applied sciences

MDPI

Article

Probabilistic Seismic Hazard Analysis of a Back Propagation Neural Network Predicting the Peak Ground Acceleration

Xin Guo [1], Hongnan Li [1,2], Hao Zhang [1,*], Qi Wang [1] and Jiran Xu [1]

[1] School of Civil Engineering, Shenyang Jianzhu University, Shenyang 110168, China; guoxin@stu.sjzu.edu.cn (X.G.); hnli@dlut.edu.cn (H.L.); wang-qi@stu.sjzu.edu.cn (Q.W.); jiran0407@163.com (J.X.)
[2] Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian 116024, China
* Correspondence: h_zhang@sjzu.edu.cn

Abstract: Probabilistic seismic hazard analysis (PSHA) has been recognized as a reasonable method for quantifying seismic threats. Traditionally, this method ignores the effect of the focal depth, in which the ground motion prediction equations (GMPEs) are applied to estimate the probability distribution associated with the possible motion levels induced by the site earthquakes, but it is limited by the unclear geological conditions, which makes it difficult to provide a uniform equation, and these equations cannot express the non-linear relationship under geological conditions. Hence, this paper proposed a method to consider the seismic focal depth for the PSHA with the example of California and used a back propagation neural network (BPNN) to predict the peak ground acceleration (PGA) instead of the GMPEs. Firstly, the measured PGA and unknown PGA seismic data applicable to this method were collected separately. Secondly, the unknown PGA data were supplemented by applying the BPNN based on the measured PGA data. Lastly, based on the full-probability equation, PSHA considering the focal depth was completed and compared with the current California seismic zoning results. The results showed that using the BPNN in the PSHA can ensure computational accuracy and universality, making it more suitable for regions with unclear geological structures and providing the possibility of adding other parameters to be considered for the influence of the PSHA.

Keywords: back propagation neural network; hazard analysis; peak ground motion; probabilistic analysis

Citation: Guo, X.; Li, H.; Zhang, H.; Wang, Q.; Xu, J. Probabilistic Seismic Hazard Analysis of a Back Propagation Neural Network Predicting the Peak Ground Acceleration. *Appl. Sci.* **2023**, *13*, 9790. https://doi.org/10.3390/app13179790

Academic Editor: Rosario Montuori

Received: 14 August 2023
Revised: 24 August 2023
Accepted: 29 August 2023
Published: 30 August 2023

1. Introduction

Earthquakes are one of the most influential disasters in the entire life cycle of engineering structures [1]. To avoid damage to engineering structures caused by earthquakes, seismic hazard analysis methods can be used to predict the probability of earthquakes of various strengths in the future and prevent them in their structural designs [2]. Seismic hazard analysis includes the deterministic and probabilistic methods [3]. The deterministic seismic hazard analysis (DSHA) requires the prior determination of the seismic source and magnitude, but neglects the uncertainties in the data and calculations, which could lead to an overestimation of the risk for the seismic source with longer return periods or an underestimation of the risk for the seismic source with shorter return periods. Since Cornell proposed the probabilistic seismic hazard analysis method in 1968, this method has been used in the context of engineering design and in the verification of the performance requirements of structures induced by earthquakes [4]. The probabilistic seismic hazard analysis (PSHA) has been generally recognized as the rational method to quantify seismic threats [5]. The PSHA includes four aspects (as shown in Figure 1): (1) the seismic source characteristics, (2) the relationship between the magnitude and the frequency of the earthquake, (3) the attenuation relationship of the ground motion parameters, and

(4) the hazard curve [6]. However, there are three uncertainties in the analysis process: the determination of the potential seismic source area, the estimation of the seismicity parameters, and the relationship between the parameters [7]. Surrounding the uncertain factors, how to improve the PSHA is the focus of scholars' continuous research.

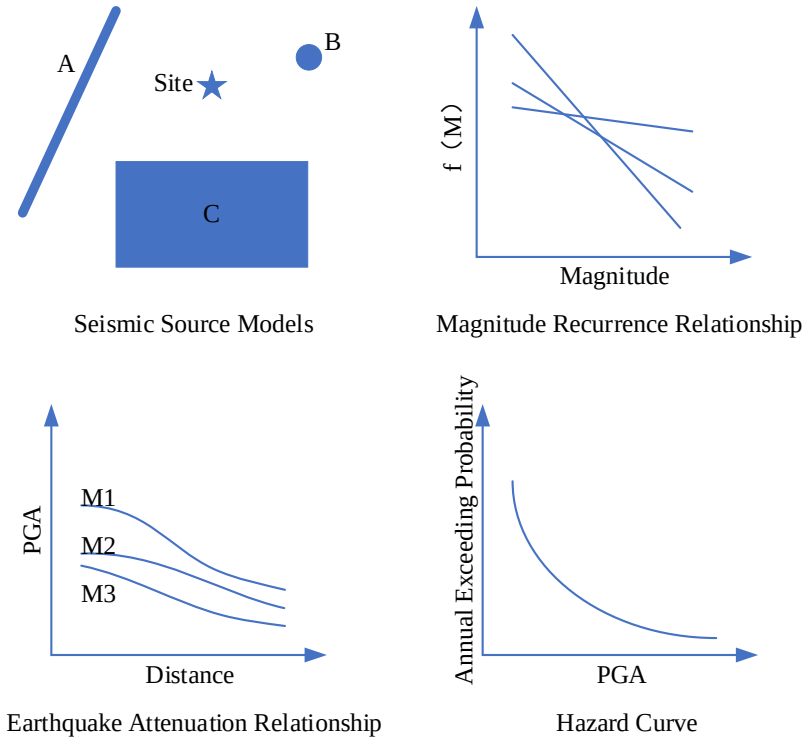

Figure 1. Steps underlying PSHA.

Researchers have refined the uncertainties in the determination of potential seismic source regions caused by an incomplete understanding of subsurface structures and earthquakes by adjusting the seismic zoning and seismic source models [8,9]. The most important method to determine the potential seismic source area is to gain an in-depth understanding of the geological structural parameters, such as the soil parameters [10] and regional tectonic distribution characteristics [11], along with the deep earth tectonic [12] and seismic zone distributions [13]. In addition, the parameters of the seismic source mechanism can be effectively solved by defining the different seismic source models. The fact that these parameters vary widely in space and time, combined with the effect that different geological formations are interacting systems, makes the accurate analysis of potential seismic source zones very difficult [14,15]. Continuous refinement of seismic zoning can precisely determine the parameters of the area and apply them to calculate the PSHA for each differential zone; however, detailed seismic zoning is not conducive to integration into region-wide, full-probability calculations [16]. Certain parameters, such as the average annual occurrence of earthquakes, earthquake magnitude values, and upper limits of earthquake magnitude, can lead to uncertainties in the valuation of the seismicity parameters. The seismic catalogue obtained from using seismic monitoring techniques can supplement this part of the parameters with a large number of actual measurements, but the number of such monitoring results is still limited by the capacity of the measuring instruments and the number of arrangements. To consider a more complete seismic catalogue, previous studies [17,18] have combined the Monte-Carlo simulations of historical seismicity catalogues based on the measured data but may affect the accuracy of the PSHA. For the uncertainty underlying the determination of the potential seismic source regions and seismic activity parameter valuations, multiple combinations of scenarios may be used to calculate the seismic hazard analysis results of different scenarios while considering the

range in the variation of the results of all calculated scenarios to provide the mean and variance of the seismic hazard analysis results.

