

Article

Estimation of Wave-Breaking Index by Learning Nonlinear Relation Using Multilayer Neural Network

Miyoung Yun ¹ , Jinah Kim ² and Kideok Do ^{3,*} 

¹ Department of Convergence Study on the Ocean Science and Technology, Korea Maritime & Ocean University, Busan 49112, Korea; yunmy@g.kmou.ac.kr

² Coastal Disaster Research Center, Korea Institute of Ocean Science and Technology, Busan 49111, Korea; jakim@kiost.ac.kr

³ Department of Ocean Engineering, Korea Maritime & Ocean University, Busan 49112, Korea

* Correspondence: kddo@kmou.ac.kr; Tel.: +82-51-410-5248

Abstract: Estimating wave-breaking indexes such as wave height and water depth is essential to understanding the location and scale of the breaking wave. Therefore, numerous wave-flume laboratory experiments have been conducted to develop empirical wave-breaking formulas. However, the nonlinearity between the parameters has not been fully incorporated into the empirical equations. Thus, this study proposes a multilayer neural network utilizing the nonlinear activation function and backpropagation to extract nonlinear relationships. Existing laboratory experiment data for the monochromatic regular wave are used to train the proposed network. Specifically, the bottom slope, deep-water wave height and wave period are plugged in as the input values that simultaneously estimate the breaking-wave height and wave-breaking location. Typical empirical equations employ deep-water wave height and length as input variables to predict the breaking-wave height and water depth. A newly proposed model directly utilizes breaking-wave height and water depth without nondimensionalization. Thus, the applicability can be significantly improved. The estimated wave-breaking index is statistically verified using the bias, root-mean-square errors, and Pearson correlation coefficient. The performance of the proposed model is better than existing breaking-wave-index formulas as well as having robust applicability to laboratory experiment conditions, such as wave condition, bottom slope, and experimental scale.

Keywords: wave breaking; breaking-wave height; breaking-water depth; multilayer neural network; nonlinear relationships; machine learning



Citation: Yun, M.; Kim, J.; Do, K. Estimation of Wave-Breaking Index by Learning Nonlinear Relation Using Multilayer Neural Network. *J. Mar. Sci. Eng.* **2022**, *10*, 50. <https://doi.org/10.3390/jmse10010050>

Academic Editor: Domenico Curto

Received: 22 November 2021

Accepted: 30 December 2021

Published: 3 January 2022

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



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

When ocean waves propagate to the shoreline, the decrease in the water depth shortens its wavelength and increases wave height. Upon reaching a physical limit, the wave breaks without maintaining its form, which is defined as the “Wave-Breaking” phenomenon [1]. Wave breaking causes various events in the process of wave-energy (transported from deep water) dispersion, such as impact waves, longshore currents, rip currents and sediment transport, and also affects vessels and the stability of coastal structures [2]. Therefore, information on wave breaking is essential for the design and maintenance of coastal structures. In addition, it is considered a necessary factor for predicting and responding to sediment transport and morphological changes in the nearshore area. Particularly among the physical quantities related to wave breaking, the most important properties are the wave height and water depth at the position of the breaking wave, which are defined as the “Breaking-Wave Height” and “Breaking-Water Depth,” respectively. These quantitative values, called wave-breaking indexes, indicate the starting point of the breaking wave and the maximum wave height in the nearshore, and numerous studies have been conducted to estimate these values [3–8]. However, owing to strong turbulent features and nonlinearity,

observing and predicting coastal breaking waves is complex, and many studies are still challenged by this subject [9,10].

Initial studies related to wave breaking began with the theoretical approach presented by Michell [11] and McCowan [12]. Michell proposed the wave-steepness limit of the deep-water wave value of 0.142 using the relationship between the water-particle motion and wave celerity at the wave crest. McCowan proposed a ratio of 0.78 of the wave height and water depth when the wave breaks while considering a solitary wave in deep-water. Subsequently, Miche [13] applied the linear wave theory to the results drawn by Michell and expressed the breaking-wave height under the monochromatic-regular-wave condition by using the wavelength and water depth at the breaking point as input variables. After that, Ippen and Kulin [14] also pointed out the inadequacy of the presented formula, wherein the wave-breaking mechanism of a solitary wave was not suitable for estimating the breaking-wave height and water depth of wind waves.

It was reported that various errors occurred in estimating the breaking-wave height and water depth using the linear wave theory due to the nonlinearity of the waves by the shallow-water-depth effect [8,15,16]. Therefore, numerous studies have attempted to estimate empirical wave-breaking formulas for the breaking-wave height and water depth in order to overcome these limitations. Hence, the breaking-wave height and water depth were estimated in subsequent studies by reproducing various wave conditions in wave-flume laboratory experiments and obtaining consecutive wave transformations depending on the water depth [17–19]. Maruyama et al. [20], Stive [21], and Smith and Kraus [15] performed wave-flume laboratory experiments by installing offshore structures resembling the actual coastal terrain or by setting a sand bar (formed by sand moved out towards the open sea by waves) as the topographical conditions. Numerous empirical wave-breaking formulas have been proposed based on wave-flume laboratory experiment data obtained under various conditions, including the above [8,9,22–30]. However, consistent performance was limited under generalized conditions as these empirical wave-breaking formulas depended on certain laboratory experiment data. Furthermore, the methods and basic forms of the breaking index used to develop the empirical wave-breaking formulas were different from each other. In particular, the empirical wave-breaking formulas suggested by Goda [8], Robertson et al. [9], Tadayan et al. [30], and others had a high degree of reproducibility because the estimated breaking-wave height and water depth were plugged in as the input values for the equations, but that makes them insufficient for coastal engineering applications.

In this study, a method based on a multilayer neural network with a nonlinear activation function and backpropagation is proposed in order to estimate the breaking-wave height and breaking-water depth by discovering the nonlinear relationship between the deep-water wave condition, bottom slope and wave-breaking index [31]. For the improvement of the applicability and usability compared to the previous wave-breaking formulas, the newly proposed method is designed to simultaneously obtain breaking-wave indexes (breaking-wave height and water depth). Additionally, there is no possibility of error from the secondary transformation of raw wave data by directly utilizing the input and output without nondimensionalization. Next, various experiment data for monochromatic regular waves published in previous studies were collected and used for training the network. The proposed neural network comprising the nonlinear activation function and backpropagation can use all of the data acquired from various laboratories without distinguishing between the experimental conditions. This gives the proposed model robust applicability to laboratory experiment conditions. Finally, the performance of our proposed method is evaluated in comparison with the existing breaking-wave-index formulas.

2. Related Works

The existing breaking-wave formulas use the deep-water wave data and bottom slope as input variables and reproduce the breaking-wave height and wave-breaking location. In this study, the accuracy and meaningfulness of these formulas were summarized, and the

laboratory experiment data for the monochromatic regular wave from the previous studies were applied to construct the neural network for estimating the breaking-wave height and water depth. The variables proposed in this study are defined as follows: (1) H_0 , h_0 , and L_0 indicate the deep-water wave height, water depth, and wavelength, respectively; (2) H_b and h_b denote the breaking-wave height and breaking-water depth, respectively; and (3) m denotes the gradient of the bottom slope.

Le Mehaute and Koh [32] considered the deep-water wave steepness and bottom slope parameters as input variables and predicted the nondimensionalized breaking-wave height with the deep-water wave height using Equation (1). This equation was derived using the data of the gentle slope, which has high utilization potential as it demonstrated sound predictive performance in the majority of slope ranges [33]. Additionally, this equation was the first trial that simultaneously considered the wave steepness and bottom slope in calculating the breaking-wave height. Subsequently, this was partially modified by Galvin [34] and Collins and Weir [35]. Data from wave-flume laboratory experiments performed by Suquet [36], Hamada [37], and Iversen [17] were used to develop the equation, and the data were within the following ranges: $0.02 < m < 0.2$ and $0.002 < H_0/L_0 < 0.093$. From these datasets, the experimental results published by Iversen [17] were collected and used in this study.

$$\frac{H_b}{H_0} = 0.76m^{1/7} \left(\frac{H_0}{L_0}\right)^{-1/4} \tag{1}$$

Rattanapitikon and Shibayama [28] used 574 experimental data in the ranges $0 \leq m \leq 0.38$ and $0.001 < H_0/L_0 < 0.100$ to validate the existing wave-breaking formulas. In addition, they modified the terms that correspond to the bottom slope in the equations developed by Komar and Gaughan [19], Goda [22], and Ostendorf and Madsen [24].

