freezing	16
Selu3	α=1.67, λ=1.05	16
Affine4	Weight updating	16
Selu4	α=1.67, λ=1.05	16
Affine5	Weight updating	5
Mean Absolute Error		5

3.3.3. Cross-validation

To verify the validity of the TL model when training data are limited, we used only the data collected at 16 sites as the training data for the target task. In addition, a separate validation set sometimes does not provide a clear enough evaluation, especially when number of sites is limited. Therefore, a *k*-fold cross-validation [47] was applied in this study. The process of cross-validation is shown in Figure 3-7. Firstly, the total 112 sites of the target dataset are divided into 7 folds, with each fold containing the MHVRs and SAFs data from 16 sites (a total of 2,688 sets for input and output). The specific sites included in each fold are given in Appendix.

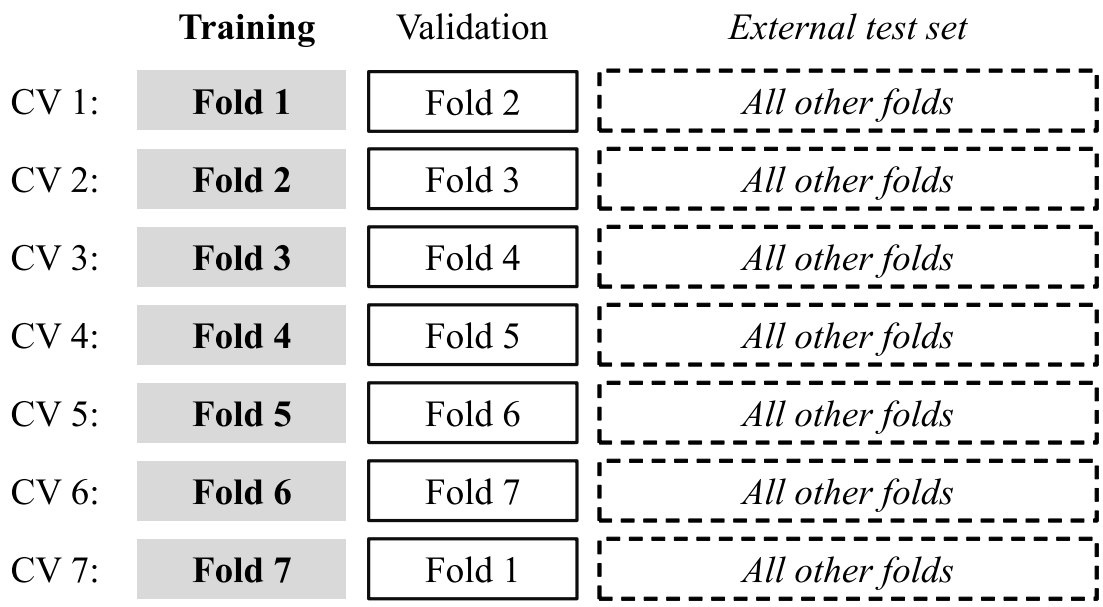

Figure 3-7. Cross-validation adopted in this study, and corresponding folds of the training, validation and external test sets.

During the training process, one of the folds was utilized as the training set, while another fold was designated as the validation set to monitor the process of the model and prevent overfitting. The generalization capability of the TL model was assessed using an

independent external test dataset that was not involved in the training or validation processes. The external test set contained a total of 80 sites all except the training and validation sets, which is much larger than the number of sites used for training and validating. The ratio of the training set, validation set, and external test set was 1:1:5. This process was repeated seven times by changing the training and validation sets. Sites with both low and high peak frequencies were equally distributed among each fold to ensure that the dataset used for training, validation, and testing was representative of the overall target dataset. This approach helped to validate the robustness of the TL model and ensure that it could generalize well to new data.

3.4. Results

3.4.1. Result of test set

The estimated pseudo-SAFs (pSAFs) by the TL model on the external test set are presented in Figure 3-8, where CV+No. is used to indicate the results of each cross-validation. It is important to note that all subsequent analyses are based solely on the results of the external test set, which was not involved in the training or validation process. Figure 3-8 shows the results of the CV 1. In addition, to demonstrate the differences between the TL model and the pre-trained DNN model in Chapter 2, we also compared the performance of the two models on the same external test set.

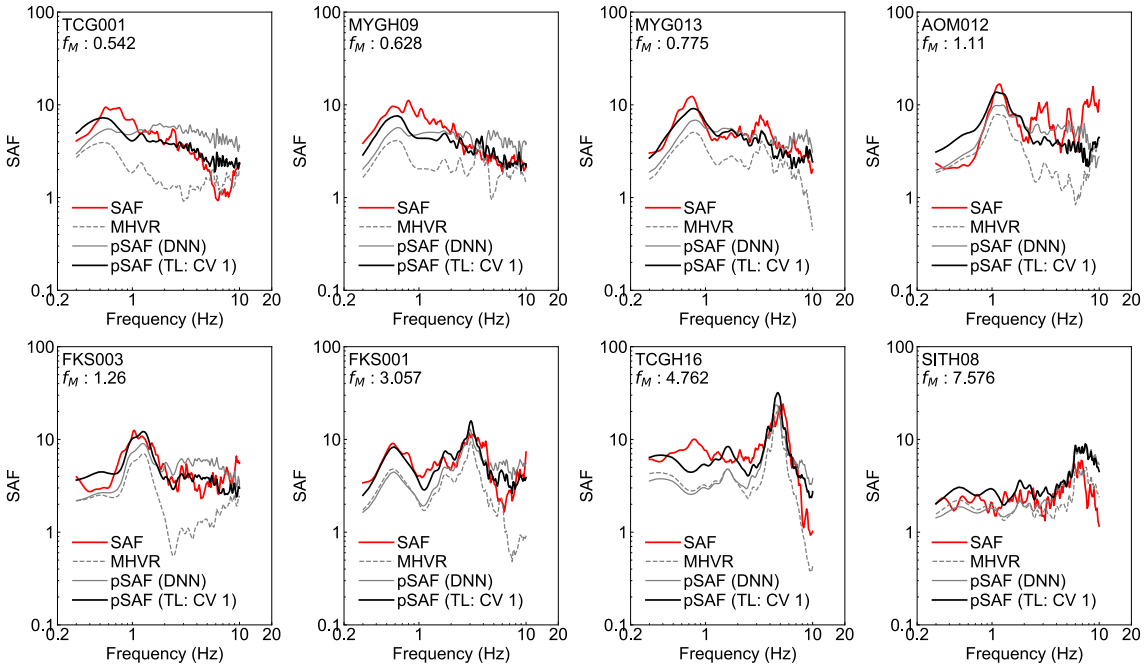

Figure 3-8. Comparison of SAFs by GIT, MHVRs and pSAFs estimated by the DNN model and TL model for test set of CV 1. The red line is the SAF estimated by using GIT, which can be considered as the ground truth for SAF. The dotted gray line is the MHVR measurements.

The results in Figure 3-8 clearly show that the TL model outperforms the original DNN model on the external test set. In particular, the pSAFs by the DNN model tend to

underestimate the observed SAF values especially in low frequency range, and the pSAFs by the DNN model are smaller than the peak values of the SAFs at the first-mode peak. The findings suggest that the influence of regional variations and methodological differences in the GIT approach should be carefully considered when employing the DNN model for MHVR-based SAF estimation.

On the other hand, the pSAFs by the TL model exhibit superior fitting performance compared to the DNN model at almost all sites, particularly at peaks with low frequencies, as shown in FKS001 and MYG013. Furthermore, the TL model only requires data from 16 sites in the training set, which is more practical than obtaining a larger dataset for constructing a new model and enhances the generalization performance of the model to target regions with limited data.

3.4.2. Residual analysis

The external test set contains data from a total of 80 sites. To quantitatively analyze the degree of fit of pSAFs for all the sites, we calculated the residuals in lognormal scale between the pSAFs and the observed SAFs using Eq. (3-4).

$$\text{Residual}(f_i) = log(SAF(f_i)) - log(pSAF(f_i)) \tag{3-4}$$

The average of the residuals was calculated for all the 80 sites along with their one standard deviation. The 7-fold cross-validation were performed using the completely unduplicated training sets. The residual of the external test set in each cross-validation are shown in Figure 3-9, including those of both the TL model and the DNN model.

The average of the residuals for each frequency of the pSAF obtained by the TL model were around the zero line, with the standard deviations in the range of 0.2-0.4. All the results of the cross-validation showed consistency. In contrast, the averages of residuals for each frequency of pSAF by the DNN model were identical to the results

exhibited in Figure 3-8. The pSAFs were clearly underestimated in the low frequency range (0.3-2.0 Hz), with the average values around 0.2 and a maximum standard deviation of 0.6. Furthermore, the residual curves of the TL model showed a similar trend or shape to those of the DNN model. This could be attributed to the utilization of the weights from the DNN model's layers in the construction of the TL model's hidden layers. These weights provided the correction effects for the MHVR-derived SAF, which still had some influence on the results of the pSAFs. Then the new fully-connected layer added at the end of the TL model built on the additional effect by learning a new representation

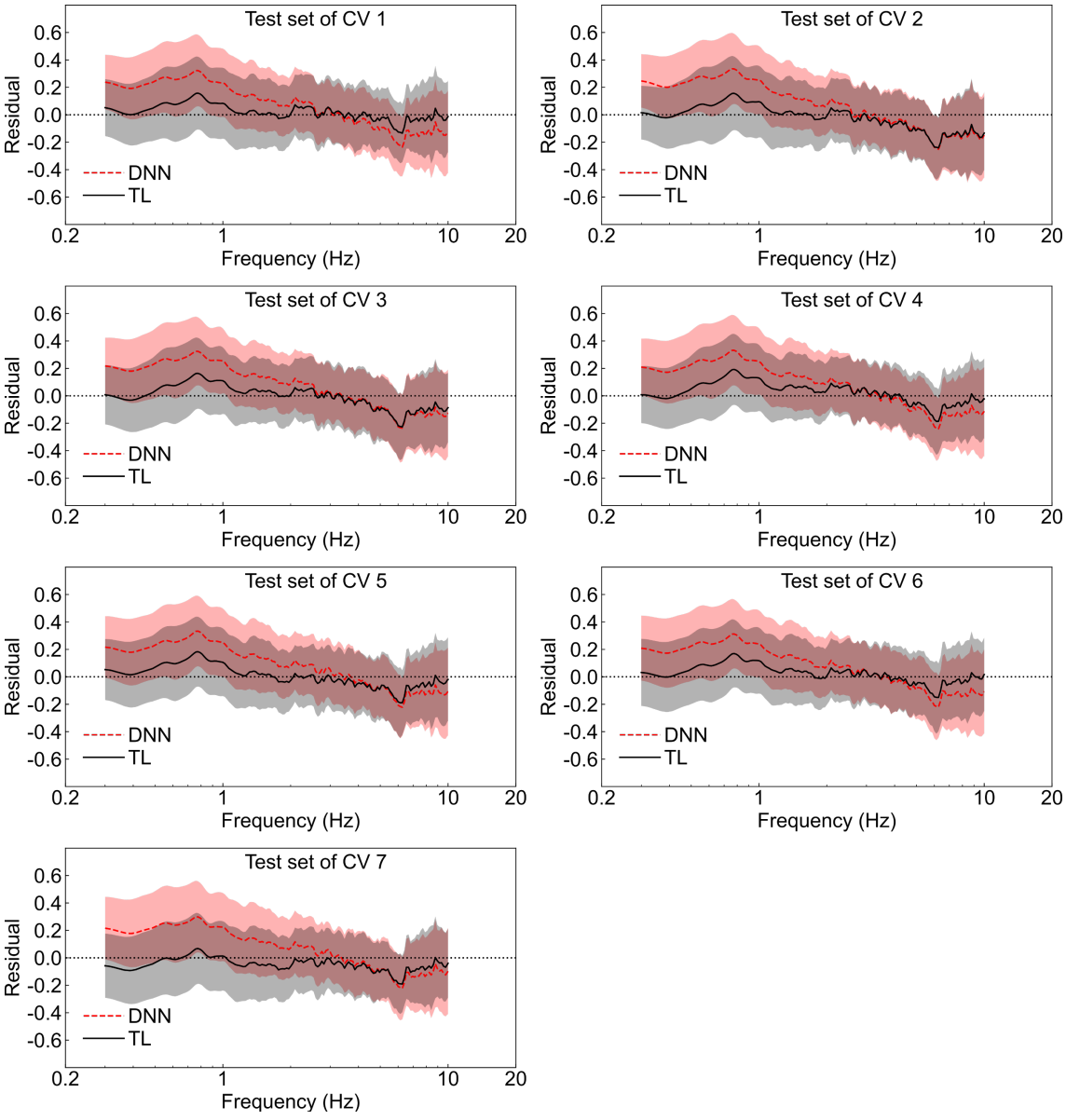

Figure 3-9. Residuals in logarithmic scale between observed SAFs and pSAFs on test sets of all CVs.

specific to the target dataset from the training set of 16 sites. This representation was intended to provide a correction effect for the previously mentioned regional variability and GIT method variability. These results were in line with our expected role for the TL model, which aimed to reuse the knowledge learned on the pre-trained model and then use the limited data to construct a new representation for the target task.

The above results demonstrate the effectiveness of the TL model proposed in this study to enhance the pre-trained DNN model for non-source data regions and to improve the accuracy of pSAFs. Additionally, the cross-validation results suggest that the TL model can achieve consistent performance regardless of the selected 16 training sites.

3.5. Discussion

3.5.1. Different number of training sets

To investigate the effect of the size of the training set on the performance of the TL model, several comparative experiments were conducted.

First, we considered whether the TL model could be successfully trained if the number of training set was further reduced. Therefore, the 16 sites in the training set were reduced by half to eight sites. The specific strategy adopted here was to divide the 112 sites equally into 14 folds (14 folds × 8 sites). Then, the same cross-validation strategy was used, with one fold used for training, one fold used for validation, and all the remaining folds used as the external test set. The pSAFs from all external test sets were employed as the primary metric for model evaluation. The details regarding the specific sites included in each fold are provided in the supplementary material.

The pSAFs obtained for the first CV (CV 1) are shown in Figure 3-10. The results indicate that the TL model for CV 1 still outperformed the DNN model, even with a reduced training set size. However, after completing all cross-validations, the results of several CVs did not perform as well as CV 1. The residual results of CV 1, CV 7, CV 8, and CV 14 are shown in Figure 3-11. The results of CV 7 and CV 8 suggest that the TL model was unable to learn any new corrections. In contrast, for CV 14, the TL model overcorrected, resulting in an average residual of -0.2, indicating that the model provided pSAFs that were larger than the observed SAF. This suggests that using only eight sites as the training set for the TL model is unreliable in practical applications, since it is often not possible to perform sufficient cross-validation to determine the most appropriate eight sites. Therefore, there is no guarantee that the new representation learned from these eight sites is correct.

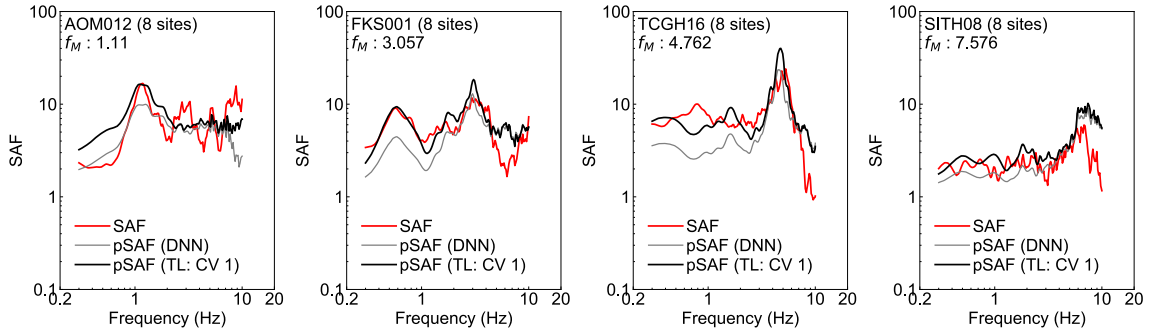

Figure 3-10. Comparison of SAFs by GIT, and pSAFs estimated by the DNN model and TL model trained by 8 sites on test set of CV 1.

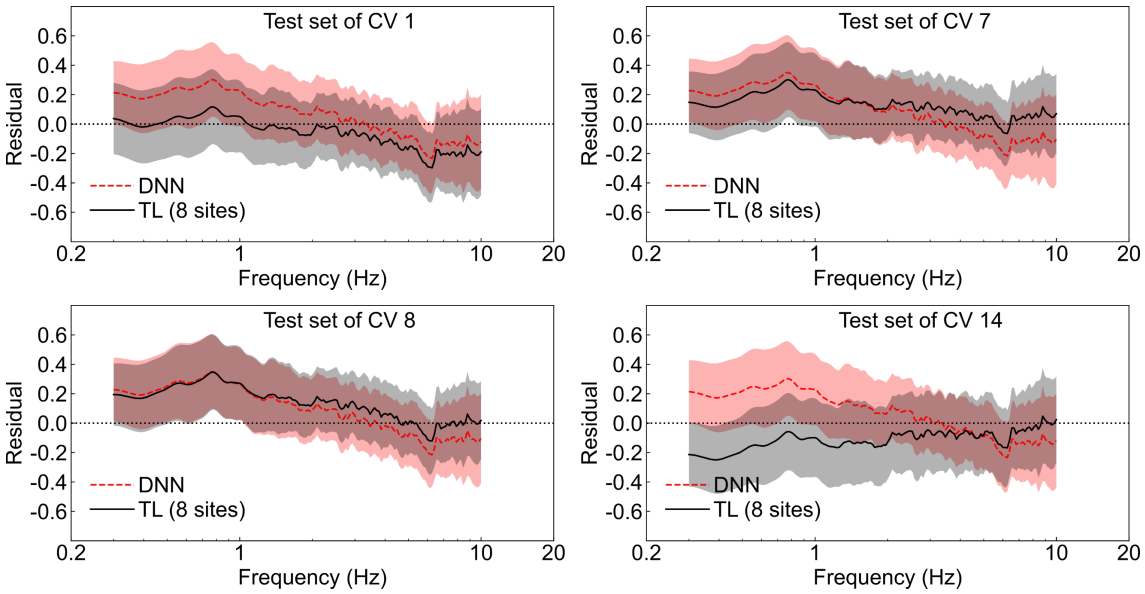

Figure 3-11. Residuals in logarithmic scale between observed SAFs and pSAFs on test sets. Here the TL models were trained by 8 sites.

Furthermore, we investigated the effect of using more sites as the training set for the TL model on its performance. We expanded the number of training sites from 16 to 32. The cross-validation strategy adopted was to use two folds from Section 3.3 as the training set, one fold as the validation set, and all remaining folds as the external test set. As shown

in Figure 3-12, the results of CV1 exhibit better performance than the DNN model. The average residuals of all test set sites in Figure 3-13 also show better performance. We confirmed that the residual results of all CVs (CV 1-7) were almost consistent. The residual results of all CVs are provided in the supplementary material.

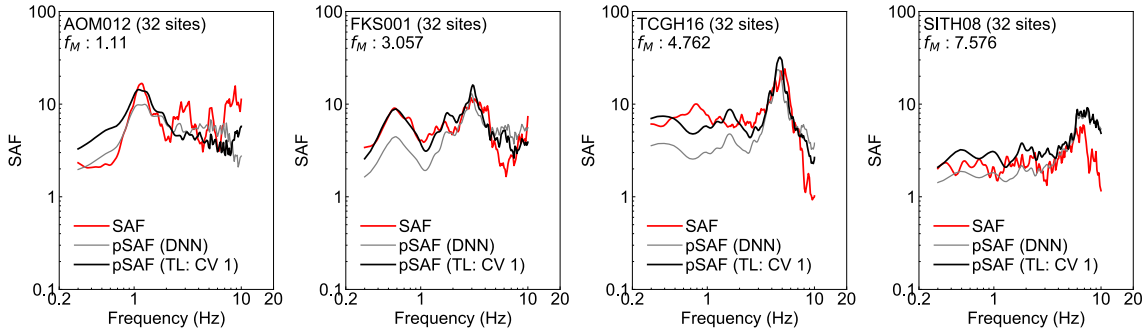

Figure 3-12. Comparison of SAFs by GIT, and pSAFs estimated by the DNN model and TL model trained by 32 sites on test set of CV 1.

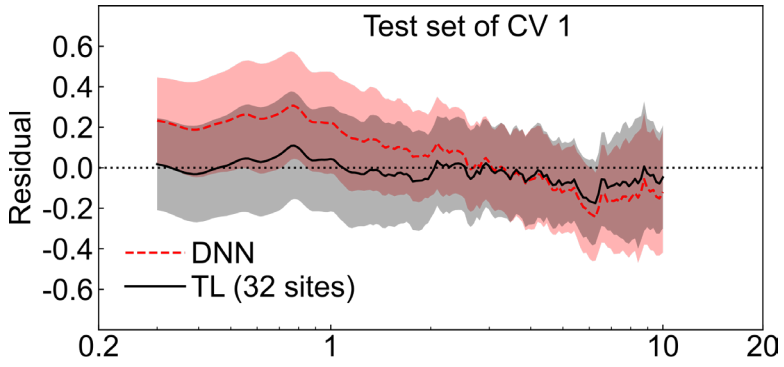

Figure 3-13. Residuals in logarithmic scale between observed SAFs and pSAFs on test sets. Here the TL models were trained by 32 sites. We confirmed that the residual results of all CVs (CV 1-7) were almost consistent.

In addition, we conducted a visual analysis on the single site, FKS001. Figure 3-14 shows the fitting performance of the TL model trained using 8, 16, and 32 sites on the pSAFs and SAF of FKS001. The results indicate that the TL model trained using eight

sites exhibits significant differences in pSAFs for different training set combinations. However, when trained using 16 sites, all pSAFs results are consistently maintained, with only a slight improvement observed when the number of sites is increased to 32. Hence, using 16 sites as the training set for the TL model in this study is reasonable. This provides useful insights into selecting the number of sites in practical applications.

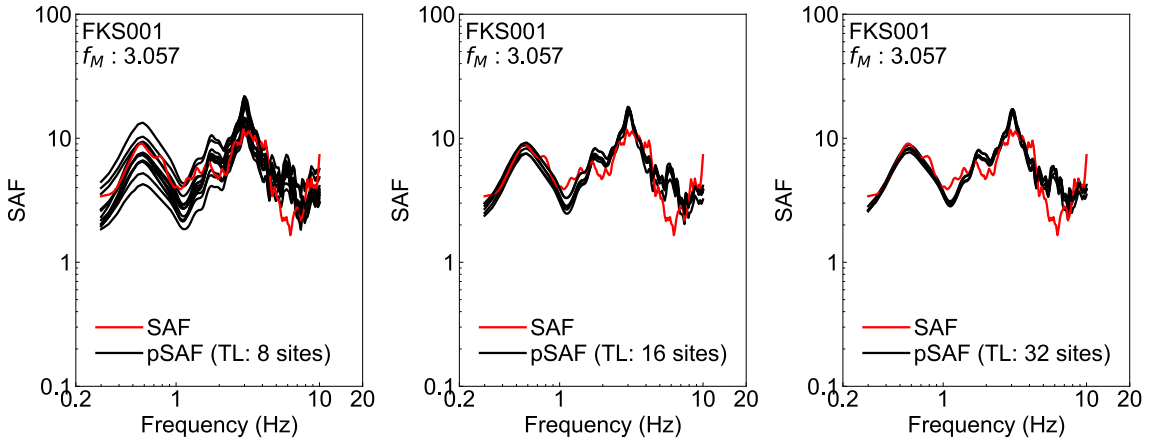

Figure 3-14. Comparison of cross-validation results obtained from different TL models (trained by 8, 16, 32 sites) for a single site according to FKS001. Here are all results when FKS001 is included as an external test set.

3.5.2. Comparison with DNN model with extended training set

In practical applications, in addition to utilizing the TL model, the performance of one DNN model can be enhanced by augmenting the existing dataset with new data to create a larger training set compared to the source dataset. To evaluate whether the enhancement provided by the TL model surpasses that of the source DNN model with an expanded training set, we conducted the following comparative experiments.

In this section, we used the training set from Chugoku district, Chapter 2, comprising 14,400 sets, along with a fold (2,688 sets) from Figure 3-7 as the new training set for the new DNN model (DNN+). The DNN+ model was constructed using the layers of the TL model, but with all layer weights unfrozen and updated during training process. The

validation and external test sets followed the same strategy as in Figure 3-7. Cross-validation was also performed using the same strategy, except for a different training set. Other settings were also kept consistent with the TL model.

The fitness and residuals of the pSAFs for the external test set of the DNN+ model are presented in Figure 3-15 and Figure 3-16, respectively. The results for all sites are nearly identical to those of the DNN model, which are completely different from the TL model. Furthermore, after completing all cross-validations, the same conclusion was drawn that the DNN+ model failed to achieve the objective of enhancing the performance of the DNN model. The fit of the pSAFs to the SAF for all CVs is illustrated by Figure 3-17. Although the DNN+ model demonstrated stability across all CVs, it was unable to learn new representations despite the addition of new data.

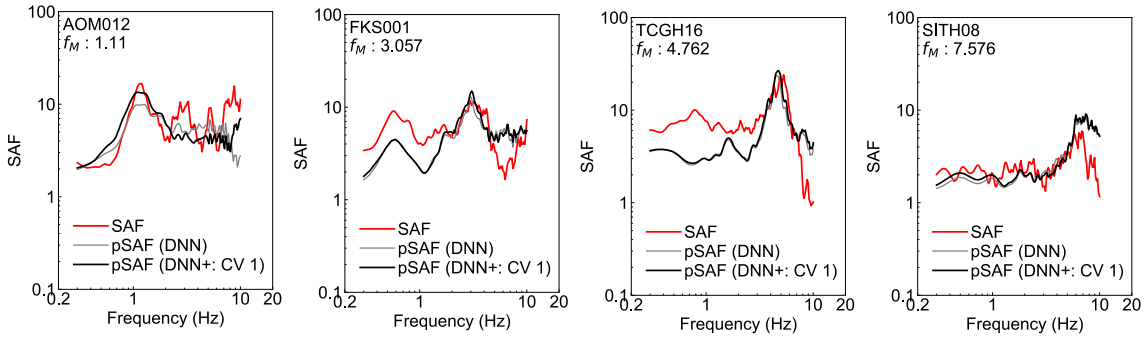

Figure 3-15. Comparison of SAFs by GIT, and pSAFs estimated by the DNN model and DNN+ model on test set of CV 1.

The reason for this is that the DNN model tends to overfit to the training data and may struggle to improve its accuracy on SAF obtained by different methods in different regions, even when new training sites are added. In contrast, the TL model has demonstrated better performance in this regard.

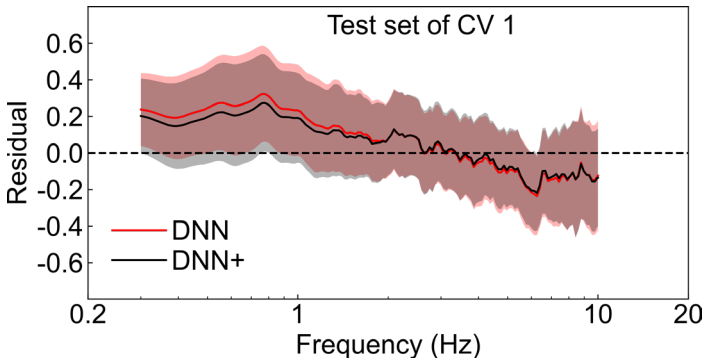

Figure 3-16. Residuals in logarithmic scale between observed SAFs and pSAFs estimated by the DNN model and DNN+ model on test sets. We confirmed that the residual results of all CVs (CV 1-7) were almost consistent.

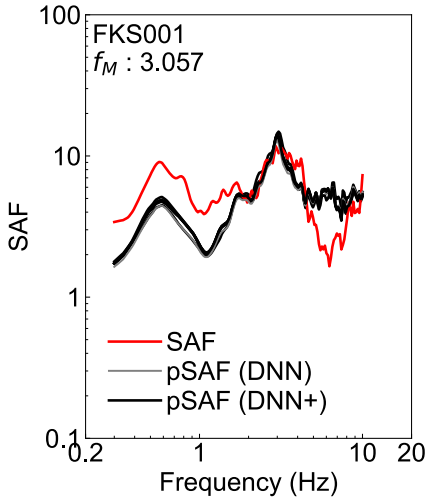

Figure 3-17. Comparison of cross-validation results obtained from the DNN+ models for a single site according to FKS001. Here are all results when FKS001 is included as an external test set.

3.5.3. Prediction of sites outside the target peak frequency

In training the TL model, we used the sites with peak frequencies of MHVRs (f_M) within 0.3-10 Hz as the data set considering the frequency range of the SAFs in Nozu et al. [55]. However, there were 17 sites with the f_M outside the range. On the other hand,

the pre-trained DNN model by Chapter 2 used the f_M up to 20 Hz. Therefore, this section discusses whether the TL model would be valid for the sites where the peak frequencies are not in the range of 0.3-10 Hz

We applied the trained TL model (16 sites) with CV 1 for estimating the pSAFs of the sites with the f_M outside the range of 0.3-10Hz. Figure 3-18 shows the comparison of the observed SAFs and pSAFs estimated from the TL model and the pre-trained DNN model. The pSAFs by the TL model shows better agreement with the SAFs at most sites than those by the DNN model. However, the pSAF of either the DNN or TL model at MYGH12 whose f_M is over 20 Hz differs significantly from the SAF. The reason for the discrepancy is speculated to be because the training dataset of both the TL and the pre-trained DNN models also does not contain sites with f_M above 20 Hz, which leads to the fact that the TL model that reuses some of the weights of the DNN model is not effective for such sites either. The averages of residuals with its standard deviation from the SAFs and pSAFs at 17 sites are shown in Figure 3-19. We can confirm that the TL model shows better accuracies in estimating SAFs than the DNN model even for the sites whose f_M are outside the target frequency. These findings suggest that the developed TL model has the potential to be utilized for evaluating site amplifications at arbitrary sites and may contribute to more detailed seismic hazard predictions.

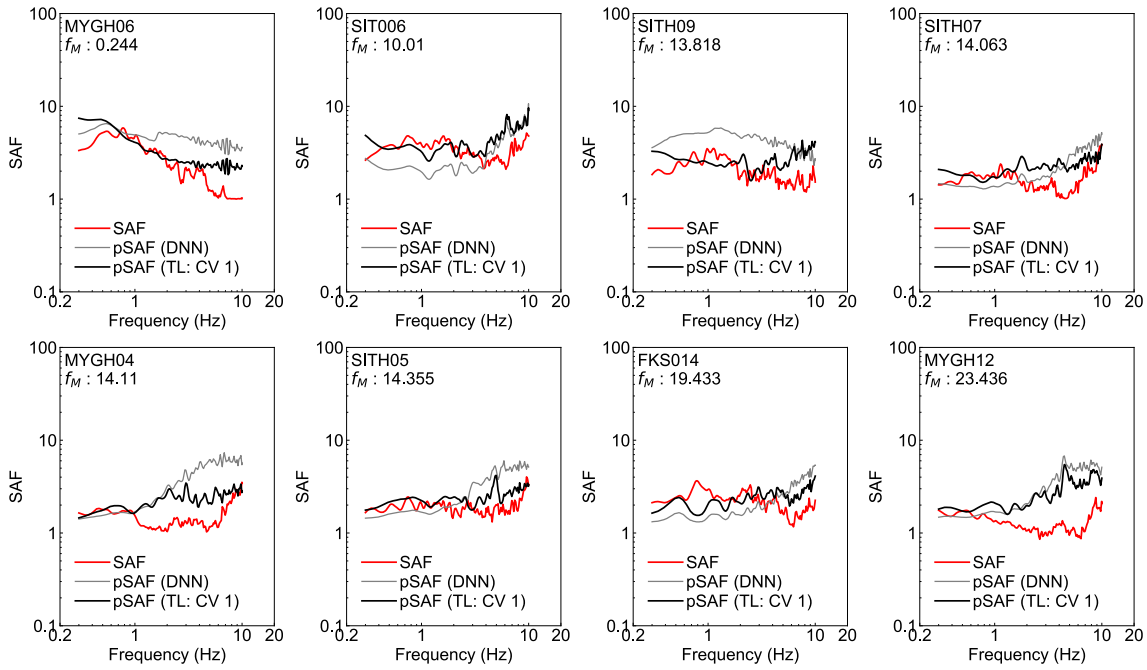

Figure 3-18. Comparison of SAFs by GIT, and pSAFs estimated by the DNN model and TL model on the sites where the f_M are not in the range of 0.3-10 Hz.

