Modeling Maximum Tsunami Heights Using Bayesian Neural Networks

Min-Jong Song and Yong-Sik Cho *

Department of Civil and Environmental Engineering, Hanyang University Seongdong-gu, Seoul 04763, Korea; thdcnd3517@gmail.com
* Correspondence: ysc59@hanyang.ac.kr; Tel.: +82-2-2200-0393

check for
updates

Abstract: Tsunamis are distinguished from ordinary waves and currents owing to their characteristic longer wavelengths. Although the occurrence frequency of tsunamis is low, it can contribute to the loss of a large number of human lives as well as property damage. To date, tsunami research has concentrated on developing numerical models to predict tsunami heights and run-up heights with improved accuracy because hydraulic experiments are associated with high costs for laboratory installation and maintenance. Recently, artificial intelligence has been developed and has revealed outstanding performance in science and engineering fields. In this study, we estimated the maximum tsunami heights for virtual tsunamis. Tsunami numerical simulation was performed to obtain tsunami height profiles for historical tsunamis and virtual tsunamis. Subsequently, Bayesian neural networks were employed to predict maximum tsunami heights for virtual tsunamis.

Keywords: tsunami; machine learning; bayesian neural networks; numerical simulation; maximum tsunami heights

1. Introduction

Tsunamis triggered by undersea earthquakes around the Pacific Ocean have frequently caused loss of human life and property damage in coastal areas. To date, two huge tsunamis in history occurred in 2004 and 2011. The Sumatra tsunami event in 2004 caused approximately 220,000 human deaths and 10 billion dollars in property damage. More recently, the East Japan tsunami event in 2011 caused over 20,000 casualties and 300 billion dollars in property damage [1] and incited nuclear power plant meltdowns at Fukushima in Japan. In the field of tsunami research, data analysis and hydraulic experiments for tsunami events are challenging because major tsunamis are rare and laboratory installations for hydraulic experiments are exceedingly expensive. For these reasons, many researchers have focused on developing numerical models to predict accurately tsunami heights defined as vertical height from the troughs to crests of tsunamis and run-up heights defined as vertical heights from the datum levels to the highest inundated areas for several decades [2–6]. Recently, artificial intelligence (AI), using big data, has been developed and applied in many science and engineering fields with significant performances. In machine learning, there are two main methods: deterministic and probabilistic methods. The representatives of the deterministic method are Multi-layer perceptron (MLP), convolution neural networks (CNNs), and recurrent neural networks (RNNs) [7]. Meanwhile, probabilistic random forest and Bayesian neural networks (BNNs) can be classified as probabilistic methods [8,9]. In this study, BNNs, which consist of neural networks (NNs) and a Bayesian inference, are proposed. One advantage of a BNN is that it enables prediction under insufficient data. Moreover, it not only considers the relationships between data but rarely causes overfitting problems, which are a major issue in NNs [10]. BNNs include three processes: training, testing, and prediction. Training is performed to search for optimized parameters with given input variables on BNNs. In the testing,

so-called model validation, the performance of the trained weights and bias are estimated by applying it to new data that were not part of the training dataset. Finally, predictions are made based on the provided input variables. This study explores the estimation of maximum tsunami heights with BNNs. Jumunjin Port, located on the eastern coast of the Korean Peninsula, was offered to target location. Inundation damage was caused by a tsunami in 1983 [11]. A numerical simulation was performed for two historical tsunamis and virtual tsunamis, which could be triggered by probable undersea earthquake in the future [12,13], to obtain tsunami height profiles and evaluate the maximum tsunami heights. BNNs were employed, and the pairs of tsunami height profiles at the site of the tsunami epicenter and Jumunjin port (37.89 N, 128.83 E) considering two historical tsunami events: tsunami events in 1983 (40.54 N, 139.02 E) and 1993 (42.34 N, 139.25 E), were used as the input variables for training and testing. Finally, BNNs were used to predict the maximum tsunami heights for the virtual tsunamis.

2. Numerical Simulation

A tsunami that is triggered along the western coast of Japan can travel to the Korean Peninsula. The propagation model for a tsunami traveling across deep water should account for frequency dispersion and the Coriolis force. Thus, the nonlinear convective terms of the momentum equations can be ignored in the numerical simulation because the free surface displacement of a tsunami is relatively lower than the water depth. However, the frequency dispersion effects are influential in distance propagation and should be considered in the model's governing equations because the wave length of a tsunami is relatively short compared to that of the tide and tsunamis are propagated over long distances. Therefore, the governing equations that address the dispersion effect, namely, the linear Boussinesq equations, should be used in the numerical simulation of tsunami propagation. However, the dispersion terms of the linear Boussinesq equations include higher-order derivatives that may cause practical numerical difficulty. To resolve this difficulty, linear shallow-water equations in a set of lower-order derivatives are used for dispersion effects [6,14].

2.1. Propagation

For a tsunami propagating far across a deep ocean, a linear shallow-water equation in a set of lower-order derivatives is used. The linear shallow-water equations can be expressed in the form as below,

$$\frac{\partial \eta}{\partial t} + \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} = 0 \tag{1}$$

$$\frac{\partial P}{\partial t} + gh\frac{\partial \eta}{\partial x} = 0 \tag{2}$$

$$\frac{\partial Q}{\partial t} + gh\frac{\partial \eta}{\partial y} = 0 \tag{3}$$

where η is the free-surface displacement, h is the mean water depth, g is the acceleration of gravity, and P and Q are the volume fluxes in the x and y directions, respectively, with depth-averaged flow velocities at the still-water level.