Rattanapitikon and Shibayama [38] supplemented approximately 100 large wave-flume experimental data ($0 \leq m \leq 0.29$ and $0.003 < H_0/L_0 < 0.112$) from Kajima et al. [39] and Smith and Kraus [40] to propose the new forms of empirical equations that can predict the breaking-wave height and water depth. The equations are presented in Equations (2) and (3).

$$\frac{H_b}{L_0} = (-0.57m^2 + 0.31m + 0.58) \left(\frac{H_0}{L_0}\right)^{0.83} \tag{2}$$

$$\frac{h_b}{L_0} = (3.86m^2 - 1.98m + 0.88) \left(\frac{H_0}{L_0}\right)^{0.84} \tag{3}$$

From the results of Rattanapitikon and Shibayama [38], the deep-water wave steepness and bottom slope were used as input variables to estimate the breaking-wave height and water depth using the nondimensionalized form by deep-water wavelength. The laboratory experiment data used to fit the equation comprised 695 data generated from 26 wave-flume laboratory experiments. In this study, 351 accessible data were used to construct a neural-network model for predicting breaking-wave indexes (breaking-wave height and water depth).

Xie et al. [29] restricted the wave-breaker type to “plunging” to improve the breaking-water-depth-prediction accuracy. The new formula was developed based on the linear wave theory, so it can significantly improve the applicability of the input wave condition. However, the equation is too complicated and applicable only under certain wave-breaker types (plunging). A semi-empirical formula in Equation (4) was proposed, which applied 242 data covering a range of $0.0125 \leq m \leq 0.2$ and $0.0016 < H_0/L_0 < 0.092$, acquired from six sources, for fitting the equation coefficient. The proposed equation was verified using their own 25 wave-flume laboratory experiments and eight plunging wave data from Lara

et al. [41]. All of the experimental data were collected and used in the learning phase of the proposed model and in its validation.

$$\frac{h_b}{L_0 m} = 0.4787 \frac{\Omega}{1 - \Omega}, \quad \Omega = \left\{ \frac{4 \left(\frac{H_0}{L_0} \right)}{4 \left(\frac{H_0}{L_0} \right) + m} \right\}^{4/3} \quad (4)$$

Lee and Cho [42] used 860 data obtained from previous experimental datasets in the ranges of $0.01 \leq m \leq 0.2$ and $0.018 < H_0/L_0 < 0.1272$. They applied them in order to derive the linear relationship using several linear regression methods and a feed-forward neural network. Moreover, the linear relationships of the breaking-wave height and water depth nondimensionalized by deep-water wavelengths were derived and expressed in terms of the deep-water wave steepness and bottom slope, as shown in Equations (5) and (6).

$$\frac{H_b}{L_0} = 0.002 + 0.064m + 0.906 \frac{H_0}{L_0}, \quad (5)$$

$$\frac{h_b}{L_0} = 0.003 + 0.011m + 1.128 \frac{H_0}{L_0} \quad (6)$$

The equations of Lee and Cho [42] fit very well on the linear relationship between deep-water wave condition, bottom slope, and wave-breaker index. However, both equations tend to overestimate the results for gentle bottom slopes ($m \leq 0.02$). In addition, the linear relation limits the performance when the nonlinearity of the wave is strong or when the scale effect of the experiment is important. In this study, 433 data were collected and used to construct the newly proposed model.

3. Methodology

3.1. Data

In this study, openly published wave-flume laboratory experimental data were used in developing the new breaking-wave-index-prediction model that applies a neural network with a nonlinear activation function and backpropagation to the estimation of the breaking-wave height and water depth. Table 1 summarizes the source and experimental conditions of 630 data collected from 31 previous studies. A substantial part of the data was obtained from Gaughan [43] and Smith and Kraus [15], and these are indicated as “*” and “**” in the Source column of Table 1, respectively. The remaining data were collected from Bowen et al. [44], Weggel and Maxwell [45], Ozaki et al. [46], Van Dorn [47], Kirgoz [48], Ishida and Yamaguchi [49], Sakai et al. [50], Ting and Kirby [51], Kakuno et al. [52], Yüksel et al. [53], Hoque [54], Shin and Cox [55], Deo and Jagdale [56], Lara et al. [41], Mori and Kakuno [57], and Xie et al. [29]. Each dataset consists of input values such as the deep-water wave height, bottom slope, period or wavelength, and breaking-wave heights and water depths obtained from wave-flume laboratory experiments with slopes distributed between 0.009–0.225. Among the data from Iversen [17], Ishida and Yamaguchi [49], Yüksel et al. [53], and Xie et al. [29], those that did not include the breaking-wave heights or water depths were excluded from this study. The breaking-wave heights and water depths of 584 data used in this study are plotted in Figure 1. In Figure 1, the data located independently in the upper right corner are the results obtained from the experiments performed by Maruyama et al. [20] and Stive [21] in a large wave flume within a movable bed condition. There are various wave-flume laboratory experimental data other than the acquired data in this study. In this study, we only utilized easily accessible data in order to facilitate the reproducibility of the process and the results of this research for estimating the breaking-wave height and water depth.

Table 1. Summary of compiled laboratory datasets on breaking waves.

No.	Source	Bottom Slope (m)	Period (T [s])	Deep-Water Wave Height (H_0 [cm])	H_0/L_0	No. of Cases
1	Munk (1949) *	0.01–0.07	0.86–1.98	4.2–11.6	0.007–0.092	16
2	Iversen (1952) **	0.02–0.10	0.74–2.67	2.7–12.4	0.003–0.091	68
3	Morison and Crooke (1953) *	0.02–0.10	0.78–2.62	3.5–12.4	0.004–0.080	6
4	Singamsetti and Wind (1967) **	0.03–0.20	1.03–1.73	6.6–16.0	0.017–0.080	95
5	Horikawa and Kuo (1967) **	0.01–0.05	1.20–2.30	4.7–17.3	0.006–0.073	97
6	Bowen et al. (1968)	0.08	0.82–2.37	3.6–9.0	0.007–0.049	11
7	Komar and Simmons (1968) *	0.04–0.11	0.81–2.37	2.8–14.4	0.003–0.071	44
8	Galvin (1969) **	0.05–0.20	1.00–6.00	2.7–9.8	0.001–0.050	19
9	Weggel and Maxwell (1970)	0.05	1.26–2.05	4.6–12.6	0.013–0.055	9
10	Saeki and Sasaki (1973) **	0.02	1.30–2.50	5.3–10.3	0.005–0.039	2
11	Iwagaki et al. (1974) **	0.03–0.10	1.00–2.00	3.1–11.4	0.005–0.073	23
12	Walker (1974) **	0.03	1.17–2.33	1.0–8.0	0.001–0.038	15
13	Ozaki et al. (1977)	0.10	0.79–1.40	0.9–5.8	0.006–0.060	20
14	Van Dorn (1978)	0.02–0.08	1.65–4.80	3.7–13.2	0.001–0.031	12
15	Mizuguchi (1981) **	0.10	1.20	10.0	0.045	1
16	Kirgoz (1982)	0.07–0.23	0.78–2.03	2.2–7.3	0.003–0.061	16
17	Visser (1982) **	0.05–0.10	0.70–2.01	6.0–10.2	0.014–0.079	7
18	Ishida and Yamaguchi (1983)	0.10	0.68–1.50	2.7–7.6	0.008–0.095	6
19	Maruyama (1983) **	0.03	3.10	137.0	0.091	1
20	Stive (1985) **	0.03	1.80–5.00	16.0–121.0	0.031–0.032	2
21	Sakai et al. (1986)	0.02–0.03	1.78–2.21	12.2–23.7	0.016–0.048	19
22	Smith and Kraus (1990) **	0.03	1.02–2.49	8.5–15.8	0.009–0.092	5
23	Ting and Kirby (1994)	0.03	2.00–5.00	8.9–12.7	0.002–0.020	2
24	Kakuno et al. (1996)	0.03–0.10	0.88–2.00	2.2–13.2	0.008–0.092	55
25	Yüksel et al. (1999)	0.10	1.10–2.05	10.0–19.1	0.022–0.065	10
26	Hoque (2002)	0.11	1.12–1.80	11.0–16.6	0.024–0.076	6
27	Shin and Cox(2003)	0.03	1.50–3.00	7.8–12.7	0.006–0.036	3
28	Deo and Jagdale (2003)	0.03–0.10	0.74–1.20	7.3–13.0	0.042–0.127	20
29	Lara et al. (2006)	0.05	1.20–4.00	5.0–15.0	0.006–0.073	12
30	Mori and Kakuno (2008)	0.03	1.60–3.80	11.6–16.3	0.012–0.046	3
31	Xie et al. (2019)	0.10	1.75–2.05	3.0–5.0	0.005–0.010	25
Total No. of Samples		0.01–0.23	0.68–6.00	0.9–137.0	0.001–0.127	630

* Data from Gaughan [43]. ** Data from Smith and Kraus [15].