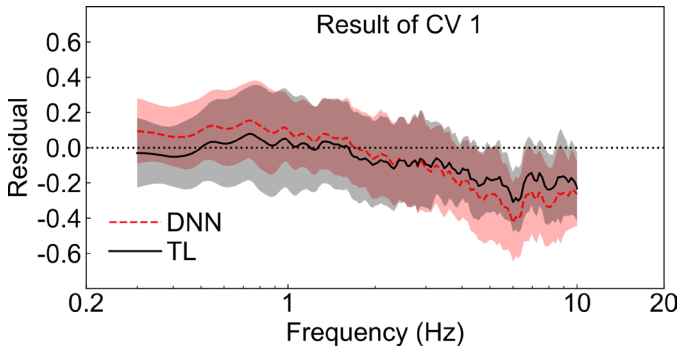

Figure 3-19. Residuals in logarithmic scale between observed SAFs and pSAFs estimated by the DNN model and TL model on the sites where the f_M are outside the range of 0.3-10 Hz.

3.6. Prospect

This section discusses geophysical applications of the proposed method, particularly its prospect for seismic ground motion evaluation and prediction. The TL technique would facilitate the model development for estimating SAFs based on MHVRs over a wider geographic area, even when site-specific data are limited. The model inputs exclusively require microtremor data obtained from sensor-based measurements, obviating the necessity for extensive site investigations or long-term seismic observations. This allows for efficient derivation SAFs for target sites with minimal data requirements, which is expected to lead to the development of a global predictive model for site effects as data are collected for a wider area.

In some countries and cities with well-developed seismic observation networks, distributions of seismic intensities can be estimated immediately after significant earthquakes, and published as ShakeMaps in order to rapidly identify severities of ground shakings and to gauge the extent of affected areas [74]. These maps are typically based on recorded ground motions at seismic stations and site amplification factors estimated from site proxy data such average S-wave velocity in upper 30 m (V_{S30}). Although the ShakeMaps are useful for evaluating seismic intensity distributions, frequency-dependent spectral amplitudes have not been rarely evaluated because of the lack of spectral information in such site proxies. Microtremor-derived pSAF can be used as a more adequate site proxy in the seismic motion evaluations instead of V_{S30} because the pSAF purely represents the spectral amplification of seismic waves at the site. If pSAFs are available by the TL model not only at seismic observation sites but also at arbitrary sites with different ground conditions, the spectral amplitudes at the arbitrary sites can be estimated from the ground motion record at the observation site and the pSAFs at both sites, which is probably more accurate than using V_{S30}.

Microtremor-derived pSAFs can be applied also in ground motion prediction at a specific site in earthquake scenarios. One of the critical tasks in the ground motion predictions is lack of detailed site condition at a target site. If pSAF is available at the

target site, the spectral amplitudes at the ground surface of the site can be estimated by multiplying the pSAF with the Fourier spectrum at the seismic bedrock simulated by the stochastic approach [75,76]. This method does not require geophysical investigations other than microtremor observation, making it a simple and cost-effective way to simulate and predict seismic ground motions.

We need to pay attention to the fact that the pSAFs obtained from this method do not include nonlinear site response during strong shakings because the SAFs analyzed in this study represent linear site amplifications during weak and moderate motions. Nonlinear site amplification can be typically expressed as shift to lower frequency of spectral peaks and low-amplitude effect [77]. Since the degree of nonlinear site effect strongly depends on seismic intensities on the bedrock, such nonlinear site amplifications need to be evaluated by observed or predicted strong shaking data with nonlinear site response analysis. However, it is beyond the scope of this study. Empirical or analytical approaches for evaluating nonlinear site responses in SAFs would be discussed in future studies.

3.7. Conclusions

In this Chapter, we have explored the applicability of TL model in improving the performance of the pre-trained DNN model for estimating SAFs in new regions. The result showed that the TL model trained using 16 sites (2,688 sets of inputs and outputs) exhibits better fitting performance than the original DNN model. Furthermore, the cross-validation results using 80 sites as the test set demonstrated that the TL model can effectively generalize the prediction of SAFs for new sites outside the original dataset.

In addition, we conducted further experiments to investigate the effects of different elements on the performance of the proposed TL model, including the size of the training sets and the comparison with the original DNN model using the extended training set. The result suggested that 16 sites can be used as the training set, and the SAFs obtained from cross-validation were stable. Even though the number of sites was increased to 32, only a slight improvement was observed. And the TL model trained with eight sites showed significant differences in pSAFs among different training set combinations. We also constructed a DNN model with an extended training set and compared its results with the TL model. The proposed TL model still outperformed the DNN model with an extended training set. The comparison of the residuals also demonstrated the effectiveness of the TL model in extending the applicability of the MHVR-estimated SAF to new regions, which is informative for practical applications. Finally, we used the trained TL model at 17 sites with peak frequencies of MHVR that were not within the peak frequency range of the training set. The results show that the TL model can still effectively improve the performance of the pre-trained DNN model at such sites.

In practical scenarios, it is anticipated that the field of MHVR-estimated SAF can be extended to countries and regions where the number of seismic observatories is inadequate. For regions with limited data or unique geological features, such as some developing countries, constructing an appropriate TL model based on limited data can improve the accuracy of SAFs estimated with MHVRs.

However, there are also limitations in this study. Firstly, the proposed TL model may

be influenced by the number of regions used. The dataset used in this study consists of sites from four different regions. More comprehensive research is needed to further investigate the impact of data between different regions on the performance of the TL model. In particular, it is worth exploring whether better fitting can be achieved by training the TL model using data from only one region. Secondly, we only considered the application of the TL model to improve the performance of a single DNN model. In addition to this DNN model, future research can explore the application of TL model to improve the performance of other machine learning models or techniques in the estimation of SAF by MHVR.

In summary, this study demonstrates the potential of TL model in improving the performance of the MHVR-estimated SAFs to new regions. Our results provide insight into the selection of training sites for the TL model, which has important implications for the practical application of the MHVR-estimated SAFs. Future research would include further investigation of the application of alternative machine learning models and methods as well as the inclusion of nonlinear site effects in the SAFs estimated by MHVRs.

3.8. Appendix

This part provides the selection of the sites for each fold during cross-validation (**Table S**). It also includes the residual results for all cross-validations when using eight and 32 sites as the training set (**Figure S**). We also illustrate the comparison of observed SAFs, MHVRs, and pSAFs by the DNN and TL models. These are beneficial to help reproduce the work of the TL model.

Table S3-2. Sites included in each fold with their peak frequencies of MHVR (f_M) when using 16 sites as the training set. Each fold included sites with high and low peak frequencies equally.

Fold 1		**Fold 2**		**Fold 3**		**Fold 4**		**Fold 5**		**Fold 6**		**Fold 7**	
site	f_M	site	f_M	site	f_M	site	f_M	site	f_M	site	f_M	site	f_M
SIT003	0.826	ISK007	0.826	ISK004	2.751	NIG022	2.091	NIG017	3.775	SIT010	1.287	SIT013	3.938
SIT003	0.826	ISKH03	4.285	IWT010	0.439	IBRH07	0.683	FKSH16	0.728	SITH08	7.576	ISKH05	6.817
IWTH20	0.487	SITH06	6.535	MYG017	1.11	AOM012	1.11	MYG006	1.02	TCG001	0.542	MYG016	0.589
IBR006	1.842	IBR013	0.386	TCG013	1.461	IBR003	3.326	FKSH14	1.342	MYG010	0.88	IWT014	0.978
HKD092	2.637	FKS011	2.091	IBR001	3.544	FKS008	3.619	FKS013	2.81	IBR005	2.637	FKS018	1.729
FKSH12	4.196	IWT013	2.993	MYG004	4.376	TCG014	5.52	IWTH02	6.265	IWT007	3.619	KSRH09	2.181
FKS031	4.762	IWT018	4.285	TCGH15	6.817	TCGH12	6.817	HKD100	6.535	FKSH09	4.47	TCGH13	4.376
IBR002	7.263	IBR017	8.419	MYGH10	7.737	FKSH10	7.576	MYG008	8.419	FKS009	6.817	IWTH15	7.418
IBRH15	8.969	SITH11	4.196	ISK005	1.234	NIG018	0.589	SZO016	0.498	SIT008	1.882	SZOH33	6.817
SZO018	1.208	ISK006	6.535	SIT009	6.675	HKD091	1.183	SIT011	1.064	HKD096	0.628	MYGH09	0.628
SIT002	1.524	FKS005	0.683	SIT014	8.419	FKS012	1.524	IBR012	1.342	FKS003	1.26	MYG013	0.775
NIG020	8.071	TKCH07	1.208	TCG006	0.439	HKD086	2.424	FKS001	3.057	FKS017	2.228	FKSH11	1.589
MYGH08	0.683	IWT020	1.623	MYG007	1.658	FKS019	3.775	TCG009	4.196	TCG005	3.856	IBRH11	2.582
FKS015	3.257	IBRH13	2.694	FKSH19	3.257	FKS006	4.565	TCGH11	4.285	TCGH10	6.134	FKS002	3.775
FKSH17	3.938	MYG014	5.292	FKSH18	4.285	FKS010	7.418	IBRH12	7.263	FKS016	6.675	IBRH18	5.292
IBR007	5.073	IWTH27	7.576	TCGH16	4.762	IWT009	9.356	IBR018	7.737	MYG003	9.356	IBRH16	7.737

Table S3-3. Sites included in each fold with their f_M when using eight sites as the training set. Each fold included sites with high and low peak frequencies equally.

Fold e1		**Fold e2**		**Fold e3**		**Fold e4**		**Fold e5**	
site	f_M	site	f_M	site	f_M	site	f_M	site	f_M
SIT003	0.826	ISK007	0.826	ISK004	2.751	NIG022	2.091	SIT013	3.938
IWTH20	0.487	ISKH03	4.285	IWT010	0.439	IBRH07	0.683	ISKH05	6.817
IBR006	1.842	SITH06	6.535	MYG017	1.11	AOM012	1.11	MYG016	0.589
HKD092	2.637	IBR013	0.386	TCG013	1.461	IBR003	3.326	IWT014	0.978
FKSH12	4.196	FKS011	2.091	IBR001	3.544	FKS008	3.619	FKS018	1.729
FKS031	4.762	IWT013	2.993	MYG004	4.376	TCG014	5.52	KSRH09	2.181
IBR002	7.263	IWT018	4.285	TCGH15	6.817	TCGH12	6.817	TCGH13	4.376
IBRH15	8.969	IBR017	8.419	MYGH10	7.737	FKSH10	7.576	IWTH15	7.418

Fold e6		**Fold e7**		**Fold e8**		**Fold e9**		**Fold e10**	
site	f_M	site	f_M	site	f_M	site	f_M	site	f_M
SZO018	1.208	NIG017	3.775	SIT010	1.287	SITH11	4.196	ISK005	1.234
SIT002	1.524	FKSH16	0.728	SITH08	7.576	ISK006	6.535	SIT009	6.675
NIG020	8.071	MYG006	1.02	TCG001	0.542	FKS005	0.683	SIT014	8.419
MYGH08	0.683	FKSH14	1.342	MYG010	0.88	TKCH07	1.208	TCG006	0.439
FKS015	3.257	FKS013	2.81	IBR005	2.637	IWT020	1.623	MYG007	1.658
FKSH17	3.938	IWTH02	6.265	IWT007	3.619	IBRH13	2.694	FKSH19	3.257
IBR007	5.073	HKD100	6.535	FKSH09	4.47	MYG014	5.292	FKSH18	4.285
IBRH17	7.418	MYG008	8.419	FKS009	6.817	IWTH27	7.576	TCGH16	4.762

Fold e11		**Fold e12**		**Fold e13**		**Fold e14**	
site	f_M	site	f_M	site	f_M	site	f_M
SZOH33	6.817	SZO016	0.498	NIG018	0.589	SIT008	1.882
MYGH09	0.628	SIT011	1.064	HKD091	1.183	HKD096	0.628
MYG013	0.775	IBR012	1.342	FKS012	1.524	FKS003	1.26
FKSH11	1.589	FKS001	3.057	HKD086	2.424	FKS017	2.228
IBRH11	2.582	TCG009	4.196	FKS019	3.775	TCG005	3.856
FKS002	3.775	TCGH11	4.285	FKS006	4.565	TCGH10	6.134
IBRH18	5.292	IBRH12	7.263	FKS010	7.418	FKS016	6.675
IBRH16	7.737	IBR018	7.737	IWT009	9.356	MYG003	9.356

Table S3-4. Cross-validation and corresponding folds of the training, validation and external test sets when using eight sites as the training set.

Cross Validation (CV)	Training set	Validation set	*External test set*
CV 1	Fold e1	Fold e14	*All other folds*
CV 2	Fold e2	Fold e13	*All other folds*
CV 3	Fold e3	Fold e12	*All other folds*
CV 4	Fold e4	Fold e11	*All other folds*
CV 5	Fold e5	Fold e10	*All other folds*
CV 6	Fold e6	Fold e9	*All other folds*
CV 7	Fold e7	Fold e8	*All other folds*
CV 8	Fold e8	Fold e7	*All other folds*
CV 9	Fold e9	Fold e6	*All other folds*
CV 10	Fold e10	Fold e5	*All other folds*
CV 11	Fold e11	Fold e4	*All other folds*
CV 12	Fold e12	Fold e3	*All other folds*
CV 13	Fold e13	Fold e2	*All other folds*
CV 14	Fold e14	Fold e1	*All other folds*

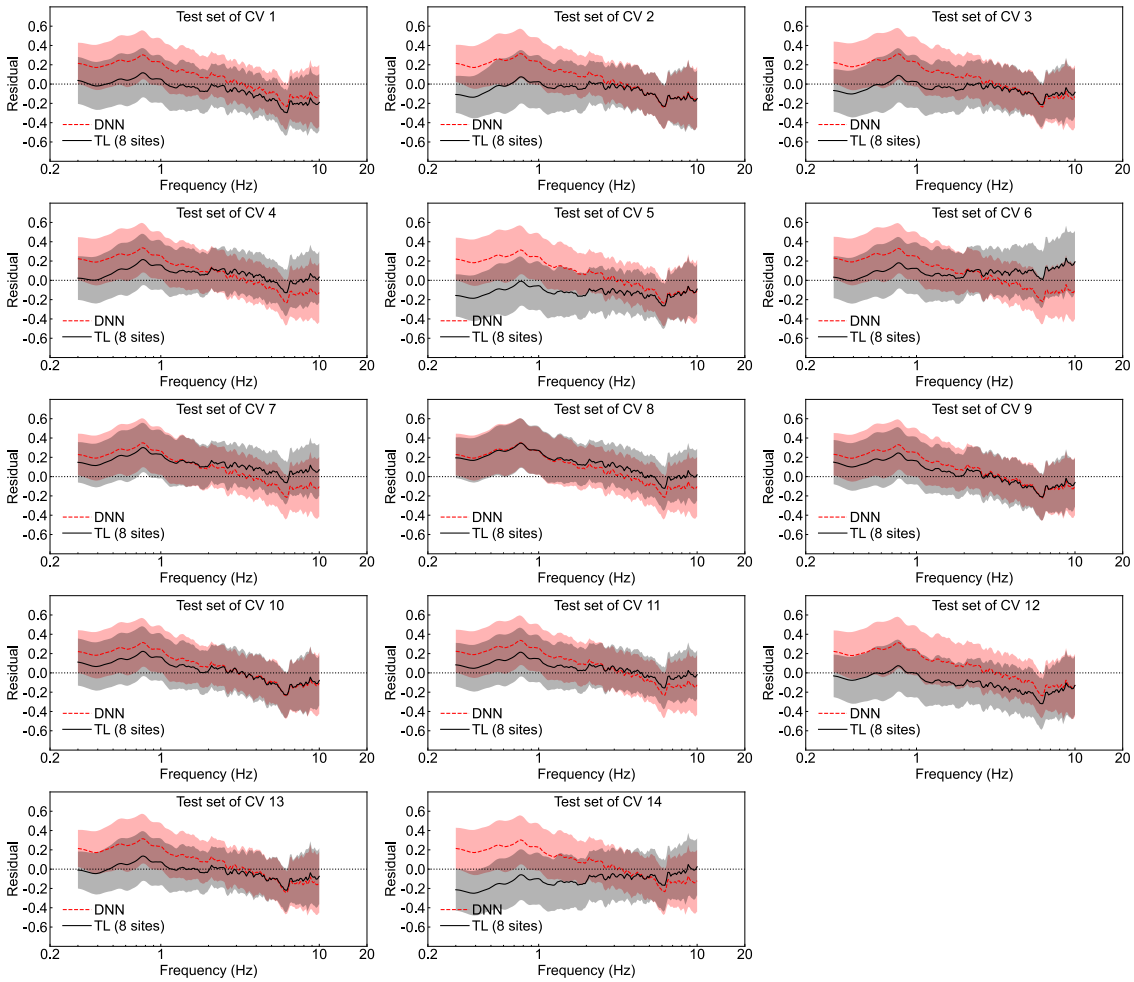

Figure S3-20. Comparison of the residuals of the corresponding external test set for cross-validation with the original DNN model. Black represents the TL model using eight sites in this study and red represents the original DNN model.

Table S3-5. Cross-validation and corresponding folds of the training, validation and external test sets when using 32 sites as the training set.

Cross Validation (CV)	Training set	Validation set	*External test set*
CV 1	Fold 7 + Fold 1	Fold 2	*All other folds*
CV 2	Fold 1 + Fold 2	Fold 3	*All other folds*
CV 3	Fold 2 + Fold 3	Fold 4	*All other folds*
CV 4	Fold 3 + Fold 4	Fold 5	*All other folds*
CV 5	Fold 4 + Fold 5	Fold 6	*All other folds*
CV 6	Fold 5 + Fold 6	Fold 7	*All other folds*
CV 7	Fold 6 + Fold 7	Fold 1	*All other folds*

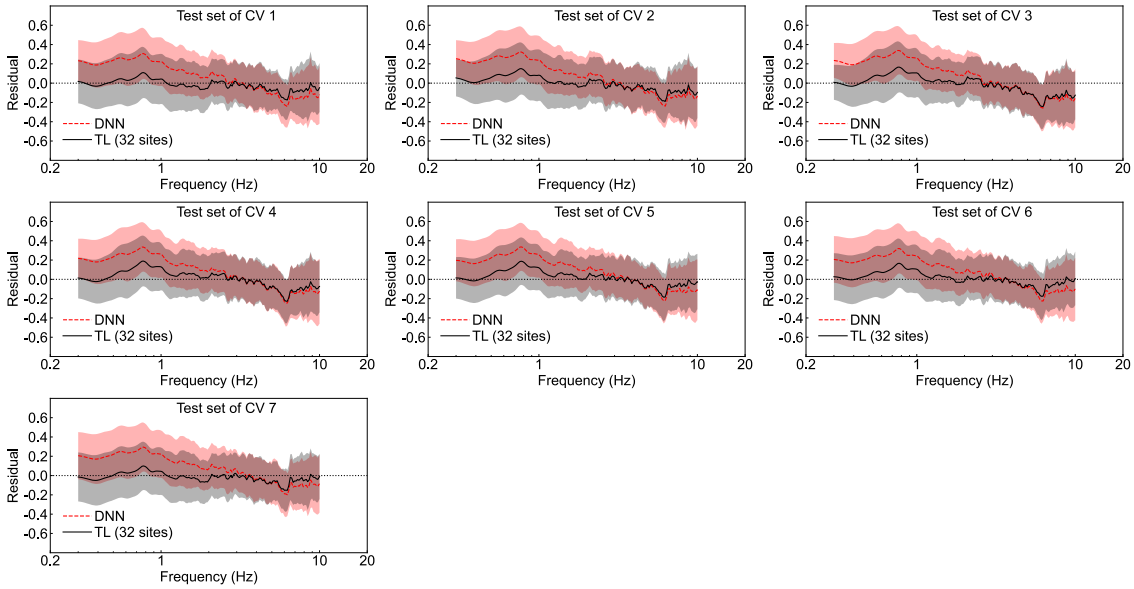

Figure S3-21. Comparison of the residuals of the corresponding external test set for cross-validation with the original DNN model. Black represents the TL model using 32 sites in this study and red represents the original DNN model (Pan et al., 2022).

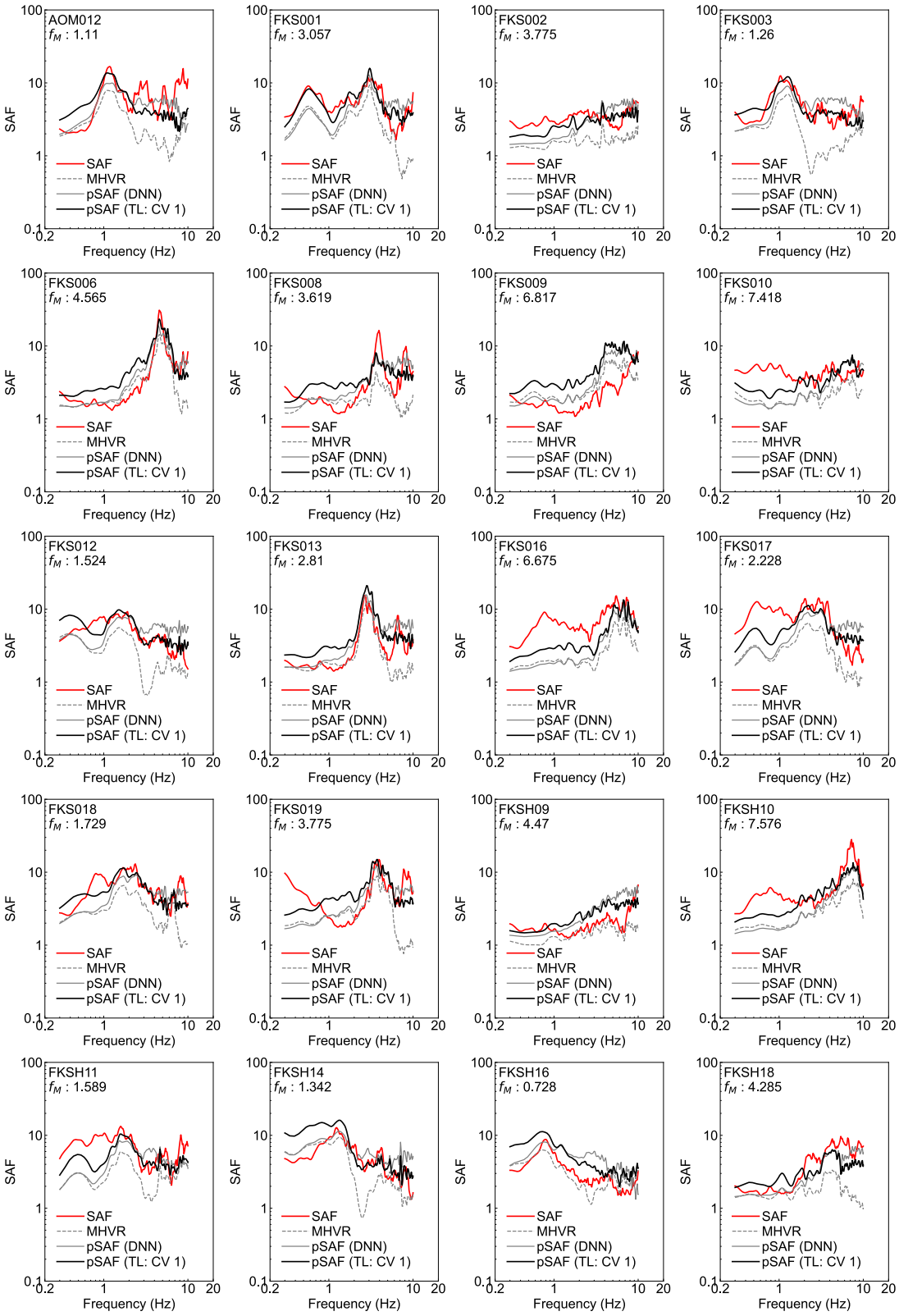

Figure S3-22. Comparison of SAFs by GIT, MHVRs, and pSAFs estimated by the DNN model and TL model for all test sites (80 sites) of CV 1.

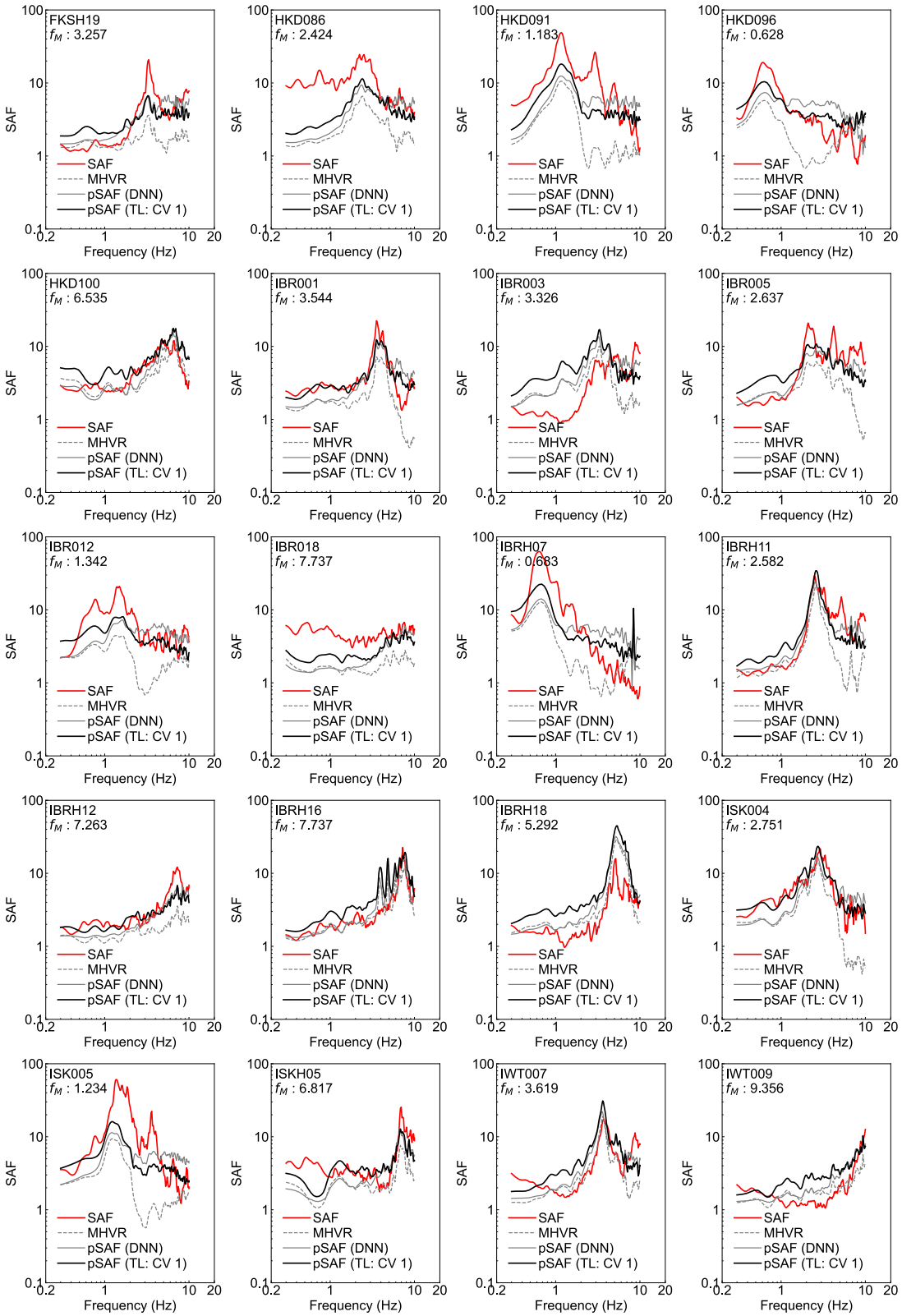

Figure S3-23. Comparison of SAFs by GIT, MHVRs, and pSAFs estimated by the DNN model and TL model for all test sites (80sites) of CV 1.

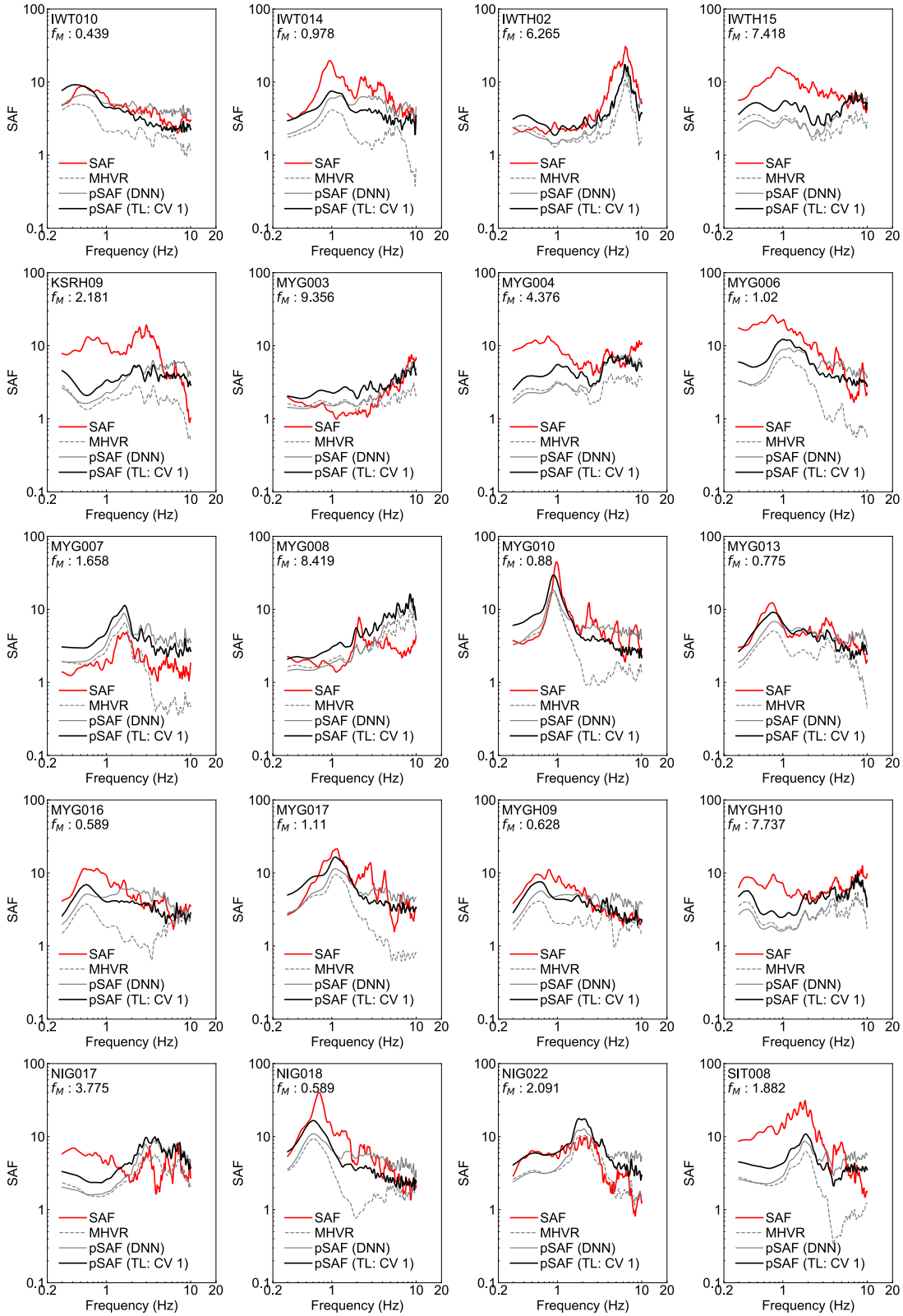

Figure S3-24. Comparison of SAFs by GIT, MHVRs, and pSAFs estimated by the DNN model and TL model for all test sites (80 sites) of CV 1.