2.2. Initial Free Surface Displacement

Two historical tsunami events—the Central East Sea tsunami of 1983 and the Hokkaido tsunami in 1993—as well as the virtual tsunamis were employed for the tsunami numerical simulation (Figure 1). The two selected historical tsunamis are representative of tsunamis that have affected Korea in the last 100 years. The virtual tsunamis were selected based on the seismic gap where potential earthquakes could strike in the vertically developed fault zone along the western coast of Japan. Table 1 presents the fault parameters of the historical tsunamis and virtual tsunamis: H is the height at the top of the fault plane, θ is the strike angle, δ is the dip angle, λ is the slip angle, L is the length, W is the width,

and D is the dislocation where the plane is displaced (Figure 2). The source locations, strike angles, heights, dip angles, and slip angles were obtained from the Korean Peninsula Energy Development Organization [12]. The length, width, and dislocation considering the magnitude (M) were calculated according to the equations presented below,

$$logL = 0.5M - 1.9 \tag{4}$$
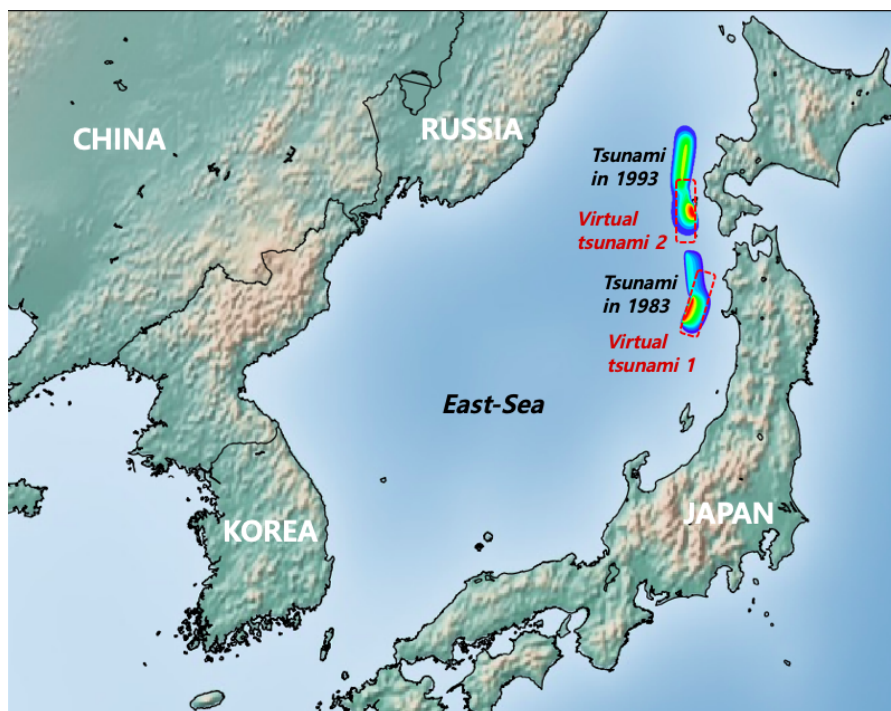
$$W = 0.5L \tag{5}$$

$$logD = 0.5M - 1.2 \tag{6}$$

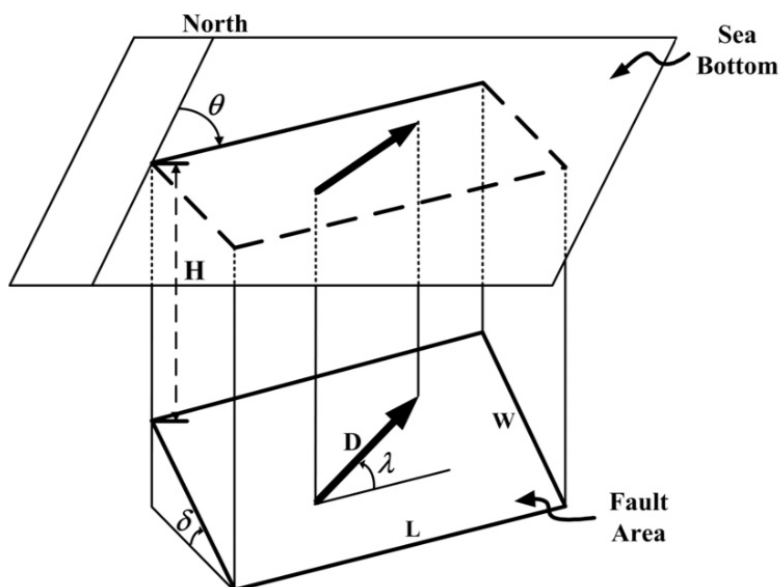**Figure 1.** Location of historical tsunamis and virtual tsunamis.

Figure 2. Definition of fault parameters.

Table 1. Fault parameters of historical tsunamis and virtual tsunamis.

| Classification	Longitude (°E)	Latitude (°N)	H (km)	θ (°)	δ (°)	λ (°)	L (km)	W (km)	D (km)	M
Tsunami in 1983	139.02	40.54	3	355	25	80	60	30	3.05	7.7
	139.30	42.10	5	163	60	105	24.5	25	12.0	
Tsunami in 1993	139.25	42.34	5	175	60	105	30	25	2.50	7.8
	139.40	43.13	10	188	35	80	90	25	5.71	
Virtual tsunami 1	138.70	40.20	1	10.0	40	90	125.9	62.9	6.31	8.0
Virtual tsunami 2	138.90	40.90	1	5	40	90	125.9	62.9	6.31	8.0

3. Bayesian Neural Networks

Neural networks are characterized by two main trends: deterministic and probabilistic methods. BNNs, based on the Bayes' theorem [15], utilize a probabilistic method that combines NNs and Bayesian inference. A robust advantage of a BNN is that it offers more information with probability distribution than deterministic method such as MLP or CNNs or RNNs.

3.1. Neural Networks

The MLP, which is a fundamental NN, is explained in this section. Given a dataset X, which is an input vector composed of several causal variables, and Y, which is an output vector composed of several resulting variables, the NNs are conducted. Most NNs consist of nodes, links, and layers with input, hidden, and output layers. The input layer receives input variables for the problem to be dealt with. In the output layer, the prediction of values is evaluated through a loss function, and the final output is represented. In the hidden layers, a trial-and-error procedure is performed. Every layer has nodes, and these nodes are connected by links, with synaptic weight assigned to each link to represent the relative connection strength [16]. In mathematics, NNs are based on the parameterized basis function as follows,