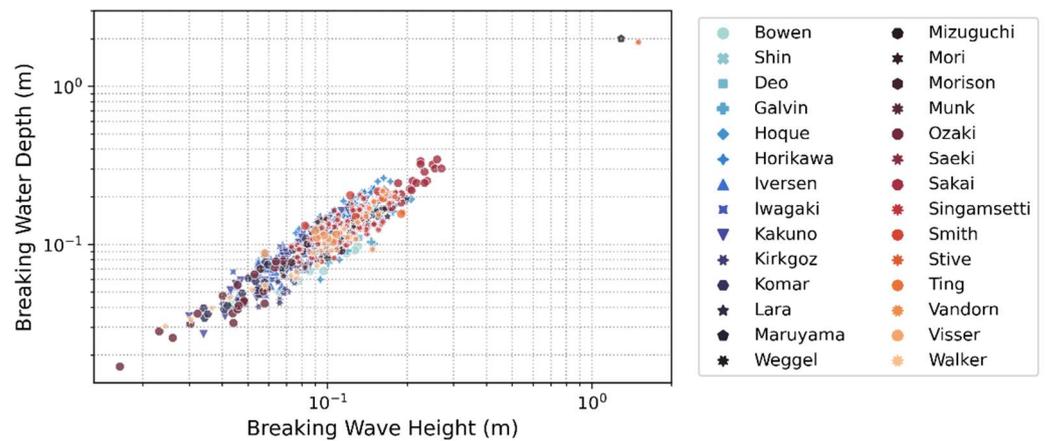


Figure 1. Breaking-water depth and breaking-wave height obtained from 584 laboratory datasets.

A total of 584 data were classified as training and test data at a ratio of approximately 3:1 such that the bottom slope and deep-water wave height and period data could be distributed evenly according to the experimental conditions. Thus, the amount of data for training and testing was approximately 455 and 129, respectively. Figure 2 presents the outer and inner circles as the training data and test data, respectively, in addition to the composition of the experimental conditions for acquiring each wave datum.

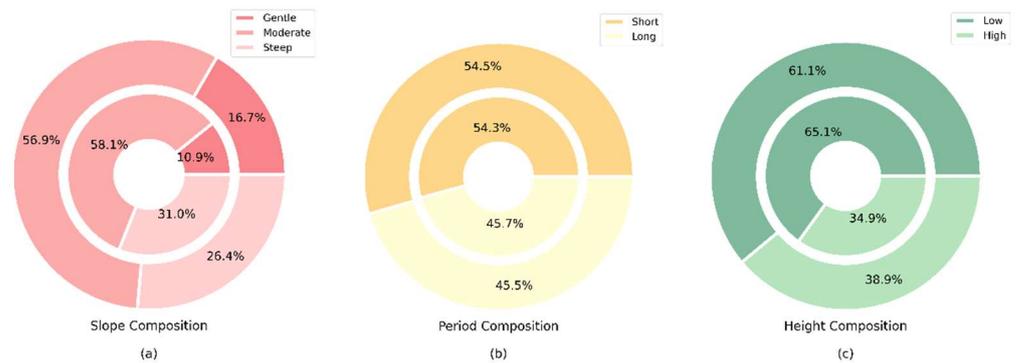


Figure 2. Composition of components according to each wave condition. Outer and inner circles represent 455 training data and 129 test data, respectively: (a) Bottom slope is classified into three groups, i.e., Gentle ($m \leq 0.02$), Moderate ($0.02 < m < 0.10$), and Steep ($m \geq 0.10$). (b) Wave period is classified into two groups, i.e., Short ($T \leq 1.50$ s), and Long ($T > 1.50$ s). (c) Deep-water wave height is classified into two groups, i.e., Low ($H_0 < 0.1$ m), and High ($H_0 \geq 0.1$ m).

3.2. Multilayer Neural Network for Estimating Wave-Breaking Index

To abstract the more useful features by creating a multi-level and multi-neuron neural network, called the fully connected Deep Neural Network (DNN), which automatically learns more appropriate weights and thresholds based on the structure of the related data on wave breaking, we proposed a multilayer neural network comprising a number of hidden layers, neurons per layer, and connections per unit with a nonlinear activation function and backpropagation.

Figure 3 illustrates a multilayer neural network with three layers, where the input layer, denoted as z , represents multiple factors (the bottom slope and deep-water wave height and period) that affect the wave breaking, and the output layer, denoted as a , refers to the breaking-wave height and breaking-water depth.

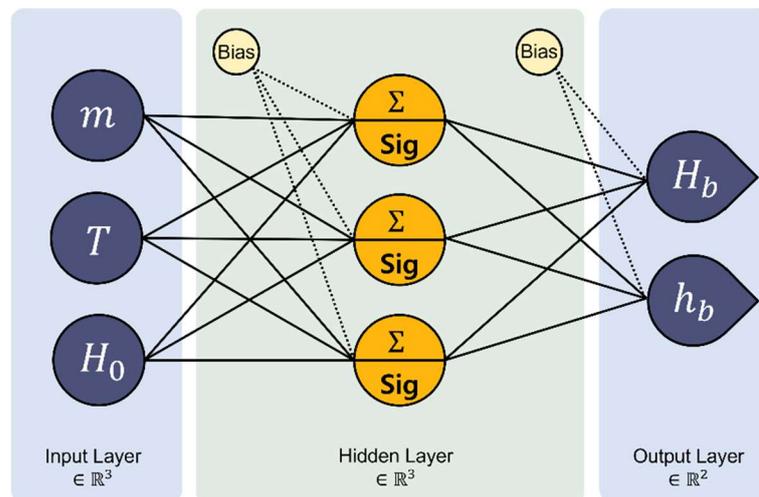


Figure 3. Architecture of the proposed fully connected multilayer neural network with the sigmoid activation function.

We used the sigmoid activation function [58], which transforms the weighted sum of the input into an output from neurons in a hidden layer of the network as follows:

$$\frac{1.0}{1.0 + e^{-x}} \tag{7}$$

The activation function is nonlinear and may be referred to as nonlinearity in the layer or the network design. The function takes any real value as an input and outputs values in

the range of 0 to 1. The larger the input (more positive), the closer the output value will be to 1.0, whereas the smaller the input (more negative), the closer the output will be to 0.0.

Backpropagation is operated in order to determine the optimal learning result by propagating the error in the reverse direction, unlike the feed-forward neural network. First, the input is transmitted to the final output during the feed-forward process. Second, the error and cost function are determined at the final output layer. Thereafter, in the backpropagation process, the errors between the expected outputs and actual values obtained from the final step are propagated in the reverse direction, and each weight and bias value of the neurons is updated.

Such optimizers for training neural networks are responsible for finding the free parameters (usually denoted as weights) of a cost function that, typically, includes a performance measure evaluated on the training set and additional regularization terms. Adaptive moment (Adam) is an update to the root-mean-square propagation (RMSProp) optimizer, wherein momentum [59] is incorporated, i.e., in addition to storing an exponentially decaying average of the previous squared gradients, Adam also employs an exponentially decaying average of the previous gradients. Loss denotes the loss function that is employed at the training time and is given by the mean-squared error (MSE) as follows:

$$\frac{1}{n} \sum (y - \hat{y})^2 \tag{8}$$

The MSE is calculated as the average of the squared differences between the estimated (\hat{y}) and actual values (y). The epoch and batch size for the experiment were 317 and 1, respectively, and the learning rate was 0.0001. All data were normalized using the z-score for standardization to put different variables on the same scale.

In contrast to the previous empirical equations, the raw wave-period value was plugged in as the input values instead of deep-water wavelength because the dimension between the input and output did not need to be considered. Moreover, that eliminated the possibility of errors resulting from the dispersion relation equation, which is used to convert deep-water wavelength to wave period, induced by the linear wave theory.