The uncertainty in the relationship between the seismic parameters originates from the unknown ground motion attenuation relation, which is currently obtained through fitting the empirical formulae to known ground motion-related data, among which include ground motion intensity indicators (e.g., PGA and spectral acceleration [19]), earthquake magnitude, and distance (epicenter distance or hypocentral distance). This formula is known as the GMPE, and currently, GMPEs can be classified into three types [20]: functional, controlling, and predictive. Functional formulas require the determination of functional relationships among the parameters; controlling formulas can estimate the independent effects among the parameters experimentally; and predictive formulas are only used for the calculation of non-deterministic relationships among the data. All physical aspects associated with the earthquake ground motion are not known [21], and even if they are known, they cannot be expressed in the form of a simple equation; the estimated ground motion relationship is the predictive model. After a large number of scholars' studies on the GMPE, the deviation value of the GMPE in these seismic areas has been reduced year by year [22,23], but it still cannot provide a single definite GMPE [24]. For those areas with low seismicity, the selection of a suitable GMPE for the PSHA is an even more challenging step [25,26].

Undoubtedly, the in-depth analysis of geological formations to solve the problems caused by uncertainty is the fundamental method to improve the PSHA, but facing the problems of many influencing factors and unclear unknown formations is not conducive to the use and design in the engineering field. The effect of the focal depth on the ground motion, which is widely known, has not been systematically considered in the PSHA [27]. This neglect is also present in the GMPE. Therefore, this paper focused on using known data to predict the PGA through the non-linear fitting ability of the BPNN, while the PSHA is based on the full probability formulation considering the effect of the focal depth. The organization of this paper is as follows. Section 2 describes the selection guidelines for the California earthquake catalogue, which includes both known PGA and unknown PGA data, in detail. The BPNN-based PGA prediction method and processes have been described in Section 3, and a preliminary judgment on the accuracy of the fitting results has been made. In Section 4, the PSHA considering the depth of the seismic source has been completed, with the data having been fitted using the BPNN. In Section 5, the resulting hazard curves are compared with the seismic zoning results in ACSE 7–16 to determine the accuracy of the method proposed in this paper. At the end of this paper, the relevant conclusions and recommendations have been summarized in Section 6.

2. Earthquake Catalogs

The BPNN-based PGA predictions need to be supported with as much data as possible, and the data mentioned here include both known PGA data and unknown PGA data. In this context, researchers examining these effects in California [28], as an earthquake-prone region, also started to record and analyze their seismic data earlier in this area; therefore, in this paper, we counted the seismic data in the region, combined with the time of database establishment and the completeness of known data, and determined the time range of data acquisition from the years 1970 to 2021.

2.1. Known PGA Data

The known PGA data were acquired and recorded to the strong motion database through the strong motion station, and the main known strong motion databases include the Next-Generation Attenuation (NGA) database [29], the Center for Engineering Strong Motion Data (CESMD) database [30], and the COSMOS (Consortium of Organizations for Strong-Motion Observation Systems) Virtual Data Center database [31]. Considering the data volume, data time range, and demand parameters, the final data set utilized was the CESMD, jointly established by the United States Geological Survey (USGS) and the

California Geological Survey (CGS), and the strong motion data and strong motion station are shown in Figure 2.

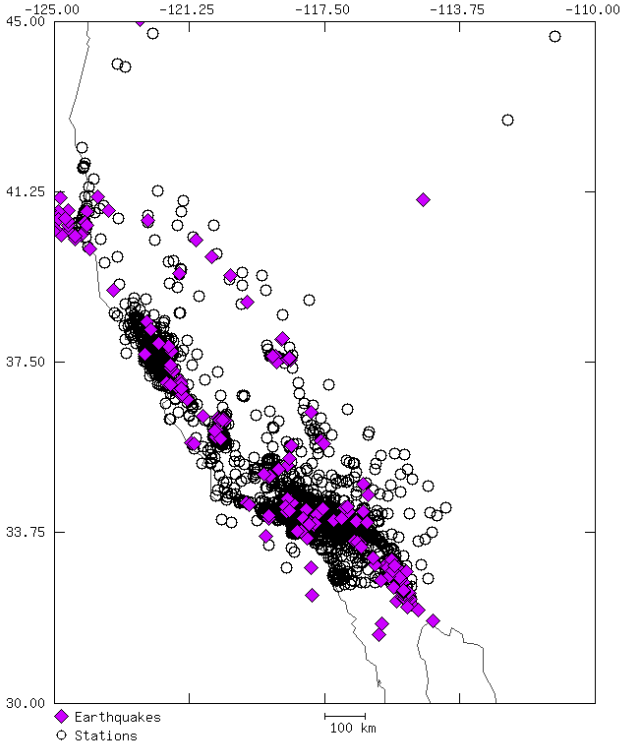

Figure 2. California strong motion database map.

The Initial screening of the obtained data was performed. Since the boxed range was controlled by the latitude and longitude, the selected area in Figure 2 is larger than the California area, and the redundant data were removed by combining the actual California range in the data. With the increase in the number of strong seismograph arrangements, more strong motion data were measured for the same strong earthquake. The increase in the amount of data can improve the accuracy of the prediction results, but the propagation of the ground motion is greatly affected by the existence of the earth's soil, and the more distant the measured locations are, the higher the loss in the confidence rate; therefore, only the data closest to the seismic source were selected to be used in this study. The measured PGAs were provided in the form of two horizontal components and one vertical component, and the largest PGA in the same earthquake was used due to the disparity between the ground motion components, which has not been considered in this paper.

The GMPE generally calculates the PGA using the magnitude and distance (epicenter distance or hypocentral distance), but in fact, the epicenter distance is the distance between the ground location of the seismic source and the seismic station, and the hypocentral distance is the distance between the seismic source and the seismic station, so the calculated PGA is affected by the focal depth. To obtain more accurate PGA calculation results, the magnitude, epicentral distance, focal depth, and PGA were obtained in this paper. In total, 107 known PGA data points were obtained through combining the above selection principles. The relationship between the magnitude and the epicenter distance of these data is shown in Figure 3.