$$f(x) = \phi(wx + b) \tag{7}$$

where ϕ is the activation function that leads to a nonlinear effect, x represents the input variables, w is the weight, and b is the bias. Tangent hyperbolic functions or sigmoid functions, expressed below, were frequently used in the beginning of research as activation functions; however, ReLU or Selu functions have been developed and widely used more recently (Figure 3).
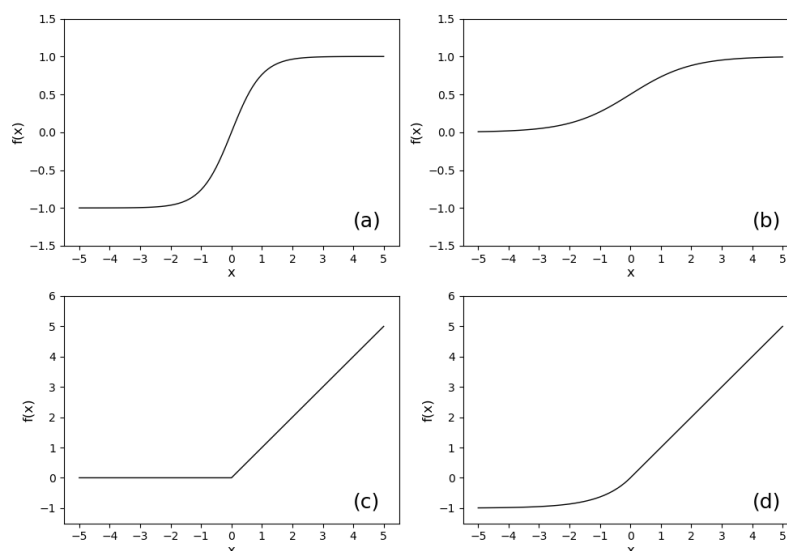
**Figure 3.** Activation functions: (**a**) Tangent hyperbolic, (**b**) Sigmoid, (**c**) ReLU, and (**d**) Selu functions.

The loss function is employed at nodes in the output layer and expressed as follows,

$$E(X,Y) = \frac{1}{2N} \sum_{i=1}^{N} (Y - y')^2 \qquad (8)$$

where N is the number of observations, Y is the output vector, and y' is the value predicted by the networks. If the result is not sufficient to objective error after the loss function is evaluated, it is necessary to adjust the weight and bias. The algorithm was introduced as follows [17],

$$\Delta w_{ij}(n) = -\eta \frac{\partial E}{\partial w_{ij}} \qquad (9)$$

where Δw_{ij} is the adjusted weight and η is the learning rate. One of the most noticeable differences between NNs and BNNs is the method of computing the weight and bias. The BNNs employ a probability distribution for the weight and bias (Figure 4).
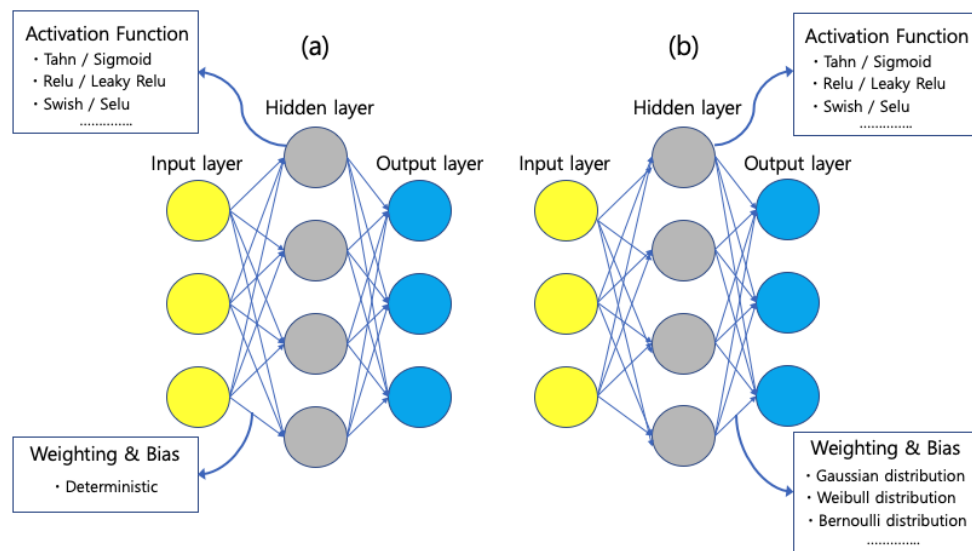
Figure 4. Architecture of (**a**) neural networks (NNs) and (**b**) Bayesian neural network (BNNs).

3.2. Bayesian Inference

Bayesian inference, which has been used in many science and engineering fields, is based on Bayes' theorem. It offers posterior probability based on evidence, prior probability, and marginal likelihood:

$$P(w|X,Y) = \frac{P(Y|X,w)P(w)}{P(Y|X)} \propto P(Y|X,w)P(w) \qquad (10)$$

where X and Y are the datasets, and w is the weight. $P(Y|X)$ is the evidence, $P(w)$ is the prior probability, $P(Y|X,w)$ is the marginal likelihood, and $P(w|X,Y)$ is the posterior probability. In the BNNs, given the data X, Y, and the initial $P(w)$, the evidence, marginal likelihood, and posterior probability can be evaluated for weight at each link in all layers. Meanwhile, it is well known that predicting an analytical solution for posterior probability is typically difficult $P(w|X,Y)$. Therefore, an approximating variational distribution $q_\pi(w)$ is used in this study [18], and an automatic differentiation variational inference (ADVI) scheme that uses a spherical Gaussian without the correlation of parameters was introduced for variational distribution. The goal was to approximate the posterior distribution $P(w|X,Y)$:

$$P(w|X,Y) \approx q_\pi^* \qquad (11)$$

where X and Y are the datasets, and q_π^* is the minimum of this optimization objective.

$$KL(q_\pi(w)|P(w|X,Y)) = \int q_\pi(w)log(\frac{q_\pi(w)}{P(w|X,Y)})dw \tag{12}$$

$$L(\pi) = \int q_\pi(w)P(Y|X,w)dw - KL(q_\pi(w)|P(w|X,Y)) \tag{13}$$

Kullback–Leibler divergence (*KL*) was performed to minimize the distance between the variational distribution $q_\pi(w)$ and posterior distribution $P(w|X,Y)$, which implies that the evidence lower bound (ELBO, $L(\pi)$) is at a maximum.