3.3. Evaluation Metrics

The bias (B), root-mean-square errors ($RMSE$), and Pearson correlation coefficient (R) were used as the evaluation metrics to verify a newly constructed breaking-wave-index-prediction model. B and $RMSE$ indicate the difference between the actual value x_i and the value estimated through the model y_i , and R numerically expresses the similarity between the two values. The model performance is considered high when the absolute value of B is small, the $RMSE$ value is small, and the R value is close to 1. There is no absolute standard, but an R value of 0.8 or higher generally implies a suitable correlation between the estimations of the prediction model and real values. \bar{x} and \bar{y} indicate the average of x_i and y_i , respectively, and n refers to the amount of data.

$$B = \frac{\sum_i^n x_i - y_i}{n}, \tag{9}$$

$$RMSE = \sqrt{\frac{\sum_i^n (x_i - \bar{y})^2}{n}}, \tag{10}$$

$$R = \frac{\sum_i^n (x_i - \bar{y})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2 \sum_i^n (y_i - \bar{y})^2}} \tag{11}$$

4. Results

In this study, 129 test data, excluding the data used to train the network, were used to evaluate wave-breaking-index-estimation performance. The results showed that B , $RMSE$, and R of the breaking-wave height and water depth were 0.004, 0.019, and 0.894 and

0.005, 0.021, and 0.921, respectively. These results were compared with the approach of modeling the linear relationship using multiple regression analysis. In contrast to previous breaking-wave formulas that nondimensionalized the wave height and water depth, this newly proposed model utilizes real values of deep-water wave height, period, bottom slope, breaking-wave height, and breaking-water depth. This makes it possible to estimate more accurate inference results by excluding the errors from the secondary transformation of the raw wave data. Therefore, the existing breaking-wave formulas were reorganized to evaluate and compare the performance of the proposed model and the previous breaking-wave formulas for estimating the breaking-wave height and water depth (Table 2).

Table 2. Reorganized previous breaking-wave formulas for the performance comparison.

Target	Authors	Modified Formulas	Abbreviation
H_b	Le Mehaute and Koh (1967)	$H_b = 0.76(H_0/L_0)^{-1/4}m^{1/7}H_0$	CA_Hb_1
	Rattanapitikon and Shibayama (2006)	$H_b = (-0.57m^2 + 0.31m + 0.58)(H_0/L_0)^{0.83}L_0$	CA_Hb_2
	Lee and Cho (2021)	$H_b = (0.002 + 0.064m + 0.906H_0/L_0)L_0$	CA_Hb_3
h_b	Rattanapitikon and Shibayama (2006)	$h_b = (3.86m^2 - 1.98m + 0.88)(H_0/L_0)^{0.84}L_0$	CA_hb_2
	Xie et al. (2019)	$h_b = (0.4787 \frac{\Omega}{1-\Omega})mL_0,$ $\Omega = \{4(H_0/L_0) / (4(H_0/L_0) + m)\}^{4/3}$	CA_hb_1
	Lee and Cho (2021)	$h_b = (0.003 + 0.011m + 1.128H_0/L_0)L_0$	CA_hb_3

The proposed model shows better performance for each breaking-wave height and water depth than the three existing breaking-wave formulas in Figure 4. The analysis showed that no significant difference in performance was observed between the proposed model and previous breaking-wave formulas. Excluding the equations of Rattanapitikon and Shibayama (CA_Hb_2 and CA_hb_2) in Figure 4, the results of the others were considered to have suitable predictive performances as the absolute values of B for both the breaking-wave height and water depth were less than or equal to 0.005.

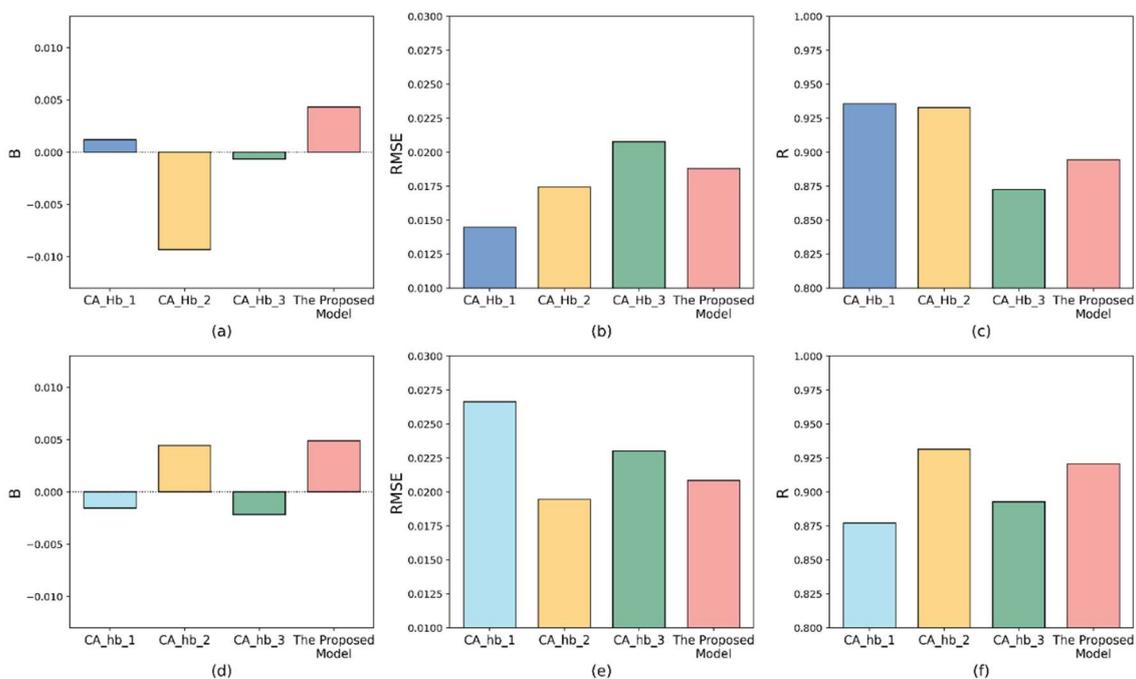


Figure 4. Results of performance analysis of the proposed model and the previous breaking-wave formulas against test data: (a) B, (b) RMSE, and (c) R for breaking-wave height; (d) B, (e) RMSE, and (f) R for breaking-water depth.

A scatterplot of the entirety of the test data is presented in Figure 5 to obtain a more precise comparison between the proposed model and the linear regression equations (CA_Hb_3 and CA_hb_3) using the feed-forward neural network with the linear activation function from Lee and Cho [42]. The equations of Lee and Cho (CA_Hb_3 and CA_hb_3) present, on average, the lowest B absolute value for all the targets, but they have a higher RMSE value and a lower R value compared to the proposed model. This is presumed to be because, as seen in Figure 5a,b, the equations of Lee and Cho (CA_Hb_3 and CA_hb_3) presented irrational estimations for some data. A total of three data were considered to correspond to this issue. Two data were obtained from the study of Galvin [34] and the other from that of Ting and Kirby [51]. These data are the low-scale condition with a period equal to or greater than 4 s. This implies that the equations of Lee and Cho (CA_Hb_3 and CA_hb_3) were ineffective in the estimation of breaking-wave parameters in certain conditions. In contrast, the proposed model produced a reasonable prediction performance for the test data under all conditions. It was presumed that this study tried to separate the training and test data as equally as possible by reflecting the data conditions so that the proposed model could properly learn the inherent irregularities in the wave-flume laboratory experimental data. Such an attempt contributed to enhancing the versatility of the proposed model.

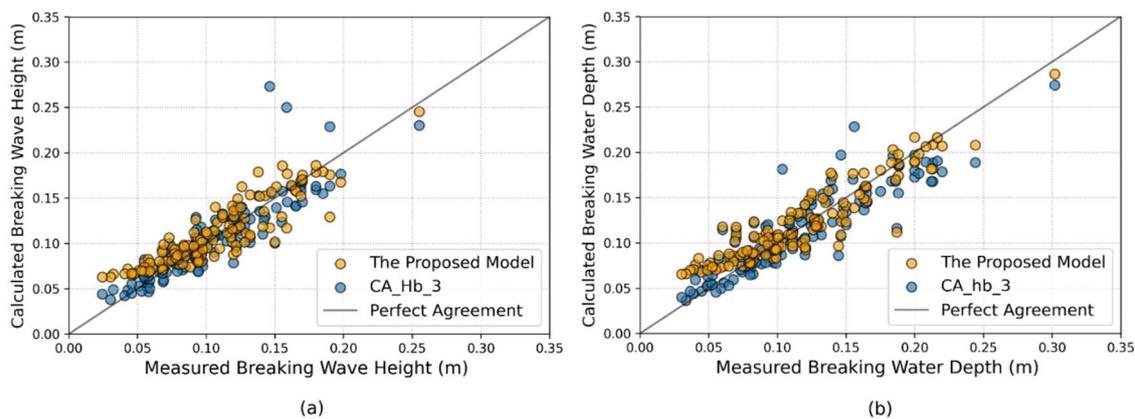


Figure 5. Comparison of the proposed model and the formulas of Lee and Cho (CA_Hb_3 and CA_hb_3) against the laboratory test data. The straight line indicates perfect agreement: (a) breaking-wave height, (b) breaking-water depth.

