water

MDPI

Article

Analysis of Factors Influencing Wave Overtopping Discharge from Breakwater Based on an MIV-BP Estimation Model

Songgui Chen [1], Hanbao Chen [1,*], Cheng Peng [1,*], Yina Wang [1] and Yuanye Hu [1,2]

[1] Tianjin Research Institute of Water Transport Engineering, Ministry of Transport, Tianjin 300456, China
[2] School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China
* Correspondence: chenhanbao@163.com (H.C.); 18630903060@163.com (C.P.)

Abstract: Aiming at the problem of calculating the overtopping of single-slope breakwaters, a mean impact value-backpropagation (MIV-BP) estimation model for predicting overtopping was established. Experimental data from the Tianjin Research Institute of Water Transport Engineering (TIWTE) were utilized to further enrich the dataset of the CLASH project for single-slope wave overtopping discharge. This paper established a comprehensive prediction model based on an ensemble learning average method combination strategy. There are 10 input parameters in the model, including the offshore effective wave height, average wave period, offshore water depth, toe submergence, toe width, slope tangent, armor rock surface roughness factor, crest height with respect to the static water level, wall height with respect to the static water level, and crest width; the output parameter is the mean overtopping discharge. Subsequently, a comparative analysis was conducted between this estimation model, the Chinese standard formula calculation model, and the European Van der Meer formula calculation model. Compared with the two formulas mentioned above, this estimation model's coefficient of correlation increased by 0.23 and 0.26, respectively. Finally, a weight evaluation analysis of the 10 main factors affecting overtopping was carried out based on a MIV-BP neural network model. In the analysis, a positive correlation was found for factors, such as the wave height, average wave period, and water depth at the structure toe; a negative correlation was found for factors, such as the slope, crest height with respect to the static water level, wall height with respect to the static water level, and crest width. Overall, the results provide a significant basis and reference for optimizing the design of the wave overtopping control.

Keywords: wave overtopping; ANN; mean impact value; ensemble learning; influencing factors

Citation: Chen, S.; Chen, H.; Peng, C.; Wang, Y.; Hu, Y. Analysis of Factors Influencing Wave Overtopping Discharge from Breakwater Based on an MIV-BP Estimation Model. *Water* **2022**, *14*, 2967. https://doi.org/10.3390/w14192967

Academic Editor: Rafael J. Bergillos

Received: 4 August 2022
Accepted: 19 September 2022
Published: 21 September 2022

1. Introduction

Breakwaters are an important hydraulic structure in many coastal areas and are vital for protecting the infrastructure of harbors. Overtopping is one of the factors that cause breakwater damage [1]. Firstly, it washes away the crest. Secondly, it washes away the back slope, which results in the instability of breakwaters [2]. Thirdly, it increases the water content of breakwaters, which will further compromise their stability. Therefore, it is of great importance to estimate the wave overtopping discharge and to analyze the various influencing factors.

The conventional method for estimating wave overtopping discharge is to use a calculation formula based on different physical modeling experiments. For example, a standard formula was proposed by Wang and Zhang et al. [3,4] based on their research results, which was included in China's "Code of Hydrology for Port and Waterway Engineering" [5]. The overtopping discharge calculation formula proposed by Van der Meer et al. [6,7] after an extensive study on slope dyke overtopping discharge is also widely used in European countries. Similarly, the overtopping discharge calculation formula included in the "Design of Wave Resistant Harbor Structure" by Goda [8] is generally used in Japan.

A comparative analysis of different overtopping discharge calculation formulae was conducted by Yu [9], who identified the applicability of each calculation formula. Most of these calculation formulae are only applicable in a small range due to the limited parameters captured. Calculation results based on different formulae may also vary substantially. Dong et al. [10–12] conducted major laboratory physical modeling, which was conducted to examine the performance in mitigating wave overtopping, and the spatial distribution of the wave overtopping volume behind a plain vertical seawall and a seawall with recurve retrofitting. Moreover, they investigated the effects of the geometrical properties of the recurve on the pressure distribution, overall force, and overturning moment at the seawall, subject to both impulsive and non-impulsive waves. Salaudin et al. [13] analyzed wave-by-wave overtopping volumes through physical model experiments on a 1V:2H sloped impermeable structure. The following year, Salaudin et al. [14] presented the results of a small-scale laboratory study of wave overtopping on artificial rough seawalls. Subsequently, the seawall was modified to include 10 further test configurations. The results of physical model tests show that increasing the length and/or density of surface protrusions can reduce the overtopping on the seawall.