4. Results

In this study, numerical simulations for historical tsunamis and virtual tsunamis were performed at Jumunjin Port, where a 1983 tsunami caused inundation (Figure 5). Tsunami height profiles used on BNNs were extracted in the epicenter (G1 or G2) and at the gage point in Jumunjin (G3) for 175 min (Figure 6).
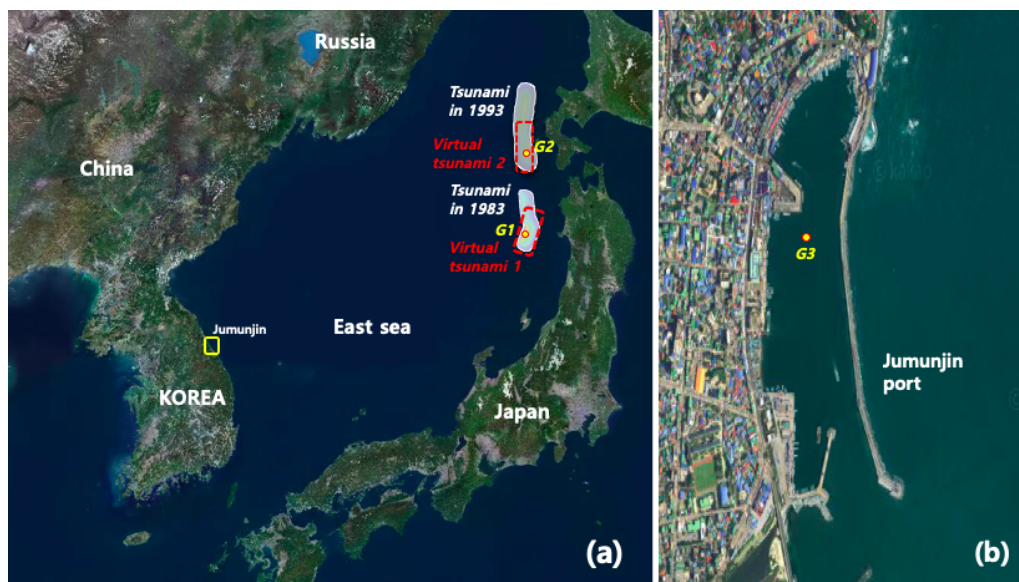

Figure 5. Location of (**a**) epicenter for historical tsunamis and virtual tsunamis and (**b**) gage in Jumunjin port.
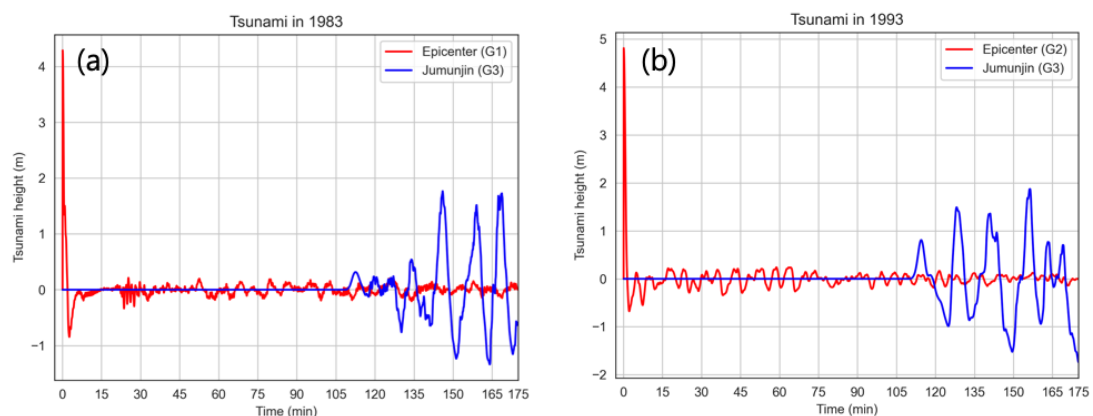

Figure 6. Time series of tsunami heights at epicenter and Jumunjin port by tsunami in (**a**) 1983 and (**b**) 1993.

BNNs were constructed with a combination of NNs and Bayesian inference. First, two nodes for tsunami height and time profiles in the input layer and arbitrary ten nodes in each of the three hidden layers are introduced. Moreover, a tangent hyperbolic function is adopted at the nodes to describe nonlinear effects in NNs. The assumption of Gaussian distribution for weights and bias, expressed in Equation (14), are accepted as prior distribution for links on BNNs [19]. Based on these conditions, BNNs were established to evaluate the maximum tsunami height for virtual tsunamis following three steps: training, testing, and prediction. Prior to estimating the maximum tsunami height, optimization over parameters of Gaussian distribution was carried out owing to effects of parameters of prior distribution as small dataset [10],

$$f(x) = \frac{1}{\sigma\sqrt{(2\pi)}} exp\left(- \frac{(x - \mu)^2}{2\sigma^2} \right) \tag{14}$$

where μ is mean value and σ^2 expresses standard deviation over tsunami height.

4.1. Training

Training plays a key role in BNN prediction. It is well known that normalization can improve performance and contribute to high-speed training of NNs [20]. Therefore, given tsunami height profiles, normalization, expressed below, was performed for tsunami height and time profiles at the epicenter and gage point prior to training:

$$\zeta_i = \frac{H_i - min(H)}{max(H) - min(H)} \tag{15}$$

where H and ζ are the original and normalized values, respectively. The training and testing datasets were collected along the arrival of the tsunamis at G3 in Jumunjin Port over historical tsunami events, which was estimated to approximately 1300 sample size for both tsunami in 1983 and in 1993. Following the works of Afshin G. et al. (2018) [21] and Amazon (2020) [22], 80% of the collected datasets were randomly selected for training (Figure 7), and these datasets are used for updating weights and bias on links. Estimated weights and bias at all links were checked for test and prediction processes.
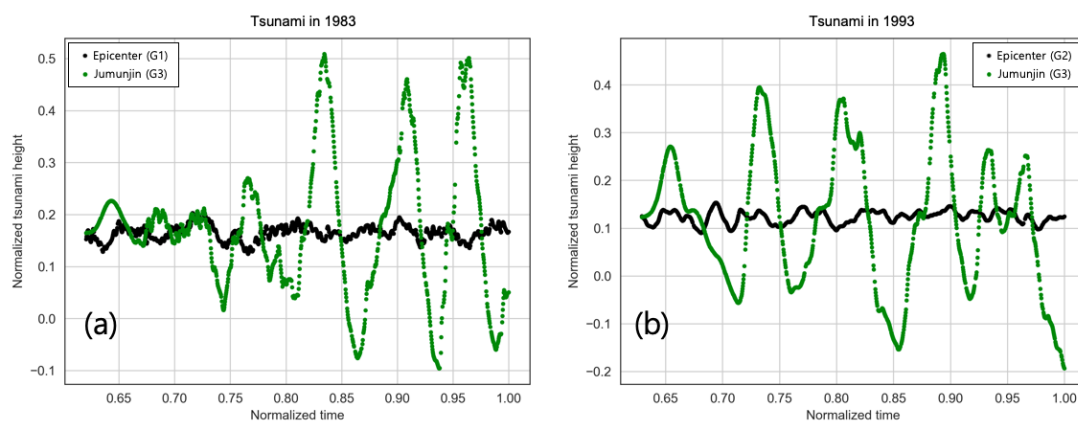