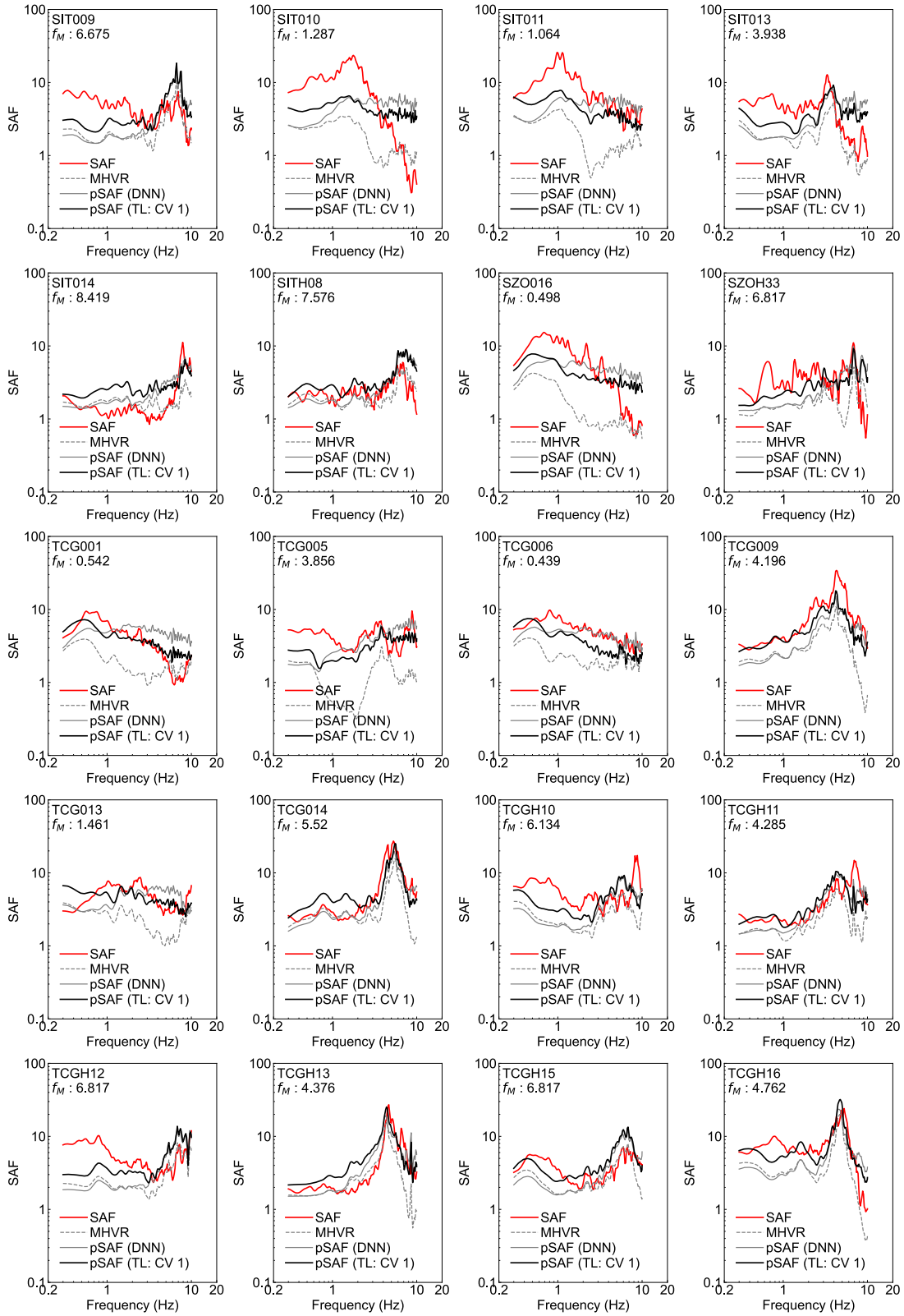

Figure S3-25. Comparison of SAFs by GIT, MHVRs, and pSAFs estimated by the DNN model and TL model for all test sites (80 sites) of CV 1.

Chapter 4. Prediction of Earthquake HVR （EHVR） from MHVR using DNN model

In previous chapters, we detailed how to use deep neural networks and microtremor horizontal-to-vertical spectral ratio (MHVR) to estimate site amplification factors (SAF), as well as how to expand the developed model. In addition to those originating from microtremors, there is another type of horizontal-to-vertical spectral ratio (EHVR) derived from earthquake motion. It is generally considered that EHVR better reflects the true site effects during earthquakes, making it a more reliable indicator for simulating specific site seismic responses and improving existing ground motion prediction models. In recent years, EHVR has been widely used in fields such as site classification, site effects, and the acquisition of underground structures. However, at locations without seismic activity, it is obviously impossible to obtain EHVR, which significantly limits its broader application. In this chapter, we will delve into the development and validation of a DNN model designed for this purpose: using MHVR alone to estimate EHVR, to facilitate the use of MHVR to obtain pseudo-EHVR at locations where seismic records are unavailable, thereby enabling further applications such as site classification, site effect assessment, and earthquake intensity prediction.

4.1. Background

The previous chapters primarily discussed the calculation of site effects using Microtremor Horizontal-to-Vertical Spectral Ratios (MHVR). Apart from originating from microtremors, another type of HVR is derived from earthquake motion, namely the Earthquake Horizontal-to-Vertical Spectral Ratio (EHVR) [27,78–80]. Both possess distinct characteristics. Microtremor measurements are non-invasive and can continuously collect data at low cost, thus offering advantages. In the previous chapters, MHVR has been used to estimate site effects[66,81], providing a passive method to infer the dynamic behavior of the ground during earthquakes. On the other hand, EHVR is measured directly during earthquakes, reflecting the actual response of the ground to seismic forces. There is also variability between MHVR and EHVR due to differences in energy and frequency content between ambient and earthquake vibrations.

Fundamentally, MHVR is derived from ambient vibrations, typically of lower energy, and can be influenced by daily environmental factors such as traffic or wind. These measurements provide information such as the peak frequency for specific site characteristics. In contrast, EHVR captures the response of the soil during actual seismic events, involving significantly higher energy levels and typically broader frequencies. Ground motion during earthquakes entails complex wave propagation phenomena.

EHVR is generally considered superior to MHVR in reflecting the true site effects during earthquakes. This advantage stems from its direct measurement during seismic events, capturing the actual dynamic properties and responses of soil layers under real earthquake stress conditions. Thus, EHVR data more closely represent true site effects, making EHVR a more reliable indicator for modeling site-specific seismic responses and for improving existing ground motion prediction models. In recent years, EHVR has been widely used in the fields of site classification, site effects, and obtaining underground structures. However, EHVR cannot be obtained at sites without seismic activity, which significantly limits its wider application.

The application of deep learning in seismology offers a new pathway to address these

challenges. DNN models can learn complex non-linear relationships from large datasets, enabling the inference of EHVR from MHVR with greater accuracy. This approach allows for the integration of various types of data, including those from different geographic regions and geological settings, thereby enhancing the model's robustness and generalizability.

In this chapter, we delve into the development and validation of a DNN model designed for the purpose of estimating EHVR solely from MHVR, to facilitate the use of MHVR for obtaining pseudo-EHVR at locations where seismic records are unavailable. This enables further applications such as site classification, site effect assessment, and earthquake intensity prediction. The model is trained on a comprehensive dataset that combines MHVR and EHVR measurements from multiple sites, encompassing a wide range of soil types and seismic activity levels. The goal is to develop a reliable predictive model that will contribute to more accurate seismic risk assessment and ground motion prediction using EHVR in future studies.

4.2. Data

4.2.1. Earthquake HVR

The EHVR data were from an open source database developed by Zhu et al. [82], which includes station information and EHVRs from a total of 1,742 sites from K-NET and KiK-net in Japan.

The EHVR curves in the discussed database are derived from earthquake recordings. Here's a detailed summary of how these EHVR curves are obtained:

(1). Data Collection: The database gathers seismograms recorded by the K-NET and KiK-net stations in Japan from a range of earthquake events. A total of 696,242 seismograms (three components each) recorded between 1996 and 2019 are initially considered.

(2). Data Filtering: Recordings with peak ground acceleration (PGA) less than 0.001 g and greater than 0.1 g are excluded to ensure that the remaining data are not significantly affected by soil nonlinearity.

(3). Fourier Amplitude Spectrum (FAS): The Fourier amplitude spectrum of each component is computed using the entire waveform. This captures the influences of P-waves, S-waves, and surface waves.

(4). Butterworth Filtering and Smoothing: Each FAS is then filtered using a Butterworth filter at cutoff frequencies of 0.1 and 30 Hz, which corresponds to the sensor high-cut frequency. The filtered spectra are smoothed using the Konno-Ohmachi smoothing function with a coefficient of 20.

(5). Calculation of EHVR: For each recording, the EHVR is calculated as the geometric mean of the smoothed FAS of the two horizontal components divided by the smoothed FAS of the vertical component. This process is repeated for all valid recordings at a site.

(6). Averaging EHVR: The final EHVR curve for a site is obtained by calculating the geometric mean of the EHVRs from all eligible seismograms recorded at that site.

(7). Peak Identification and Resonant Frequency Determination: The EHVR curve is

analyzed to identify significant peaks which represent potential site resonant frequencies. This involves determining the amplitude, frequency, width, and prominence of each peak. Peaks are selected based on a set of criteria including significance, prominence, and sharpness.

(8). Automated Peak-Picking Scheme: An automated peak-picking scheme is employed to identify significant peaks from the EHVR curves. This method considers several statistical properties of the peaks and employs a threshold-based approach to ensure that only significant peaks are selected.

4.2.2. Comparison of EHVR and MHVR

We collected microtremor records from 248 sites in Japan and calculated MHVR. After resampling the data, we compared MHVR and EHVR, as shown in Figure 4-1. The primary peaks of MHVR and EHVR align closely, suggesting that MHVR can effectively capture the fundamental site frequency observed during earthquakes. However, the overall shape of the spectra may differ, highlighting the complex interaction of seismic waves with local geology. These difference may be due to the different nature of seismic energy involved in microtremors versus earthquake motions, or variations in local site conditions affecting the seismic wave propagation [15,27].

The comparison generally supports the reliability of MHVR in estimating site resonant frequencies, a crucial parameter in seismic hazard assessment and structural design. However, the differences in peak amplitudes and the presence of unique peaks in EHVR not mirrored in MHVR underscore the importance of using EHVR when available for precise seismic analysis.

In Figure 4-2, the statistical assessment of the relationship between MHVR and EHVR at the same site is quantitatively analyzed using various goodness-of-fit (GOF) metrics. This comparative analysis aims to elucidate the degree of consistency between MHVR and EHVR. The correlation coefficients predominantly range between 0.25 and 0.75, highlighting a moderate to strong positive linear relationship between MHVR and

EHVR measurements. The mean absolute error (MAE) values, mainly confined between 0.1 and 0.3, represent the average magnitude of absolute errors between MHVR and EHVR. The relatively low MAE values suggest minor deviations between MHVR and EHVR, underscoring their comparative reliability. The mean squared error (MSE) and root mean squared error (*RMSE*) values, observed mostly between 0.1 and 0.3, measure the average of the squares of the errors and the square root of these squares, respectively. Similar to MAE, these low values indicate small squared differences, which point to the accuracy of MHVR in mirroring EHVR. A high index of agreement (*d*) suggests that MHVR not only correlates with EHVR but also agrees closely in terms of the magnitude.

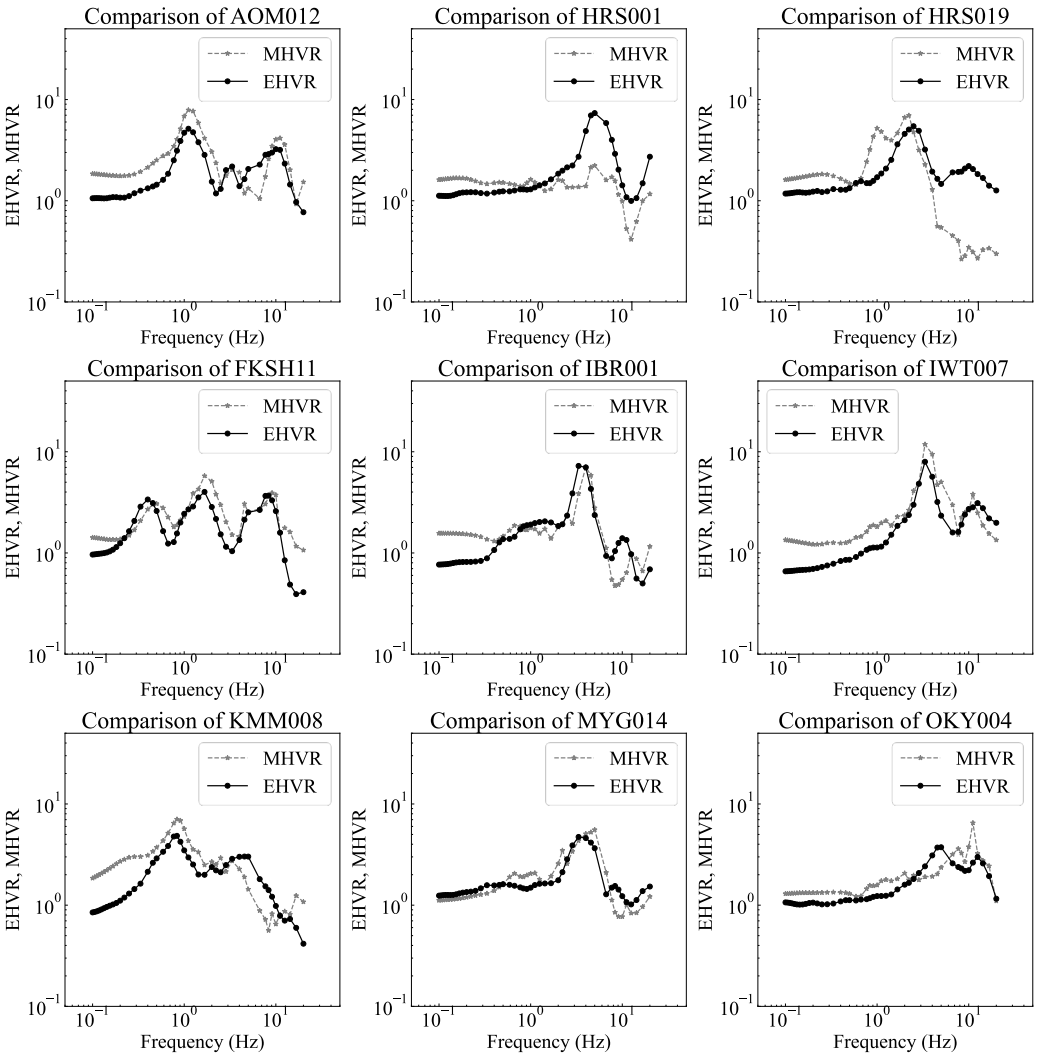

Figure 4-1. Comparison of EHVR and MHVR across multiple sites.

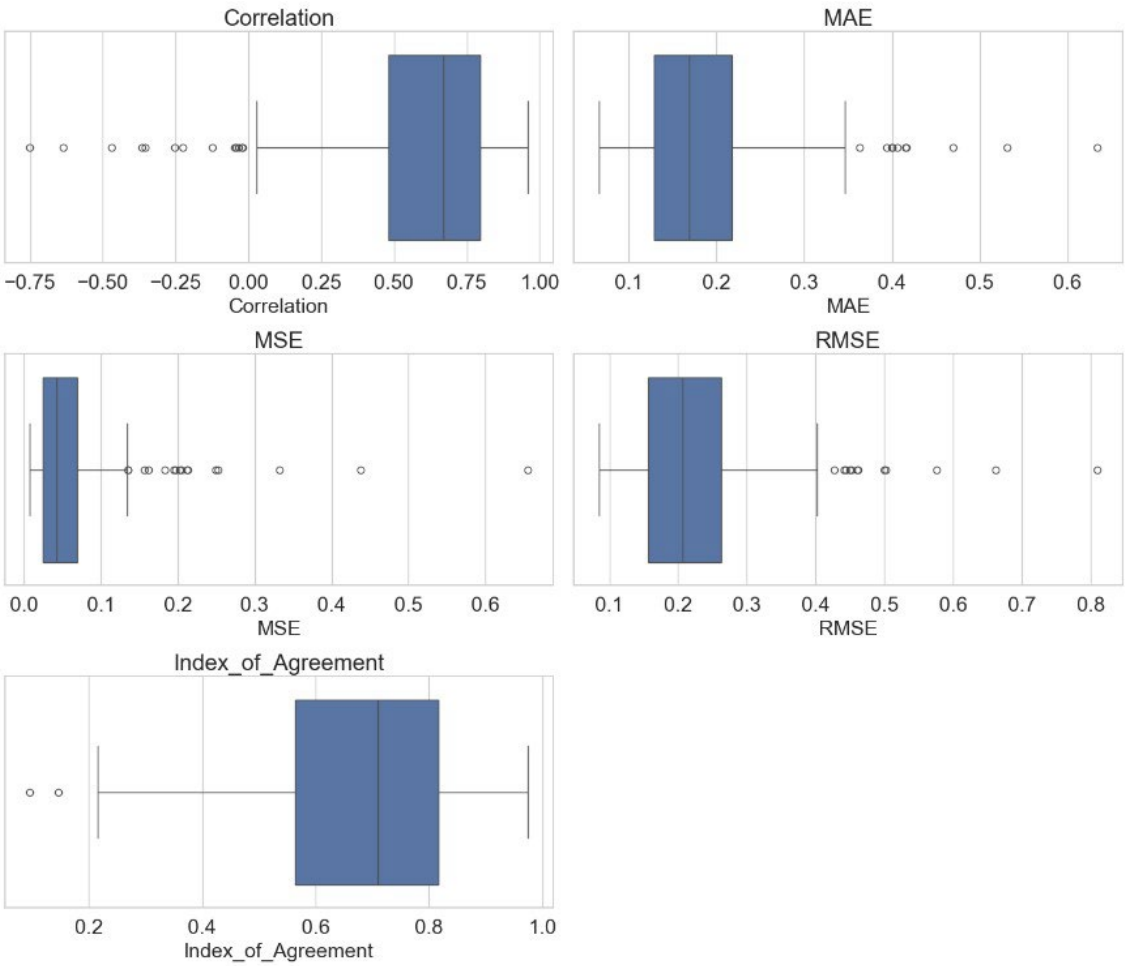

Figure 4-2. Distribution of goodness-of-fit (GOF) metrics between EHVR and MHVR at the same site.

However, considering the potential errors that may arise during the collection of microtremor data, we have adopted the following two data selection strategies:

[1]. Sites where the amplitudes of EHVR and MHVR are less than 2 have been excluded; refer to Chapter 2.

[2]. Sites with peak frequencies below 0.2 Hz have been excluded.

After implementing these strategies, we excluded data from a total of 18 sites. All

18 sites are listed in Figure 4-3.

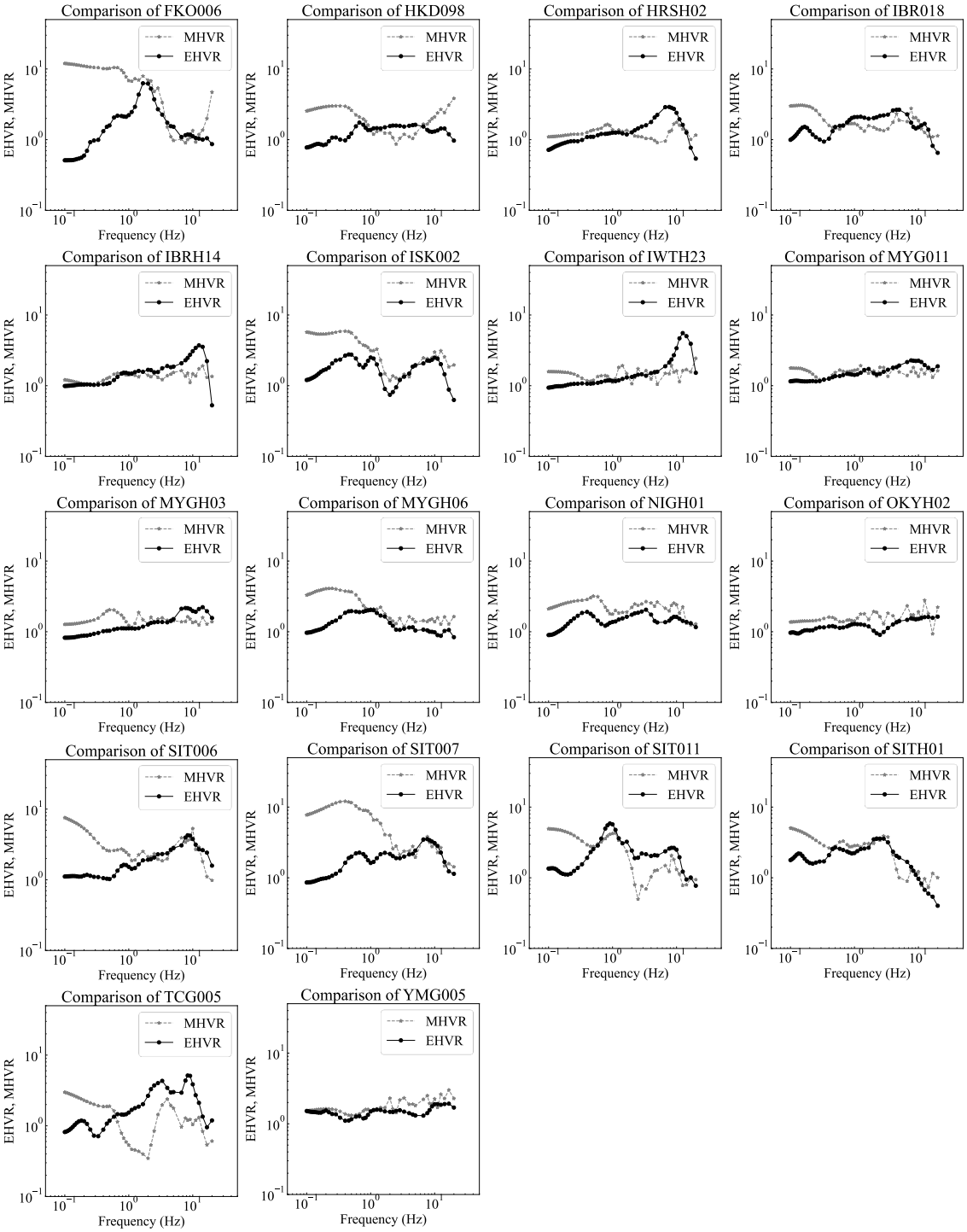

Figure 4-3. Sites excluded from subsequent analyses.

4.3. Construction of DNN model for prediction

After excluding the above 18 sites, we retained a total of 230 sites as a database for subsequent analysis. In this chapter, we still use the cross-validation strategy to split all data. In order to make each fold of data more balanced, we divided the data according to the peak frequency of MHVR (f_M). As shown in Figure 4-4, all data are divided equally into Fold1-5 and an independent test set.

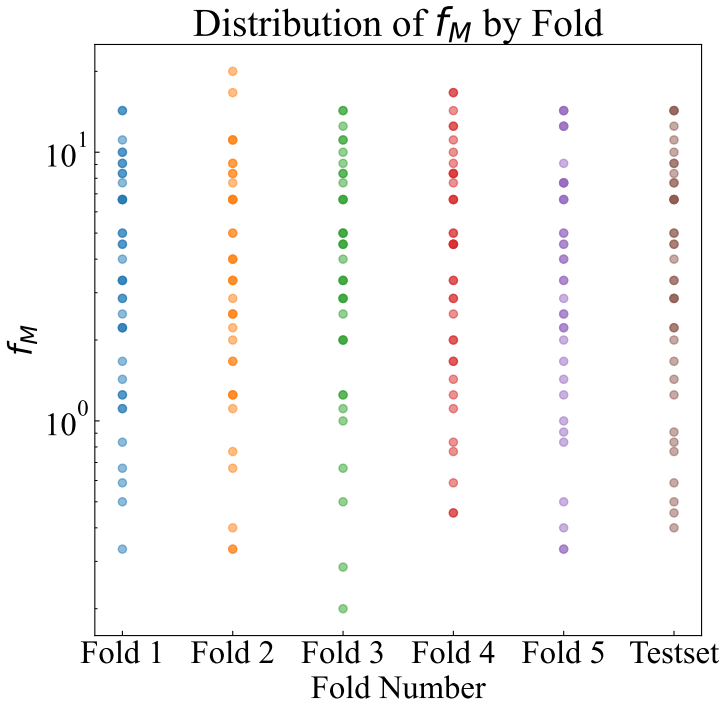

Figure 4-4. Distribution of cross validation folds based on MHVR peak frequencies.

Subsequently, the construction of the model first considers the same model as in Chapter 2, as shown in Figure 4-5. Using the same input and output strategy as in Chapter 2, the final pseudo-EHVR (pEHVR) is then obtained using Eq. (4-1).

$$pEHVR = MHVR \times output \tag{4-1}$$

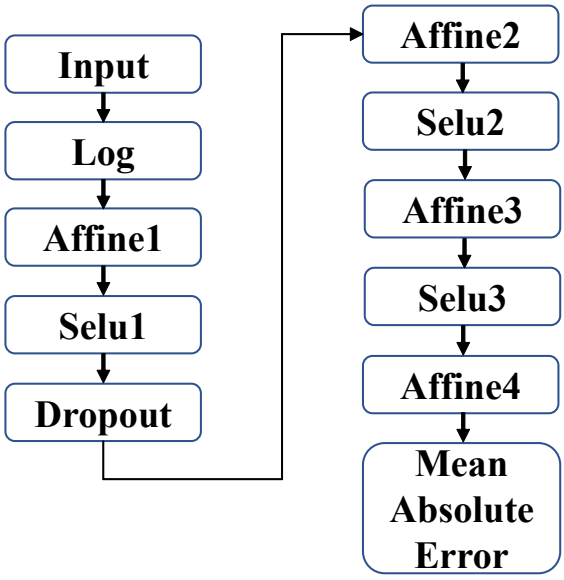

Figure 4-5. Architecture of the DNN model for pEHVR from MHVR. This is a preliminary model design and was not the final model adopted for the study.

The model is validated using the training-validation set of CV1. After confirming that the model is not overfitting, all sites in the CV1 validation set are calculated, and the residuals are calculated using Eq. (4-2).

$$Residual(f_i) = \log\big(EHVR(f_i)\big) - \log\left(pEHVR(f_i)\right) \tag{4-2}$$

The average and standard deviation of residuals for all sites in the CV1 validation set are shown in Figure 4-6. The red portion in the graph represents the average residual between the true values of MHVR and EHVR, while the gray portion represents the average residual between the EHVR and the pEHVR derived from the above model. In the low-frequency range, the model's estimated accuracy is higher, with the average residual tending towards zero, and the standard deviation also decreasing. However, in the high-frequency range, the improvement effect of the model is not ideal, almost

approaching the average residual between the two true values. Similar results were obtained for all CV validations, indicating that the model is not very adaptable to this topic.

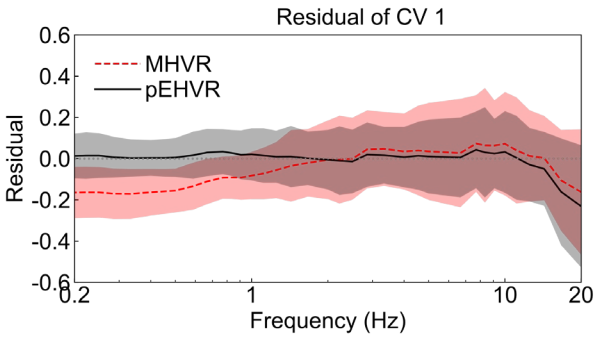

Figure 4-6. Comparison of average residuals and standard deviation between predicted and true HVR. Red regions depict the residuals between MHVR and EHVR values, while gray regions indicate residuals between the predicted pseudo-EHVR (pEHVR) and true EHVR. This figure corresponds to the model output of Figure 4-5.

To address this issue, several alternative model constructions were attempted and compared. Ultimately, a neural network incorporating Batch Normalization was selected as the most effective solution.

Batch Normalization is a technique designed to improve the training of deep neural networks by standardizing the inputs to a layer for each mini batch. This addresses the problem of internal covariate shift where the distribution of each layer's inputs changes during training, as the parameters of the previous layers change [83,84]. This can make training slow and requires careful parameter initialization and a smaller learning rate. Batch Normalization helps to mitigate these issues by making the network training more stable and faster.

The architecture of this improved DNN model is shown in Figure 4-7:

Input Layer: The model begins with an input layer that receives the MHVR data,

with a feature size of 7, refers to Section 2.4.3 and Table 2-3.

Affine Layers: These fully connected layers linearly transform the input features, with neuron counts decreasing progressively (64, 32, 16).

Batch Normalization Layers: Following each affine layer, batch normalization layers normalize the inputs, stabilizing the learning process.

SELU Activation Functions: The Scaled Exponential Linear Unit (SELU) activations are used to maintain the self-normalizing property of the network.

Dropout Layer: A dropout layer with a rate of 0.5 is included to prevent overfitting

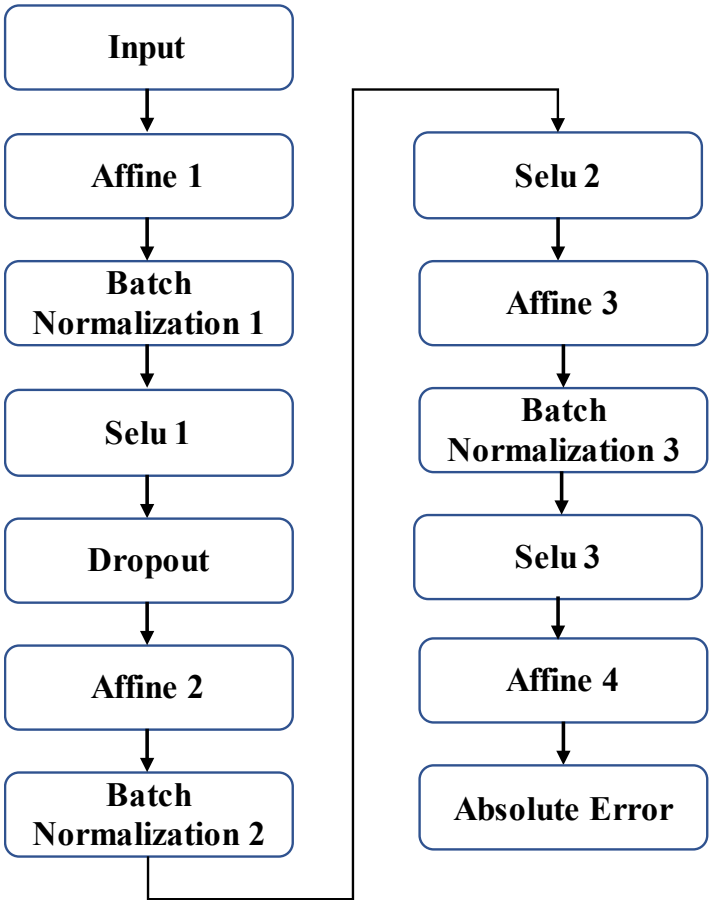

Figure 4-7. Architecture of the improved DNN model for predicting pEHVR from MHVR. The model incorporates Affine transformations, Batch Normalization, SELU activations, and a Dropout layer to enhance prediction accuracy and stability. This architecture was selected as the final model.

by randomly setting a fraction of input units to zero during training.

Output Layer: The final output layer computes the absolute error between the predicted pseudo-EHVR and the true EHVR values.

The number of neurons in each layer is shown in Table 4-1. The other important parameters are set as follows: Max epoch is 1000, learning rate is 0.001, and Batch size is set to full batch size.