Meanwhile, the equations of Le Mehaute and Koh (CA_Hb_1) and Xie et al. (CA_hb_1) present an acceptable performance in predicting the breaking-wave height and water depth, respectively. However, only one wave-breaking parameter could be obtained from these equations, which have less practicality and applicability than the proposed model. In contrast, the equations of Rattanapitikon and Shibayama (CA_Hb_2 and CA_hb_2) show a lower RMSE value and higher R value, but an equal or larger B absolute value as compared to those of the proposed model for breaking-wave parameters. In conclusion, regarding B as the criteria with the highest priority for assessing the equation accuracy, since the amount of training data is small, and reflecting the practical applicability, the proposed model is better than the existing breaking-wave formulas.

A new dataset was constructed from the test data while excluding the data used for the formation of the previous breaking-wave formula, and the performance analysis was performed using the same method as before for a fair comparison and evaluation of the proposed and conventional methods. In contrast to the above case, wherein the entire test dataset was applied, only the statistics-based, nonlinear equations (CA_Hb_2 and CA_hb_2) and the linear regression equations (CA_Hb_3 and CA_hb_3) using the feed-forward neural network, which can calculate both breaking-wave parameters, were compared. The amount of newly structured data for comparatively evaluating the methods was 55 and 34 for the studies of Rattanapitikon and Shibayama [38] and Lee and Cho [42],

respectively. Their compositions according to the bottom slope and deep-water wave height and period are presented in Figure 6.

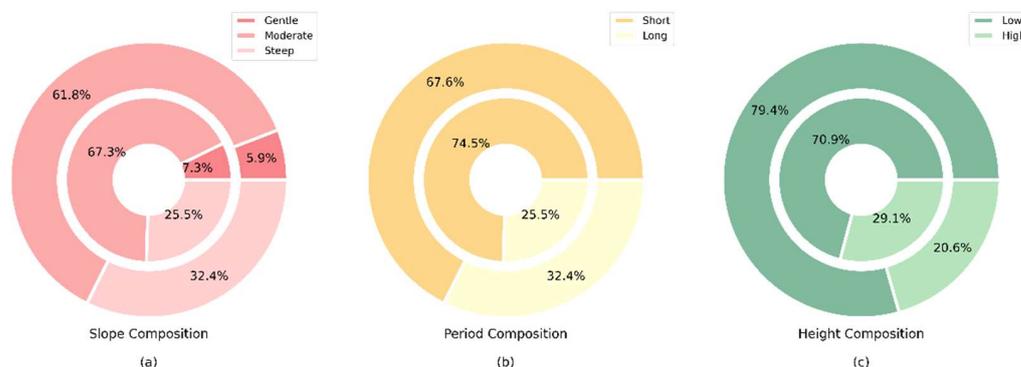


Figure 6. Composition of components according to each wave condition. The outer circle indicates 55 test data to compare the performance of the proposed model and the formulas of Rattanapitikon and Shibayama (CA_Hb_2 and CA_hb_2). The inner circle indicates the 34 test data used to compare the performance of the proposed model and the formulas of Lee and Cho (CA_Hb_3 and CA_hb_3): (a) bottom slope; (b) wave period; and (c) deep-water wave height.

The performances of a multilayer neural network and the equations of Rattanapitikon and Shibayama (CA_Hb_2 and CA_hb_2) for 55 test data out of 129 (data used to calculate the Rattanapitikon and Shibayama equations excluded) are provided in Figure 7. Figure 8 presents the breaking-wave height and water depth predictions obtained using the proposed model and the Rattanapitikon and Shibayama equations by applying the new dataset specified above. First, the performance for estimating breaking-wave height was compared in Figure 7a–c. The proposed model demonstrated a higher performance for B compared to the empirical equation of Rattanapitikon and Shibayama (CA_Hb_2). This was identified in the majority of the predicted values from the empirical equation that remained below the perfect agreement line, while the predicted values of the proposed model in Figure 8a were similar to the perfect agreement line in all areas. The proposed model and empirical equation had similar results for RMSE and R. The performances for the breaking-water depth were evaluated using Figure 7d–f. Unlike the breaking-wave height results, the values of B obtained using the proposed model and the empirical equation (CA_hb_2) were similar. In contrast, the empirical equation exhibited higher performance in terms of the RMSE and R. This is presumably because, as shown in Figure 8b, the predicted values of the empirical equation did not significantly deviate from the perfect agreement line in the majority of the areas, but the proposed model tended to overestimate when the observation value was less than or equal to 0.1 m.

Figure 9 presents the statistics of the error metric for the proposed model and the equations of Lee and Cho (CA_Hb_3 and CA_hb_3), which were calculated using 34 data out of 129 (the data used to calculate the Lee and Cho equations were excluded). Figure 10 provides a comparison between the estimations for the breaking-wave height and water depth that were obtained with the proposed model and the empirical equations by employing the previously introduced new dataset. First, the estimations for the breaking-wave height were compared in Figure 9a–c. The proposed model exhibited a slightly higher performance for B than the equation of Lee and Cho (CA_Hb_3). For the remaining evaluation indicators, the empirical equation demonstrated a higher performance. Next, the predicted results for the breaking-water depth were compared in Figure 9d–f. Unlike the results mentioned earlier, the equation of Lee and Cho (CA_hb_3) provided a smaller B absolute value than the proposed model. However, the RMSE and R values of the proposed model were superior to those of the former.

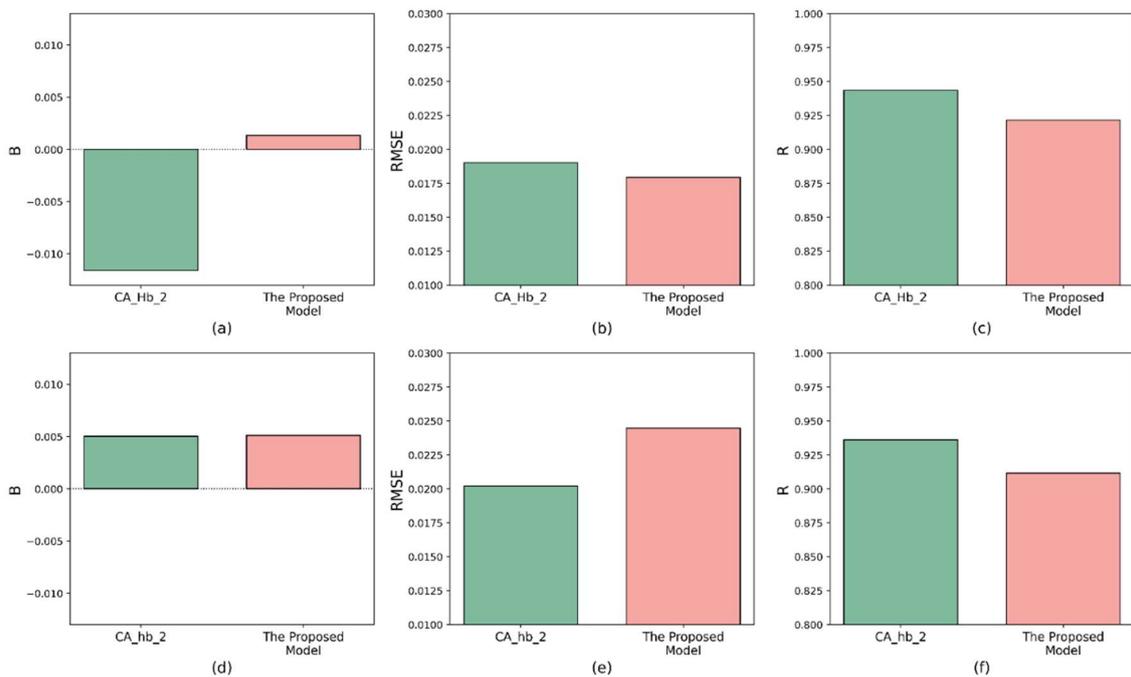


Figure 7. Results of performance analysis of the proposed model and the Rattanapitikon and Shibayama formulas (CA_Hb_2 and CA_hb_2) against the relative complement of dataset used in the formulas of Rattanapitikon and Shibayama in the test dataset: (a) B , (b) $RMSE$, and (c) R for breaking-wave height; (d) B , (e) $RMSE$ and (f) R for breaking-water depth.