Figure 7. Training dataset of tsunami in (**a**) 1983 and (**b**) 1993.

Moreover, it is necessary to optimize parameters of Gaussian distribution for maximum tsunami heights. In this study, given training datasets, optimization estimation for parameters over historical tsunamis is conducted beforehand. Figure 8 shows the mean difference fields, which describe the absolute difference between numerical simulation and BNNs according to mean (μ) and standard deviation (σ^2) for maximum tsunami height. Figure 8a,b represents tsunamis in 1983. Figure 8a reveals that the minimum of mean difference field is concentrated in the range of 2.6 to 3.5 standard deviation. In this study, through the estimation of each standard deviation where mean difference

demonstrates the minimum, the parameters of Gaussian distribution is estimated to 0.0 mean at 2.9 standard deviation as a prior distribution (Figure 8c). Figure 7b,d shows the 1993 tsunami data. Unlike the tsunami in 1983, the minimum mean difference was distributed from 2.0 to 2.6 (Figure 8b), and the parameters are considered for 0.2 and 0.3 mean at 2.5 standard deviations, where the mean difference becomes small (Figure 8d), as a prior distribution to determine the best performance. Based on training datasets and prior distribution for both tsunamis in 1983 and 1993, training is examined with ADVI. Figure 9 shows how ELBO reached a stable state where KL was minimized. In this study, the ELBO approached stable state after 35,000 of the 50,000 over historical tsunami events.
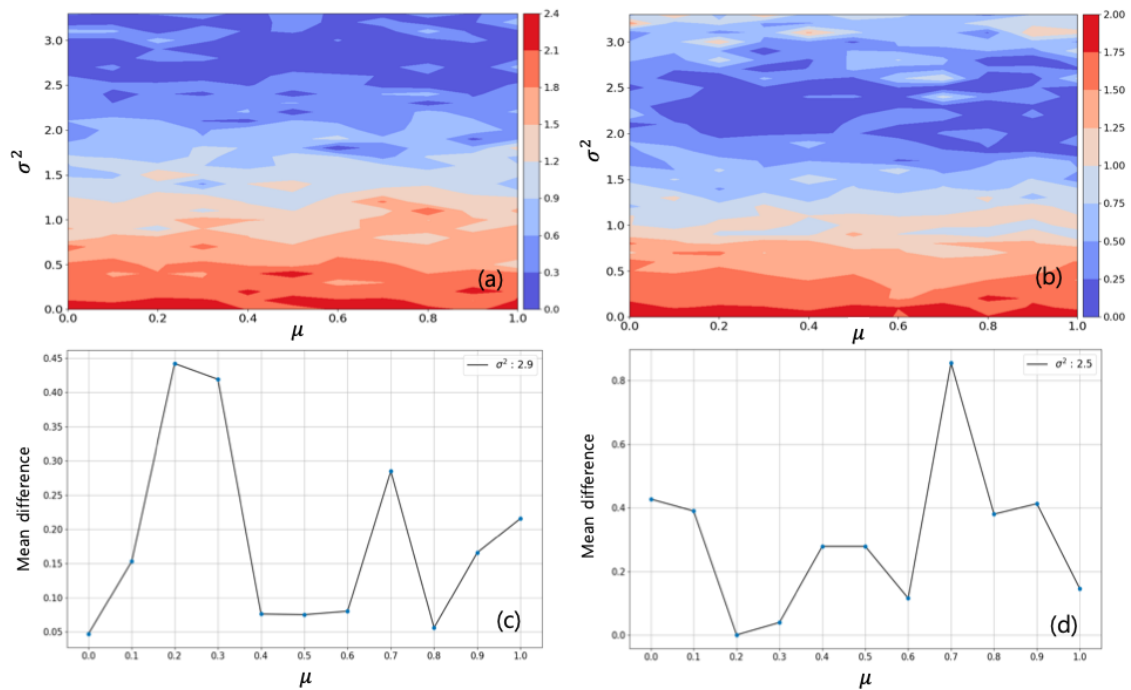
Figure 8. Mean difference fields of tsunami in (**a**) 1983 and (**b**) 1993, and the variation of mean difference along mean at stationary standard deviation over tsunami in (**c**) 1983 and (**d**) 1993.
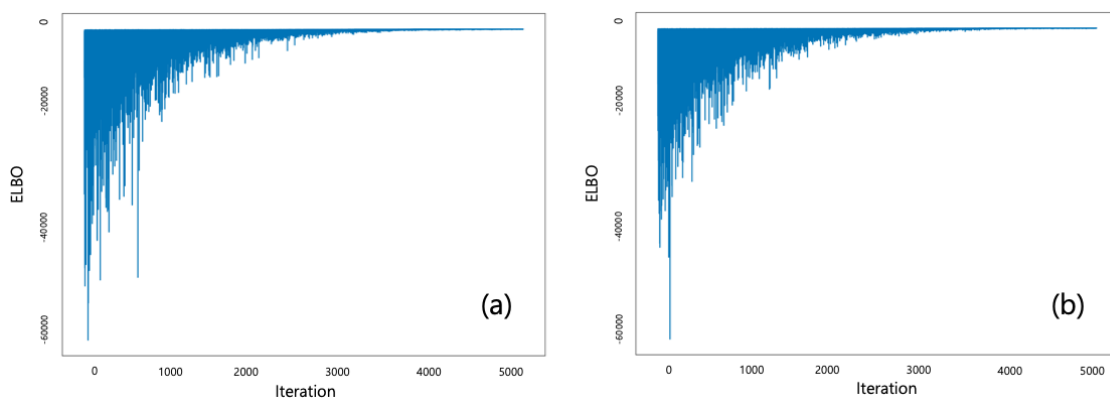

Figure 9. Evidence lower bound for tsunamis in (**a**) 1983 and (**b**) 1993 at a mean of 0.3.