Table 4-1. Number of neurons in the DNN model

Layer	Number of neurons
Input	7
Affine 1 + BN 1+ Selu 1	64
Affine 2 + BN 2+ Selu 2	32
Affine 3 + BN 3+ Selu 3	16
Affine 4	5

To evaluate the performance of the improved DNN model incorporating Batch Normalization, we conducted a residual analysis comparing the predicted pEHVR with the true EHVR values. Additionally, we compared the residuals between the true MHVR and EHVR values. The average and standard deviation of residuals for all sites in the CV1

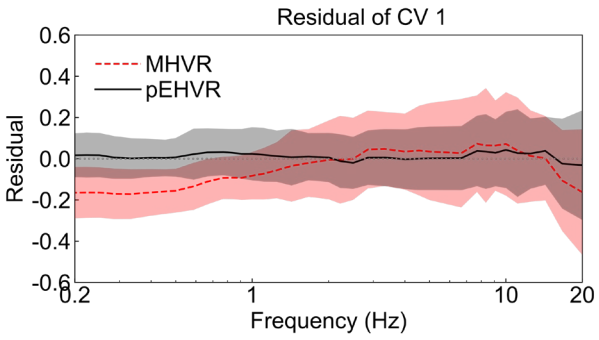

Figure 4-8. Comparison of average residuals and standard deviation between pEHVR and EHVR, MHVR and EHVR. This figure corresponds to the model output of Figure 4-7.

validation set are shown in Figure 4-8.

The black line shows that the model's residuals tend towards zero, indicating high prediction accuracy. The standard deviation is also low, indicating consistent performance. This also solves the problem of inaccuracy in the high frequency range shown in Figure 4-9. The results of all cross-validation sets, and the test set will be discussed in the next section.

4.4. Cross-validation and results

4.4.1. Prediction of EHVR for cross-validation sets

The DNN model in Figure 4-7 was adopted as the final model for predicting pEHVR using MHVR. Figure 4-9 presents the residual analysis of the final DNN model for predicting pEHVR from MHVR across all cross-validation (CV) folds. The analysis compares the residuals between the predicted pEHVR and the true EHVR values, as well as the residuals between the true MHVR and EHVR values for each CV fold.

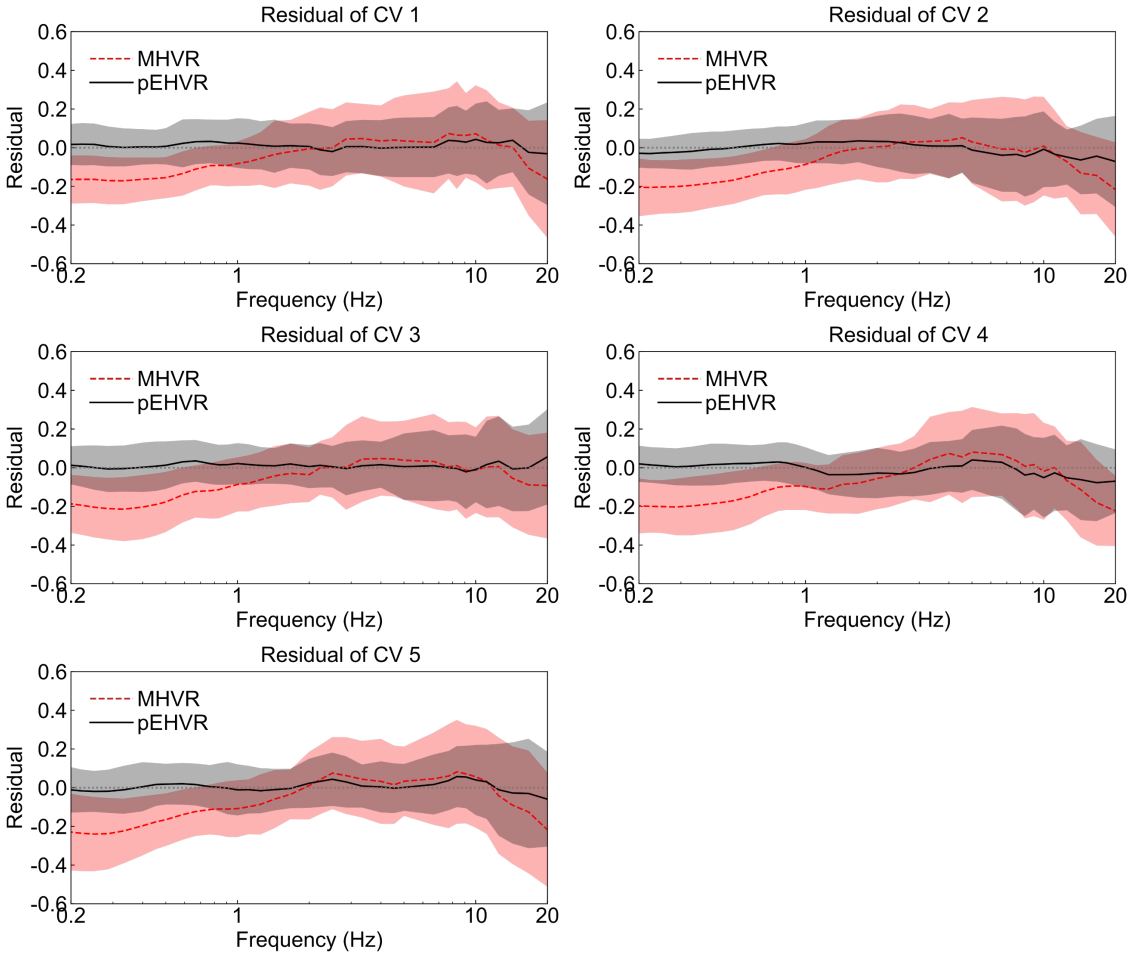

Figure 4-9. Residual analysis across all cross-validation folds

In the low-frequency range, the black line shows that the residuals tend towards zero, indicating high prediction accuracy. The standard deviation also shows a decreasing trend, indicating consistent performance across different CV folds.

Similarly, in the high-frequency range, the average residuals also tend towards zero, and the standard deviation decreases, suggesting good model performance.

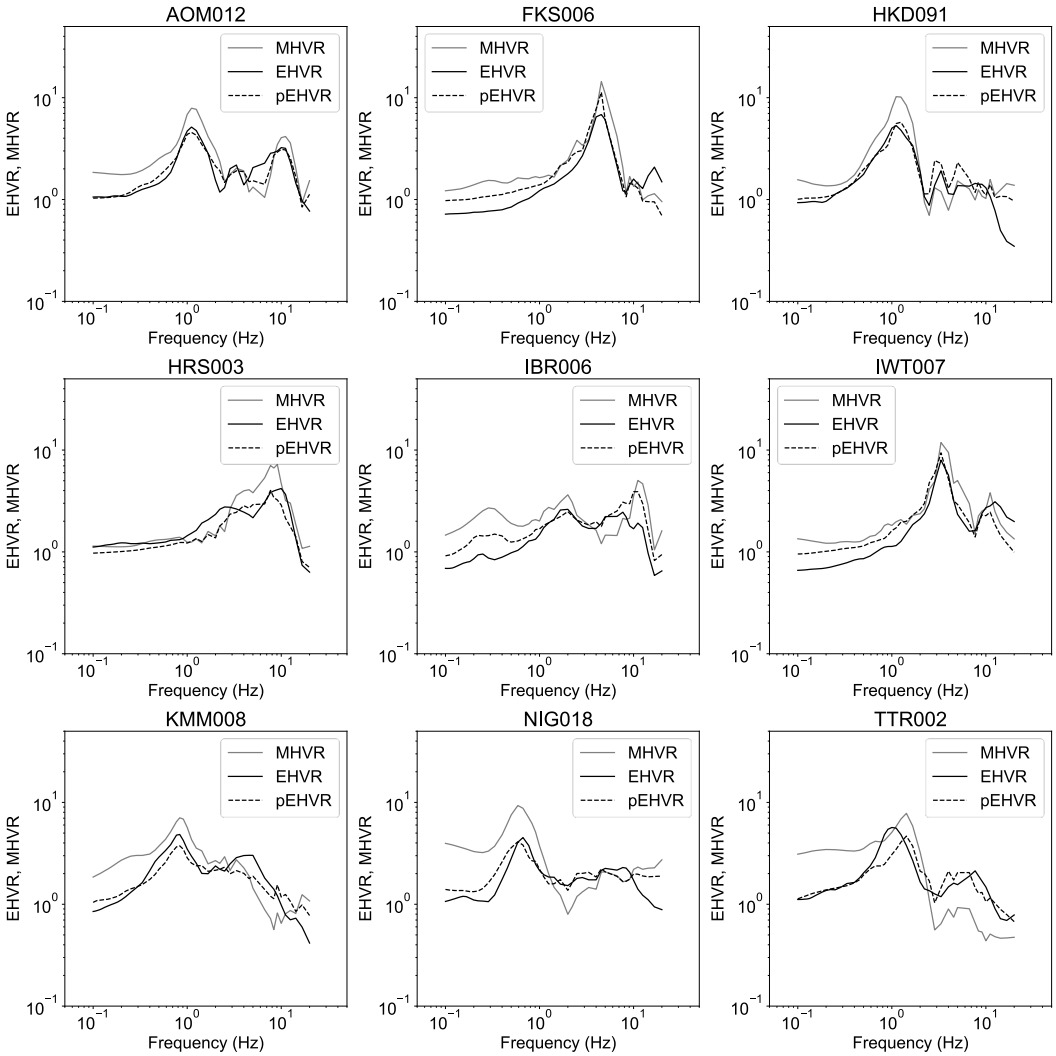

Figure 4-10. Comparison of MHVR, EHVR, and predicted pEHVR across multiple sites on CV 1.

Figure 4-10 presents a series of sites comparing the MHVR, EHVR, and the predicted pEHVR generated by the DNN model across multiple sites. Each subplot represents a different site, providing a visual comparison of the three spectra.

The alignment of the pEHVR (dashed line) with the EHVR (solid black line) indicates the accuracy of the model in predicting the spectral ratios at each site. In many sites, the pEHVR closely follows the EHVR, especially in the low-frequency range, indicating the model's high prediction accuracy in this range. Some discrepancies are observed in the high-frequency range, where the pEHVR may deviate from the EHVR, highlighting areas where the model's performance could be improved.

In addition, the Goodness-of-Fit (GOF) of each site was also calculated. The validation sets of all CVs were summarized according to the average values of each metric. Table 4-2 summarizes the GOF metrics for the cross-validation (CV) sets, comparing the performance of Microtremor Horizontal-to-Vertical Spectral Ratios (MHVR) and the predicted pseudo-EHVR (pEHVR) against the true Earthquake Horizontal-to-Vertical Spectral Ratios (EHVR). The metrics include Correlation (R), Determination Coefficient (R^2, using Eq. (4-3)), Root Mean Squared Error ($RMSE$), and Index of Agreement (d). The delta values indicate the improvements achieved by using the predicted pEHVR over MHVR.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(obs_i - pre_i)^2}{\sum_{i=1}^{n}(obs_i - \overline{obs_i})^2} \tag{4-3}$$

Here n represents the number of periodic terms, which is equal to 46. Upper bar indicates a mean value of observed values. Higher R^2 values indicate a stronger alignment between the predicted and observed spectral amplitude, indicating a better match in terms of spectral shape.

The summarized GOF metrics across all CV sets show that the proposed DNN model

significantly enhances the prediction of EHVR from MHVR. The model consistently demonstrates higher correlation, higher determination coefficients, lower *RMSE*, and higher Index of Agreement, indicating improved accuracy and reliability in the predictions. These improvements validate the model's robustness and effectiveness in predicting EHVR using MHVR data.

Table 4-2. Summary of Goodness-of-Fit (GOF) metrics for cross-validation sets comparing MHVR, pEHVR, and EHVR. The values present the GOF metrics averaged across all validation sites within each CV set.

		Correlation R	Determination R^2	*RMSE*	Index of Agreement d
CV 1	MHVR vs EHVR	0.695	0.535	0.197	0.750
	pEHVR vs EHVR	0.854	0.750	0.124	0.880
	Delta	0.159	0.215	-0.073	0.131
CV 2	MHVR vs EHVR	0.619	0.433	0.217	0.689
	pEHVR vs EHVR	0.823	0.691	0.125	0.869
	Delta	0.204	0.258	-0.092	0.180
CV 3	MHVR vs EHVR	0.657	0.483	0.208	0.730
	pEHVR vs EHVR	0.814	0.698	0.132	0.851
	Delta	0.157	0.215	-0.076	0.120
CV 4	MHVR vs EHVR	0.579	0.418	0.215	0.669
	pEHVR vs EHVR	0.806	0.667	0.128	0.852
	Delta	**0.226**	**0.249**	**-0.086**	**0.183**
CV 5	MHVR vs EHVR	0.630	0.448	0.232	0.692
	pEHVR vs EHVR	0.815	0.695	0.137	0.851
	Delta	0.185	0.248	-0.096	0.159

4.4.2. Prediction of EHVR for Test sets

Given these significant improvements in CV 4, the model trained on CV 4 data was selected for further testing. This trained model was applied to the test set, which consists of all sites that were not included in the training or validation phases. The performance on the test set provides a robust assessment of the model's generalizability and its ability to predict EHVR from MHVR across different, unseen sites.

As shown in Figure 4-11, the residual analysis on the test set demonstrates that the average residuals tend towards zero across both low and high-frequency ranges. The standard deviation also shows a decreasing trend, indicating that the model generalizes well to unseen data and does not exhibit overfitting. Combined with Table 4-3, this robust performance confirms the model's effectiveness in predicting EHVR from MHVR. A comparison of pEHVR, EHVR, and MHVR for some of the sites is shown in Figure 4-12.

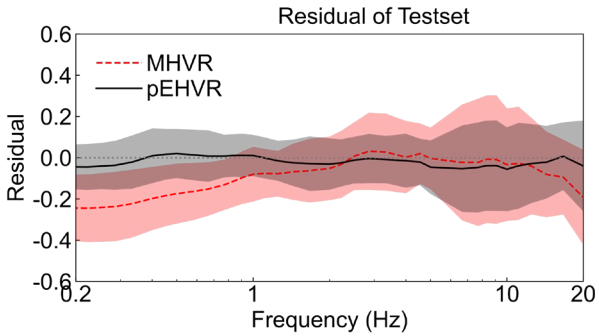

Figure 4-11. Residual of pEHVR and EHVR on the Test Set.

Table 4-3. GOF metrics for test set comparing MHVR, pEHVR, and EHVR.

		Correlation R	Determination R^2	*RMSE*	Index of Agreement d
Test set	MHVR vs EHVR	0.580	0.410	0.223	0.650

pEHVR vs EHVR	0.798	0.658	0.132	0.852
Delta	0.217	0.248	-0.091	0.202

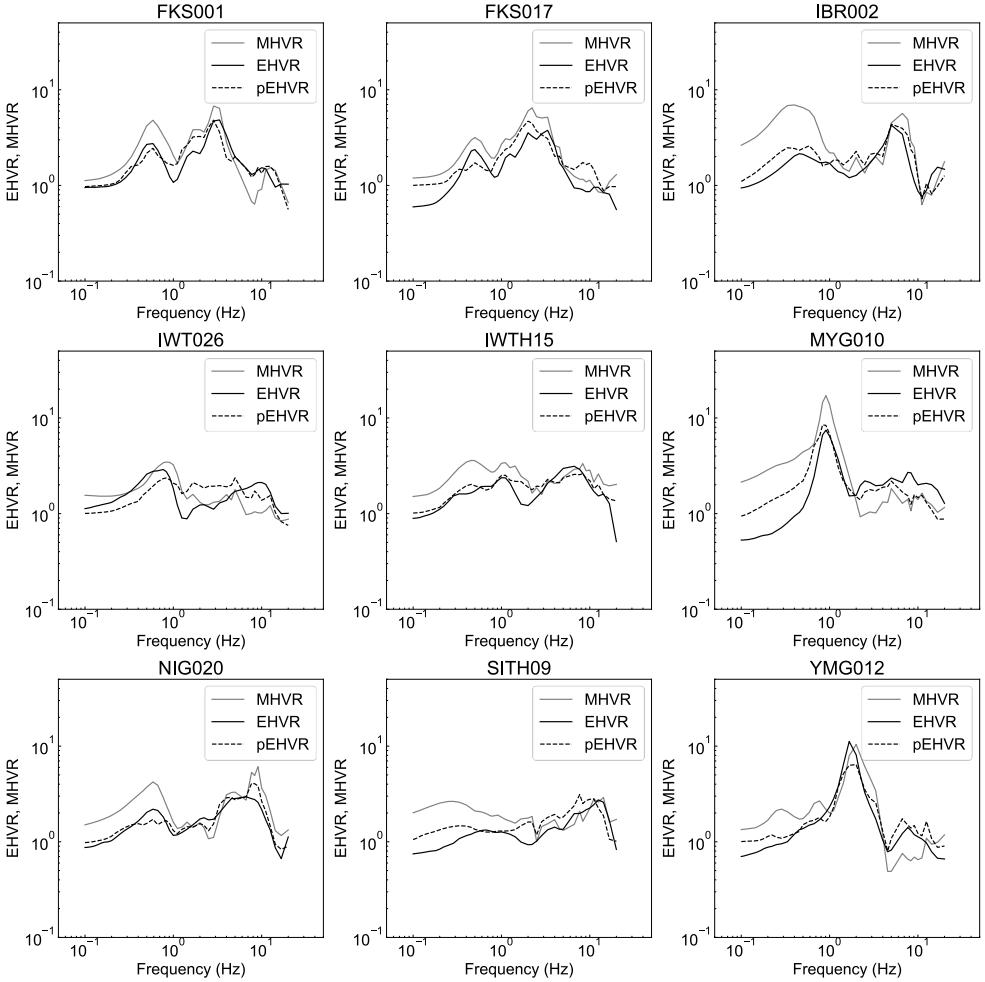

Figure 4-12. Comparison of MHVR, EHVR, and predicted pEHVR across multiple sites on test set

4.5. Conclusions

This study presented a comprehensive approach to predicting EHVR using MHVR. The primary goal was to develop a robust and reliable model capable of accurately estimating EHVR from MHVR data, thus providing a valuable tool for seismic risk assessment and structural design, particularly in regions where earthquake data are sparse.

A DNN model incorporating Batch Normalization and SELU activation functions was developed and validated using cross-validation techniques. The model demonstrated significant improvements in prediction accuracy, as evidenced by the various goodness-of-fit (GOF) metrics such as Correlation (R), Determination Coefficient (R^2), Root Mean Squared Error ($RMSE$), and Index of Agreement (d).

The trained model was applied to an independent test set to assess its generalizability. The residual analysis on the test set confirmed that the average residuals tend towards zero across both low and high-frequency ranges, with a decreasing trend in standard deviation. This suggests that the model generalizes well to unseen data and does not exhibit overfitting.

The findings of this study underscore the effectiveness of using MHVR data to predict EHVR, offering a cost-effective and non-invasive method for seismic risk assessment. The developed model provides a reliable means to estimate site-specific seismic responses.

Although the current model demonstrates strong performance, future research could focus on the following areas:

Further improving the prediction accuracy in the high-frequency range where minor discrepancies still exist. Expanding the dataset to include a wider variety of geological conditions to validate the generalizability. Verifying the effectiveness of the predicted pEHVR for other applications in the field of seismology by combining it with other uses of EHVR.

Chapter 5. Ground Motion Prediction by DNN model using EHVR

The previous chapters have detailed the relationships between microtremor horizontal-to-vertical spectral ratios (MHVR), earthquake horizontal-to-vertical spectral ratios (EHVR), and site effects, as well as the corresponding prediction models. Site effects play a crucial role in the prediction of ground motion during earthquakes. This chapter will focus on how EHVR can be used to enhance the accuracy of existing ground motion prediction model. This chapter proposed a deep-neural-network (DNN) model for seismic ground motion prediction by utilizing a unified strong motion database by the National Research Institute for Earth Science and Disaster Resilience, and EHVR database in Japan. The model aims to enhance the accuracy of predictions by incorporating the EHVRs for complementing site effects, and utilizing existing ground motion prediction equation (GMPE) as the base model for source and propagation path effects. The hybrid approach enables the prediction of peak ground accelerations (PGAs), peak ground velocities (PGVs) and 5% damped absolute acceleration response spectra (SAs). After classifying the training and test sets from the database, the trained DNN models were applied on the test set to evaluate the performance of the predicted results. The accuracy assessment by the residuals, R-squared (R^2) and Root Mean Square Error (*RMSE*) between the predicted and observed values in the test set revealed the superior performance of the proposed model compared to the traditional GMPE with proxy-based site effects such as V_{S30} especially in predicting both the spectral amplitude and shape of SAs.

5.1. Background

Ground motion models (GMMs) are key tools for predicting seismic ground motions in future earthquakes, and are also critical for regional disaster planning and seismic-resistant building design. Most existing GMMs have been derived through empirical regression analysis of ground motion records induced by past earthquake events. This type of regression analysis, based on pre-defined equations, has been referred to as ground motion prediction equations (GMPEs). The functional form of GMPEs varies depending on the choice of the modeler, but typically includes factors such as moment magnitude M_W, focal depth, earthquake type (crustal or subduction earthquake), and considering site effects and propagation path effects [85–91]. Reviewing the existing GMPEs [92], many seismologists have introduced proxies such as the V_{S30} (time-averaged shear-wave velocity in the upper 30 m) or site classification into the constructed GMPEs to evaluate the site effects on ground motion predictions. Additionally, these prediction models have been highly relevant to seismic safety assessment and design of local building structures and infrastructures. Therefore, researchers often use local or regional databases [93–95], as well as a global ground motion database NGA-West2 [96,97], to construct GMPEs. Recently, the National Research Institute for Earth Science and Disaster Resilience (NIED) has been developing a unified strong motion database to improve ground motion prediction equations in Japan [98] with the aim of developing models based on this unified database for further research.

On the other hand, in recent years, ground motion prediction methods based on machine learning and deep learning models, also known as non-parametric models, have also been proposed. Non-parametric model is one of constructing prediction models by learning and analyzing large amounts of data that do not rely on specific mathematical equations. Derras et al. [99] proposed an artificial neural network model that used M_W, focal depth, epicentral distance, site resonance frequency (f_0), and V_{S30} as input variables. These five parameters correspond to the physical meanings of the source characteristics, propagation path effects, and site effects, respectively. Recently, researchers have

constructed non-parametric models using locations of epicenters and observation sites as input variables. Oana et al. [100,101] used parameters such as longitudes, latitudes, and source azimuths for epicenters and sites to construct machine learning models. Lilienkamp et al. [102] viewed ground motion prediction as an image processing task and used latitudes and longitudes of epicenters and observation sites as two-dimensional array information. They combined this information with other key parameters to construct a U-Net model. In these studies, geographical information such as longitude and latitude are considered a convenient definition of propagation path effects. There are studies using source mechanisms or fault of earthquake events as input variables to represent source characteristics, and combined site and path characteristics to construct GMM models by machine learning technology [61,101,103,104]. All these studies indicate that the data-driven models have better prediction accuracy than existing GMPEs.

The advantage of non-parametric models is that they can more accurately reflect the interaction between different physical factors, and the prediction accuracy can be continuously improved as data accumulates [105,106]. Additionally, Kubo et al. [107] pointed out that using the ground motion database with significant bias can lead to substantial distortion in the trained machine learning model, particularly regarding the underestimation of the prediction of strong motions. This is because there are relatively fewer strong ground motion records available, which leads to the model training being biased towards weaker ground motion data. Kubo et al. [107] proposed a hybrid approach of combining machine learning and physics-based models to address this issue. Specifically, they combined physics-based model proposed by Morikawa and Fujiwara [90] with machine learning to ensure predictive performance for infrequent strong motion events, and the hybrid approach outperforms any single method applied alone.

In general, the parameters such as V_{S30} [60,108], depth of seismic bedrock surface D_{base}, and depth to the S-wave velocity layer of 1400 m/s (D_{1400}) [101,107] have been often used as input for many GMMs to describe the site effects. However, considering that the site effects are frequency-dependent [109], such proxy values are difficult to

represent the effects of the entire frequency range comprehensively. Midorikawa [110] also indicated that the accuracy of GMPE strongly depends on site amplification effects.

On the other hand, since the proposal of the horizontal-to-vertical spectral ratio (HVR) by Nakamura [9,20], the relationships between site effects and HVR have been discussed based on microtremor and earthquake observation data in various regions such as Japan and Mexico [15,21–24]. Recently, our team [66,81,111] proposed a deep-neural-network model for direct estimation of the site effects based on the property that the microtremor H/V spectral ratios (MHVRs) approximates the site amplification factors. Kawase et al. [27,32] showed that although the MHVRs and earthquake H/V spectral ratios (EHVRs) are similar, the EHVRs are more representative of site effects and better reflect subsurface structure identification. Zaker et al. [112] developed data-driven models to predict site amplification using EHVRs, noting that EHVR proxies could replace common site condition proxies like V_{S30}. Additionally, Zhu et al. [65] through comparison between traditional site effect assessment methods and EHVR-based site effect machine models, further validated the effectiveness of combining EHVR with machine learning models in predicting site amplification effects. These studies inspired us to explore whether EHVR could also improve the accuracy of prediction models for seismic parameters (e.g., peak ground accelerations), given the relationship between EHVR and site effects. As several studies have pointed out the usefulness of EHVRs as descriptor variables for site effects [31,65,113–115], the site-specific EHVR can be used as the input variables for GMMs. Compared to single proxy values such as V_{S30} and D_{base}, the EHVRs have the potential to dramatically improve the accuracy of existing GMMs because of the full frequency-dependence to describe site effects. Recently, the EHVR database for K-NET and KiK-net stations across Japan has been established [82], providing a user-friendly analysis environment.

In this study, we developed deep-neural-network (DNN) models for ground motion prediction. Leveraging two primary databases, the unified Strong Motion Database introduced by Morikawa et al. [98] and the EHVR database provided by Zhu et al. [82],

we aimed to utilize the inherent comprehensive frequency dependency within EHVRs as the supplement site terms for amplification factors to construct hybrid models. Specifically, we developed two distinct DNN models: one aimed at predicting peak ground accelerations (PGAs) and peak ground velocity (PGVs) at sites, and another focused on predicting 5% damped absolute acceleration response spectra (SAs) values across 46 periods ranging from 0.05s to 10s. Our methodology involved the existing GMPE by Morikawa and Fujiwara [90] abbreviated as MF13, as the base model. Initially, calculations based on the base model of MF13 are performed to obtain predictions considering source and propagation effect. Then the DNN models supplement the computed results by accounting for site effects. The description of source and propagation path effects was given by the existing GMPE. The utilization of the DNN models with EHVRs as input effectively characterizes site effects, overcoming the limitations of traditional site effect proxies in adequately capturing site complexity. To validate our proposed method, we conducted standard deviation analysis and compared observed and estimated SAs, achieving two commendable matching scores.

5.2. Data

5.2.1. Ground motion database

The ground motion database used herein was built on the basis of the unified strong motion database in Japan (Morikawa et al., 2020), which includes the public ground motion records of K-NET and KiK-net from 1996 to 2018 and consists of three sets of files: earthquake event data, station information data and ground motion record data. The main information included in the NIED strong motion database is summarized in Table 5-1. The earthquake event data utilized in the database comprises source parameters obtained from JMA (Japan Meteorological Agency) and seismograph network of F-net, finite-fault rupture models, earthquake type, and fault parameters, and so on. Regarding the station information data, it encompasses V_{S10}, V_{S20}, and V_{S30} values. These values were derived from the station location, PS logging data, geomorphological classification map. The information also includes the top-surface depths of the main layers in the deep ground, and so on. The ground motion record data consists of various parameters, including PGAs, PGVs, measured seismic intensity in Japanese scale, and absolute acceleration response spectra (SAs). The SAs comprise 59 points covering periods ranging from 0.02 to 20 s with a damping factor of 5%. Furthermore, in addition to the horizontal motion components such as NS, EW, RotD000 to RotD100 [116], the vertical motion components (UD) were also incorporated. The distribution of the number of records by fault distance *FD* (also known as rupture distance, R_{rup}) and M_W of the registered strong-motion data is shown in Figure 5-1. Here, the fault distance is defined as the minimum distance between the observation point and the earthquake fault. There is a large number of records where the M_W is less than 5.0 and the *FD* is around 100 km.

Table 5-1. Main information included in the NIED strong motion database.

Earthquake event data
• Date and time of event • Location of source • Magnitude (JMA scale and Moment magnitude) • Fault length and width • Strike and dip • Depth • Earthquake type (Crustal, Inter-plate, Intra-plate) • etc.
Station information data
• Location of the observation site • Distance from volcanic front • Average shear-wave velocity in upper 10m, 20m and 30m (V_{S10}, V_{S20}, V_{S30}) • Depth of seismic bedrock surface D_{base} • Depth to layer with S-wave velocity of 1100 m/s D_{1100} (and D_{1400} , D_{1700} , D_{2100}) • etc.
Ground motion record data
• PGA (Peak ground acceleration) • PGV (Peak ground velocity) • SA (5% damped spectral response acceleration; periods 0.02~20 s; 59 points) • Fault distance (*FD*) • etc. * PGAs and PGVs were obtained for NS, EW, UD, RotD000, RotD025, RotD050, RotD075, and RotD100 components. SAs of horizontal two components were obtained for RotD000, RotD025, RotD050, RotD075, and RotD100.

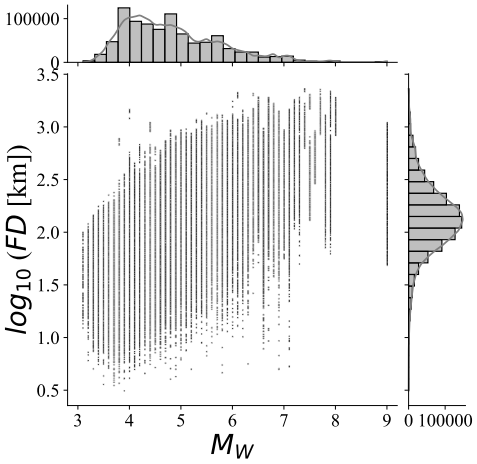

Figure 5-1. Distribution of the number of records included in the strong motion database.

More specifically, five record selection strategies listed below were used to extract a subset from this database to serve as the basis dataset for developing the predictive model proposed in this study. The criteria of $M_W \geq 5.5$, $FD \leq 200$ km and RotD100 were determined based on consistency with the base model selection strategy (Morikawa and Fujiwara 2013). This ensures that the range of data used in the proposed methodology is consistent with that of the base model database, thereby improving the comparability and reliability of the study.

(1). Earthquake events with M_W larger than or equal to 5.5.

(2). FD within 200 km.

(3). At least five stations are available for ground motion recording of earthquake events.

(4). Ground motion records only from surface free-field condition at K-NET and KiK-net stations.

(5). Excluding records from subduction intra-plate earthquake events.

We focused on the impact of site effects on the GMM model. In the case of intra-plate earthquake events in Japan, ground motion records through the volcanic front in the propagation path are strongly influenced by anomalous seismic intensity distribution

[117,118]. Figure 5-2 displays the data distribution of the impact of different earthquake types and volcanic regions on ground motion. The upper part represents earthquake records where the propagation path from the source to the observation site does not pass through volcanic regions. Different colors represent different types of earthquakes. It can be observed that, regardless of whether it is a crustal, interplate, or intraplate earthquake, the source distance shows a high correlation with PGA. However, intraplate earthquakes exhibit deviations.

The lower portion of the figure presents three plots that represent earthquake records where the propagation path from the epicenter to the observation site passes through volcanic regions. Crustal and interplate earthquakes still show high correlation. In contrast, intraplate earthquakes exhibit significant deviations, with a correlation coefficient of only -0.36. This indicates that the propagation path passing through volcanic regions has a significant impact on intraplate earthquakes.