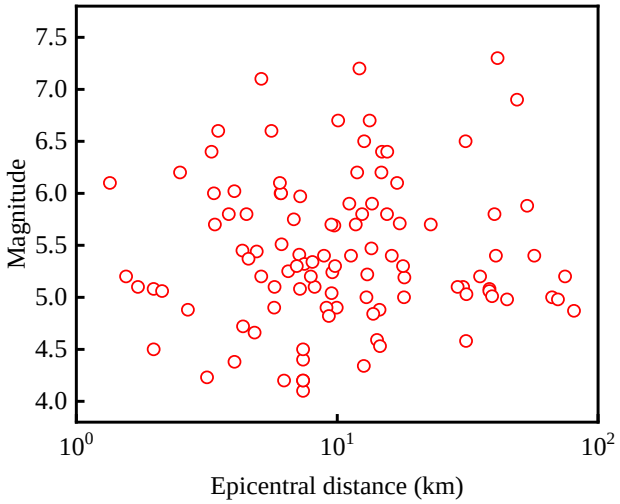

Figure 3. Known PGA data.

2.2. Unknown PGA Data

In addition to the strong motion database, the earthquake database mainly records the magnitude, epicenter distance, and focal depth. There is no PGA (unknown PGA), but the PGA data is a necessary part of PSHA. The unknown PGA database can complete the earthquake catalogues to reduce the problem of overestimating the annual PGA exceedance probability due to the strong motion data. According to the basic requirements for strong motion data selection, this part of the data needs to ensure at least three kinds of information with the magnitude, epicenter distance, and focal depth, which can be provided by the United States Geological Survey (USGS) [32], and the data acquisition process is shown in Figure 4.

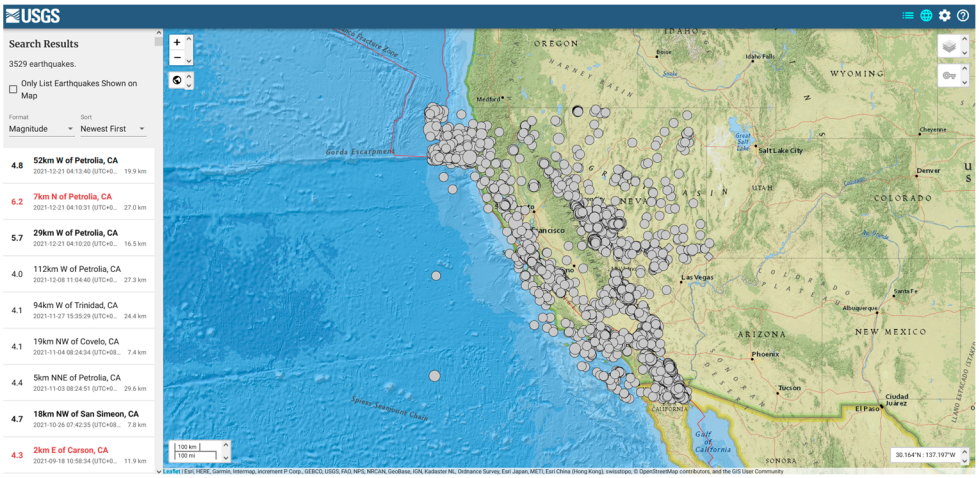

Figure 4. The California earthquake database map (https://earthquake.usgs.gov/earthquakes/map (accessed on 30 August 2022)).

The initial screening of the obtained data was performed. Consistent with the screening of the known strong motion data, the data for the non-California regions and the presence of the required parameters (magnitude, epicenter distance, and focal depth) were removed. Since the seismic hazard is calculated for destructive earthquakes, the magnitude of the destructive earthquakes is generally taken to be about 4.5 or higher and considering that the minimum value of earthquake magnitude in the strong motion data is 4.1, some data with an earthquake magnitude of 4.0 or higher have been retained. A total of 2645 data points were obtained from the California area. The frequency distribution histogram (Figure 5) was obtained with a magnitude difference of 0.1. The data displayed in Figure 5

are concentrated in lower magnitudes, which is consistent with the occurrence pattern of earthquakes.

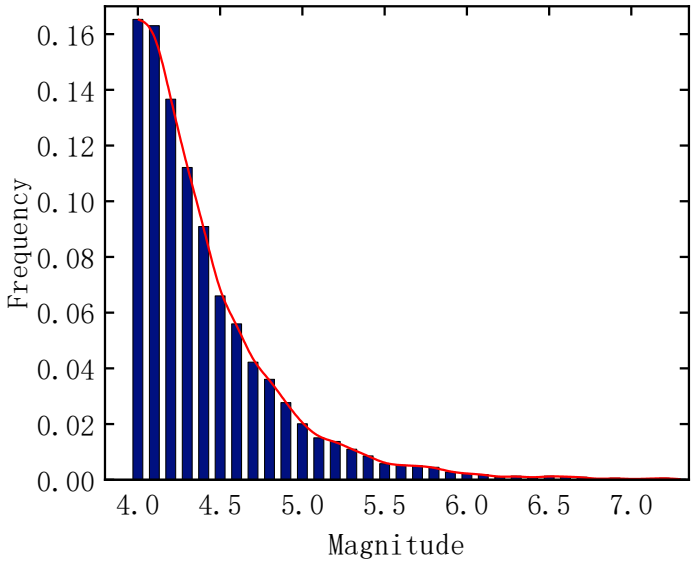

Figure 5. Histogram of magnitude frequency distribution.

3. BPNN-Based PGA Prediction

The application of neural networks to compute multi-parameter fitting problems has been widely used in structural engineering [33], where a non-linear computational power can effectively reduce the time to analyze the mechanism while ensuring computational accuracy. Lin Jyh-Woei [34] used the BPNN to predict the probability of earthquakes in Taiwan, a method that can be commercialized at a relatively low cost and with minimal resources and equipment using only the earthquake catalogues.

Combined with the statistical data, it is easy to find that the PGA, as an important part of the strong motion data, is affected by the propagation speed and the location of the strong motion station and has not yet achieved the purpose of obtaining effective real motion data for each earthquake; the amount of recorded seismic data is much larger than the strong motion data, and the supplemental PGA, through seismic data, has also become an important part of the PSHA.