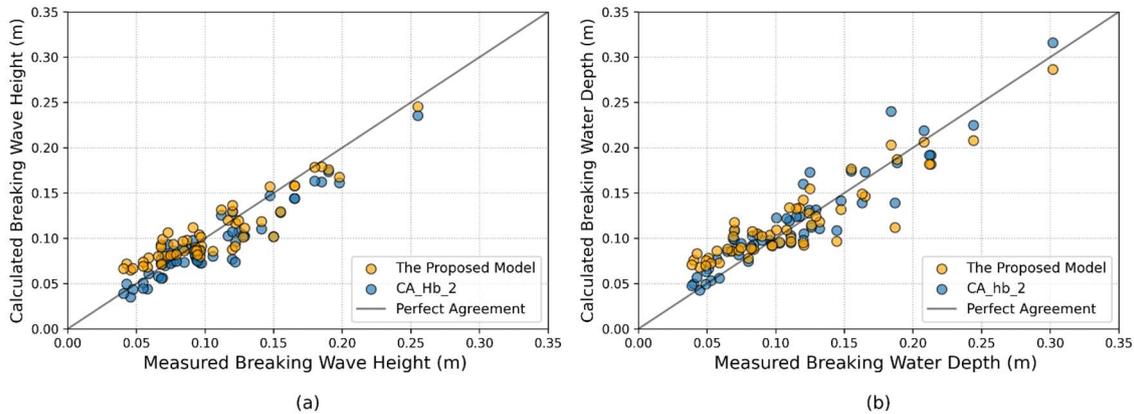


Figure 8. Comparison of the proposed model and Rattanapitikon and Shibayama formulas (CA_Hb_2 and CA_hb_2) against the relative complement of the dataset used in the Rattanapitikon and Shibayama formulas in the test dataset: (a) breaking-wave height, (b) breaking-water depth.

The proposed model did not demonstrate a significantly higher performance than the previous breaking-wave formulas as the amount of data used for the training was noticeably small, and the targets with observation values less than or equal to 0.1 m were overestimated, as pointed out above. That can be confirmed through Figures 8 and 10.

The results in the case where the sigmoid activation function is not applied (denoted as NonAF) for the modeling of the linear relationship in the same network architecture in order to compare the performances with and without the consideration of the nonlinearity between the floor slope, deep-water wave height, and period and the breaking-wave height and water depth, which correspond to the results of training the linear relationship, are listed in Table 3. Based on the aggregate statistical values presented in Table 3, the linear model derived from our proposed network architecture (NonAF) can be observed to exhibit a slightly higher performance compared to the proposed model. Such a result was obtained

because the proposed model tended to overestimate all of the targets with observation values less than or equal to 0.1 m, as shown in Figure 11a,b. In contrast, for observation values greater than 0.1 m, no significant difference in performance between the models was observed. However, this does not imply a linear relationship between the input wave characteristics and wave-breaking parameters. The main reason for this is that most of the input wave characteristics were disproportionately from a one-dimensional, small-scale wave-flume laboratory experiment rather than field conditions. In addition, the results of previous studies have demonstrated that wave breaking is an extremely unpredictable phenomenon. Most of all, the amount of data used for training and testing was insufficient. These aspects are expected to improve as data acquisition is updated in the future.

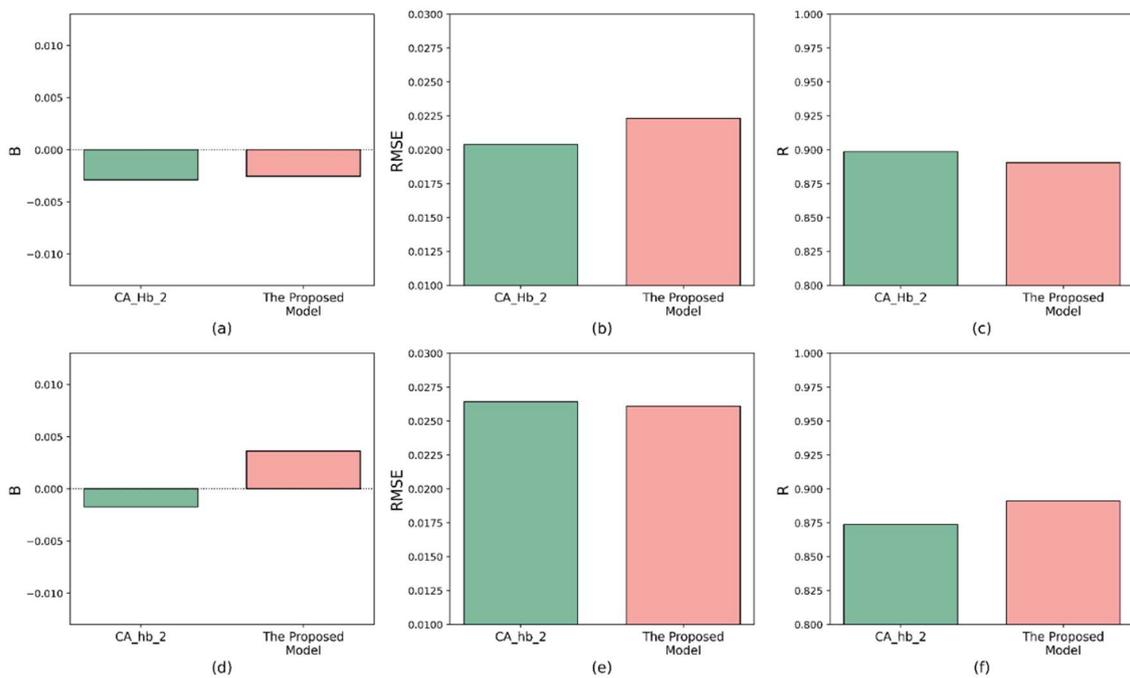


Figure 9. Results of the performance analysis of the proposed model and the Lee and Cho formulas (CA_Hb_2 and CA_hb_2) against the relative complement of the dataset used in the Lee and Cho formulas in the test dataset: (a) B, (b) RMSE, and (c) R for breaking-wave height; (d) B, (e) RMSE, and (f) R for breaking-water depth.

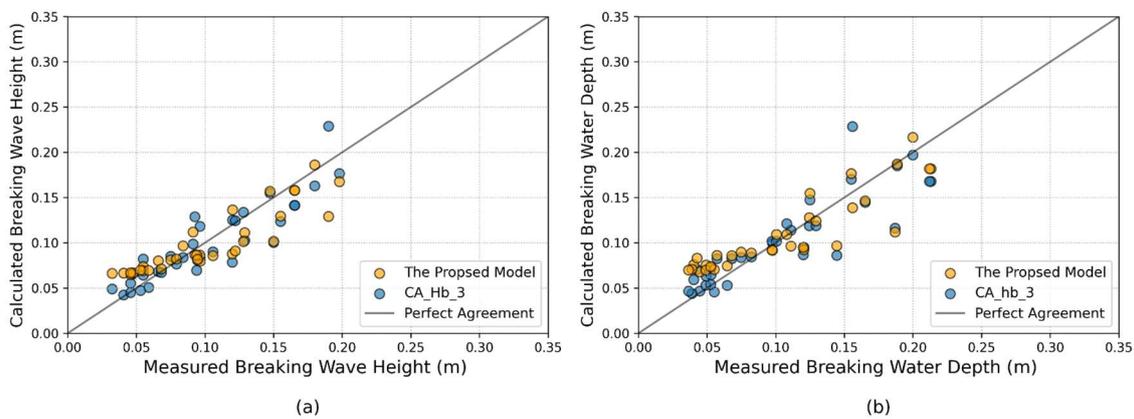


Figure 10. Comparison of the proposed model and the Lee and Cho formulas (CA_Hb_3 and CA_hb_3) against the relative complement of the dataset used in the Lee and Cho formulas in the test dataset: (a) breaking-wave height, (b) breaking-water depth.

Table 3. Performance of the proposed model and the NonAF model.