4.2. Testing

Testing was performed to determine how training should be conducted. Figures 10 and 11 demonstrate several of the trained weight and bias distribution at links from the input layer to the first hidden layer for the tsunamis in 1983 and 1993. It can confirm that both the weights and bias at each link depict Gaussian-like distributions, and the bias has broader standard deviation bands than weights as a result of training (Figure 10). Contrary to the tsunami in 1983, it shows that the standard

deviation of weights at links expanded wider than the bias for the tsunami in 1993 (Figure 11). Given trained weights and bias, testing is carried out with a rest of 20% of the normalized tsunami height profiles (Figure 12).
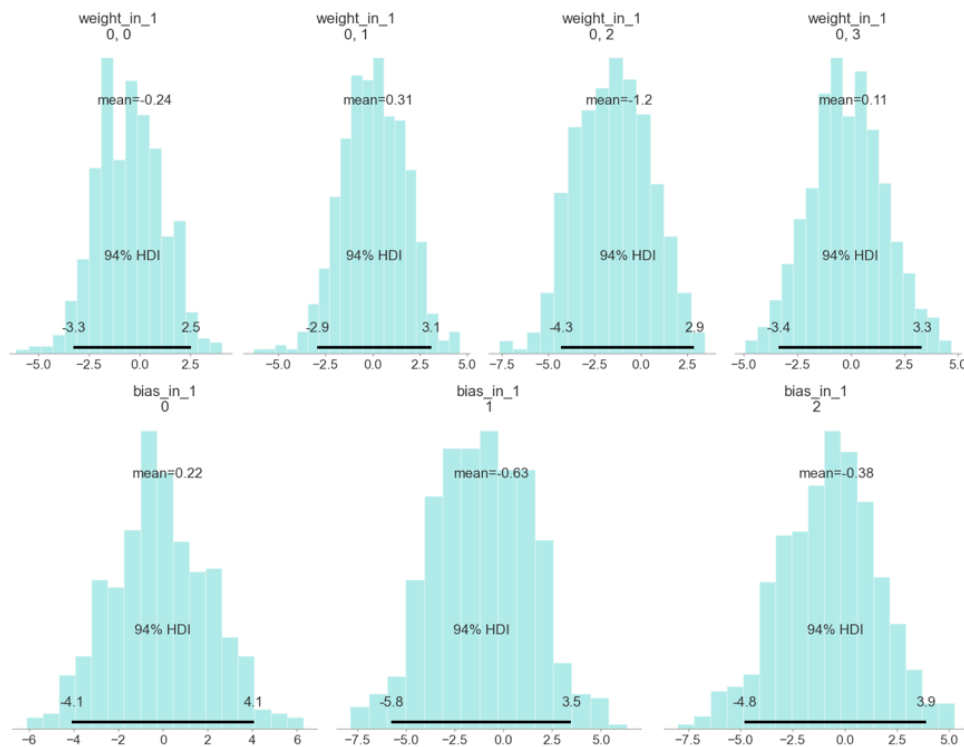
Figure 10. Trained weights and bias at links that connect input layer to the first hidden layer for tsunami in 1983.
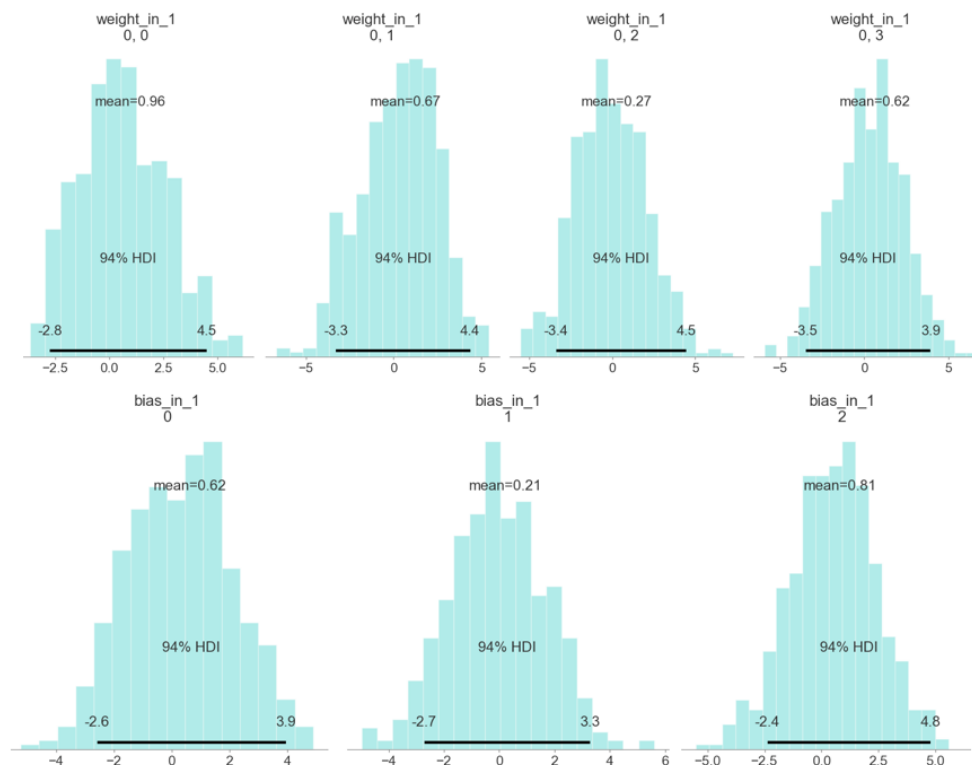
Figure 11. Trained weights and bias at links that connect input layer to first hidden layer at a 0.3 mean for tsunami in 1993.