Previously, the GMPE was used to supplement the PGA, and this method generally fitted the seismic magnitude separately according to the polar difference of one (e.g., Richter scales four, five, and six). The fitted equations also differed due to the distinct geological considerations, and the goodness-of-fit was low. The GMPE is an empirical prediction relationship, as shown in Equation (1) [35] and Equation (2) [36], that represents the ground motion parameters (including the PGA, spectral acceleration at a given frequency, peak velocity, or peak displacement) as a function of other parameters (such as the earthquake magnitude and epicentral distance):

$$\ln PGA = C_1 + C_2 m + C_3 \ln R + C_4 R \tag{1}$$

$$\ln PGA = C_1 + C_2 m - C_3 \ln(R + C_4 \exp(C_5 m)) \tag{2}$$

where m is the earthquake magnitude; R is the epicentral distance or hypocentral distance; and C_1, C_2, C_3, C_4, and C_5 are constants that are independent of m and R.

As a traditional PGA prediction method, the GMPE varies greatly under different regions and is influenced by the geological soil and other conditions; however, there is still no uniform model. This paper makes use of the better non-linear computational ability of the BPNN to compensate for the computational errors caused by the lack of understanding of geology in the engineering field and to improve the problem that multiple GPMEs need to be considered for supplementing PGAs across different regions. Due to the complex

non-linear correlation between the input and the output quantities, the accuracy of the previous primary function fitting was deemed to be poor, while the neural network, as a tool to describe the correlation between the multiple quantities, has a strong non-linear fitting ability, and the above mapping relationship has been considered by the neural network. The process of predicting the PGA using the BPNN is as follows:

(1) The known PGA data are randomly split into a training set, a validation set, and a test set in the ratio of 7:1.5:1.5 to avoid errors in the prediction results caused by large differences in the order of magnitude using data normalization.

(2) The BPNN (Figure 6) is constructed with three elements (a magnitude of Richter scale, focal depth, and the epicentral distance) in the input layer and one element (PGA) in the output layer. To avoid underfitting and overfitting, the number of hidden layers is set to two, each layer has four neurons, and the number of neurons is obtained by considering the number of elements in the input and output layers in combination with several attempts, where the transfer function (transig) and the training function (trainlm) are non-linear functions [37].

(3) The data set is trained with the determinants of the coefficients that are greater than 0.95 as the index to obtain the fitted network, where the regression fitting results are shown in Figure 7.

(4) After saving the network, the normalized unknown PGA data are inputted into the network. The predicted data are obtained and inverse normalized to obtain the final predicted PGA.

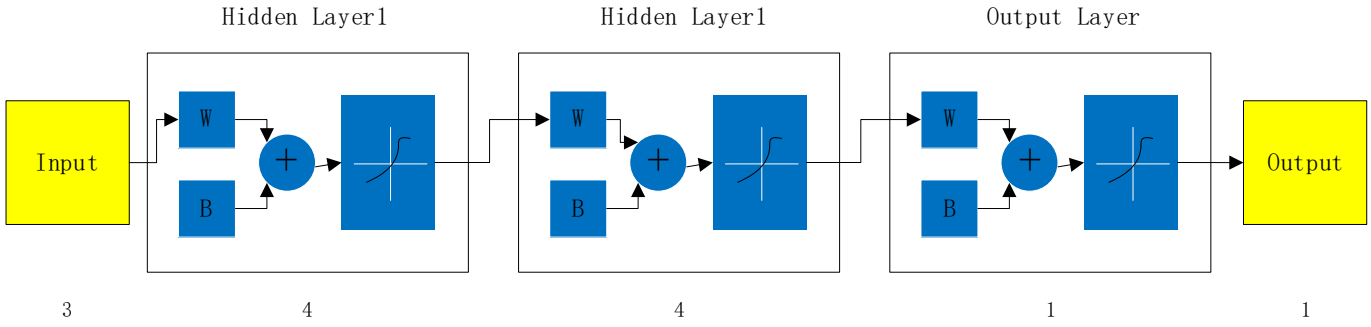

Figure 6. The BPNN.

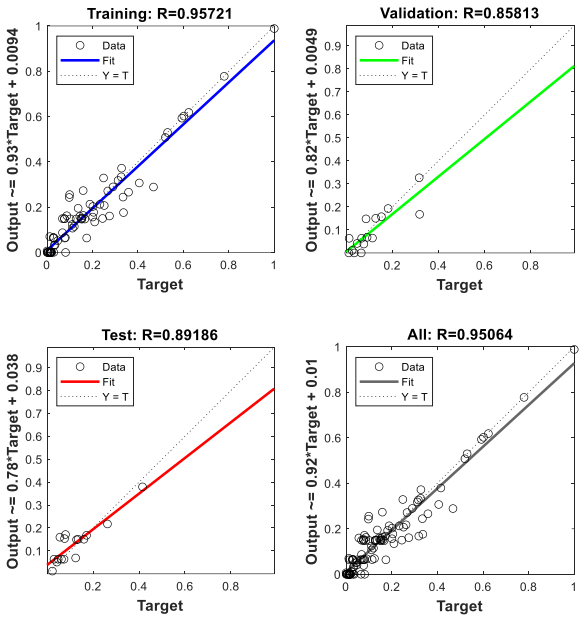

Figure 7. Regression results.

The regression fitting results of the California data are shown in Figure 7. There is randomness in the calculation process of the BPNN, and a 0.95 goodness-of-fit value of the total data was selected as the calculation index in this study, although, this index can still be stricter. As the total data was divided into different data sets, the fit superiority between each data set differs, and the superiority of the total data set is easier to reach the index when the training data set accounts for the most and its superiority is higher.

4. PSHA

4.1. The PSHA Method

The PSHA is obtained based on the full probability formula [38]; to consider the influence of the focal depth on the ground motion, this paper combined the existing data available for regression analysis to select the best-fit formula. Adding the probability density function about the depth of the earthquake, in this regard, the annual average transcendence probability of a seismic event that meets the PGA greater than a given value is shown in Equations (3)–(8):

$$P(PGA > x) = \lambda(m > m_{\min}) \iiint P(PGA > x|m, R, D) f(m) f(R) f(D) dm dR dD \tag{3}$$

$$\lambda(m > m_{\min}) = (A_D / A)\lambda \tag{4}$$

$$f(m) = \frac{b \ln(10) 10^{-b(m - m_{\min})}}{1 - 10^{-b(m_{\max} - m_{\min})}} \qquad m_{\min} < m < m_{\max} \tag{5}$$

$$\log_{10}(N_m / T) = a - bm \tag{6}$$

$$f(R) = \frac{d}{dR} F(R) = \frac{d}{dR}\left(\frac{\pi R^2}{\pi R_t^2}\right) = \frac{2R}{R_t^2} \quad 0 < R < R_t \tag{7}$$

$$f(D) = a_1 e^{(a_2 D)} + a_3 e^{(a_4 D)} \tag{8}$$

where $P(PGA > x|m, R, D)$ denotes the conditional probability of the PGA exceeding x given m, R, and D; $\lambda(m > m_{\min})$ is the average annual occurrence rate of earthquake events with a magnitude greater than the minimum magnitude [39], based on the California earthquake catalogue [40]; A_D is the area of the analyzed region; A is the area of the analyzed region and its 300 km radius; λ is the average annual occurrence rate; $f(m)$ is the probability density function of the magnitude; $m_{\min}$ and $m_{\max}$ are the minimum and maximum earthquakes that may occur in the seismic source, respectively [41]; b denotes the Gutenberg–Richter coefficient [42]; N_m is the annual number of seismic events in a given magnitude range; T is the earthquake catalogue period; $f(R)$ is the probability density function of the epicentral distance; R is the epicentral distance [43]; R_t denotes the limit of the epicentral distance; $f(D)$ is the probability density function of the focal depth; D is the focal depth; and a_1, a_2, a_3, a_4 represent the fitting coefficient.