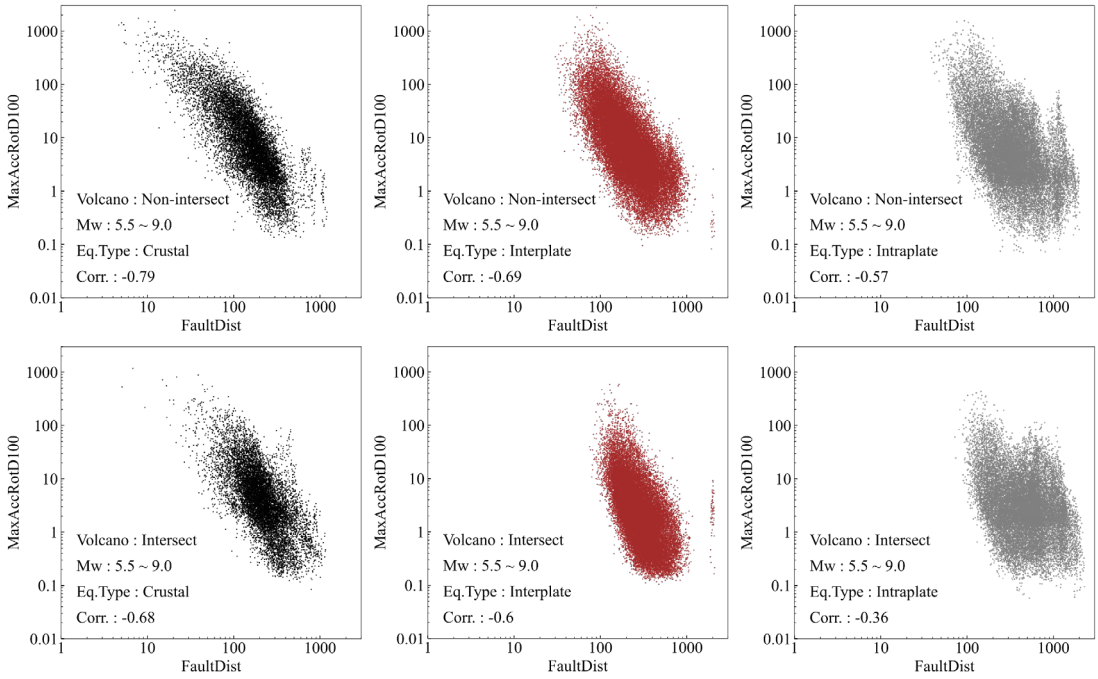

Figure 5-2. Distribution of data on the effects of different earthquake types and Volcano on ground motions.

To avoid learning errors caused by such anomalous seismic records during the learning process, only records of crustal and subduction inter-plate earthquakes were used. In the following analyses, we used the ground motion parameters corresponding to the two horizontal time domain components of the maximum rotated component (RotD100). The 352 earthquake events and 36,914 ground motion records obtained at 1,634 seismic observation stations were used as a database from the above selection strategy.

5.2.2. EHVR database

The EHVRs data were from an open source database developed by Zhu et al. [82], which includes station information and EHVRs from a total of 1,742 sites from K-NET and KiK-net in Japan. Figure 5-3 shows a comparison between the EHVRs in Zhu et al. [82] and the site amplification factors (SAFs) in Pan et al. [66] obtained from generalized

spectral inversion technique (GIT) at same sites. As noted in the previous studies [3,32,66], the peak frequencies and spectral shapes of the EHVRs are similar to those of the SAFs. Although SAFs are not readily available at a single observation site, EHVR can be obtained relatively easily from ground motion records. For this reason, EHVRs were used as the representative parameters for site effects.

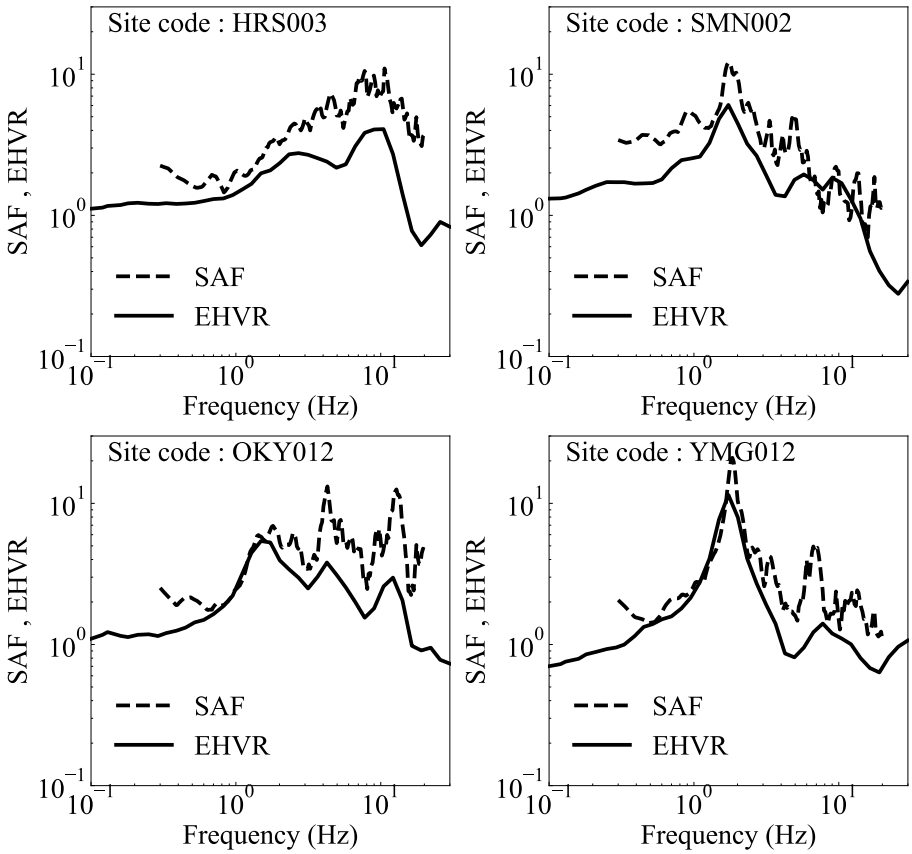

Figure 5-3. Comparison of the EHVRs in Zhu et al. [82] and site amplification factors (SAFs) in Pan et al. [66].

5.3. Methods

5.3.1. Ground motion prediction equation

We aim to construct a hybrid model by exploiting the full frequency dependence of EHVRs as site proxies for spectral amplification factors. The approach proposed involves using the existing GMPE (Morikawa and Fujiwara [90], hereafter we referred as MF13) as base model. The base model was used to perform initial calculations, and the calculated values are subsequently corrected using a DNN model. The DNN model was based on EHVRs as input to accurately represent the site effects, which compensates for the inability of proxy values in fully describing the site effects, especially frequency-dependent amplitudes. In contrast, the base model by MF13 was used to describe the effects of source and propagation paths. This approach was adopted due to the inherent pre-event uncertainties of explanatory variables such as M_W and FD that are commonly used as source and propagation path effects. Deep learning technologies, which are often criticized for being a "black box", are more susceptible to uncertainty in input variables than traditional GMPEs. This is mainly attributed to the complexity of deep learning model. Several factors contribute to this increased sensitivity: complex model architectures, nonlinearities and limited understanding of the inner workings. Deep learning models often have complex architectures containing many parameters and inherently involve nonlinear activation functions, enabling them to capture complex patterns and relationships within the data. However, the intricate interconnections and nonlinearity among these parameters make the models more sensitive to input variables. Furthermore, as noted by Kubo et al. [107], non-parametric models can be biased by the imbalanced distribution of M_W and FD.

The GMPE in MF13 was specifically designed to accommodate seismic records with M_W values ranging from 5.5 to 9.0 in Japan. As shown in Eqs. (5-1) and (5-2), it predicts PGAs, PGVs and SAs for 46 periods from 0.05 to 10 seconds (0.1 to 20 Hz). These ground motion parameters correspond to the values of RotD100.

$$log(pre_{basemodel}) = a(M_{W1}{}' - M_{W1})^2 + b_k \cdot FD + c_k - log(FD + d \cdot 10^{e \cdot M_{W1}{}'}) \quad (5\text{-}1)$$

$$M_{W1}{}' = min(M_W, M_{W01}) \quad (5\text{-}2)$$

Here, *pre* is the predicted value, *FD* is the shortest distance to the fault, a, b_k, c_k, d and e are the regression coefficients, and M_{W1} (=16.0) and M_{W01} (=8.2) are the default correction values. The database used to construct the MF13 model contains earthquake records up to the end of 2011. To verify that the MF13 model maintains a stable predictive capability for the new database, we applied the MF13 model to the database used in our study. Figure 5-4 shows the predicted PGAs calculated from the base model in Eq. (5-1); the existing GMPE has good applicability to new ground motion record database compared to the standard deviation results in MF13. It is important to note that there is an underestimation of the predictions in Figure 5-4 which is partially due to the fact that the base model here only used the parameters representing the source and propagation characteristics, and did not yet include a correction for site effects.

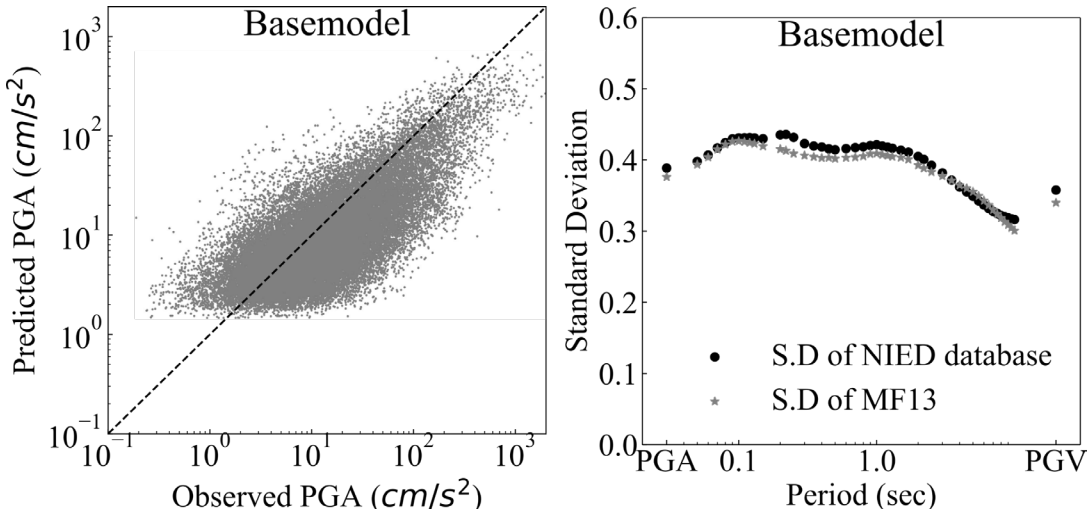

Figure 5-4. Comparison of observed and predicted PGAs (left) and standard deviations calculated from the base model (right). Here, NIED database is the new database used in this study. The standard deviation of MF13 is sourced from Morikawa and Fujiwara (2013).

5.3.2. Leveraging EHVRs as the representative parameters for site effects

Figure 5-5 shows the flow chart of the analysis adopted in this study. The data for the following analyses were selected from the unified ground motion database. The EHVRs were resampled to the same period range as SAs. Predictions based on M_W and *FD* were then calculated using the base model as shown in Eq. (5-1), which considers source and propagation path effects. A DNN model was constructed using the EHVRs and $pre_{basemodel}$ as input data to complement the effect of site effects on the predicted values of ground motions. Additionally, a complementary calculation was performed using the traditional correcting equation by V_{S30} and D_{1400} (MF13). The standard deviations and the goodness of fit of the SAs were then compared to verify the feasibility of the DNN model based on EHVRs.

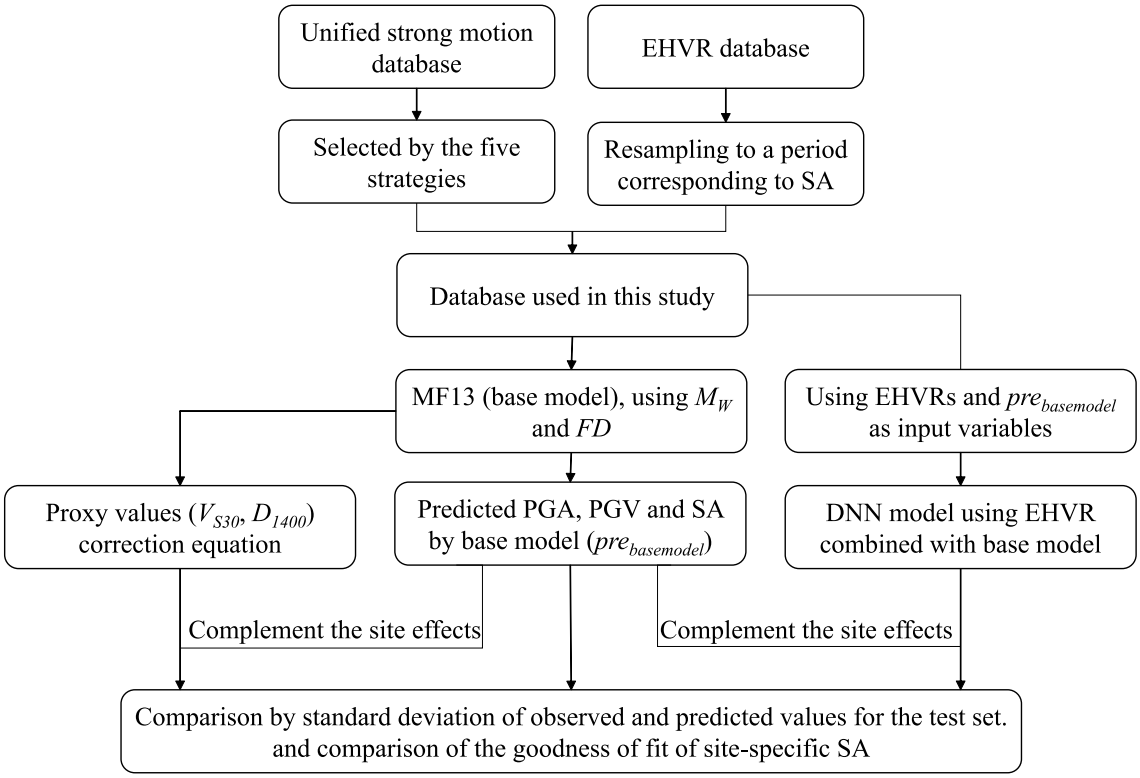

Figure 5-5. Flowchart of analysis adopted in this study.

5.4. Structure of EHVR-based DNN model

5.4.1. DNN model

A DNN model is composed of multiple layers of artificial neurons that are trained using large datasets to recognize patterns and relationships within the data. The DNN model contains many hidden layers, which allows it to learn complex nonlinear relationships between the input and output variables. The DNN model uses a process called backpropagation to adjust the weights of the neurons in each layer, based on the error between the predicted output and the actual output, in order to improve the accuracy of its predictions.

In the proposing method, the input variables consist of two components. One part of the input consists the base prediction values ($pre_{basemodel}$ in log10 logarithms) which were calculated using Eq. (5-1). The physical meaning of the base values represents the predicted values of ground motion considering the source effects and propagation path effects. Then, to account for site-specific effects on ground motion predictions, we used the site-specific EHVRs as another important input variables. The EHVRs were obtained by resampling the data to match the period (1/frequency) range of SA values in our database, spanning from 0.05s to 10 s with a total of 46 points. This selection of input variables aims to exclude proxy values such as M_W and FD that represent source and propagation path effects from being directly used as inputs in the DNN model. Instead, the traditional and reliable GMPE was employed to determine the impact of source and propagation path effects on the final predicted values. The output variables of the DNN model are the correction values, denoted as $Correct_{DNN}$, applied to the base values (in log10 logarithms) to account for site effects. Hence, the final predictions of the proposed method, denoted as pre_{DNN}, were obtained by adding the $Correct_{DNN}$ values to the base values as shown in Eq. (5-3).

$$log(pre_{DNN}) = log(pre_{basemodel}) + Correct_{DNN} \tag{5-3}$$

Then, nonlinear behaviors of ground motions can be found in the observed records of the large and near-source events in the database. Nonlinear site effect can be typically represented as increase of predominant period and change of site amplification compared with linear site effect in frequency domain due to decrease of shear modulus and increase of damping factor in the ground induced by strong shaking. Although such nonlinear effects were not explicitly considered in our DNN model shown in Eq. (5-3), both the training and test datasets include ground motion records influenced by nonlinear effects. This is because the input data of the developed model includes the $pre_{basemodel}$ from the base prediction equation, as well as EHVRs strongly related to site amplifications. It indicates that our DNN model can implicitly evaluate the nonlinear effects of strong ground motions obtained in the database by adjusting the EHVRs according to the intensities in the $pre_{basemodel}$.

The architecture of the DNN model was developed by trial-and-error approach considering the accuracies of the predicted values in preliminary analyses. We employed automated hyperparameter tuning techniques to acquire the optimal model architecture and hyperparameter settings. Batch size and Max epoch were set initially by adjusting them several times at the beginning of model construction, observing whether the loss function curve was steadily decreasing. Detailed hyperparameters for DNN construction, such as intermediate layer architecture, the number of neurons, activation function types, etc., were obtained through automatic hyperparameter tuning. By automatically changing a series of hyperparameter settings for DNN construction, we obtained the training error for each training session, as shown in Figure 5-6. The horizontal axis CostMultiplyAdd is an indicator used to represent the number of multiplication and addition operations (Multiply and Add) included in the entire edited neural network. CostMultiplyAdd can serve as a reference indicator for the computational complexity of the entire network. Computational complexity is a measure of the network's computational intensity. If the computational complexity is large, it indicates that the network may be handling a more

complex, computationally intensive task. However, this also comes with an increase in computational resource requirements. When CostMultiplyAdd increases, it indicates an increase in the computational complexity of the network, suggesting that more computational resources may be needed to execute the model. It can be observed that the architecture of the model used in this study is the one with the minimum training error, and even with further increases in computational complexity, it is challenging to further reduce the training error.

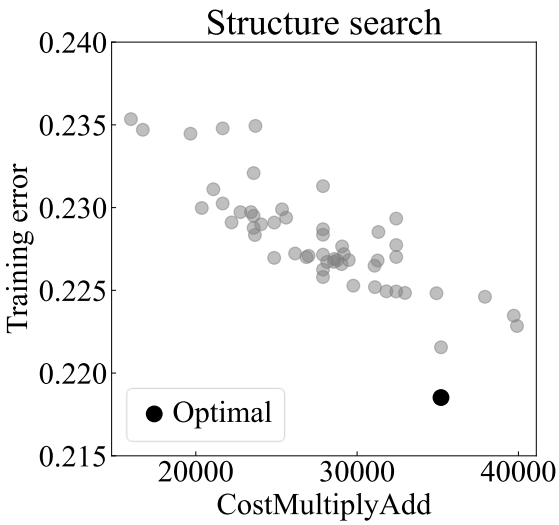

Figure 5-6. Structure search and training error.

Figure 5-7 shows the DNN model adopted in this study. The constructed DNN model consists of an input layer, hidden layers, and an output layer. The hidden layer follows the form proposed by Ioffe and Szegedy [83], such as Affine, Batch Normalization and ReLU (Rectified Linear Unit) layers. The Affine layer refers to a layer that performs an affine transformation on the input, with its output obtained through matrix multiplication and bias offset of the input. The Affine layer is commonly used in fully connected layers of neural networks, with the aim of predicting data by performing linear transformations and offsets on the input. In this layer, the weights are trainable parameters that can be learned

through backpropagation algorithm to minimize the loss function. The advantage of batch normalization is that it can effectively improve the accuracy of the model and increase the learning speed. ReLU was used as the activation function. The ReLU layer is a commonly used nonlinear activation function layer, with its output being the input value itself when the input value is non-negative, and zero when the input value is negative. The role of the ReLU layer is to introduce nonlinearity into the neural network, enabling the network to learn more complex and abstract feature representations. Compared with other activation functions, the ReLU layer has the advantages of fast computation speed and less gradient vanishing problem during training, and thus has been widely used in deep learning. Mean Absolute Error (MAE) was used as the loss function, which is the sum of the absolute values of the deviation between the target values.

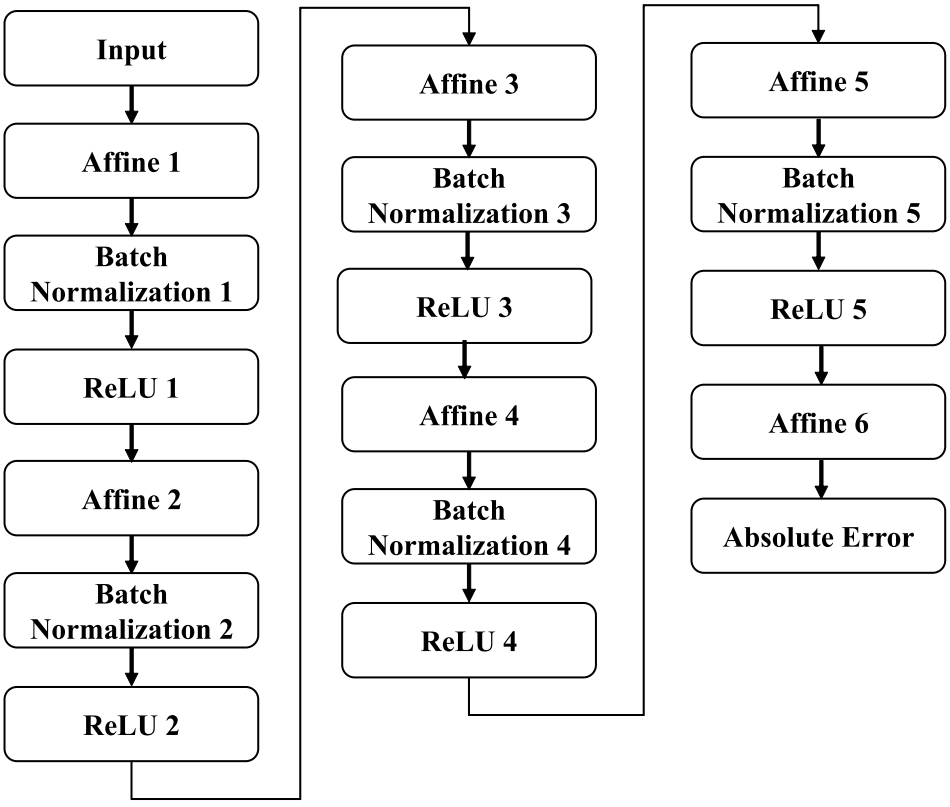

Figure 5-7. DNN architecture adopted in this study.

Considering that the influence of the site effects is different when predicting strong motion parameters, we constructed two models for predicting PG (PGA and PGV) and SA, respectively, with the number of neurons in each model shown in Table 5-2.

Table 5-2. Number of neurons in the DNN model

Layer	Number of neurons
Input	48 for PG / 92 for SA
Affine 1 + BN 1+ ReLU 1	128
Affine 2 + BN 2+ ReLU 2	64
Affine 3 + BN 3+ ReLU 3	64
Affine 4 + BN 4+ ReLU 4	64
Affine 5 + BN 5+ ReLU 5	64
Affine 6	2 for PG / 46 for SA

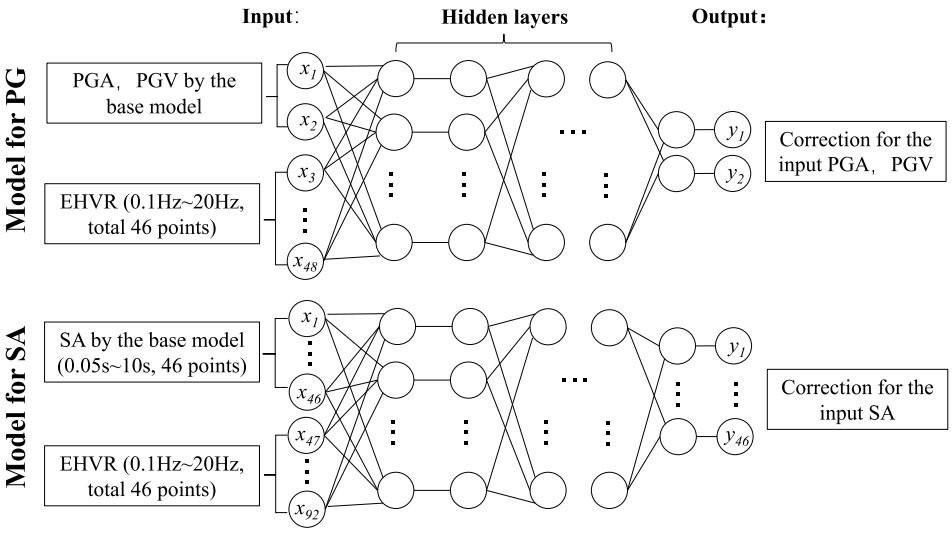

Figure 5-8. Input and output patterns of the proposal models for PG (upper) and SA (lower), respectively. The outputs are the values $Correct_{DNN}$ in Eq. (5-3).

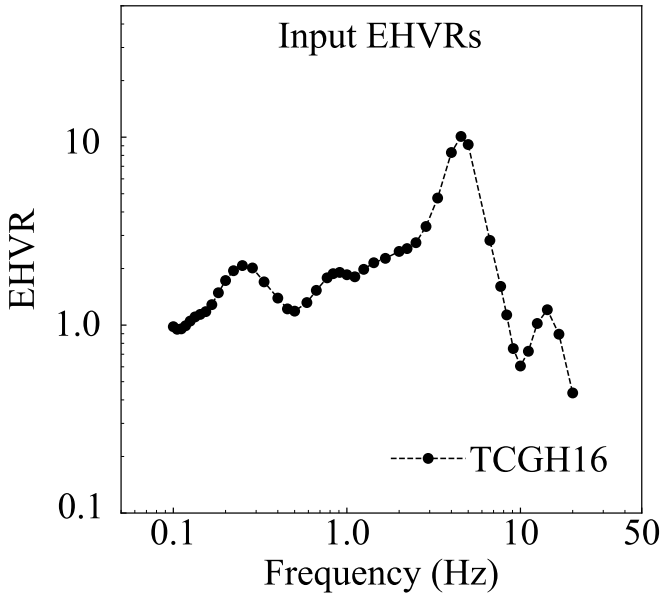

Figure 5-9. Examples of EHVRs used for input, totaling 46 points.

The input and output patterns correspond to Figure 5-8. Regarding as the DNN model for PG, x_1 and x_2 are the predicted PGA and PGV of the base model, respectively, and x_3 to x_{48} this represents the EHVRs of the site effects (0.1 Hz~20 Hz, 46 points). The output y_1 and y_2 are the corrections for the input PGA and PGV ($Correct_{DNN}$ in Eq. (5-3)), respectively. Regarding as the model for SAs, x_1 to x_{46} are the predicted SAs for a total of 46 points from 0.05s to 10s. The values of x_{47} to x_{92} are the EHVRs. The outputs are the corrections of the corresponding SA ($Correct_{DNN}$ in Eq. (5-3)). Figure 5-9 shows an example of the EHVRs for TCGH16 as inputs.

5.4.2. Dataset division and loss functions

As shown in Figure 5-10, the target data were divided into a training set and a test set by record year: 29,861 records from 1997 to 2013 (80.9% of the total data) were used as the training set and 7,053 records from 2014 to 2018 (19.1% of the total data) were

used as the test set. Figure 5-11 depicts the distribution of the motion records in the training and test sets, presenting histograms illustrating the distribution of source depth and M_W. Notably, there is no significant imbalance observed between the training and test sets in terms of these parameters. Batch size, which refers to the number of data samples used in each iteration of a machine learning algorithm during the training process, was set to 2,048. The number of training epochs was set to 500 and the learning rate to 0.0001. Figure 5-12 depicts the development of the loss function over the course of the training process for both the training and test sets. The loss functions for both sets exhibit good convergence to a stable value, suggesting that the learning process was successful without encountering underfitting or overfitting.

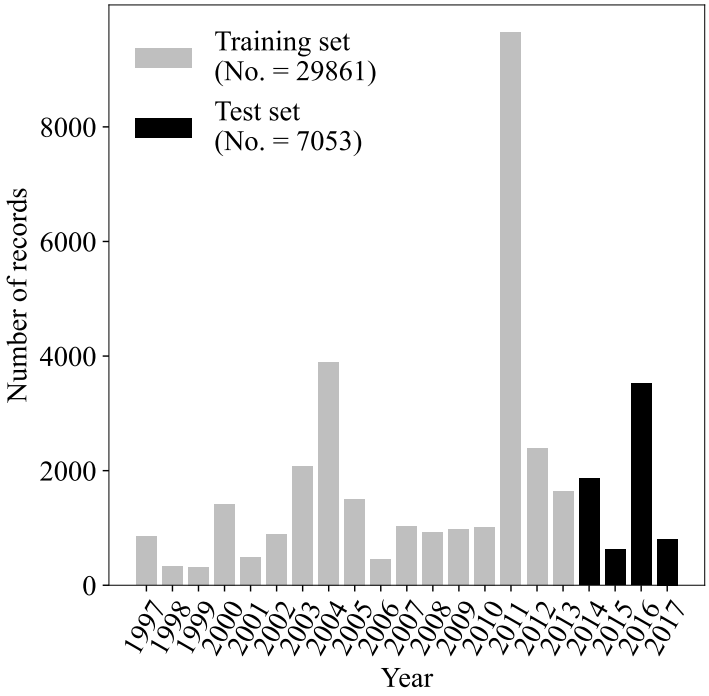

Figure 5-10. Training and test sets divided by year of earthquake events.

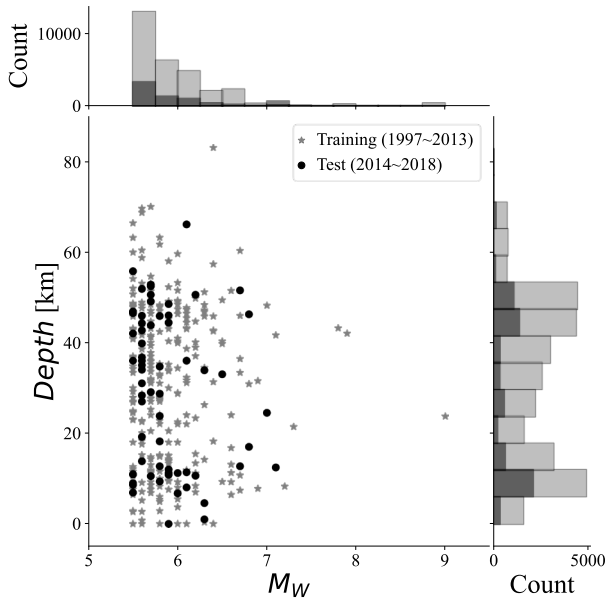

Figure 5-11. Distribution of the number of ground motion records in the training and test sets, with histograms of source depth and M_W showing no significant imbalance between the training and test sets.

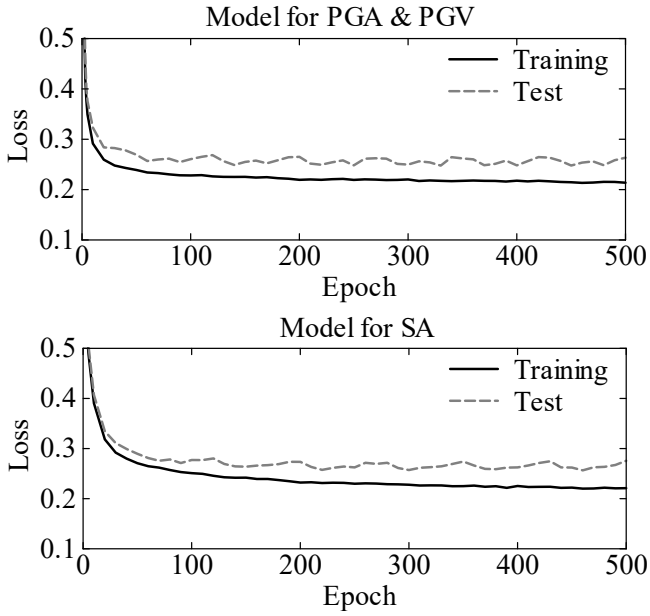

Figure 5-12. Loss function for the training and test sets. Upper figure shows the model for predicting PGA and PGV, and lower figure shows the model for predicting SA.

5.5. Results of test set

5.5.1. Predicted results of PGAs and PGVs for the test set

For the DNN model predicting PGAs and PGVs, the input variables consist of two components. Firstly, the base prediction values ($pre_{basemodel}$) for PGAs and PGVs, which were calculated using Eq. **(5-1)**. And we used the site-specific EHVRs as another important input variables. Hence, the final predictions pre_{DNN}, were obtained by adding the $Correct_{DNN}$ values to the base values as shown in Eq. **(5-3)**.

Subsequently, we applied the trained DNN model to all 7,053 records in the test set to derive the corresponding complementary values for PGAs and PGVs. The predicted values for both were then obtained separately using Eq. (3). A comparison of the predicted values (*PrePGA* (DNN), *PrePGV* (DNN)) with the observed values of the test set is shown in Figure 5-13. Additionally, we calculated the standard deviations of the residuals for both PGAs and PGVs. The standard deviation for PGAs was 0.339 and for PGVs was 0.314. In comparison, the standard deviation of the base model was 0.403 for PGAs and 0.362 for PGVs, indicating an improvement in the predicted accuracy when utilizing the DNN model. Here, the standard deviations of the base model were also the results from the same test set. Furthermore, Figure 5-13 also illustrates the normal distribution of the residuals between the observed and predicted values. The probability density function curves of the residual distributions reveal a notable tendency for the predicted values of the base model to underestimate the observed values, as no site amplification model was used for correction here. However, when corrected using the DNN model, this underestimation tendency is effectively reduced.