Target	Relationship	<i>B</i>	<i>RMSE</i>	<i>R</i>
H_b	Proposed Model	0.004	0.019	0.894
	NonAF	0.000	0.017	0.912
h_b	Proposed Model	0.005	0.021	0.921
	NonAF	0.000	0.021	0.916

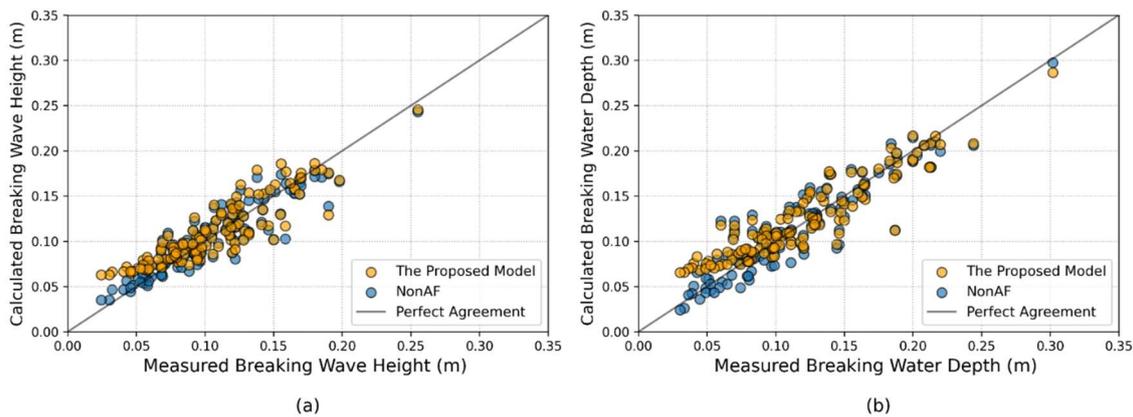


Figure 11. Comparison of the proposed model and the NonAF model against the test data: (a) breaking-wave height, (b) breaking-water depth.

5. Discussion and Conclusions

The breaking wave is a significant factor for predicting and responding to the coastal hydrodynamic and environmental issues, so numerous theoretical studies and wave-flume laboratory experiments were conducted in order to acquire the prediction accuracy and knowledge of wave breaking. However, the nonlinearity between parameters, such as the bottom slope, deep-water wave height, and period, has not been fully incorporated into the existing empirical equations. Therefore, this study proposed a multilayer neural network utilizing a nonlinear activation function and backpropagation in order to investigate the effects of nonlinearity. From 31 sources, 630 data were collected to construct the proposed network for estimating the breaking-wave index. After excluding the data that were inapplicable for analysis, 584 data were used for training and testing the newly proposed model. Approximately 80% of the total data (455) were randomly selected for training the model, and the remaining data (129) were used to evaluate the performance. The bottom slope, deep-water wave height, and period were plugged in as the input variables that simultaneously estimated the breaking-wave height and wave-breaking location. The estimated breaking-wave-index performances were evaluated using error metrics and compared with the existing wave-breaking formulas.

Consequently, the performance of the proposed model was better than the existing breaking-wave-index formulas as well as being robustly applicable to laboratory experiment conditions, such as wave condition, bottom slope, and experimental scale. Furthermore, the proposed method directly predicted the breaking-wave height and water depth with nondimensionalization. The input and target variables of the proposed model were not nondimensionalized and directly estimated the breaking-wave height and water depth. So, it can exclude errors from the secondary transformation of raw wave data and improve the prediction performance. However, it tended to overestimate the breaking-wave height and water depth in the case of observation values less than or equal to 0.1 m. The proposed method is expected to show significantly better performance compared to the existing method if there is sufficient data. Therefore, in future studies, more wave-flume laboratory experiments will be performed under various conditions in order to obtain the data of the bottom slope, deep-water wave height and period, and breaking-wave height and

water depth. In addition, the spatial distribution of continuous water surface elevation is important for accurately estimating the breaking-wave height and water depth. However, it is difficult to obtain that information using conventional acoustic sensors. Thus, if continuous wave measurement in space and time is possible through visual intelligence of experimental video, more data can be obtained, so more accurate wave-breaker-index estimation and high applicability can be expected.

Contribution Points

- The accuracy of estimating the breaking-wave height and water depth was improved by fully incorporating the nonlinear relationship between deep water wave condition, bottom slope, and wave-breaker index.
- Furthermore, a single model was proposed for simultaneously estimating the breaking-wave height and water depth by setting the input variable as deep-water wave data. This gave invaluable usability to the proposed model in this study.
- The performance of the proposed model is robustly applicable to laboratory experiment conditions, such as wave condition, bottom slope, and experimental scale.
- The newly proposed model directly utilizes breaking-wave height and water depth without nondimensionalization; thus, applicability can be significantly improved and excludes errors from the secondary transformation of raw wave data.

Author Contributions: Conceptualization, J.K. and K.D.; methodology, J.K.; software, J.K.; validation, M.Y., J.K. and K.D.; formal analysis, M.Y., J.K. and K.D.; investigation, M.Y.; resources, M.Y.; data curation, M.Y.; writing—original draft preparation, M.Y., J.K. and K.D.; writing—review and editing, M.Y., J.K. and K.D.; visualization, M.Y.; supervision, J.K. and K.D.; project administration, J.K. and K.D.; funding acquisition, J.K. and K.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly supported by the National Research Foundation of Korea (NRF-2019R1C1C1003160) and the project titled ‘Establishment of the Ocean Research Station in the Jurisdiction Zone and Convergence Research’, funded by the Ministry of Oceans and Fisheries.

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

References

1. Davidson-Arnott, R.; Bauer, B.; Houser, C. *Introduction to Coastal Processes and Geomorphology*; Cambridge University Press: Cambridge, UK, 2019; p. 458. ISBN 9781108424271.
2. López, I.; López, M.; Iglesias, G. Artificial neural networks applied to port operability assessment. *Ocean Eng.* **2015**, *109*, 298–308. [[CrossRef](#)]
3. Stokes, G.G. On the theory of oscillatory waves. *Trans. Camb. Philos. Soc.* **1847**, *8*, 441–473.
4. Galvin, C.J. Wave breaking in shallow water. In *Waves on Beaches and Resulting Sediment Transport*; Meyer, R.E., Ed.; Academic Press: New York, NY, USA, 1972; pp. 413–455.
5. Yamada, H.; Shiotani, T. On the highest water waves of permanent type. *Bull. Disaster Prev. Res. Inst.* **1968**, *18*, 1–22.
6. Goda, Y. A synthesis of breaker indices. *Porc. Jpn. Soc. Civ. Eng.* **1970**, *180*, 39–49. (In Japanese) [[CrossRef](#)]
7. Camenen, B.; Larson, M. Predictive formulas for breaker depth index and breaker type. *J. Coast. Res.* **2007**, *234*, 1028–1041. [[CrossRef](#)]
8. Goda, Y. Reanalysis of regular and random breaking wave statistics. *Coast. Eng. J.* **2010**, *52*, 71–106. [[CrossRef](#)]
9. Robertson, B.; Gharabaghi, B.; Hall, K. Prediction of incipient breaking wave-heights using artificial neural networks and empirical relationships. *Coast. Eng. J.* **2015**, *57*, 348–363. [[CrossRef](#)]
10. Choi, B.J.; Park, C.W.; Cho, Y.H.; Kim, D.S.; Lee, K.H. A proposal of new breaker index formula using supervised machine learning. *J. Korean Soc. Coast. Ocean Eng.* **2020**, *32*, 384–395. [[CrossRef](#)]
11. Michell, J.H. The highest waves in water. *Philos. Mag.* **1893**, *36*, 430–437. [[CrossRef](#)]
12. McCowan, J. On the highest wave of permanent type. *Philos. Mag.* **1894**, *38*, 351–358. [[CrossRef](#)]
13. Miche, M. Mouvements ondulatoires de la mer en profondeur constante ou décroissante. *Ann. Ponts Chaussées* **1944**, *2*, 25–78. (In French)
14. Ippen, A.T.; Kulin, G. The shoaling and breaking of the solitary wave. *Coast. Eng. Proc.* **1954**, *1*, 4. [[CrossRef](#)]
15. Smith, E.R.; Kraus, N.C. *Laboratory Study on Macro-Features of Wave Breaking over Bars and Artificial Reefs*; TR-CERC-90-12; USACE-WES: Vicksburg, MS, USA, 1990.
16. Munk, W.H. The Solitary wave theory and its application to surf problems. *Ann. N. Y. Acad. Sci.* **1949**, *51*, 376–424. [[CrossRef](#)]