4.2. Cumulative Probability Distribution of the PGA

It is necessary to determine the probability distribution of the PGA in the PSHA and judge the accuracy of the BPNN in supplementing the PGA data through the probability distribution results that were previously obtained through fitting. In this section, all the California data were divided into two categories: Sample A only contains the measured PGA and Sample B contains the measured data and the BPNN prediction data.

Four commonly used asymmetric probability models, namely the Gumbel, Frechet, Weibull, and generalized Pareto distribution (GPD), were selected to model the PGA probability distribution for analysis. The cumulative distribution function (CDF) of the

above four models is shown in Equations (9)–(12), and the values of the parameters in the probability distribution function through regression analysis are listed in Table 1.

$$\text{Gumbel}: \ F(x) = \exp\left[-\exp\left(-\frac{x-a}{b}\right)\right] \tag{9}$$

$$\text{Frechet}: \ F(x) = \exp\left[-\left(-\frac{x-a}{b}\right)^{-c}\right] \tag{10}$$

$$\text{Weibull}: \ F(x) = 1-\exp\left[-\left(\frac{x-a}{b}\right)^{c}\right] \tag{11}$$

$$\text{GPD}: \ F(x) = 1-[1+c(x-a)/b]^{-\frac{1}{c}} \tag{12}$$

Table 1. Regression results of the CDF of the PGA.

Type	CDF	*a*	*b*	*c*	R^2
Sample A	Gumbel	122.4	145.5	——	0.9797
	Frechet	−157.6	268.8	2.179	0.9942
	Weibull	10.17	223.5	0.9294	0.9971
	GPD	6.825	215.6	0.09101	0.9969
Sample B	Gumbel	46.7	42.6	——	0.9724
	Frechet	-3.254×10^4	3.259×10^4	765.3	0.9723
	Weibull	55.5	136	2.958	0.9755
	GPD	6.944	73.03	2.676×10^{-7}	0.9591

The closer that the determinants of the coefficients (R^2) in Table 1 are to one, the better their goodness-of-fit values. Weibull displayed the best goodness-of-fit value in both samples. BPNN prediction had less impact on the regularity of the data, which still conforms to the original probability distribution. The CDFs of the two samples are plotted in Figure 8.

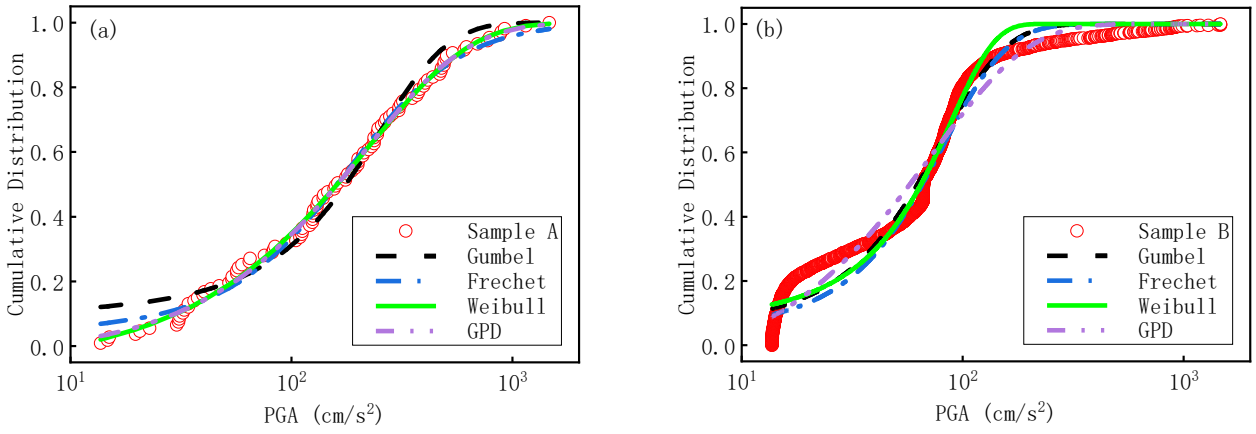

Figure 8. Cumulative distribution of the PGA. (**a**) Sample A. (**b**) Sample B.

It is not difficult to interpret that inflexion points appeared in the regions with more concentrated PGA values. Combining Figures 3 and 5, the known PGAs were often larger ground motion events, while the more frequent earthquakes were relatively concentrated at lower magnitudes. This indicates that the PGA results obtained via the BPNN prediction will increase the number of medium and low PGAs, and the change of this data volume eventually affects the CDF of the PGAs.

4.3. Logic Tree

In the PSHA, the uncertainties underlying the magnitude, focal depth, epicenter distance, *b*-value, PGA data prediction, and the CDF of the PGA have significant effects on the accuracy of the results. Section 3 adjusted the PGA data prediction method through the use of the BPNN, while improving the uncertainty of the magnitude and the uncertainty of the focal depth through the regression analysis. The logic tree method can be used to refine the uncertainty caused by other factors.

Branch I (Figure 9) is used to consider the uncertainty of the *b*-value, whose value can be determined from the seismic data. According to Roshan [44], three values of the *b*-value, namely $b - \varepsilon$, b, and $b + \varepsilon$, as calculated from the maximum likelihood method, are considered. For the current database, the *b*-value that was estimated using the maximum likelihood method was calculated to be 0.9125 ($\pm$0.0718). A 75% weightage was assigned to the estimated value that was estimated from the database and the bounds on the *b*-value following the accounting error were each assigned a weightage of 12.5%.