Table 2 shows that the maximum tsunami heights, evaluated over historical tsunami events in BNNs, were compared with the numerical model, and inverted normalized tsunami heights to true tsunami heights before comparison. The tsunami in 1983 was evaluated to be overestimation with +2% bias. In contrast, tsunami in 1993 exhibits non-bias as a mean of 0.2, and it is overestimated little with +2% bias as mean of 0.3 at a standard deviation of 2.5. The case of a 0.2 mean is estimated to be less biased than the 0.3 mean in testing for tsunamis in 1993.
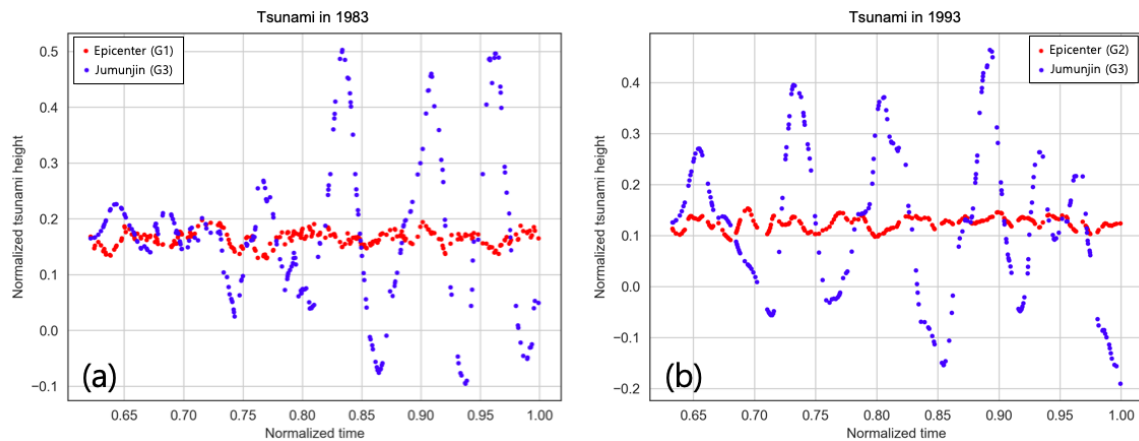

Figure 12. Testing dataset of tsunami in (**a**) 1983 and (**b**) 1993.

Table 2. Comparison of numerical model and BNNs for maximum tsunami height at gage in Jumunjin port over historical tsunamis in testing.

Location	Historical Tsunamis	Numerical Model			BNNs	BIAS (m)
		Tsunami Height (m)	μ	σ^2	Tsunami Height (m)	
Jumunjin	1983	2.26	0.0	2.9	2.31	+0.05
	1993	1.86	0.2	2.5	1.86	0.00
		1.86	0.3	2.5	1.90	+0.04

4.3. Prediction

Given trained information on weights and bias for historical tsunamis as described in Figures 10 and 11, BNNs are introduced to virtual tsunamis, which are similar to historical tsunamis in terms of magnitude and occurrence locations at the epicenter. Figure 13 shows the tsunami height profiles of the virtual tsunamis at the epicenter, and it was converted to normalized tsunami heights and time as performed in the training process. Afterward, prediction was performed to determine how well the BNNs predicted virtual tsunamis. Table 3 present the results for the maximum tsunami heights of the numerical model and BNNs. In the case of virtual tsunami 1, although the tsunami height of BNNs is underestimated as compared to those of numerical model approximately 0.07 m, it shows a similar tendency to the tsunami height of the numerical model with a 3% difference. Regarding virtual tsunami 2, on the condition of a stationary standard deviation of 2.5, the BNNs model with a mean of 0.3 indicates better performance. The BNNs model with a mean of 0.2, which yielded more accurate results in testing, did not capture the maximum tsunami height, showing a 35% difference. In contrast, although the BNNs model with a mean of 0.3 has minimal bias, which is considered acceptable, it exhibits better performance. The reason why BNNs model with a mean of 0.2 showed worse performance than those with 0.3 mean is the issue of overfitting, which occurs as data fit into the trained dataset well, not considering other datasets [10]. Although overfitting is not a significant concern for BNNs, sensitivity to parameters of prior distribution was observed in this study. In the case of virtual tsunami 2, the 0.3 mean yields the best performance.
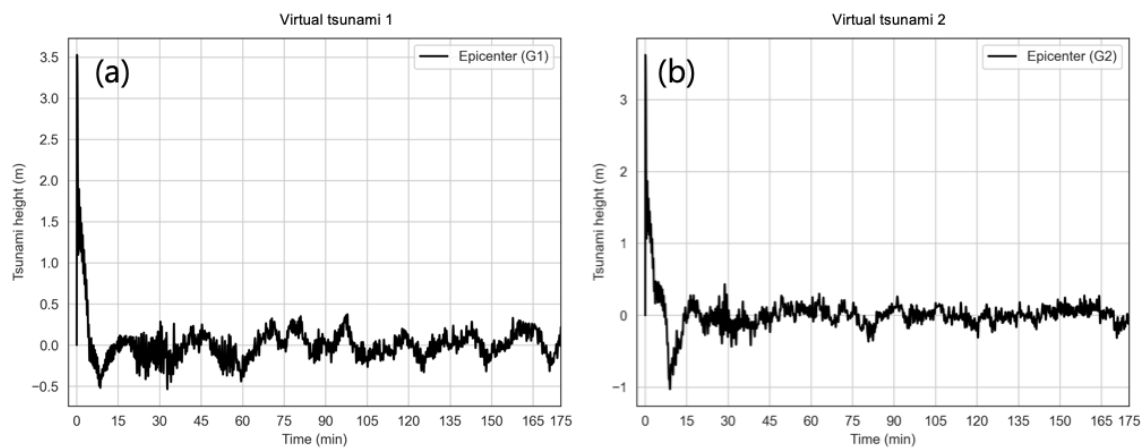
Figure 13. Time series of tsunami heights at epicenter by (**a**) virtual tsunami 1 and (**b**) virtual tsunami 2.

Table 3. Comparison of numerical model and BNNs for predicted maximum tsunami heights at gage in Jumunjin port over virtual tsunamis in prediction.