17. Iversen, H.W. Laboratory study of breakers. *NBS Circular*. **1952**, *521*, 9–32.
18. Weggel, J.R. Maximum breaker height. *J. Waterw. Harb. Coast. Eng. Div.* **1972**, *98*, 529–548. [[CrossRef](#)]
19. Komar, P.D.; Gaughan, M.K. Airy wave theory and breaker height prediction. In Proceedings of the 13th International Conference on Coastal Engineering, Vancouver, BC, Canada, 10–14 July 1972; pp. 405–418.
20. Maruyama, K.; Sakakiyama, T.; Kajima, R.; Saito, S.; Shimizu, T. *Experiment Study on Wave Height and Water Particle Velocity Near the Surf Zone Using a Large Wave Flume*; Report No. 382034; The Central Research Institute of Electric Power Industry: Chiba, Japan, 1983. (In Japanese)
21. Stive, M.J.F. A scale comparison of waves breaking on a beach. *Coast. Eng.* **1985**, *9*, 151–158. [[CrossRef](#)]
22. Goda, Y. New wave pressure formulae for composite breakwaters. In Proceedings of the 14th International Conference on Coastal Engineering, Copenhagen, Denmark, 24–28 June 1974; pp. 1702–1720. [[CrossRef](#)]
23. Sunamura, T.; Horikawa, K. Two-dimensional beach transformation due to waves. In Proceedings of the 14th International Conference on Coastal Engineering, Copenhagen, Denmark, 24–28 June 1974; Volume 2, pp. 920–938. [[CrossRef](#)]
24. Ostendorf, D.W.; Madsen, O.S. *An Analysis of Longshore Currents and Associated Sediment Transport in the Surf Zone*; Report No. 241; Massachusetts Institute of Technology: Cambridge, MA, USA, 1979; p. 169.
25. Ogawa, Y.; Shuto, N. Run-up of periodic waves on beaches of non-uniform slope. In Proceedings of the 19th International Conference on Coastal Engineering, Houston, TX, USA, 3–7 September 1984; pp. 328–344. [[CrossRef](#)]
26. Seyama, A.; Kimura, A. The measured properties of irregular wave breaking and wave height change after breaking on the slope. In Proceedings of the 21st International Conference on Coastal Engineering, Malaga, Spain, 20–25 June 1988; pp. 419–432. [[CrossRef](#)]
27. Muttray, M.; Oumeraci, H. Wave transformation on the foreshore of coastal structures. In Proceedings of the 27th International Conference on Coastal Engineering, Sidney, Australia, 16–21 July 2000; pp. 2178–2191. [[CrossRef](#)]
28. Rattanapitikon, W.; Shibayama, T. Verification and modification of breaker height formulas. *Coast. Eng. J.* **2000**, *42*, 389–406. [[CrossRef](#)]
29. Xie, W.; Shibayama, T.; Esteban, M. A semi-empirical formula for calculating the breaking depth of plunging waves. *Coast. Eng. J.* **2019**, *61*, 199–209. [[CrossRef](#)]
30. Tadayon, B.; Dehghani, H.; Ershadi, C. Proposing new breaking wave height prediction formulae using gene expression programming. *Ocean Eng.* **2021**, *228*, 108952. [[CrossRef](#)]
31. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
32. Le Méhauté, B.; Koh, R.C.Y. On the breaking of waves arriving at an angle to the shore. *J. Hydraul. Res.* **1967**, *5*, 67–88. [[CrossRef](#)]
33. Liu, Y.; Niu, X.; Yu, X. A new predictive formula for inception of regular wave breaking. *Coast. Eng.* **2011**, *58*, 877–889. [[CrossRef](#)]
34. Galvin, C.J. Breaker travel and choice of design wave height. *J. Waterw. Harb. Div.* **1969**, *95*, 175–200. [[CrossRef](#)]
35. Collins, J.I.; Weir, W. *Probabilities of Wave Characteristics in the Surf Zone*; Tetra Tech Report No. TC. 149; Tetra Tech Inc.: Pasadena, CA, USA, 1969; p. 122.
36. Suquet, F. Experimental study on the breaking of waves. *La Houille Blanche* **1950**, *36*, 342–361. [[CrossRef](#)]
37. Hamada, T. *Breakers and Beach Erosion*; Port and Harbor Research Institute, Ministry of Transportation: Yokosuka, Japan, 1963; p. 165.
38. Rattanapitikon, W.; Shibayama, T. Breaking wave formulas for breaking depth and orbital to phase velocity ratio. *Coast. Eng. J.* **2006**, *48*, 395–416. [[CrossRef](#)]
39. Kajima, R.; Shimizu, T.; Maruyama, K.; Saito, S. Experiments on beach profile change with a large wave flume. In Proceedings of the 18th International Conference on Coastal Engineering, Cape Town, South Africa, 14–19 July 1982; pp. 1385–1404. [[CrossRef](#)]
40. Kraus, N.C.; Smith, J.M. *SUPERTANK Laboratory Data Collection Project*; TR-CERC-94-3; USACE-WES: Vicksburg, MS, USA, 1994.
41. Lara, J.L.; Losada, I.J.; Liu, P.L.F. Breaking waves over a mild gravel slope: Experimental and numerical analysis. *J. Geophys. Res.* **2006**, *111*, C11019. [[CrossRef](#)]
42. Lee, K.-H.; Cho, Y.-H. Simple breaker index formula using linear model. *J. Mar. Sci. Eng.* **2021**, *9*, 731. [[CrossRef](#)]
43. Gaughan, M.K. *Breaking Waves: A Review of Theory and Measurements*. Master's Thesis, Oregon State University, Corvallis, OR, USA, 1973.
44. Bowen, A.J.; Inman, D.L.; Simmons, V.P. Wave 'Set-down' and Set-up. *J. Geophys. Res.* **1968**, *73*, 2569–2577. [[CrossRef](#)]
45. Weggel, J.R.; Maxwell, W.H.C. Experimental study of Breaking Wave Pressures. In Proceedings of the Offshore Technology Conference, Houston, TX, USA, 21–23 April 1970.
46. Ozaki, A.; Sasaki, M.; Usui, Y. Study on rip currents: Experimental observation of nearshore circulation on a sloping bottom. *Coast. Eng. J.* **1977**, *20*, 147–158. [[CrossRef](#)]
47. Van Dorn, W.G. Breaking invariants in shoaling waves. *J. Geophys. Res.* **1978**, *83*, 2981–2988. [[CrossRef](#)]
48. Kirkgoz, M.S. Shock pressure of breaking waves on vertical walls. *J. Waterw. Port Coast. Ocean Div.* **1982**, *108*, 81–95. [[CrossRef](#)]
49. Ishida, H.; Yamaguchi, N. A theory for wave breaking on slopes and its application. In Proceedings of the 30th Japanese Conference on Coastal Engineering, Hokkaido, Japan, 1 November 1983; pp. 34–38. (In Japanese).
50. Sakai, S.; Kazumi, S.; Ono, T.; Yamashita, T.; Saeki, H. Study on wave breaking and its resulting entrainment of air. In Proceedings of the 33rd Japanese Conference on Coastal Engineering, Nagasaki, Japan, 5 November 1986; pp. 16–20. (In Japanese).
51. Ting, F.C.K.; Kirby, J.T. Observation of undertow and turbulence in a laboratory surf zone. *Coast. Eng.* **1994**, *24*, 51–80. [[CrossRef](#)]

52. Kakuno, S.; Sugita, T.; Goda, T. Effects of wave breaking on entrainment of oxygen, a review. In Proceedings of the 43rd Japanese Conference on Coastal Engineering, Wakayama, Japan, 13–15 November 1996; pp. 1211–1215. (In Japanese).
53. Yüksel, Y.; Bostan, T.; Çevik, E.; Çelikoglu, Y.; Günal, M. Two-phase flow structure in breaking waves. In Proceedings of the 9th International Offshore and Polar Engineering Conference, Brest, France, 30 May–4 June 1999; pp. 231–235.
54. Hoque, A. Air Bubble Entrainment by Breaking Waves and Associated Energy Dissipation. Ph.D. Thesis, Toyohashi University of Technology, Toyohashi, Japan, 2002.
55. Shin, S.; Cox, D. Laboratory observations of inner surf and swash-zone hydrodynamics on a steep slope. *Cont. Shelf Res.* **2006**, *26*, 561–573. [[CrossRef](#)]
56. Deo, M.C.; Jagdale, S.S. Prediction of breaking waves with neural networks. *Ocean Eng.* **2003**, *30*, 1163–1178. [[CrossRef](#)]
57. Mori, N.; Kakuno, S. Aeration and bubble measurements of coastal breaking waves. *Fluid Dyn. Res.* **2008**, *40*, 616–626. [[CrossRef](#)]
58. Sibi, P.; Jones, S.A.; Siddarth, P. Analysis of different activation functions using back propagation neural networks. *J. Theor. Appl. Inf. Technol.* **2013**, *47*, 1264–1268.
59. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.