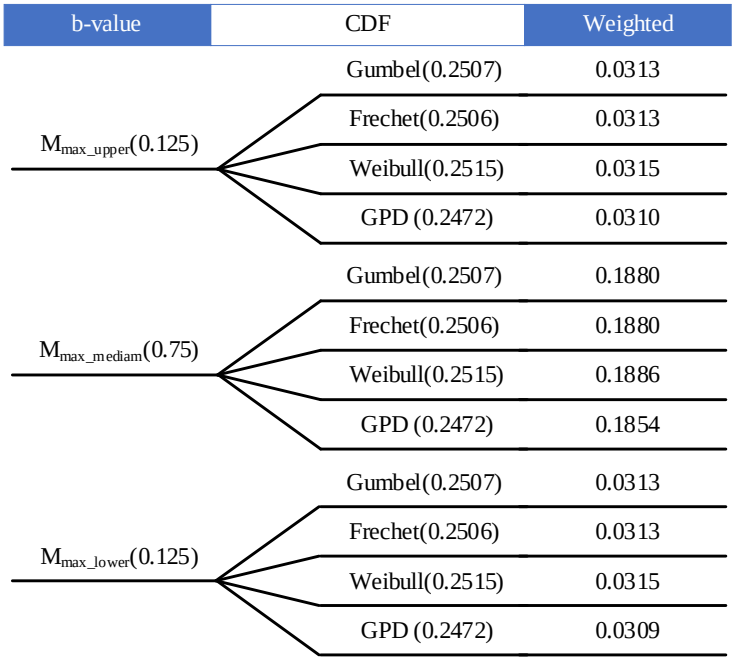

Figure 9. Logic tree used for PSHA calculations.

Branch II is used to consider the uncertainty of the CDF. Ayman Shama [35] took equal weights for each model. Villani Manuela [45] combined a stochastic model weight of 35% and an empirical model weight of 65% to fit the ground motion results. DU W [46] assigned the 10 published GMPEs to three zones: the subduction zone, the strike-slip fault, and the deformation background zone, with different weights. The models that were selected in this study took generality into account, meaning that each model had a good fit, but there was also a certain gap, and using equal weights will pay too much attention to the models with a lower fit. Thus, this paper assigned weights according to the weighted average of the fit coefficients in Table 1. The branch III of the logic tree is the final weight of each item.

5. Results

After considering the influence of the uncertainty factors on the PSHA using the logic tree, the annual average occurrence of seismic events regarding the PGA can be determined according to Equation (3). In Section 4.2, the Sample B database had a similar fitting result as Sample A. Meanwhile, in order to determine the advancement in the BPNN prediction results, this paper plotted the sample B database in Figure 10 and compared it with the results in the specification.

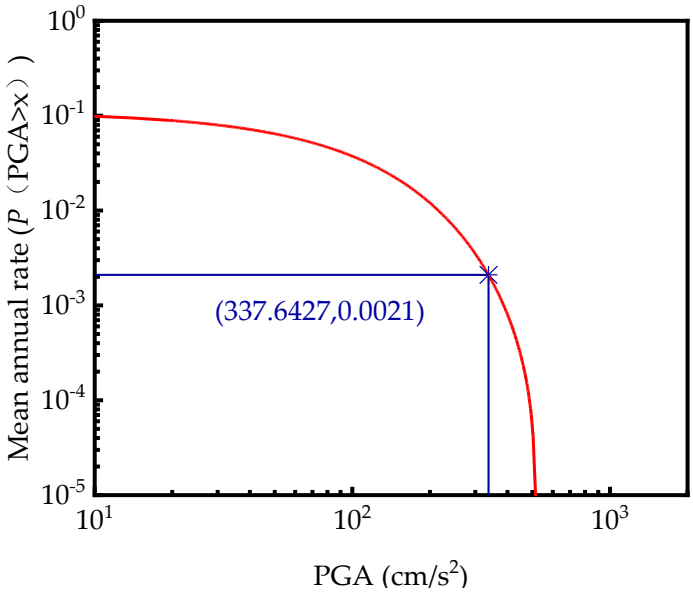

Figure 10. Hazard curves.

Assuming that the number of seismic events in a given time obeys the Poisson distribution [47], the probability of k occurrences of a seismic event can be expressed as:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{13}$$

The probability that no earthquake occurs in a given time can be expressed by taking a value of zero for k. Therefore, the probability that an earthquake event satisfying the PGA is greater than a given limit, x, which occurs at least once a year, t, can be expressed as:

$$P(\text{PGA} > x \text{ in } t \text{ years}) = 1 - e^{-P(\text{PGA}>x)t} \tag{14}$$

The inverse of Equation (14) yields the formula for the average annual transcendence rate of a certain intensity seismic event:

$$P(\text{PGA} > x) = \frac{\ln[1 - P(\text{PGA} > x \text{ in } t \text{ years})]}{t} \tag{15}$$

Comparing the PGAs for the California zone and ASCE 7-16 [48], the probability of exceeding the ground motion in 50 years was found to be equal to 10%. With t = 50, the average annual occurrence corresponding to a given intensity seismic event was determined to be 2.10×10^{-3}. Based on the seismic hazard curves determined in this section, the PGA value of the corresponding average annual occurrence seismic event was calculated as 0.345 g. ASCE 7-16 provides two different frequencies of spectral acceleration across the California area, where the short-period PGA was found to be larger compared to the long-period PGA. Hence, the 0.2 s spectral acceleration for the main seismic zone located in California was selected and converted for the Class D sites, resulting in about 0.353 g. The error between the two was about 2.27%. The calculation results obtained in this study underestimate the seismic hazard of the region, probably due to the fact that the data used in this paper covered the whole region of California, which contains some areas with the lower data, and the data used as the comparison terms are the results of the denser seismic areas in the California region. Moreover, the certain gap observed in the initial data caused the final error.

6. Conclusions

In this paper, the BPNN prediction PGA approach has been proposed instead of the GMPE, which takes into account the effect of the focal depth and is based on the full

probability formula, while adding the focal depth to the analysis of the PSHA. Following the calculation with data from the California area as an example, the results are compared with the ASCE results, and the main conclusions are drawn as follows:

(1) The results of the seismic hazard curves show that the probability of exceeding the ground motion equal to 10% in 50 years for the California region is in close agreement with the data in ASCE 7-16, with an error of 2.27%.
(2) The BPNN method of predicting the PGA is convenient for the PSHA in areas with unknown geological conditions, which can be only performed using the seismic catalogues alone and can ensure the accuracy and generalizability of the analysis results while having the ability to take into account the non-linear relationships that exist between multiple parameters.
(3) Based on the full probability formula, adding the effect of the focal depth on the annual average transcendental probability formula for the seismic events that satisfy the PGA greater than a given value can improve the PSHA.
(4) When determining the weights of the CDFs through the logic tree, the use of weighting the goodness-of-fit coefficients enables more emphasis to be placed on the CDFs with better fitting results, but without biasing the results due to overemphasis.