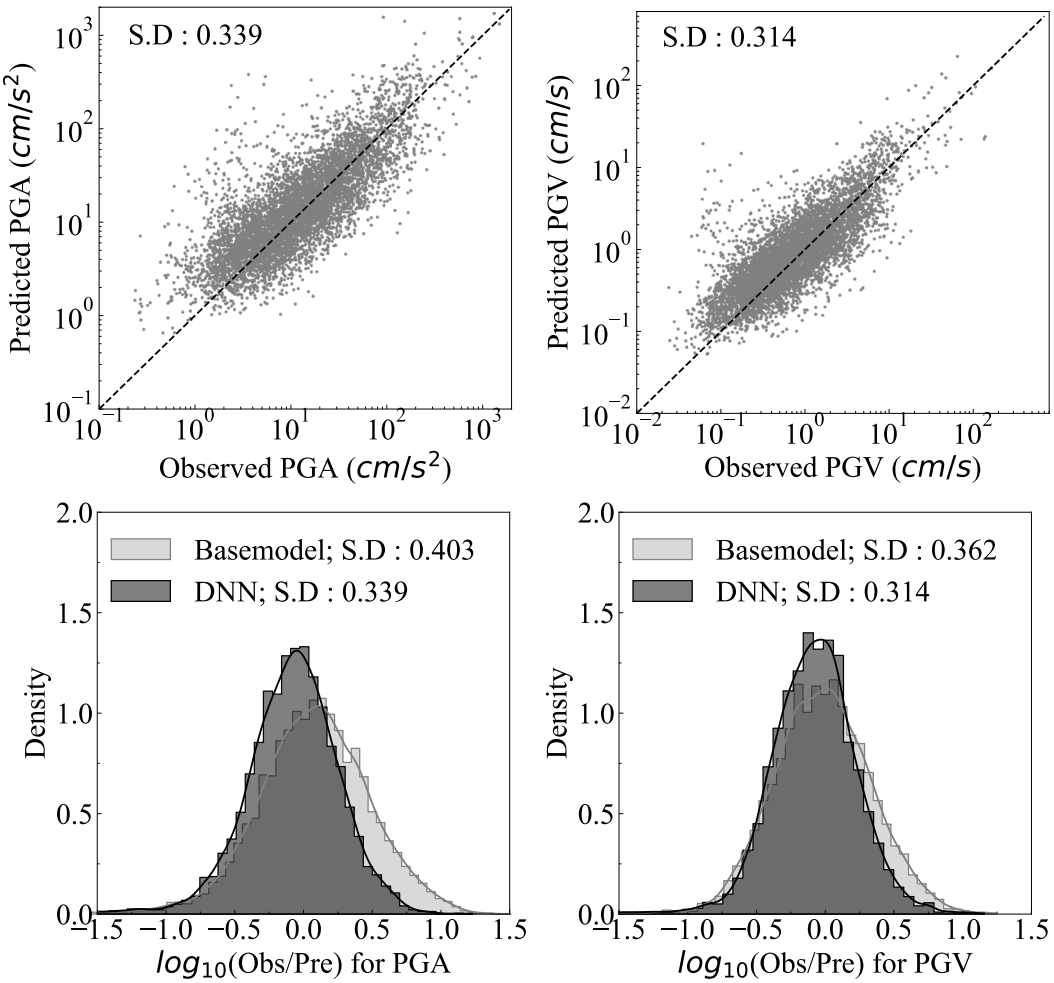

Figure 5-13. Comparison of observed and predicted values for test set. And the residual distribution of the predicted values based on the base model and the DNN model, respectively.

5.5.2. Predicted results of SAs for the test set

For the model predicting SAs, the input variables consisted of the base prediction values $pre_{basemodel}$ (SAs for 46 points from 0.05 to 10 s) and the site-specific EHVRs. Each input contains 92 values, and the output consisted of 46 corrected values for the base values. Figure 5-14 illustrates the comparison between the predicted and observed values for SA at specific periods (0.05 s, 1.0 s, and 8.0 s) in the test set along with the normal

distribution of the residuals. Figures for the all 46 points are provided in Appendix. As depicted in Figure 5-14, the DNN model exhibited improved prediction accuracy compared to the base model across all periods. Particularly in the short-period range, the standard deviation results indicated the significant effectiveness of the corrections made by the DNN model. The base model shows a trend of decreasing standard deviation in the long-period components (around 8.0 s) compared to the short-period components (0.05 s and 1.0 s), Even so, the DNN model still made a corrective effect.

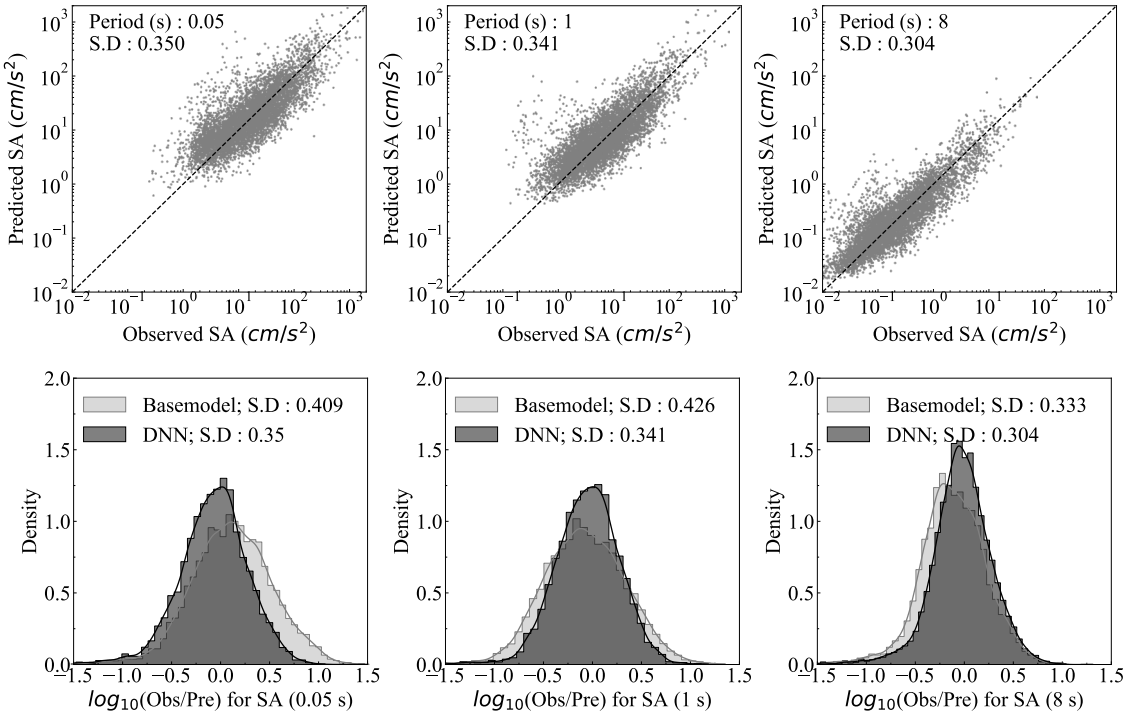

Figure 5-14. Predicted results of SAs in 0.05 s, 1.0 s and 8.0 s for the test set.

Then, we explored the effects of five different ranges of EHVRs as inputs on the results. For a detailed view of the complete input combinations, please refer to Figure 5-8 in the main text:

- Case 1: Using EHVRs from 0.1 Hz to 0.25 Hz as inputs.
- Case 2: Using EHVRs from 0.25 Hz to 1.0 Hz as inputs.

- Case 3: Using EHVRs from 1.0 Hz to 5.0 Hz as inputs.
- Case 4: Using EHVRs from 5.0 Hz to 20.0 Hz as inputs.
- Case 5: Using EHVRs from 0.1 Hz to 30.0 Hz as inputs.

These cases were evaluated against a consistent DNN model structure and a shared test set in the main text. The analysis, detailed in Figure 5-16, highlighted notable impacts of different EHVR ranges on the results. For instance, Case 1 influenced longer periods of SAs because of the input of low-frequency EHVRs. While Case 4, which uses EHVRs from 5.0 Hz to 20.0 Hz, notably improved accuracy for shorter-period SAs. The results suggest that the choice of EHVR ranges influences the accuracy of corresponding period numbers in the output SA. Notably, the superior characteristics of the proposed DNN model in the main text may be attributed to its holistic coverage of the entire spectral range within the input EHVRs, aligning well with predicted SA outputs. Even we extended the length of EHVRs up to 30.0 Hz in Case 5, there was no significant change in the results. Therefore, we selected the input combination shown in Figure 5-8 of the main text as the final solution.

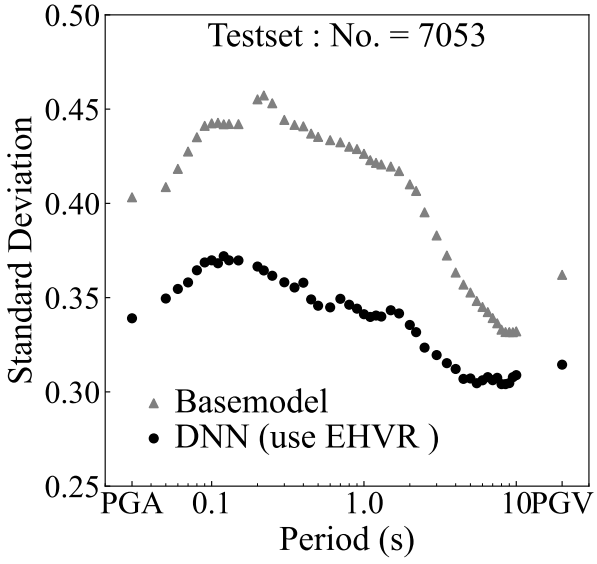

Figure 5-15. Comparison of the standard deviation of the residuals of the DNN models with those of the base model, including PGA, PGV and SA.

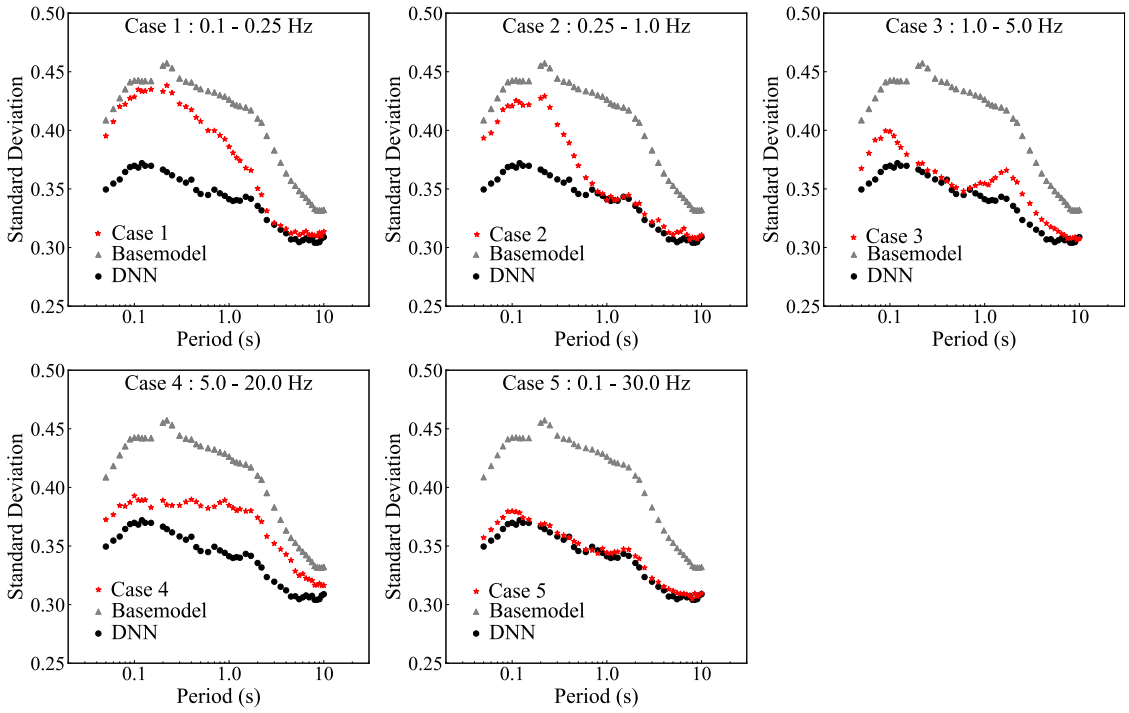

Figure 5-16. Comparison of the standard deviation of the residuals of the DNN model with those of Case 1 to 5.

5.6. Comparison

5.6.1. Comparison of the results based on GMPE

The GMPE proposed in MF13 introduced a complementary formulation (Eq. (5-4)) to account for site effects based on proxy values, namely D_{1400} and V_{S30}.

$$log(pre_{proxy}) = p_d \cdot log\left(\frac{max(D_{lmin}, D_{1400})}{D_0}\right) + p_s \cdot log\left(\frac{min(V_{Smax}, V_{S30})}{V_0}\right) \quad (5\text{-}4)$$

Here, p_d, p_s, D_0 and V_0 represent regression coefficients and D_{lmin}, V_{Smax} are constant values. In accordance with Eq. (5-4), the parameter V_{S30} was a prerequisite for the analysis. However, in cases where V_{S30} was unavailable such as most of the K-NET sites, the estimated V_{S30} values derived from Eq. (5-5) by MF13 were utilized. Among the records included in the test set, a total of 4,807 instances had either V_{S20} or V_{S30} available.

$$V_{S30} = 1.13V_{S20} + 19.5 \quad (5\text{-}5)$$

To further evaluate the efficacy of the EHVR-based DNN models, ground motion predictions incorporating site effects were calculated based on proxy values using Eq. (5-6).

$$log(pre_G) = log(pre_{basemodel}) + log(pre_{proxy}) \quad (5\text{-}6)$$

The comparison results with the DNN models are shown in Figure 5-17, which includes the standard deviation of the residuals for PGAs, PGVs, and SAs as well as the normal distribution. As depicted in the figure, the EHVR-based DNN models

demonstrates superior correction capabilities for short-period components compared to the Eq. (5-4), Eq. (5-5), Eq. (5-6) based on proxy values (D_{1400} and V_{S30}). Morikawa and Fujiwara [90] have pointed out the limitations and ineffectiveness of proxy-based site effect correction for short periods (shorter than 0.3 s). The concentration of the correction effect from the proxy values primarily occurs around the 1.0 s. This concentration arises from the inherent limitations of the site proxies in extracting the full-period characteristics, resulting in a more pronounced correction within the sensitive period range (i.e., around 1.0 s). For the long-period components, both the EHVR-based and proxy-based correction approaches yield comparable standard residuals. Furthermore, the standard deviation of the long-period components obtained using the base model is relatively smaller compared to the short-period components. This may be attributed to the dominant influence of source and propagation path effects on the long-period SAs, with the site effects playing a relatively smaller role. On the other hands, the SA values of the long-period component (8.0 s in Figure 5-14) exhibit significantly smaller amplitudes, thereby emphasizing the effect of the correction effect by the EHVR-based DNN models on the short-period component characterized by larger SA values.

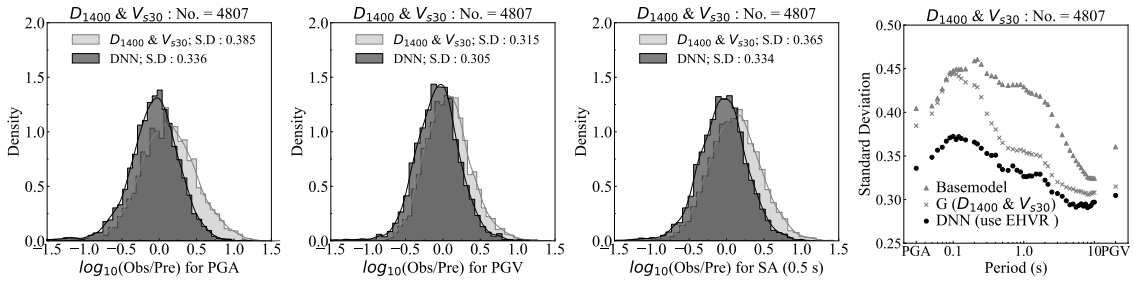

Figure 5-17. Comparison of the results based on the complementary formulation using D_{1400} and V_{S30} with those of the EHVR-based DNN models.

In addition, we generated box plots to visualize the distribution of errors between the true and predicted values at various periods. As shown in Figure 5-18, this figure includes the distribution results of prediction errors for PGA, PGV, and several periods.

The detailed explanation of the box plot is provided on the right side of the figure. The height of the box represents the degree of dispersion, with larger heights indicating greater dispersion. From Figure 5-18, it can be observed that the black boxes representing the DNN models exhibit noticeable trend of reducing dispersion compared to the other two groups. Moreover, the median values of the black box consistently tend towards a value closer to zero. The results validate the effectiveness of the DNN models in reducing the dispersion of prediction errors and confirm that EHVRs can serve as proxies for site effects and play a role in ground motion prediction.

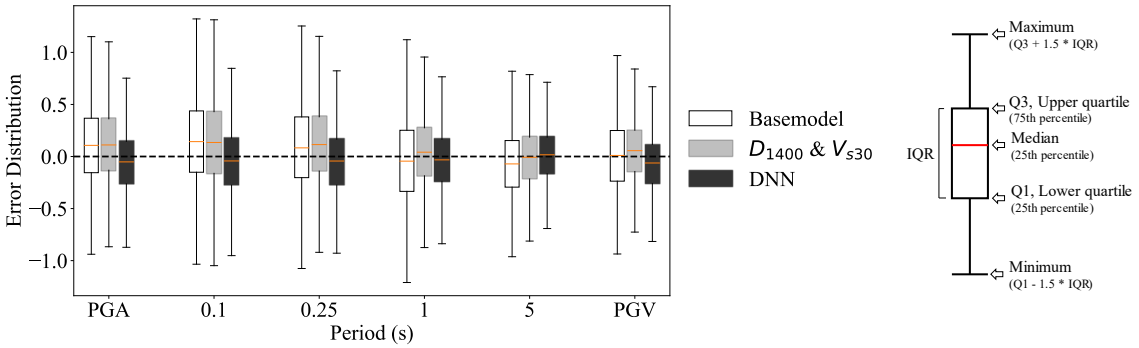

Figure 5-18. Comparison of the error distribution.

5.6.2. Comparison of the site-specific SA

In addition to analyzing the standard deviations of SA values for individual period components, it is crucial to consider the predicted SA for the entire period of seismic records at each site, particularly focusing on the peak values and spectral shapes. Accurate estimations of peaks and shapes are essential for assessing the seismic safety of buildings, especially in the presence of potentially strong shakings. To evaluate the accuracy of SA prediction for the entire period, we conducted a comparison between the observed SAs of the Kumamoto earthquake (M_W 7.1, occurred on April 16, 2016) and the prediction obtained from both the DNN model, the base model and the proxy-based correction.

Figure 5-19 presents the comparison between the observed and predicted SAs for

the Kumamoto earthquake in and around Kumamoto prefecture. The base values (represented by the grey dotted line in Figure 5-19) were calculated using Eq. **(5-1)**. The corrected values for the base values were determined using the EHVR-based DNN model and subsequently incorporated using Eq. **(5-3)** to obtain the final prediction (depicted by the red solid line in Figure 5-19). Eq. (5-4) was employed to obtain predictions based on the proxy values (shown by the grey dashed line in Figure 5-19). It is worth noting that none of the records used in this comparison were involved in the training process of the DNN model. Instrumental seismic intensities of 5.0 or greater were observed at all sites mentioned in the figure.

As shown in Figure 5-19, the comparisons of the observed and predicted SA values

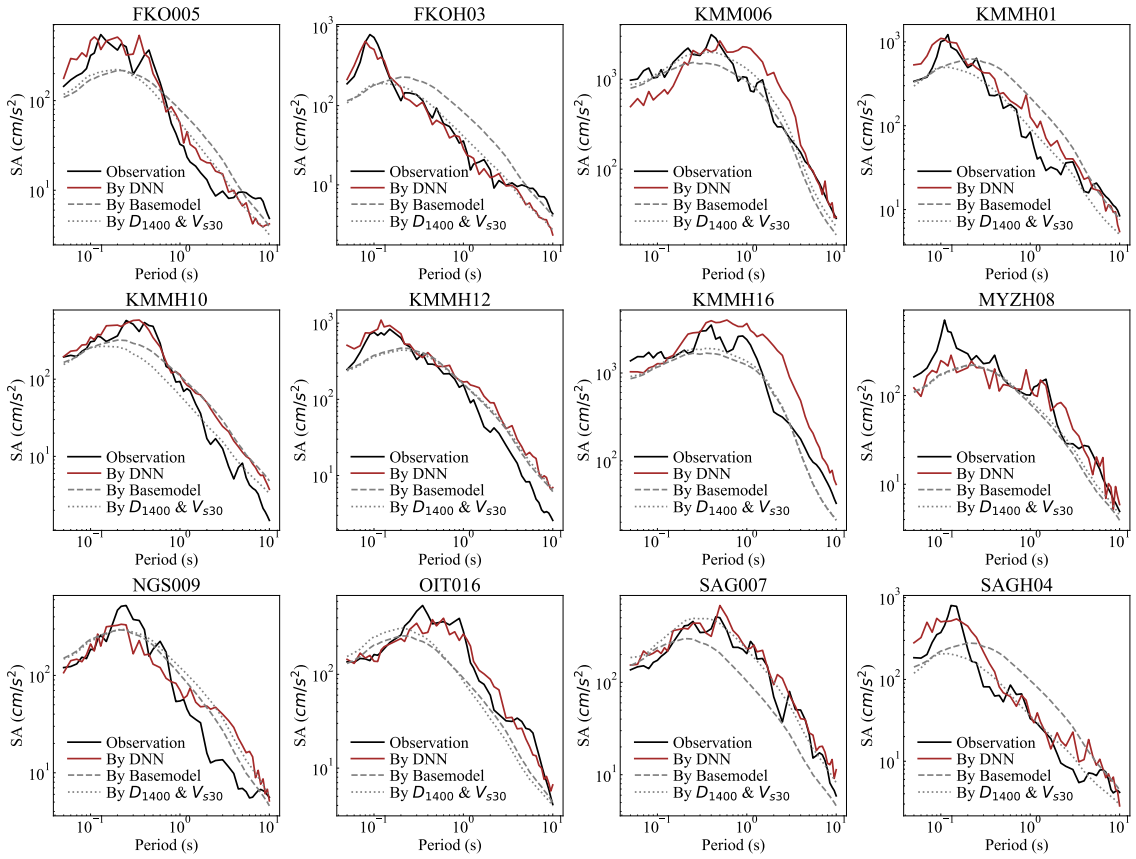

Figure 5-19. Comparisons of observed and predicted SAs observed in the April 16, 2016 Kumamoto earthquake (M_W 7.1). Here, FKO, KMM, MYZ, NGS, OIT and SAG represent Fukuoka, Kumamoto, Miyazaki, Nagasaki, Oita, and Saga prefecture, Japan, respectively.

reveal the improved performance of the corrected predictions obtained from the EHVR-based DNN model. Notably, both the peak SA values and the overall shape of the spectral response exhibit a significantly better fit compared to the base model and the correction formulation based on the site proxies. This finding highlights the effectiveness of utilizing EHVR as an input variable in the DNN model to accurately extract the full-period spectral shape of SAs. Because by incorporating EHVR as the input variables, the model effectively provides complementary corrections to ground motion predictions that consider site effects.

In addition, as depicted in Figure 5-19, the SA values associated with the long-period components (>2.0 s) are relatively small. The predictions obtained from the base model, the proxy-based correction formulation, and the DNN model demonstrate a considerable degree of consistency. On the short-period components, in the cases of FKO005, KMMH12 and SAGH04, the predictions by the DNN model exhibits a favorable alignment with the observations for the short- period components (particularly shorter than 1.0 s), and notable discrepancies emerge with the predictions by the base model or the proxy-based correction formulation.

5.6.3. Good matching for the predicted SAs across all sites

Good matching often implies that the predicted values closely align with the observed values. In this section, two evaluation metrics were employed to assess the quality of the predicted SAs by the models and determine the degree of good matching. The coefficient of determination, denoted as R-squared (R^2), was computed for each record across all sites by comparing the predictions with the corresponding observations using Eq. (5-7).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(obs_i - pre_i)^2}{\sum_{i=1}^{n}(obs_i - \overline{obs_i})^2} \tag{5-7}$$

Here n represents the number of periodic terms, which is equal to 46. Upper bar

indicates a mean value of observed values. Higher R^2 values indicate a stronger alignment between the predicted and observed spectral amplitude, indicating a better match in terms of spectral shape.

Additionally, the Root Mean Square Error (*RMSE*) for each record across all sites was calculated to quantify the mean absolute difference between the predicted and observed amplitude values using Eq. (5-8).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(obs_i - pre_i)^2} \tag{5-8}$$

A smaller *RMSE* indicates higher accuracy in extracting the amplitude variations. To achieve a comprehensive and accurate evaluation, it is recommended to utilize both evaluation metrics; one for assessing the spectral shape and the other for evaluating the amplitude differences, thus providing a more comprehensive assessment of the performance of the model. For the test set, evaluation metrics were calculated for a total of 7,053 site-specific SAs. Among these, a total of 4,807 items were available for the proxy-based correction formulation.

Table 5-3 presents the percentages of the predicted SAs that satisfied the evaluation metrics. These percentages represent good matching scores. By the threshold for reducing the *RMSE* and increasing the R^2, the comparison of good matches among the three methods is presented in the table. The DNN model consistently demonstrated superior performance compared to both the base model and the proxy-values correction formulation. Additionally, the utilization of the proxy-values correction also exhibited slightly better results than the base model. In particular, the high prediction accuracy percentage on the right-hand side of the table indicates that the DNN model achieves a good match for approximately half of the sites, providing further evidence of the

effectiveness of EHVR in predicting SAs.

Table 5-3. Good matching for the SAs of each record. Here No. is number of records.

V_{S30}/V_{S20} is known: No. = 4,807				
Good matching	$RMSE \leq 0.6$ & $R^2 \geq 0.6$	$RMSE \leq 0.5$ & $R^2 \geq 0.7$	$RMSE \leq 0.4$ & $R^2 \geq 0.8$	$RMSE \leq 0.3$ & $R^2 \geq 0.9$
By DNN	85.56 %	79.09 %	69.38 %	48.41 %
By Base model	78.14 %	67.05 %	49.53 %	22.40 %
By D_{1400} and V_{S30}	83.73 %	74.62 %	58.19 %	29.00 %

5.6.4. Assessment of nonlinear effects of strong ground motion

It is important to evaluate nonlinear behaviors of ground motions during strong shakings for accurate ground motion predictions. As mentioned earlier, terms for nonlinear effects were not explicitly formulated in our proposed model. However, our model learned the characteristics of observed ground records not only for weak motions but also for strong motions affected by nonlinear effects. It means that such nonlinear effects during strong shakings were expected to be implicitly included in the model. To verify this, we analyzed the predicted SAs at KMMH16 where significant nonlinear ground response was observed in the 2016 Kumamoto, Japan earthquake [119].

Figure 5-20 (a) shows the comparisons of the observed and predicted SAs at KMMH16 for several earthquake events occurred after 2014, indicating that these observed data were not included in the training dataset. We can confirm that the predicted SAs by the DNN show good agreement with the observed SAs not only for weaker motions in M_W 5.5 and 5.8 events but also strong motions in M_W 7.1 event, the mainshock of the 2016 Kumamoto earthquake. Figure 5-20 (b) represents the predicted SAs at KMMH16 for eight different scale events. It can be seen that as the amplitudes of SAs

increase, the peak period shift to the longer period side as shown by gray dotted line. Such amplitude-dependent transition of peak periods in the predicted SAs could be caused by the nonlinear effect trained from the database. This result proved that the proposed DNN model can capture the long-period phenomena during strong shakings.

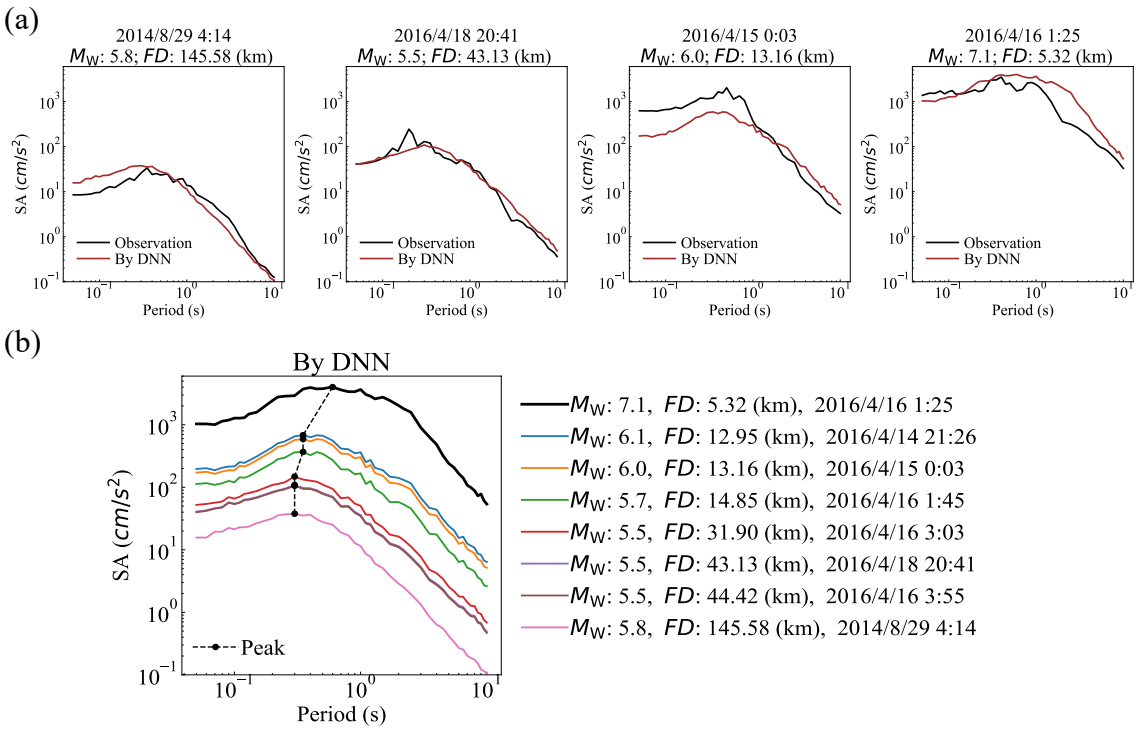

Figure 5-20. (a) Comparison of predicted and observed and predicted SAs at KMMH16 for four different earthquake events. M_W 7.1 event indicate the mainshock of the 2016 Kumamoto, Japan earthquake. (b) Predicted SAs at KMMH16 for different events. Gray dotted line indicates transition of peak periods in the predicted SAs.

5.7. Using predicted EHVRs from MHVRs as the representative parameters

5.7.1. Using predicted EHVRs as input of the proposed DNN model

In Chapter 4, a model is proposed to predict pseudo-EHVR (pEHVR) using MHVR. It is worth exploring whether the pseudo-EHVR can be used in the GMM model proposed in this chapter. For this reason, this section will discuss the accuracy and validity of using pEHVR instead of EHVR in this GMM model.

In terms of data selection, the following two strategies need to be satisfied firstly.

1. Sites are those in Testset of Chapter 4, Section 4.4.2
2. Ground motion records are those during the years 2014-2017, corresponding to Figure 5-10.

Based on these two points, a total of 282 data from 38 sites were used for subsequent analyses.

Subsequently, the trained GMM model was used to replace the EHVR in it using pEHVR according to the input pattern in Figure 5-8. And the final results were calculated according to Eq. **(5-3)**. All the predicted results containing the predicted values of PGA, PGV, and SA were obtained.