Location	Virtual Tsunamis	Numerical Model			BNNs	Difference (m)
		Tsunami Height (m)	μ	σ^2	Tsunami Height (m)	
Jumunjin	Virtual tsunami 1	2.21	0.0	2.9	2.14	-0.07
	Virtual tsunami 2	1.63	0.2	2.5	1.05	-0.57
		1.63	0.3	2.5	1.54	-0.08

5. Summary and Conclusions

AI has been applied to many science and engineering fields, exhibiting remarkable performance. In this study, we attempted to discern maximum tsunami heights for virtual tsunami at the Jumunjin Port, where the 1983 tsunami caused inundation in some areas. BNNs, combining Bayesian inference and NNs, have the advantage of enabling prediction with a small dataset and were introduced to evaluate maximum tsunami height. Two nodes in the input layer and ten nodes in each of the three hidden layers were constructed with a tangent hyperbolic function to describe nonlinear effects. In addition, variational inference was proposed in Bayesian inference. Tsunami height profiles at the epicenter and gage in port were used as fundamental datasets, and normalization for tsunami height and time was performed in advance. Approximately 80% of the tsunami height profiles were classified for the training dataset, and the remainder of the dataset was applied to the testing dataset. A Gaussian distribution was assumed as a prior distribution in BNNs, and parameters for mean (μ) and standard deviation (σ^2) were optimized. Based on this work, the training, testing, and prediction processes were examined. In training, it was determined that posterior probabilities of weights and bias at links on BNNs were well trained for historical tsunami events. Then, good agreement between the numerical model and BNNs for maximum tsunami heights over historical tsunami events were confirmed in testing. Finally, maximum tsunami height over virtual tsunamis were predicted based on trained weights and bias of BNNs. In particular, depending on distribution parameters, different results were obtained such that 0.2 mean meets overfitting, whereas a mean of 0.3 indicated the best performance. Although there were disparities in maximum tsunami heights between the BNNs and numerical model, the BNNs could reasonably predict the maximum tsunami heights for virtual tsunamis.

References

1. Cho, Y.-S.; Hwang, H.-S.; Kim, J.-Y.; Kwan, H.-K. Development of Hazard Map with probable maximum tsunamis. *J. Coast. Res.* **2016**, *75*, 1057–1061. [CrossRef]

2. Imamura, F.; Shuto, N.; Goto, C. Numerical simulations of the transoceanic propagation of tsunamis. In Proceedings of the Congress Asian and Pacific Regional Division, Yogyakarta, Indonesia, 8–10 September 1998; pp. 265–272.

3. Cho, Y.-S.; Liu, P.L.-F. Crest length effects in nearshore tsunami run-up around islands. *J. Geophys. Res.* **1999**, *104*, 7907–7913. [CrossRef]

4. Yoon, S.B. Propagation of tsunamis over slowly varying topography. *J. Geophys. Res.* **2002**, *107*, 4-1–4-11. [CrossRef]

5. Choi, B.H.; Pelinovsky, E.; Hong, S.J.; Woo, S.B. Computation of tsunamis in the East (Japan) Sea using dynamically interfaced nested model. *Pure Appl. Geophys.* **2003**, *160*, 1383–1414. [CrossRef]

6. Cho, Y.-S.; Sohn, D.-H.; Lee, S.-O. Practical modified scheme of linear shallow-water equations for distant propagation of tsunamis. *Ocean. Eng.* **2007**, *34*, 1769–1777. [CrossRef]

7. Hideki, A. Outline of deep learning. In *Deep Learning*; Muneki, Y., Shinitsi, M., Daisuke, O., Takayuki, O., Bollegala, D.K., Eds.; Kindaikakagusha: Tokyo, Japan, 2015; pp. 9–13.

8. Itamar, R.; Dalya, B.; Sahar, S. Probabilistic Random Forest: A Machine Learning Algorithm for Noisy Data sets. *Astron. J.* **2019**, *157*, 1–12.

9. Zoubin, G. Probabilistic machine learning and artificial intelligence. *Nature* **2015**, *521*, 452–459.

10. Suyama, A. Machine learning and bayesian learning. In *Introduction to Machine Learning by Bayesian Inference*; Sugiyama, M., Ed.; Kodansha: Tokyo, Japan, 2017; pp. 82–85, 121–123.

11. Back, U.-S. *1983 Tsunami in East-Sea*; Korea Meteorology Administration: Seoul, Korea, 1983; p. 44.

12. Korean Peninsula Energy Development Organization. *Estimation of Tsunami Height for KEDO LWR ProjecT Report*; Korea Power Engineering Company: Seoul, Korea, 1999.

13. Kim, K.-H.; Cho, Y.-S.; Kwon, H.-H. An integrated Bayesian approach to the probabilistic tsunami risk model for the location and magnitude of earthquakes: Application to the Eastern Coast of the Korean Peninsula. *Stoch. Environ. Res. Risk Assess.* **2018**, *32*, 1243–1257. [CrossRef]

14. Cho, Y.-S. Numerical Simulation of Tsunami Propagation and Run-Up. Ph.D. Dissertation, University of Cornell, NewYork, NY, USA, 1995; pp. 11–52.

15. Shinichi, N. Probability and bayesian theorem. In *Variational Bayesian Learning*; Kodansya: Tokyo, Japan, 2019; pp. 79–80.

16. ASCE Committee. Aritifical Neural Networks in Hydrology: Preliminary concepts. *J. Hydrol. Eng.* **2000**, *5*, 115–123. [CrossRef]

17. Tsoukalas, L.H. Fuzzy Methods in Neural Networks. In *Fuzzy and Neural Approaches in Engineering*; Uhrig, R.E., Ed.; Wiley: Hoboken, NJ, USA, 1997; pp. 32–58.

18. Yarin, G. Uncertainty in Deep Learning. Ph.D. Dissertation, University of Cambridge, Cambridge, UK, 2016; pp. 17–20.

19. Song, M.-J.; Cho, Y.-S. Probabilistic Tsunami Heights Model using Bayesian Machine Learning. *J. Coast. Res.* **2020**, *95*, 1291–1296. [CrossRef]

20.　Koki, S. Neural networks. In *Deep Learning from Scratch*; O'Reilly Japan: Tokyo, Japan, 2016; p. 77.

21.　Afshin, G.; Vladik, K.; Olga, K. *Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation*; Technical Reports; University of Texas: El Paso, TX, USA, 2018; pp. 1–6.

22.　*Amazon Machine Learning: Developer Guide*; Amazon: Seattle, WA, USA, 2020; p. 12.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.