The choice of using the BPNN to predict the PGA in this paper raises the possibility of adding more influencing factors to the PSHA. This method can be used in more areas that have more earthquake data, with a better accuracy and generalizability, making a tentative step for future scholars' research.

Author Contributions: Conceptualization, X.G., H.L. and H.Z.; software, X.G. and Q.W.; data curation, X.G. and J.X.; writing—original draft preparation, X.G.; writing—review and editing, H.L. and H.Z.; funding acquisition, H.L. and H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China (Grant No. 2016YFC0701108), the Program of the Educational Department of Liaoning Province (Grant No. LJKMZ20220949), and the Program of Shenyang Bureau of Science and Technology (Grant No. RC220171).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, H.N.; Liu, Y.; Li, C.; Zheng, X.W. Multihazard fragility assessment of steel-concrete composite frame structures with buckling-restrained braces subjected to combined earthquake and wind. *Struct. Des. Tall Spec. Build.* **2020**, *29*, e1746. [CrossRef]
2. Gerstenberger, M.C.; Marzocchi, W.; Allen, T.; Pagani, M.; Adams, J.; Danciu, L.; Field, E.H.; Fujiwara, H.; Luco, N.; Ma, K.-F.; et al. Probabilistic Seismic Hazard Analysis at Regional and National Scales: State of the Art and Future Challenges. *Rev. Geophys.* **2020**, *58*, e2019RG000653. [CrossRef]
3. Mehta, P.; Thaker, T.P. Seismic Hazard Analysis of Vadodara Region, Gujarat, India: Probabilistic & Deterministic Approach. *J. Earthq. Eng.* **2020**, *26*, 1438–1460.
4. Cornell, C.A. Engineering seismic risk analysis. *Bull. Seismol. Soc. Am.* **1968**, *58*, 1583–1606. [CrossRef]
5. Chioccarelli, E.; Cito, P.; Iervolino, I.; Giorgio, M. REASSESS V2.0: Software for single- and multi-site probabilistic seismic hazard analysis. *Bull. Earthq. Eng.* **2018**, *17*, 1769–1793. [CrossRef]
6. Ebrahimian, H.; Jalayer, F.; Forte, G.; Convertito, V.; Licata, V.; d'Onofrio, A.; Santo, A.; Silvestri, F. Site-specific probabilistic seismic hazard analysis for the western area of Naples, Italy. *Bull. Earthq. Eng.* **2019**, *17*, 4743–4796. [CrossRef]
7. Zanini, M.A.; Hofer, L.; Toska, K. A semi-analytical formulation for accounting uncertainties of hazard parameters in structural seismic reliability analysis. *Eng. Struct.* **2019**, *192*, 18–29. [CrossRef]
8. Falero, V.M.; Bommer, J.J. Virtual Fault Ruptures in Area-Source Zones for PSHA: Are They Always Needed? *Seismol. Res. Lett.* **2020**, *91*, 2310–2319. [CrossRef]
9. Pagani, M.; Hao, K.X.; Fujiwara, H.; Gerstenberger, M.; Ma, K.-F. Appraising the PSHA Earthquake Source Models of Japan, New Zealand, and Taiwan. *Seismol. Res. Lett.* **2016**, *87*, 1240–1253. [CrossRef]