5.7.2. Comparison of results

We calculated the mean and distribution of the errors and plotted box plots to visualize the distribution of the errors between the true and predicted values for different periods. Figure 5-21 show the results of the base model, the DNN model using EHVR, and the DNN model using pEHVR.

From the comparison of the three models, it can be observed that the mean errors of SA for most periods by using pEHVR are lower than those of the base model, although not as low as the DNN model using EHVR. On the other hand, for PGA and PGV, even

though the mean errors are lower than those of the base model, the deviation in their distribution is greater. This is a critical point that needs attention. In practical applications, caution should be exercised when predicting PGA and PGV. One possible reason for this deviation is that certain sites exhibit larger errors in the prediction of pEHVR. When such pEHVR values are used as inputs, the errors in the resulting ground motion parameters may also be significant.

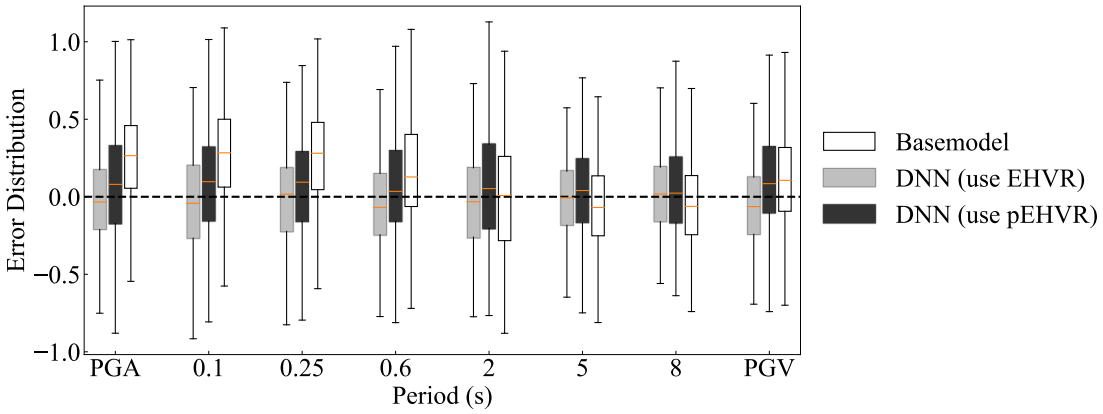

Figure 5-21. Distribution of errors between predicted and true values for the DNN model using EHVR, the DNN model using pEHVR and the basemodel.

In addition, the GOF between the true and predicted values of SA at each site was calculated using the good matching evaluation method outlined in Section 5.6.3. The distributions of R^2 and *RMSE* for the three models are shown in Figure 5-22. The R^2 and *RMSE* values for the DNN model using EHVR show the highest distribution, indicating that this model explains the variance in the true SA values most effectively. The DNN model using pEHVR also demonstrates a higher R^2 and *RMSE* distribution compared to the base model, though it does not reach the level of the DNN model using EHVR. The comparison of R^2 and *RMSE* distributions clearly demonstrates that while the DNN model using pEHVR also improves prediction accuracy over the base model, it does not match the performance of the model using EHVR. In practical terms, while the pEHVR-based

model offers a viable alternative when EHVR data is not available, the best results are achieved when EHVR data is used directly.

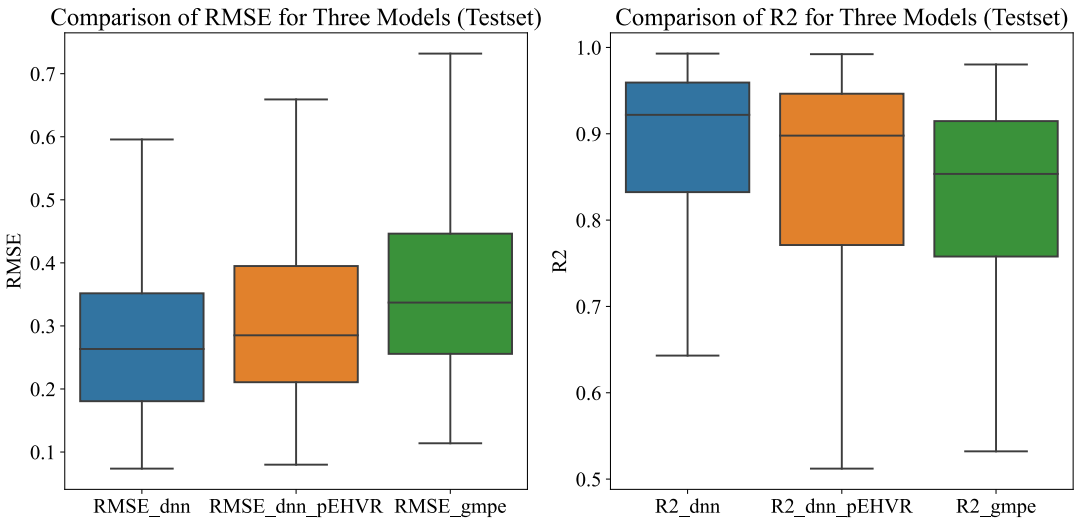

Figure 5-22. Distribution of R^2 and *RMSE* for SA predictions using different models.

As shown in Figure 5-23, the results indicate that for most sites, using pEHVR as the input for the DNN model produces an output SA that fits very closely with both the observed values and the results when using EHVR. However, there are still some sites where the fit is poor, particularly in the short-period range, such as IWTH15. The correction to the base model is not significant in these cases. Additionally, the good match rate for a total of 281 records in the test set was calculated, and the results are shown in Table 5-4. For the highest accuracy threshold, the DNN model using pEHVR, although not as accurate as the DNN model using EHVR directly, shows a significant improvement compared to the base model. This finding highlights the effectiveness of using pEHVR.

Future research should focus more on improving the prediction accuracy of pEHVR in the short-period range. Additionally, it would be beneficial to discuss more about the types of sites where the accuracy of pEHVR predictions is lower.

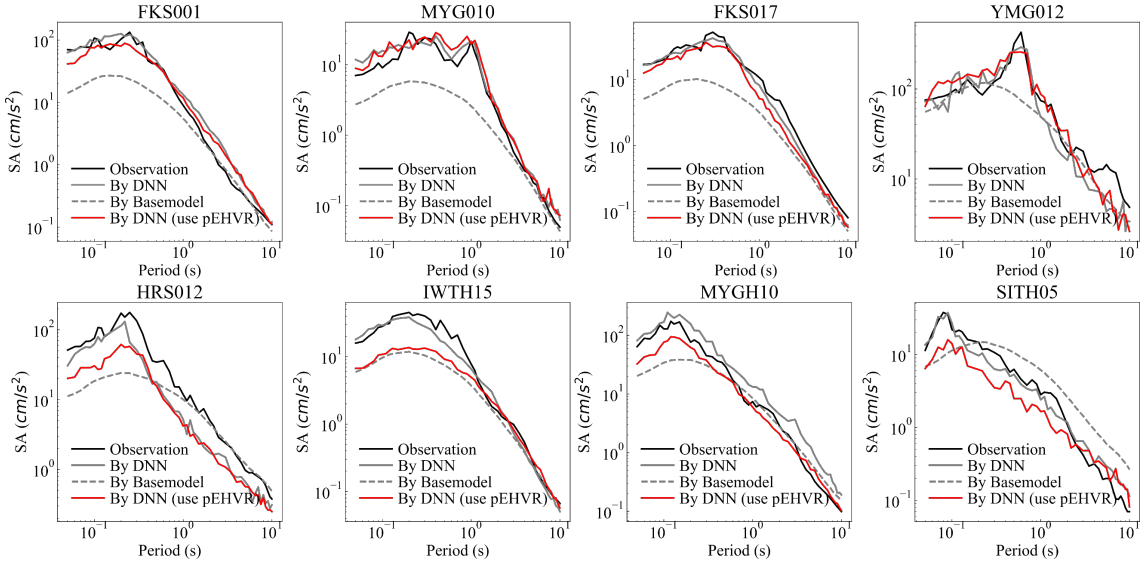

Figure 5-23. Comparison of SA results by DNN model using EHVR, pEHVR and base model.

Table 5-4. Good match rates for a total of 281 records in the test set, using pEHVR as input to the DNN model.

Good matching	$RMSE \leq 0.6$ & $R^2 \geq 0.6$	$RMSE \leq 0.5$ & $R^2 \geq 0.7$	$RMSE \leq 0.4$ & $R^2 \geq 0.8$	$RMSE \leq 0.3$ & $R^2 \geq 0.9$
By DNN	90.39 %	85.41 %	74.02 %	55.87 %
By DNN (pEHVR)	**85.77 %**	**77.58 %**	**69.04 %**	**44.48 %**
By Base model	86.12 %	78.29 %	58.36 %	26.33 %

5.8. Conclusions

This study proposed a deep neural network (DNN) model for novel seismic ground motion prediction based on two powerful databases: a unified strong ground motion database by NIED and an earthquake H/V spectral ratio (EHVR) database. A hybrid approach combining the DNN models and traditional GMPE was employed. The proposed model utilizes EHVRs as input variables and produces correction values for ground motion prediction considering site effects. The comparisons were conducted with the correction formulation that consider site effects using D_{1400} and V_{S30}. The following main findings can be drawn:

1. The DNN models constructed using EHVRs outperforms the traditional GMPE in predicting PGAs, PGVs, and SAs. It also outperforms the site proxy-based correction formulation with D_{1400} and V_{S30}.
2. The correction formulation based on D_{1400} and V_{S30} exhibit limited ability in extracting full-period features of site effects, resulting in inadequate correction of SAs particularly for short-period components. In contrast, the DNN models based on full-period EHVRs as input, demonstrates superior performance in extracting full-period features. Predictions for the SAs in each period component were significantly improved. Furthermore, we confirmed that the proposed DNN models can implicitly evaluate nonlinear effects of strong ground motions.
3. The quality and good matching for the predicted SAs were comprehensively assessed using two evaluation metrics. The DNN models exhibits better performance than the base model and the proxy-based formulation. This indicates that the DNN models performs well in predicting the spectral amplitudes and shapes of SAs.
4. The use of the pseudo-EHVR (pEHVR) obtained in Chapter 4 was equally effective in improving the prediction accuracy of the base model, especially for SA. However, there was a large bias on the prediction of PGA, which needs to be treated with caution in practical application.

Nevertheless, it should be noted that there still exists a discrepancy between the

observed and predicted ground motion parameters, particularly with the proposed model showing less significant improvement in prediction accuracy for the long-period components of SA compared to the base model. Furthermore, the availability of EHVR data is restricted to locations where seismic observatories are operated and predictions from the base model (MF13) are required as inputs, thereby limiting the applicability of the model to arbitrary locations or outside the Japan region. However, considering the calculation method of EHVR and the fact that we used a date-based split for the training and test sets in this study, the EHVRs from the same locations appear in both the training and test sets, which may lead to data leakage. Considering that the training and validation sets used in this study are all from Japan, it is important to explore the applicability of this model in non-Japanese regions globally (such as China, Europe) in future research to further validate the model's applicability and to check for the potential data leakage.

Chapter 6. Conclusions

6.1. Conclusions of previous chapters

This thesis has presented a detailed study on the use of Horizontal-to-Vertical Spectral Ratios (HVR) to improve seismic risk assessment and ground motion prediction. The research aimed to develop robust models that leverage both Microtremor HVR (MHVR) and Earthquake HVR (EHVR) for enhanced predictive capabilities.

In Chapter 2, a novel deep neural network (DNN) model was developed to estimate site amplification factors (SAFs) directly from MHVR data. We analyzed the site amplification factor derived by the generalized spectral inversion technique (GIT) and observed MHVRs at K-NET and KiK-net sites in Chugoku district, Japan. The performance of the DNN model was validated using a comprehensive dataset, demonstrating significant improvements over traditional methods. The developed DNN model does not require any hard-to-get data such as seismic velocity structures and damping models thus providing a significant cost-benefit. This established the foundation for using MHVR data in seismic assessments.

Chapter 3 expanded on this foundation by incorporating transfer learning techniques to adapt the pre-trained DNN model for new regions with varying geological conditions. In practical scenarios, it is anticipated that the field of MHVR-estimated SAF can be extended to countries and regions where the number of seismic observatories is inadequate. For regions with limited data or unique geological features, such as some developing countries, constructing an appropriate TL model based on limited data can improve the accuracy of SAFs estimated with MHVRs highlighting the adaptability and potential of the proposed DNN model for broader application.

Chapter 4 introduced a model to predict pseudo-EHVR (pEHVR) from MHVR, addressing the challenge of obtaining reliable EHVR data in regions without direct seismic measurements. The pEHVR predictions showed high accuracy, validating the

model's utility in seismic risk assessments and emphasizing the feasibility of using MHVR data to infer EHVR.

In Chapter 5, a more sophisticated DNN model was proposed to predict ground motion parameters, such as peak ground acceleration (PGA), peak ground velocity (PGV), and 5% damped absolute acceleration response spectra (SA), by integrating EHVR into existing ground motion prediction equations (GMPE). The model demonstrated superior performance in predicting both the amplitude and shape of the spectral response compared to traditional GMPEs that rely on proxy-based site effects like V_{S30}. The use of the pseudo-EHVR (pEHVR) obtained in Chapter 4 was equally effective in improving the prediction accuracy of the base model, especially for SA. However, there was a large bias on the prediction of PGA, which needs to be treated with caution in practical application.

6.2. Future work and prospect

Although the DNN models developed in this study have shown excellent performance, several avenues remain for improving their accuracy, applicability, and robustness in the future. For the models discussed in Chapter 2 and Chapter 3, one key area of improvement is expanding the dataset to include a broader range of geological conditions. This will help verify the generalizability of the models and ensure their applicability in various regional environments.

Additionally, validating the effectiveness of the predicted pEHVR for other applications in seismology, such as site classification, site effect assessment, and earthquake intensity prediction, as discussed in Chapter 4, will provide a more comprehensive evaluation of the model's utility.

Addressing the limitation of EHVR data availability in Chapter 5 is crucial for the wider adoption of these models. Innovative methods to generate reliable proxy data where direct measurements are not possible will be essential. Beyond using pEHVR as a proxy, it is worth exploring whether transfer learning techniques can further expand the model's applicability.

Furthermore, extending the applicability of the models to global regions beyond Japan, such as China and Europe, will be essential to test their robustness and identify any potential issues related to data leakage or regional specificities. This global perspective will enhance the models' relevance and utility in international seismic risk assessment practices. By focusing on these areas, future research can build on the findings of this thesis, incorporating new technologies and exploring a wider range of applications. This will contribute to more accurate and reliable site effect assessments and ground motion predictions.

Chapter 7. Appendix

7.1. Publicly available codes

In this chapter, we will provide the usage methods for the deep neural network (DNN) models developed in this thesis, primarily focusing on utilizing several publicly available pre-trained models. These models will be implemented using Python, with code snippets guiding readers on how to use and evaluate the models on their respective datasets. The primary focus will be on the following DNN models: the DNN model from Chapter 2 for predicting site amplification factors (SAF) from microtremor horizontal-to-vertical spectral ratios (MHVR) and the transfer learning model from Chapter 3 for extending the applicability of the DNN model from Chapter 2; and the DNN model from Chapter 5 for improving ground motion prediction model (GMM) using EHVR. More detailed tutorials and examples can be found on my GitHub page.

(1). The DNN model from Chapter 2 for predicting SAF from MHVR.

https://github.com/pandafhg/MHVR-to-SAF.

And the transfer learning model from Chapter 3 for extending the applicability of the DNN model from Chapter 2.

https://github.com/pandafhg/MHVR-to-SAF-TL.

```
import onnx
import numpy as np
import pandas as pd
import onnxruntime as rt
import os

# Iterate through paths of all files in the folder (fname). And save as a
 list (s_file).
def get_file(fname):
    s_file = []
```

```
    for path,dirs,files in os.walk(fname):
        for filename in files:
            s_file.append(os.path.join(path,filename))
    return s_file

# Retrieve DNN model file (model.onnx) from the current directory.
if os.path.exists("%s/model.onnx" % os.getcwd()):
    model_file = "%s/model.onnx" % os.getcwd()
# If the model file does not exist in the current directory,
# please input the full path of the model.onnx.
else:
    model_file = input("input the file path of model.onnx (/*.onnx):")

# Set the save path. The default is the MHVR_SAF folder in the current di
rectory.
if os.path.exists("%s/MHVR_SAF" % os.getcwd()):
    y_folder = "%s/MHVR_SAF" % os.getcwd()
    print("The result will be saved in this path: %s" % y_folder)
else:
    os.mkdir("%s/MHVR_SAF" % os.getcwd())
    y_folder = "%s/MHVR_SAF" % os.getcwd()
    print("The result will be saved in this path: %s" % y_folder)

# Input the path of folder containing MHVR data files.
mhvr_loc = input("input the path of the folder containing MHVR data files
: ")


for test_file in get_file(mhvr_loc):
    X_test = pd.read_csv(test_file ,skiprows=1, sep = '¥s+|,', header =No
ne, engine='python') # encoding = "utf-8"
    X_test = X_test.values

    sess = rt.InferenceSession(model_file)
    input_name = sess.get_inputs()[0].name
    label_name = sess.get_outputs()[0].name
```

```
43
44      Y_result = pd.DataFrame()
45      for row in X_test:
46          pred_onx = sess.run([label_name], {input_name: row.reshape(1,7).a
stype(np.float32)})[0]
47          Y_result = pd.concat([Y_result,pd.DataFrame(pred_onx)], ignore_in
dex= True)
48
49      Y_result.columns=["AMR(fi-2)", "AMR(fi-
1)", "AMR(fi)", "AMR(fi+1)", "AMR(fi+2)"]
50
51      X_temp = pd.read_csv(test_file, sep = '¥s+|,', engine='python') # enc
oding = "utf-8"
52
53      result = pd.concat([X_temp, Y_result], axis=1)
54      result["pSAF"] = result["%s" % X_temp.columns[4]] * result["AMR(fi)"]
55      result.round(6)
56
57      # Save by source filename.
58      name_file = os.path.basename(test_file)
59
60      if os.path.splitext(name_file)[1] == '.csv':
61          result.to_csv("%s/SAF_%s" % (y_folder, name_file), float_format =
 '%.6f', sep = ',',index=None)
62
63      else:
64          result.to_csv("%s/SAF_%s" % (y_folder, name_file), float_format =
 '%.6f', sep = '¥t',index=None)
65
66      print("Saved as %s/SAF_%s" % (y_folder, name_file))
67  input("   ")
```

(2). The DNN model from Chapter 5 for improving ground motion prediction model (GMM) using EHVR.

https://github.com/pandafhg/Ground-motion-prediction-using-EHVR.

```python
import pandas as pd
import numpy as np
import os
import onnxruntime as rt

# Define a function to load an ONNX model from a specified file path
def load_onnx_model(onnx_file):
    # Get the current directory path
    current_dir = os.getcwd()
    # Create the full path to the ONNX model file
    model_path = os.path.join(current_dir, onnx_file)

    # Check if the model file exists at the path
    if not os.path.exists(model_path):
        # Prompt the user to enter the correct model path if not found
        model_path = input("Please enter the path to the model.onnx: ")

        # Raise an error if the model file still can't be found
        if not os.path.exists(model_path):
            raise FileNotFoundError(f"ONNX model not found at {model_path}")

    # Return the valid model path
    return model_path

# Prompt the user to input the name of the data file
dt_file = input('Input the file:')

# Read the input data file into a DataFrame
dt = pd.read_csv(dt_file)
```

```
# Read another model parameter file into a DataFrame
model1 = pd.read_csv('model1.csv')

# Try to perform calculations on the DataFrame
try:
    # Iterate through each row of the model parameter DataFrame
    for index_shallow, row_model1 in model1.iterrows():
        # Extract parameters needed for the calculation
        pre = row_model1.parameter
        sigma1 = row_model1.sigma1

        # Iterate through each row of the data DataFrame
        for index, row in dt.iterrows():
            # Extract parameters and perform calculations based on model formulas
            a1 = row_model1['a1']
            Mw1_min = min(row['mw'], row_model1['Mw01'])
            Mw1 = row_model1['Mw1']
            eq_type = row['eq_location_type_id']
            b1_k = row_model1[f'b1_{eq_type}']
            c1_k = row_model1[f'c1_{eq_type}']
            X = row['fault_dist']
            d1 = row_model1['d1']
            e1 = row_model1['e1']

            # Store calculated result back into the DataFrame
            dt.loc[index, f'Basemodel_{pre}'] = a1 * ((Mw1_min - Mw1) ** 2) + b1_k * X + c1_k - np.log10(X + d1 * (10 ** (e1 * Mw1_min)))

# Catch any exceptions that occur during the process
except Exception as e:
    print("Error reading specific columns:", e)

# Ensure the output directory exists
output_folder = os.path.join(os.getcwd(), 'result')
os.makedirs(output_folder, exist_ok=True)
```

```
# Define the filename for the ONNX model
onnx_file_pg = ('EHVR_RES_pg_MF2013.onnx')

# Define the path to the input CSV file
input_folder = ("MTE Database/GMM/all/pEHVR_test_pg.csv")

# Load the ONNX model
model_file = load_onnx_model(onnx_file_pg)
sess = rt.InferenceSession(model_file)
# Retrieve the model's input and output node names
input_name = sess.get_inputs()[0].name
output_name = sess.get_outputs()[0].name

# Prepare the test data from the DataFrame
X_test = dt[['Basemodel_PGA','Basemodel_PGV',
      '0.100Hz', '0.105Hz', '0.111Hz', '0.118Hz', '0.125Hz', '0.133Hz',
       '0.143Hz', '0.154Hz', '0.167Hz', '0.182Hz', '0.200Hz', '0.222Hz',
       '0.250Hz', '0.286Hz', '0.333Hz', '0.400Hz', '0.455Hz', '0.500Hz',
       '0.588Hz', '0.667Hz', '0.769Hz', '0.833Hz', '0.909Hz', '1.000Hz',
       '1.111Hz', '1.250Hz', '1.429Hz', '1.667Hz', '2.000Hz', '2.222Hz',
       '2.500Hz', '2.857Hz', '3.333Hz', '4.000Hz', '4.545Hz', '5.000Hz',
       '6.667Hz', '7.692Hz', '8.333Hz', '9.091Hz', '10.000Hz', '11.111Hz'
,
       '12.500Hz', '14.286Hz', '16.667Hz', '20.000Hz']].values

# Create an empty DataFrame to store prediction results
Y_result_PG = pd.DataFrame()

# Iterate through each row of test data and make predictions using the ON
NX model
for row in X_test:
    # Run prediction and reshape the data as required by the model
    pred_onx = sess.run([output_name], {input_name: row.reshape(1, 48).as
type(np.float32)})[0]
    # Append predictions to the result DataFrame
```

```
96      Y_result_PG = pd.concat([Y_result_PG, pd.DataFrame(pred_onx)], ignore
    _index=True)
97
98  # Set the column names for the prediction results
99  Y_result_PG.columns = ['pre_DNN_PGA', 'pre_DNN_PGV']
100
101 # Re-read the complete CSV file to ensure all data is available
102 X_temp = dt
103
104 # Merge the prediction results with the original data
105 result = pd.concat([X_temp, Y_result_PG], axis=1)
106
107 # Specify another model file
108 onnx_file_sa = ('EHVR_RES_sa_MF2013.onnx')
109
110 # Load the second ONNX model
111 model_file = load_onnx_model(onnx_file_sa)
112 sess = rt.InferenceSession(model_file)
113 input_name = sess.get_inputs()[0].name
114 output_name = sess.get_outputs()[0].name
115
116 # Define the test data file
117 test_file = input_folder
118 # Read the CSV file and prepare the test data
119 X_test_sa = dt[['Basemodel_0.05', 'Basemodel_0.06', 'Basemodel_0.07', 'Ba
    semodel_0.08', 'Basemodel_0.09', 'Basemodel_0.1',
120                 'Basemodel_0.11', 'Basemodel_0.12', 'Basemodel_0.13', 'Ba
    semodel_0.15',  'Basemodel_0.2', 'Basemodel_0.22',
121                 'Basemodel_0.25', 'Basemodel_0.3', 'Basemodel_0.35', 'Bas
    emodel_0.4', 'Basemodel_0.45', 'Basemodel_0.5',
122                 'Basemodel_0.6', 'Basemodel_0.7', 'Basemodel_0.8', 'Basem
    odel_0.9', 'Basemodel_1', 'Basemodel_1.1',
123                 'Basemodel_1.2', 'Basemodel_1.3', 'Basemodel_1.5', 'Basem
    odel_1.7', 'Basemodel_2', 'Basemodel_2.2',
124                 'Basemodel_2.5', 'Basemodel_3', 'Basemodel_3.5', 'Basemod
    el_4', 'Basemodel_4.5', 'Basemodel_5',
```

```
                'Basemodel_5.5', 'Basemodel_6', 'Basemodel_6.5', 'Basemodel_7', 'Basemodel_7.5', 'Basemodel_8',
                'Basemodel_8.5', 'Basemodel_9', 'Basemodel_9.5', 'Basemodel_10',

                  '0.100Hz', '0.105Hz', '0.111Hz', '0.118Hz', '0.125Hz', '0.133Hz',
                   '0.143Hz', '0.154Hz', '0.167Hz', '0.182Hz', '0.200Hz', '0.222Hz',
                   '0.250Hz', '0.286Hz', '0.333Hz', '0.400Hz', '0.455Hz', '0.500Hz',
                   '0.588Hz', '0.667Hz', '0.769Hz', '0.833Hz', '0.909Hz', '1.000Hz',
                   '1.111Hz', '1.250Hz', '1.429Hz', '1.667Hz', '2.000Hz', '2.222Hz',
                   '2.500Hz', '2.857Hz', '3.333Hz', '4.000Hz', '4.545Hz', '5.000Hz',
                   '6.667Hz', '7.692Hz', '8.333Hz', '9.091Hz', '10.000Hz', '11.111Hz',
                   '12.500Hz', '14.286Hz', '16.667Hz', '20.000Hz']].values

# Create an empty DataFrame to store the prediction results for spectral accelerations
Y_result_sa = pd.DataFrame()

# Iterate through each row of test data for spectral accelerations and make predictions using the ONNX model
for row in X_test_sa:
    # Run prediction and reshape the data as required by the model
    pred_onx = sess.run([output_name], {input_name: row.reshape(1, 92).astype(np.float32)})[0]
    # Append predictions to the result DataFrame
    Y_result_sa = pd.concat([Y_result_sa, pd.DataFrame(pred_onx)], ignore_index=True)

```

```
147 # Set the column names for the prediction results for spectral accelerati
    ons
148 Y_result_sa.columns = ['pre_DNN_0.05', 'pre_DNN_0.06', 'pre_DNN_0.07', 'p
    re_DNN_0.08', 'pre_DNN_0.09', 'pre_DNN_0.1',
149                        'pre_DNN_0.11', 'pre_DNN_0.12', 'pre_DNN_0.13', '
    pre_DNN_0.15',  'pre_DNN_0.2', 'pre_DNN_0.22',
150                        'pre_DNN_0.25', 'pre_DNN_0.3', 'pre_DNN_0.35', 'p
    re_DNN_0.4', 'pre_DNN_0.45', 'pre_DNN_0.5',
151                        'pre_DNN_0.6', 'pre_DNN_0.7', 'pre_DNN_0.8', 'pre
    _DNN_0.9', 'pre_DNN_1', 'pre_DNN_1.1',
152                        'pre_DNN_1.2', 'pre_DNN_1.3', 'pre_DNN_1.5', 'pre
    _DNN_1.7', 'pre_DNN_2', 'pre_DNN_2.2',
153                        'pre_DNN_2.5', 'pre_DNN_3', 'pre_DNN_3.5', 'pre_D
    NN_4', 'pre_DNN_4.5', 'pre_DNN_5',
154                        'pre_DNN_5.5', 'pre_DNN_6', 'pre_DNN_6.5', 'pre_D
    NN_7', 'pre_DNN_7.5', 'pre_DNN_8',
155                        'pre_DNN_8.5', 'pre_DNN_9', 'pre_DNN_9.5', 'pre_D
    NN_10'  ]
156
157 # Merge the prediction results with the original data to get the final re
    sult
158 result = pd.concat([result, Y_result_sa], axis=1)
159
160 # Prepare a list of frequencies to be used in final calculations
161 fre_list = ['PGA', 'PGV',
162             '0.05', '0.06', '0.07', '0.08', '0.09', '0.1', '0.11', '0.12'
    , '0.13', '0.15',
163             '0.2', '0.22', '0.25', '0.3', '0.35', '0.4', '0.45', '0.5', '
    0.6',
164             '0.7', '0.8', '0.9', '1', '1.1', '1.2', '1.3', '1.5', '1.7',
    '2', '2.2',
165             '2.5', '3', '3.5', '4', '4.5', '5',
166             '5.5', '6', '6.5', '7', '7.5', '8', '8.5', '9', '9.5', '10']
167
168 # Perform the final calculation by adding the base model result to the pr
    ediction result
```

```
for fre in fre_list:
    result[f'Final_{fre}'] = result[f'Basemodel_{fre}'] + result[f'pre_DNN_{fre}']

# Generate the output file name from the original data file name
base_name = os.path.basename(dt_file)
output_file_name = f"result_{base_name}"
# Save the final result DataFrame to a CSV file in the specified output folder
result.to_csv(os.path.join(output_folder, output_file_name), index=False, float_format='%.6f')

# Define the frequencies for the data transformation
frequencies = [
    'PGA', 'PGV', '0.05', '0.06', '0.07', '0.08', '0.09', '0.1', '0.11', '0.12', '0.13', '0.15',
    '0.2', '0.22', '0.25', '0.3', '0.35', '0.4', '0.45', '0.5', '0.6',
    '0.7', '0.8', '0.9', '1', '1.1', '1.2', '1.3', '1.5', '1.7', '2', '2.2',
    '2.5', '3', '3.5', '4', '4.5', '5',
    '5.5', '6', '6.5', '7', '7.5', '8', '8.5', '9', '9.5', '10'
]

# Prepare the transformed data
output_data = []
for index, row in result.iterrows():
    base_info = row[['site_code', 'mw', 'fault_dist', 'eq_location_type_id']]
    for freq in frequencies:
        new_row = {
            'site_code': base_info['site_code'],
            'mw': base_info['mw'],
            'fault_dist': base_info['fault_dist'],
            'eq_location_type_id': base_info['eq_location_type_id'],
            'fre': freq,
            'Basemodel': row[f'Basemodel_{freq}'],
```

```
            'pre_DNN': row[f'pre_DNN_{freq}'],
            'Final': row[f'Final_{freq}']
        }
        output_data.append(new_row)

# Convert to DataFrame
output_df = pd.DataFrame(output_data)
output_df_name = f"result_all_{base_name}"
output_df.to_csv(os.path.join(output_folder, output_df_name), index=False, float_format='%.6f')

print(f"Results have been saved as {output_folder}/{output_file_name}")
print(f"Results have been saved as {output_folder}/{output_df_name}")
input()
```

References

[1] Mihalić S, Oštrić M, Krkač M. Seismic microzonation: A review of principles and practice. Geofizika 2011;28:5–20.

[2] Haskell NA. Crustal reflection of plane SH waves. Journal of Geophysical Research 1960;65:4147–50. https://doi.org/10.1029/JZ065I012P04147.

[3] Zhu C, Pilz M, Cotton F. Evaluation of a novel application of earthquake HVSR in site-specific amplification estimation. Soil Dynamics and Earthquake Engineering 2020;139:106301. https://doi.org/10.1016/J.SOILDYN.2020.106301.

[4] Andrews D.J. Separation of source and propagation spectra of seven Mammoth Lakes after-shocks. Proceedings of Workshop 16 Dynamic Characteristics of Faulting, 1981 1982:82–591.

[5] Iwata T, Irikura K. Source parameters of the 1983 Japan Sea earthquake sequence. Journal of Physics of the Earth 1988;36:155–84. https://doi.org/10.4294/jpe1952.36.155.

[6] Parolai S, Bindi D, Augliera P. Application of the Generalized Inversion Technique (GIT) to a microzonation study: Numerical simulations and comparison with different site-estimation techniques. Bulletin of the Seismological Society of America 2000;90:286–97. https://doi.org/10.1785/0119990041.