10. Li, B.; Cai, Z.; Xie, W.-C.; Pandey, M. Probabilistic seismic hazard analysis considering site-specific soil effects. *Soil Dyn. Earthq. Eng.* **2018**, *105*, 103–113. [CrossRef]
11. Jones, R.H.; Stewart, R.C. A method for determining significant structures in a cloud of earthquakes. *J. Geophys. Res. Solid Earth* **1997**, *102*, 8245–8254. [CrossRef]
12. Wang, Z.; Kao, H. The significance of tomographic edge zones for large earthquakes in Taiwan. *J. Geophys. Res. Solid Earth* **2019**, *124*, 11822–11839. [CrossRef]
13. Arora, B.R. Introduction to Seismology. *Episodes* **2010**, *33*, 125–126. [CrossRef]
14. Gómez-Novell, O.; García-Mayordomo, J.; Ortuño, M.; Masana, E.; Chartier, T. Fault System-Based Probabilistic Seismic Hazard Assessment of a Moderate Seismicity Region: The Eastern Betics Shear Zone (SE Spain). *Front. Earth Sci.* **2020**, *8*. [CrossRef]
15. Valentini, A.; Pace, B.; Boncio, P.; Visini, F.; Pagliaroli, A.; Pergalani, F. Definition of Seismic Input From Fault-Based PSHA: Remarks After the 2016 Central Italy Earthquake Sequence. *Tecton* **2019**, *38*, 595–620. [CrossRef]
16. Barani, S.; Ferretti, G.; De Ferrari, R. Incorporating results from seismic microzonation into probabilistic seismic hazard analysis: An example in western Liguria (Italy). *Eng. Geol.* **2020**, *267*, 105479. [CrossRef]
17. Assatourians, K.; Atkinson, G.M. Implementation of a Smoothed-Seismicity Algorithm in Monte Carlo PSHA Software EqHaz and Implications for Localization of Hazard in the Western Canada Sedimentary Basin. *Seismol. Res. Lett.* **2019**, *90*, 1407–1419. [CrossRef]
18. Li, B.; Sørensen, M.B.; Atakan, K.; Li, Y.; Li, Z. Probabilistic Seismic Hazard Assessment for the Shanxi Rift System, North China. *Bull. Seismol. Soc. Am.* **2020**, *110*, 127–153. [CrossRef]
19. Ahmed, M.; Lodi, S.H.; Rafi, M.M. Probabilistic Seismic Hazard Analysis Based Zoning Map of Pakistan. *J. Earthq. Eng.* **2019**, *26*, 1–36. [CrossRef]
20. Douglas, J. Earthquake ground motion estimation using strong-motion records: A review of equations for the estimation of peak ground acceleration and response spectral ordinates. *Earth Sci. Rev.* **2003**, *61*, 43–104. [CrossRef]
21. Phung, V.-B.; Loh, C.H.; Chao, S.H.; Abrahamson, N.A. Ground motion prediction equation for Taiwan subduction zone earthquakes. *Earthq. Spectra* **2020**, *36*, 1331–1358. [CrossRef]
22. Barani, S.; Albarello, D.; Massa, M.; Spallarossa, D. Influence of Twenty Years of Research on Ground-Motion Prediction Equations on Probabilistic Seismic Hazard in Italy. *Bull. Seismol. Soc. Am.* **2017**, *107*, 240–255. [CrossRef]
23. Milner, K.R.; Shaw, B.E.; Goulet, C.A.; Richards-Dinger, K.B.; Callaghan, S.; Jordan, T.H.; Dieterich, J.H.; Field, E.H. Toward Physics-Based Nonergodic PSHA: A Prototype Fully Deterministic Seismic Hazard Model for Southern California. *Bull. Seismol. Soc. Am.* **2021**, *111*, 898–915. [CrossRef]
24. Yazdani, A.; Shahidzadeh, M.-S.; Takada, T. Merging data and experts' knowledge-based weights for ranking GMPEs. *Earthq. Spectra* **2020**, *37*, 857–875. [CrossRef]
25. Villani, M.; Polidoro, B.; McCully, R.; Ader, T.; Edwards, B.; Rietbrock, A.; Lubkowski, Z.; Courtney, T.J.; Walsh, M. A Selection of GMPEs for the United Kingdom Based on Instrumental and Macroseismic Datasets. *Bull. Seismol. Soc. Am.* **2019**, *109*, 1378–1400. [CrossRef]
26. Anbazhagan, P.; Sreenivas, M.; Bajaj, K.; Moustafa, S.S.R.; Al-Arifi, N. Selection of Ground Motion Prediction Equations for Seismic Hazard Analysis of Peninsular India. *J. Earthq. Eng.* **2015**, *20*, 699–737. [CrossRef]
27. Herrmann, R.B.; Nuttli, O.W. Ground-motion modelling at regional distances for earthquakes in a continental interior, II. Effect of focal depth, azimuth and attenuation. *Earthq. Eng. Struct. Dyn.* **1975**, *4*, 59–72. [CrossRef]
28. Salditch, L.; Gallahue, M.M.; Lucas, M.C.; Hough, S.E.; Neely, J.S.; Stein, S. California Historical Intensity Mapping Project (CHIMP): A Consistently Reinterpreted Dataset of Seismic Intensities for the Past 162 Yr and Implications for Seismic Hazard Maps. *Seismol. Res. Lett.* **2020**, *91*, 2631–2650. [CrossRef]
29. Chiou, B.; Darragh, R.; Gregor, N.; Silva, W. NGA Project Strong-Motion Database. *Earthq. Spectra* **2008**, *24*, 23–44. [CrossRef]
30. Haddadi, H.R.; Shakal, A.F.; Stephens, C.D.; Savage, W.; Huang, M.; Leith, W.; Parrish, J. *Center for Engineering Strong Motion Data*; USGS: Reston, VA, USA, 2010.
31. Archuleta, R.J.; Steidl, J.; Squibb, M. The COSMOS Virtual Data Center: A Web Portal for Strong Motion Data Dissemination. *Seismol. Res. Lett.* **2006**, *77*, 651–658. [CrossRef]
32. Fernández-Nóvoa, D.; Costoya, X.; Decastro, M.; Gesteira, M.G. Influence of the mightiest rivers worldwide on coastal sea surface temperature warming. *Sci. Total Environ.* **2021**, *768*, 144915. [CrossRef] [PubMed]
33. Wei, M.; Hu, X.; Yuan, H. Residual displacement estimation of the bilinear SDOF systems under the near-fault ground motions using the BP neural network. *Adv. Struct. Eng.* **2022**, *25*, 552–571. [CrossRef]
34. Lin, J.-W.; Chiou, J.-S. Active Probability Backpropagation Neural Network Model for Monthly Prediction of Probabilistic Seismic Hazard Analysis in Taiwan. *IEEE Access* **2019**, *7*, 108990–109014. [CrossRef]
35. Shama, A. Site specific probabilistic seismic hazard analysis at Dubai Creek on the west coast of UAE. *Earthq. Eng. Eng. Vib.* **2011**, *10*, 143–152. [CrossRef]
36. Wu, J.; Gao, M.; Huang, B.; Chen, K. Discussion on the influence of truncation of ground motion residual distribution on probabilistic seismic hazard assessment. *Earthq. Eng. Eng. Vib.* **2011**, *10*, 379–392. [CrossRef]
37. Ismailov, V.E. On the approximation by neural networks with bounded number of neurons in hidden layers. *J. Math. Anal. Appl.* **2014**, *417*, 963–969. [CrossRef]

38. Zheng, X.-W.; Li, H.-N.; Yang, Y.-B.; Li, G.; Huo, L.-S.; Liu, Y. Damage risk assessment of a high-rise building against multihazard of earthquake and strong wind with recorded data. *Eng. Struct.* **2019**, *200*, 109697. [CrossRef]
39. Kiureghian, A.D.; Ang, H.S. A fault-rupture model for seismic risk analysis. *Bull. Seismol. Soc. Am.* **1977**, *67*, 1173–1194.
40. Frankel, A. Mapping Seismic Hazard in the Central and Eastern United States. *Seismol. Res. Lett.* **1995**, *66*, 8–21. [CrossRef]
41. Bommer, J.J.; Crowley, H. The Purpose and Definition of the Minimum Magnitude Limit in PSHA Calculations. *Seismol. Res. Lett.* **2017**, *88*, 1097–1106. [CrossRef]
42. Gutenberg, B.; Richter, C.F. Frequency of earthquakes in California. *Bull. Seismol. Soc. Am.* **1944**, *34*, 185–188. [CrossRef]
43. Ordaz, M.; Salgado-Gálvez, M.A.; Giraldo, S. R-CRISIS: 35 years of continuous developments and improvements for probabilistic seismic hazard analysis. *Bull. Earthq. Eng.* **2021**, *19*, 2797–2816. [CrossRef]
44. Roshan, A.D.; Basu, P.C. Application of PSHA in low seismic region: A case study on NPP site in peninsular India. *Nucl. Eng. Des.* **2010**, *240*, 3443–3454. [CrossRef]
45. Villani, M.; Lubkowski, Z.; Free, M.; Musson, R.M.W.; Polidoro, B.; McCully, R.; Koskosidi, A.; Oakman, C.; Courtney, T.; Walsh, M. A probabilistic seismic hazard assessment for Wylfa Newydd, a new nuclear site in the United Kingdom. *Bull. Earthq. Eng.* **2020**, *18*, 4061–4089. [CrossRef]
46. Du, W.; Pan, T.-C. Probabilistic seismic hazard assessment for Singapore. *Nat. Hazards* **2020**, *103*, 2883–2903. [CrossRef]
47. Pothon, A.; Gueguen, P.; Buisine, S.; Bard, P.-Y. Comparing Probabilistic Seismic Hazard Maps with ShakeMap Footprints for Indonesia. *Seismol. Res. Lett.* **2020**, *91*, 847–858. [CrossRef]
48. American Society of Civil Engineers. *Minimum Design Loads for Buildings and Other Structures*; ASCE: Reston, VA, USA, 2017.