[7] Ren Y, Wen R, Yamanaka H, Kashima T. Site effects by generalized inversion technique using strong motion recordings of the 2008 Wenchuan earthquake. Earthquake Engineering and Engineering Vibration 2013 12:2 2013;12:165–84. https://doi.org/10.1007/S11803-013-0160-6.

[8] Nakano K, Matsushima S, Kawase H. Statistical properties of strong ground motions from the generalized spectral inversion of data observed by K-NET, KiK-net, and the JMA Shindokei network in Japan. Bulletin of the Seismological Society of America 2015;105:2662–80. https://doi.org/10.1785/0120140349.

[9] Nakamura Y. A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. Railway Technical Research Institute/Tetsudo Gijutsu Kenkyujo 1989;30:25–33.

[10] Ji K, Ren Y, Wen R. Site classification for National Strong Motion Observation Network System (NSMONS) stations in China using an empirical H/V spectral ratio method. Journal of Asian Earth Sciences 2017;147:79–94. https://doi.org/10.1016/j.jseaes.2017.07.032.

[11] Sardiña VHR, Midorikawa S. Site classification based on spectral amplification patterns for microtremor H/V ratios. Proceedings of 13th World Conference on Earthquake Engineering

VancouverCiteseer 2004.

[12] Yaghmaei-Sabegh S, Rupakhety R. A new method of seismic site classification using HVSR curves: A case study of the 12 November 2017 Mw 7.3 Ezgeleh earthquake in Iran. Engineering Geology 2020;270:105574. https://doi.org/10.1016/j.enggeo.2020.105574.

[13] Tao Y, Rathje E. Taxonomy for evaluating the site-specific applicability of one-dimensional ground response analysis. Soil Dynamics and Earthquake Engineering 2020;128:105865. https://doi.org/10.1016/j.soildyn.2019.105865.

[14] Oliveira L, Teves-Costa P, Pinto C, Gomes RC, Almeida IM, Ferreira C, et al. Seismic microzonation based on large geotechnical database: Application to Lisbon. Engineering Geology 2020;265:105417. https://doi.org/10.1016/j.enggeo.2019.105417.

[15] Satoh T, Kawase H, Matsushima S. Differences between site characteristics obtained from microtremors, S-waves, P-waves, and codas. Bulletin of the Seismological Society of America 2001;91:313–34. https://doi.org/10.1785/0119990149.

[16] Kassaras I, Papadimitriou P, Kapetanidis V, Voulgaris N. Seismic site characterization at the western Cephalonia Island in the aftermath of the 2014 earthquake series. Geo-Engineering 2017;8:7. https://doi.org/10.1186/s40703-017-0045-z.

[17] Carpenter NS, Wang Z, Woolery EW, Rong M. Estimating Site Response with Recordings from Deep Boreholes and HVSR: Examples from the Mississippi Embayment of the Central United States. Bulletin of the Seismological Society of America 2018;108:1199–209. https://doi.org/10.1785/0120170156.

[18] Herak M. ModelHVSR—A Matlab® tool to model horizontal-to-vertical spectral ratio of ambient noise. Computers & Geosciences 2008;34:1514–26. https://doi.org/10.1016/j.cageo.2007.07.009.

[19] Bignardi S, Mantovani A, Abu Zeid N. OpenHVSR: imaging the subsurface 2D/3D elastic properties through multiple HVSR modeling and inversion. Computers & Geosciences 2016;93:103–13. https://doi.org/10.1016/j.cageo.2016.05.009.

[20] Nakamura Y. What Is the Nakamura Method? Seismological Research Letters 2019;90:1437–43. https://doi.org/10.1785/0220180376.

[21] Lachetl C, Bard P-Y. Numerical and Theoretical Investigations on the Possibilities and Limitations of Nakamura's Technique. Journal of Physics of the Earth 1994;42:377–97. https://doi.org/10.4294/JPE1952.42.377.

[22] Lermo J, Chavez-Garcia FJ. Are microtremors useful in site response evaluation?: J. Lermo & F. J. Chavez-Garcia, Bulletin — Seismological Society of America, 84(5), 1994, pp 1350–1364.

International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts 1995;32:A228. https://doi.org/10.1016/0148-9062(95)93340-U.

[23] Konno K, Ohmachi T. Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor. Bulletin of the Seismological Society of America 1998;88:228–41. https://doi.org/10.1785/BSSA0880010228.

[24] Rodríguez VHS, Midorikawa S. Comparison of spectral ratio techniques for estimation of site effects using microtremor data and earthquake motions recorded at the surface and in boreholes. Earthquake Engineering & Structural Dynamics 2003;32:1691–714. https://doi.org/10.1002/EQE.296.

[25] Arai H, Tokimatsu K. S-wave velocity profiling by joint inversion of microtremor dispersion curve and horizontal-to-vertical (H/V) spectrum. Bulletin of the Seismological Society of America 2005;95:1766–78. https://doi.org/10.1785/0120040243.

[26] Sánchez-Sesma FJ, Rodríguez M, Iturrarán-Viveros U, Luzón F, Campillo M, Margerin L, et al. A theory for microtremor H/V spectral ratio: application for a layered medium. Geophysical Journal International 2011;186:221–5. https://doi.org/10.1111/J.1365-246X.2011.05064.X.

[27] Kawase H, Mori Y, Nagashima F. Difference of horizontal-to-vertical spectral ratios of observed earthquakes and microtremors and its application to S-wave velocity inversion based on the diffuse field concept. Earth, Planets and Space 2017 70:1 2018;70:1–32. https://doi.org/10.1186/S40623-017-0766-4.

[28] Miura H, Fujita H, Than KSS, Hibino Y. Estimation of site response during the 2016 Chauk, Myanmar earthquake based on microtremor-derived S-wave velocity structures. Soil Dynamics and Earthquake Engineering 2019;126:105781. https://doi.org/10.1016/J.SOILDYN.2019.105781.

[29] Sánchez-Sesma FJ. Modeling and inversion of the microtremor H/V spectral ratio: physical basis behind the diffuse field approach. Earth, Planets and Space 2017 69:1 2017;69:1–9. https://doi.org/10.1186/S40623-017-0667-6.

[30] Kawase H, Sánchez-Sesma FJ, Matsushima S. The optimal use of horizontal-to-vertical spectral ratios of earthquake motions for velocity inversions based on diffuse-field theory for plane waves. Bulletin of the Seismological Society of America 2011;101:2001–14. https://doi.org/10.1785/0120100263.

[31] Ito E, Nakano K, Nagashima F, Kawase H. A Method to Directly Estimate S-Wave Site Amplification Factor from Horizontal-to-Vertical Spectral Ratio of Earthquakes (eHVSRs). Bulletin of the Seismological Society of America 2020;110:2892–911.

https://doi.org/10.1785/0120190315.

[32] Kawase H, Nagashima F, Nakano K, Mori Y. Direct evaluation of S-wave amplification factors from microtremor H/V ratios: Double empirical corrections to “Nakamura” method. Soil Dynamics and Earthquake Engineering 2019;126:105067. https://doi.org/10.1016/J.SOILDYN.2018.01.049.

[33] Miura H, Aridome T, Matsuoka M. Deep Learning-Based Identification of Collapsed, Non-Collapsed and Blue Tarp-Covered Buildings from Post-Disaster Aerial Images. Remote Sensing 2020, Vol 12, Page 1924 2020;12:1924. https://doi.org/10.3390/RS12121924.

[34] Al Banna MH, Taher KA, Kaiser MS, Mahmud M, Rahman MS, Hosen ASMS, et al. Application of Artificial Intelligence in Predicting Earthquakes: State-of-the-Art and Future Challenges. IEEE Access 2020;8:192880–923. https://doi.org/10.1109/ACCESS.2020.3029859.

[35] Mignan A, Broccardo M. Neural Network Applications in Earthquake Prediction (1994–2019): Meta-Analytic and Statistical Insights on Their Limitations. Seismological Research Letters 2020;91:2330–42. https://doi.org/10.1785/0220200021.

[36] Zhang J, Zheng Y, Qi D, Li R, Yi X. DNN-based prediction model for spatio-temporal data. GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems 2016. https://doi.org/10.1145/2996913.2997016.

[37] Montavon G, Samek W, Müller KR. Methods for interpreting and understanding deep neural networks. Digital Signal Processing 2018;73:1–15. https://doi.org/10.1016/J.DSP.2017.10.011.

[38] Takeda J, Kanno T, Miura K. Evaluation of Site Amplification Factors at Strong Ground Motion Observation Sites in Hiroshima Prefecture Based on Spectral Inversion Technique, 2010, p. 2375–81.

[39] Takeda J, Kanno T, Miura K. Estimation of site amplification factors for earthquake motion observation sites in the Chugoku district based on spectral inversion technique Part 3 Expanding target sites and estimation of site effect distribution. vol. 34, 2011, p. 929–33.

[40] Kawase H, Matsuo H. Separation of Source, Path, and Site Effects based on the Observed Data by K-NET, KiK-net, and JMA Strong Motion Network. Journal of Japan Association for Earthquake Engineering 2004;4:33–52. https://doi.org/10.5610/JAEE.4.33.

[41] Salazar W, Sardina V, de Cortina J. A hybrid inversion technique for the evaluation of source, path, and site effects employing S-wave spectra for subduction and upper-crustal earthquakes in El Salvador. Bulletin of the Seismological Society of America 2007;97:208–21. https://doi.org/10.1785/0120060076.

[42] Bard PY. Foreword: The H/V technique: Capabilities and limitations based on the results of the

SESAME project. Bulletin of Earthquake Engineering 2008;6:1–2. https://doi.org/10.1007/S10518-008-9059-4.

[43] Nagashima F, Matsushima S, Kawase H, Sánchez-Sesma FJ, Hayakawa T, Satoh T, et al. Application of horizontal-to-vertical spectral ratios of earthquake ground motions to identify subsurface structures at and around the K-NET site in Tohoku, Japan. Bulletin of the Seismological Society of America 2014;104:2288–302. https://doi.org/10.1785/0120130219.

[44] García-Jerez A, Piña-Flores J, Sánchez-Sesma FJ, Luzón F, Perton M. A computer code for forward calculation and inversion of the H/V spectral ratio under the diffuse field assumption. Computers and Geosciences 2016;97:67–78. https://doi.org/10.1016/J.CAGEO.2016.06.016.

[45] Klambauer G, Unterthiner T, Mayr A, Hochreiter S. Self-Normalizing Neural Networks. Advances in Neural Information Processing Systems, vol. 30, Curran Associates, Inc.; 2017.

[46] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research 2014;15:1929–58.

[47] Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection 1995.

[48] Consonni V, Ballabio D, Todeschini R. Evaluation of model predictive ability by external validation techniques. Journal of Chemometrics 2010;24:194–201. https://doi.org/10.1002/CEM.1290.

[49] Martin PP, Seed HB. One-Dimensional Dynamic Ground Response Analyses. Journal of the Geotechnical Engineering Division 1982;108:935–52. https://doi.org/10.1061/AJGEB6.0001316.

[50] Johnson SW, Chambers DJA, Boltz MS, Koper KD. Application of a convolutional neural network for seismic phase picking of mining-induced seismicity. Geophysical Journal International 2021;224:230–40. https://doi.org/10.1093/gji/ggaa449.

[51] Thompson EM, Baise LG, Tanaka Y, Kayen RE. A taxonomy of site response complexity. Soil Dynamics and Earthquake Engineering 2012;41:32–43. https://doi.org/10.1016/j.soildyn.2012.04.005.

[52] Kaklamanos J, Bradley BA. Challenges in Predicting Seismic Site Response with 1D Analyses: Conclusions from 114 KiK-net Vertical Seismometer Arrays. Bulletin of the Seismological Society of America 2018;108:2816–38. https://doi.org/10.1785/0120180062.

[53] Laurendeau A, Bard P-Y, Hollender F, Perron V, Foundotos L, Ktenidou O-J, et al. Derivation of consistent hard rock (1000 < VS < 3000 m/s) GMPEs from surface and down-hole recordings:

analysis of KiK-net data. Bull Earthquake Eng 2018;16:2253–84. https://doi.org/10.1007/s10518-017-0142-6.

[54] Pilz M, Parolai S, Petrovic B, Silacheva N, Abakanov T, Orunbaev S, et al. Basin-edge generated Rayleigh waves in the Almaty basin and corresponding consequences for ground motion amplification. Geophysical Journal International 2018;213:301–16. https://doi.org/10.1093/gji/ggx555.

[55] Nozu A, Takashi Nagao, Masayuki Yamada. Site Amplification Factors for Strong-Motion Sites in Japan Based on Spectral Inversion Technique and Their Use for Strong-Motion Evaluation. Journal of Japan Association for Earthquake Engineering 2007;7:215–34. https://doi.org/10.5610/jaee.7.2_215.

[56] Bindi D, Pacor F, Luzi L, Massa M, Ameri G. The Mw 6.3, 2009 L'Aquila earthquake: source, path and site effects from spectral analysis of strong motion data. Geophysical Journal International 2009;179:1573–9. https://doi.org/10.1111/j.1365-246X.2009.04392.x.

[57] Davatgari Fami Tafreshi M, Bora SS, Mirzaei N, Ghofrani H, Kazemian J. Spectral models for seismological source parameters, path attenuation and site-effects in Alborz region of northern Iran. Geophysical Journal International 2021;227:350–67. https://doi.org/10.1093/gji/ggab227.

[58] Tramelli A, Galluzzo D, Del Pezzo E, Di Vito MA. A detailed study of the site effects in the volcanic area of Campi Flegrei using empirical approaches. Geophysical Journal International 2010;182:1073–86. https://doi.org/10.1111/j.1365-246X.2010.04675.x.

[59] Haghshenas E, Bard P-Y, Theodulidis N, SESAME WP04 Team. Empirical evaluation of microtremor H/V spectral ratio. Bull Earthquake Eng 2008;6:75–108. https://doi.org/10.1007/s10518-007-9058-x.

[60] Derras B, Bard P-Y, Cotton F, Bekkouche A. Adapting the Neural Network Approach to PGA Prediction: An Example Based on the KiK-net Data. Bulletin of the Seismological Society of America 2012;102:1446–61. https://doi.org/10.1785/0120110088.

[61] Derakhshani A, Foruzan AH. Predicting the principal strong ground motion parameters: A deep learning approach. Applied Soft Computing 2019;80:192–201. https://doi.org/10.1016/J.ASOC.2019.03.029.

[62] Li Z, Meier M-A, Hauksson E, Zhan Z, Andrews J. Machine Learning Seismic Wave Discrimination: Application to Earthquake Early Warning. Geophysical Research Letters 2018;45:4773–9. https://doi.org/10.1029/2018GL077870.

[63] Mousavi SM, Beroza GC. Deep-learning seismology. Science 2022;377:eabm4470. https://doi.org/10.1126/science.abm4470.

[64] Mousavi SM, Beroza GC. Machine Learning in Earthquake Seismology. Annual Review of Earth and Planetary Sciences 2023;51:105–29. https://doi.org/10.1146/annurev-earth-071822-100323.

[65] Zhu C, Cotton F, Kawase H, Nakano K. How well can we predict earthquake site response so far? Machine learning vs physics-based modeling. Earthquake Spectra 2023;39:478–504. https://doi.org/10.1177/87552930221116399.

[66] Pan D, Miura H, Kanno T, Shigefuji M, Abiru T. Deep-Neural-Network-Based Estimation of Site Amplification Factor from Microtremor H/V Spectral Ratio. Bulletin of the Seismological Society of America 2022;112:1630–46. https://doi.org/10.1785/0120210300.

[67] Pan SJ, Yang Q. A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering 2010;22:1345–59. https://doi.org/10.1109/TKDE.2009.191.

[68] Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, et al. A Comprehensive Survey on Transfer Learning. Proceedings of the IEEE 2021;109:43–76. https://doi.org/10.1109/JPROC.2020.3004555.

[69] Simon J, Fabien-Ouellet G, Gloaguen E, Khurjekar I. Hierarchical transfer learning for deep learning velocity model building. GEOPHYSICS 2023;88:R79–93. https://doi.org/10.1190/geo2021-0470.1.

[70] Wang H, Wang L, Zhang L. Transfer learning improves landslide susceptibility assessment. Gondwana Research 2023;123:238–54. https://doi.org/10.1016/j.gr.2022.07.008.

[71] Jozinović D, Lomax A, Štajduhar I, Michelini A. Transfer learning: improving neural network based prediction of earthquake ground shaking for an area with insufficient training data. Geophysical Journal International 2022;229:704–18. https://doi.org/10.1093/gji/ggab488.

[72] Hatayama K, Matsunami K, Iwata T, Irikura K. Basin-Induced Love Waves in the Eastern Part of the Osaka Basin. Journal of Physics of the Earth 1995;43:131–55. https://doi.org/10.4294/jpe1952.43.131.

[73] Furumura M, Sasatani T, Furumura T. Generation of Basin-Induced Surface Waves Observed in the Tokachi Basin, Hokkaido, Japan. Journal of Physics of the Earth 1997;45:287–305. https://doi.org/10.4294/jpe1952.45.287.

[74] Wald DJ, Worden CB, Thompson EM, Hearne M. ShakeMap operations, policies, and procedures. Earthquake Spectra 2022;38:756–77. https://doi.org/10.1177/87552930211030298.

[75] Boore DM. Simulation of Ground Motion Using the Stochastic Method. Pure Appl Geophys 2003;160:635–76. https://doi.org/10.1007/PL00012553.

[76] Hisada Y. Broadband strong motion simulation in layered half-space using stochastic Green's

function technique. J Seismol 2008;12:265–79. https://doi.org/10.1007/s10950-008-9090-6.

[77] Derras B, Bard P-Y, Régnier J, Cadet H. Non-linear modulation of site response: Sensitivity to various surface ground-motion intensity measures and site-condition proxies using a neural network approach. Engineering Geology 2020;269:105500. https://doi.org/10.1016/j.enggeo.2020.105500.

[78] Mucciarelli M. The Stability of the Horizontal-to-Vertical Spectral Ratio of Triggered Noise and Earthquake Recordings. Bulletin of the Seismological Society of America 2003;93:1407–12. https://doi.org/10.1785/0120020213.

[79] Sarmadi MA, Heidari R, Mirzaei N, Siahkoohi HR. The improvement of the earthquake and microseismic Horizontal-to-Vertical Spectral Ratio (HVSR) in estimating site effects. Acta Geophys 2021;69:1177–88. https://doi.org/10.1007/s11600-021-00619-0.

[80] Cultrera G, Cornou C, Di Giulio G, Bard PY. Indicators for site characterization at seismic station: recommendation from a dedicated survey. Bulletin of Earthquake Engineering 2021;19:4171–95. https://doi.org/10.1007/s10518-021-01136-7.

[81] Pan D, Miura H, Kwan C. Transfer learning model for estimating site amplification factors from limited microtremor H/V spectral ratios. Geophysical Journal International 2024;237:622–35. https://doi.org/10.1093/gji/ggae065.

[82] Zhu C, Weatherill G, Cotton F, Pilz M, Kwak DY, Kawase H. An open-source site database of strong-motion stations in Japan: K-NET and KiK-net (v1.0.0). Earthquake Spectra 2021;37:2126–49. https://doi.org/10.1177/8755293020988028.

[83] Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. Proceedings of the 32nd International Conference on Machine Learning, PMLR; 2015, p. 448–56.

[84] Santurkar S, Tsipras D, Ilyas A, Madry A. How Does Batch Normalization Help Optimization? Advances in Neural Information Processing Systems, vol. 31, Curran Associates, Inc.; 2018.

[85] Atkinson GM, Boore DM. Earthquake Ground-Motion Prediction Equations for Eastern North America 2006;96:2181–205. https://doi.org/10.1785/0120050245.

[86] Atkinson GM, Boore DM. Ground-motion relations for eastern North America. Bulletin of the Seismological Society of America 1995;85:17–30. https://doi.org/10.1785/BSSA0850010017.

[87] Bindi D. The Predictive Power of Ground-Motion Prediction Equations. Bulletin of the Seismological Society of America 2017;107:1005–11. https://doi.org/10.1785/0120160224.

[88] Castellaro S, Mulargia F, Rossi PL. VS30: Proxy for seismic amplification? Seismological

Research Letters 2008;79:540–3.

[89] Kanno T. A New Attenuation Relation for Strong Ground Motion in Japan Based on Recorded Data. Bulletin of the Seismological Society of America 2006;96:879–97. https://doi.org/10.1785/0120050138.

[90] Morikawa N, Fujiwara H. A new ground motion prediction equation for Japan applicable up to M9 mega-earthquake. Journal of Disaster Research 2013;8:878–88. https://doi.org/10.20965/jdr.2013.p0878.

[91] Zhao JX, Jiang F, Shi P, Xing H, Huang H, Hou R, et al. Ground-motion prediction equations for subduction slab earthquakes in Japan using site class and simple geometric attenuation functions. Bulletin of the Seismological Society of America 2016;106:1535–51. https://doi.org/10.1785/0120150056.

[92] Douglas J. Ground Motion Prediction Equations. Department of Civil and Environmental Engineering University of Strathclyde, Glasgow, United Kingdom 2021.

[93] Akkar S, Çağnan Z. A local ground-motion predictive model for Turkey, and its comparison with other regional and global ground-motion models. Bulletin of the Seismological Society of America 2010;100:2978–95. https://doi.org/10.1785/0120090367.

[94] Bindi D, Pacor F, Luzi L, Puglia R, Massa M, Ameri G, et al. Ground motion prediction equations derived from the Italian strong motion database. Bulletin of Earthquake Engineering 2011;9:1899–920.

[95] Zhang B, Yu Y, Li X, Wang Y. Ground motion prediction equation for the average horizontal component of PGA, PGV, and 5% damped acceleration response spectra at periods ranging from 0.033 to 8.0 s in southwest China. Soil Dynamics and Earthquake Engineering 2022;159:107297.

[96] Bozorgnia Y, Abrahamson NA, Atik LA, Ancheta TD, Atkinson GM, Baker JW, et al. NGA-West2 research project. Earthquake Spectra 2014;30:973–87.

[97] Campbell KW, Bozorgnia Y. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. Earthquake Spectra 2014;30:1087–115.

[98] Morikawa N, Fujiwara H, Iwaki A, Maeda T, Kubo H, Aoi S, et al. Toward to Construction of Strong-Motion Database for Seismic Hazard Assessment in Japan. vol. SSS04-P01, Japan: 2020.

[99] Derras B, Bard P-Y, Cotton F, Bekkouche A. Adapting the Neural Network Approach to PGA Prediction: An Example Based on the KiK-net Data. Bulletin of the Seismological Society of America 2012;102:1446–61. https://doi.org/10.1785/0120110088.

[100] Oana A, Ishii T, Wada K. Applicability of features to ground motion evaluation models utilizing machine learning. vol. 3Rin4- 03, 2020.

[101] Oana A, Ishii T, Miyashita Y, Furukawa K. Construction of Ground Motion Evaluation Models Using Supervised Machine Learning Based on Strong Motion Database. Journal of Japan Association for Earthquake Engineering 2022;22:6_39-6_56. https://doi.org/10.5610/jaee.22.6_39.

[102] Lilienkamp H, von Specht S, Weatherill G, Caire G, Cotton F. Ground-Motion Modeling as an Image Processing Task: Introducing a Neural Network Based, Fully Data-Driven, and Nonergodic Approach. Bulletin of the Seismological Society of America 2022;112:1565–82. https://doi.org/10.1785/0120220008.

[103] Dhanya J, Raghukanth STG. Ground Motion Prediction Model Using Artificial Neural Network. Pure and Applied Geophysics 2018;175:1035–64. https://doi.org/10.1007/S00024-017-1751-3/FIGURES/22.

[104] Sreenath V, Raghukanth STG. Stochastic ground motion models to NGA-West2 and NGA-Sub databases using Bayesian neural network. Earthquake Engineering & Structural Dynamics 2023;52:248–67. https://doi.org/10.1002/eqe.3759.

[105] Kuehn NM, Bozorgnia Y, Campbell KW, Gregor N. A regionalized partially nonergodic ground-motion model for subduction earthquakes using the NGA-Sub database. Earthquake Spectra 2023;39:1625–57. https://doi.org/10.1177/87552930231180906.

[106] Stafford PJ. Continuous integration of data into ground-motion models using Bayesian updating. J Seismol 2019;23:39–57. https://doi.org/10.1007/s10950-018-9792-3.

[107] Kubo H, Kunugi T, Suzuki W, Suzuki S, Aoi S. Hybrid predictor for ground-motion intensity with machine learning and conventional ground motion prediction equation. Scientific Reports 2020;10:1–12. https://doi.org/10.1038/s41598-020-68630-x.

[108] Khosravikia F, Zeinali Y, Nagy Z, Clayton P, Rathje EM. Neural Network-Based Equations for Predicting PGA and PGV in Texas, Oklahoma, and Kansas 2018:538–49. https://doi.org/10.1061/9780784481462.052.

[109] Satoh T, Kawase H. Site Responses of Sediments. Zisin (Journal of the Seismological Society of Japan 2nd ser) 2009;61:455–70. https://doi.org/10.4294/zisin.61.455.

[110] Midorikawa S. Ground Motion Attenuation Relations. Zisin 2009;61:471–7. https://doi.org/10.4294/zisin.61.471.

[111] Miura H, Gonzales C, Diaz M, Estrada M, Lazares F, Aguilar Z, et al. Assessment of Site Amplification Factors in Southern Lima, Peru Based on Microtremor H/V Spectral Ratios and

Deep Neural Network. JDR 2023;18:298–307. https://doi.org/10.20965/jdr.2023.p0298.

[112] Zaker Esteghamati M, Kottke AR, Rodriguez-Marek A. A Data-Driven Approach to Evaluate Site Amplification of Ground-Motion Models Using Vector Proxies Derived from Horizontal-to-Vertical Spectral Ratios. Bulletin of the Seismological Society of America 2022;112:3001–15. https://doi.org/10.1785/0120220106.

[113] Chao S-H, Lin C-M, Kuo C-H, Huang J-Y, Wen K-L, Chen Y-H. Implementing horizontal-to-vertical Fourier spectral ratios and spatial correlation in a ground-motion prediction equation to predict site effects. Earthquake Spectra 2021;37:827–56. https://doi.org/10.1177/8755293020952449.

[114] Ghofrani H, Atkinson GM. Site condition evaluation using horizontal-to-vertical response spectral ratios of earthquakes in the NGA-West 2 and Japanese databases. Soil Dynamics and Earthquake Engineering 2014;67:30–43. https://doi.org/10.1016/j.soildyn.2014.08.015.

[115] Yazdi M, Anderson JG, Motamed R. Reducing the uncertainties in the NGA-West2 ground motion models by incorporating the frequency and amplitude of the fundamental peak of the horizontal-to-vertical spectral ratio of surface ground motions. Earthquake Spectra 2023:87552930221146807. https://doi.org/10.1177/87552930221146807.

[116] Boore DM. Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. Bulletin of the Seismological Society of America 2010;100:1830–5. https://doi.org/10.1785/0120090400.

[117] Furumura T, Kennett BLN. Subduction zone guided waves and the heterogeneity structure of the subducted plate: Intensity anomalies in northern Japan. Journal of Geophysical Research: Solid Earth 2005;110.

[118] Ghofrani H, Atkinson GM. Forearc versus backarc attenuation of earthquake ground motion. Bulletin of the Seismological Society of America 2011;101:3032–45.

[119] Fukutake K, Yoshida K, Kawase H, Nagashima F, Sun J. Quantitative strong motion simulations at three locations in downtown Mashiki during the 2016 Kumamoto Earthquake (Mw 7.0) based on the nonlinear ground response analyses with soil liquefaction. Earth, Planets and Space 2023;75:87. https://doi.org/10.1186/s40623-023-01821-8.

List of Publications

Peer-reviewed papers:

1. **Pan, D.**, Miura, H., Kanno, T., Shigefuji, M., & Abiru, T. (2022). Deep-Neural-Network-Based Estimation of Site Amplification Factor from Microtremor H/V Spectral Ratio. *Bulletin of the Seismological Society of America*, *112*(3), 1630-1646. https://doi.org/10.1785/0120210300.
2. Miura, H., Gonzales, C., Diaz, M., Estrada, M., Lazares, F., Aguilar, Z., **Pan, D.**, & Matsuoka, M. (2023): Assessment of site amplification factors in southern Lima, Peru based on microtremor H/V spectral ratios and deep neural network, Journal of Disaster Research, 18(4), 298-307, https://doi.org/10.20965/jdr.2023.p0298.
3. **Pan, D.**, Miura, H., & Kwan, C. (2024). Transfer learning model for estimating site amplification factors from limited microtremor H/V spectral ratios. *Geophysical Journal International*, *237*(1), 622-635. https://doi.org/10.1093/gji/ggae065.
4. **Pan, D.**, Miura, H. Deep-neural-network Model for Predicting Ground Motion Parameters Using Earthquake Horizontal-to-Vertical Spectral Ratios. *Earthquake Spectra*. (Accepted)

Conference presentation (peer-reviewed with abstracts):

1. **Pan, D.**, Miura, H., Kanno, T., Shigefuji, M. and Abiru, T.: Development of Deep Neural Network Model for Estimating S-wave Amplification from Microtremor H/V Spectral Ratio, *17th World Conference on Earthquake Engineering (17WCEE)*, At: Sendai, Japan, September 2021.
2. **Pan, D.**, Miura.: Transfer-Learning-Based Estimation of Site Amplification Factors from Microtremor H/V Spectral Ratios, *16th Japan Earthquake Engineering Symposium (16JEES)*, At: Yokohama, Japan, November 2023. **(The 16th Japan Earthquake Engineering Symposium Excellent Presentation Award)**.
3. **Pan, D.**, Miura.: Estimating Site Amplification Factor from Microtremor H/V Spectral Ratio by Transfer Learning, *18th World Conference on Earthquake Engineering (WCEE2024)*, At: Milan, Italia, July 2024.

Conference presentation (non-peer-reviewed):

1. **Pan, D.**, 三浦弘之, 神野達夫, 重藤迪子, 阿比留哲生: 微動 H/V スペクトル比による地盤増幅特性の簡易推定法に関する予備的検討, 日本建築学会大会学術講演梗概集（関東）, 構造 II, pp.92-93, September 2020.

2. 三浦弘之, **Pan, D.**, 神野達夫, 重藤迪子, 阿比留哲生: AI 技術を利用した微動 H/V スペクトル比による地盤増幅特性の経験的推定の試行, 第 48 回地盤震動シンポジウム, 日本建築学会, pp.53-56, November 2020.
3. **Pan, D.**, 三浦弘之, 神野達夫, 重藤迪子, 阿比留哲生：深層学習に基づく微動 H/V スペクトル比からの地盤増幅特性の推定, 日本建築学会大会学術講演梗概集（中部）, 構造 II, pp.317-318, September 2021.
4. Miura, H., **Pan, D.**: Application of AI Technology for Estimating Site Amplification Factor from Microtremor Data, Proceedings of 15th Research Conference on Geotechnical Issues on Buildings and Structures, pp.1-5, March 2022.
5. **Pan, D.**, 三浦弘之, 神野達夫, 重藤迪子, 阿比留哲生：深層学習に基づく微動 H/V スペクトル比からの地盤増幅特性の推定手法の適用性検討, 日本建築学会大会学術講演梗概集（北海道）, 構造 II, pp.79-80, September 2022.
6. Miura, H., **Pan, D.**: Application of AI Technology for Estimating Site Amplification Factor from Microtremor H/V Spectral Ratio, 4th Kenji Ishihara Colloquium Series on Earthquake Engineering, *Artificial Intelligence Applications in Earthquake Engineering,* EERI San Diego Regional Chapter, University of California, San Diego, September 2022.
7. **Pan D.**, 三浦弘之：強震動データベースと地震動 H/V スペクトル比に基づく地震動予測のための深層学習モデルの構築, 日本地震工学会・大会－2022, Paper No. TS20220037, December 2022.
8. **Pan, D.**, 三浦弘之：強震動データベースと地震動 H/V スペクトル比に基づく地震動予測のための深層学習モデルの構築, 日本建築学会中国支部研究報告集, 第 46 巻, pp.219-222, March 2023.
9. **Pan, D.**, Miura, H.: Transfer learning model for expanding applicability in estimating site amplification from microtremor H/V spectral ratio, 日本建築学会全国大会学術講演梗概集 (近畿), 構造 II, pp.791-792, September 2023. (**構造部門振動分野・若手優秀発表賞・受賞**)