With the fast development of computer technology, computer numerical calculation has been used in the field of water engineering, and it is possible to calculate the wave overtopping discharge with numerical simulation[15]. For example, Losada [16] established the BRAS (Cornell Breaking Waves and Structure) model to analyze the interaction between waves and breakwaters and simulated the overtopping process. Using CFD software, Guan [17] developed a 3D numerical model related to wave tanks. This software can be used to analyze overtopping-related issues by exercising overtopping differential equations. Based on the SPH hydrodynamic mathematical model, Ye [18] and Ren [19] analyzed the overtopping process by studying the interaction between regular waves and breakwaters. Dasineh et al. [20] researched the characteristics of free and submerged hydraulic jumps on the triangular bed roughness with various T/I ratios using CFD techniques. The accuracy of the numerical simulation results was checked and compared using an artificial intelligence method. Moreover, Bagherzadeh et al. [21] simulated the energy dissipation by a serrated-edge drop downstream of this structure using FLOW-3D software. Abolfathi et al. [22] developed a weakly compressible smoothed particle hydrodynamic (WCSPH) model to quantify the overtopping volume due to the impacts of wave actions. There are many drawbacks to numerical simulation methods, such as the complex modeling process and the huge number of calculations required to improve the calculation accuracy.

Artificial intelligence algorithms represented by machine learning can take a large number of parameter variables into consideration, which provides a new methodology for the estimation of overtopping discharge. Daneshfaraz et al. [23,24] investigated the application of an SVM (support vector machine) for predicting vertical-drop hydraulic parameters in the presence of horizontal screens. The European CLASH project started the development of an overtopping discharge databank a long time ago and has established an overtopping estimation method based on artificial intelligence. For example, Van Gent [25] developed an overtopping estimation model based on a neural network algorithm. Formentin [26] introduced extra input parameters to the neural network and further improved the network model. Van der Meer and Bruce further studied the early Dutch overtopping experiment in EurOtop [27] and deduced the calculation formula of overtopping.

Though there are many methods to estimate overtopping discharge, previous studies paid less attention to the influencing factors of overtopping. However, our proposed approach, i.e., MIV-BP, based on artificial intelligence, can not only evaluate overtopping but can also discuss the factors affecting the overtopping. Furthermore, considering that the performance of the neural network depends on the size and range of the training set, we further supplied the European dataset based on the experimental data from TIWTE.

2. Dataset

2.1. European Overtopping Discharge Dataset

The European CLASH dataset covers the vast majority of common breakwater structures. There are 17,942 experimental records. Each experimental record includes 40 parameters, in which there are 3 common variables, 14 wave element parameters, and 23 breakwater structure parameters (see Van der Meer et al. [27] for the detailed description of the dataset).

In this work, the following sections were analyzed as a dataset for the single-slope wave overtopping discharge study, with the following experimental labels: A. permeable armor rock single-slope embankment; B. impermeable armor rock single-slope embankment; C. artificial armor rock single-slope embankment; D. smooth impermeable single-slope embankment. Erroneous and missing data can degrade the performance of the prediction model, so the original dataset was reduced by removing the data marked as "non-core data". The data that were viewed as very complex (CF = 4) and non-reliable (RF = 4) were deleted. We considered the errors of data with overtopping discharge $q < 10^{-6}$ $m^3/m/s$, which were eliminated. The operation of the model cannot have null values, so the data with missing values needed to be deleted. After data processing, the dataset was distributed, as shown in Figure 1.

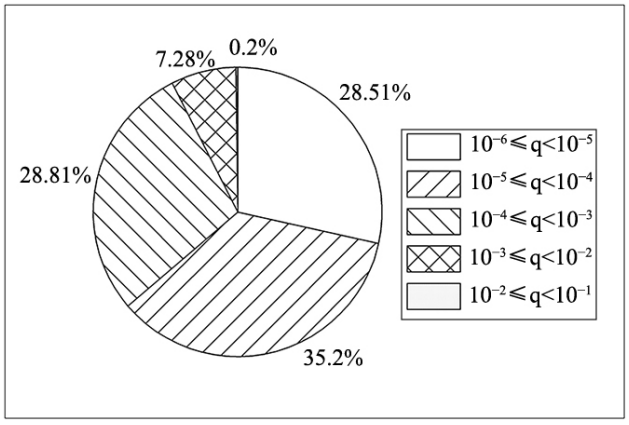

Figure 1. European dataset distribution.

The dataset covers 2020 experimental data with overtopping discharge ranging from $10^{-6} \leq q < 10^{-1}$. There were 711 experimental data with overtopping discharge ranging from $10^{-5} \leq q < 10^{-4}$, which accounted for the largest portion of the dataset, at 35.2%. There were four experimental data points with overtopping discharge ranging from $10^{-2} \leq q < 10^{-1}$, which accounted for the smallest portion of the dataset, at 0.2%.

2.2. TIWTE Overtopping Discharge Dataset

The training result of neural networks largely depends on the training set. A training set with an extensive amount of data may produce a better training result. Based on experimental reports of TIWTE's mass physical model, this paper selected 236 sets of experimental data. The distribution of these data is shown in Figure 2.

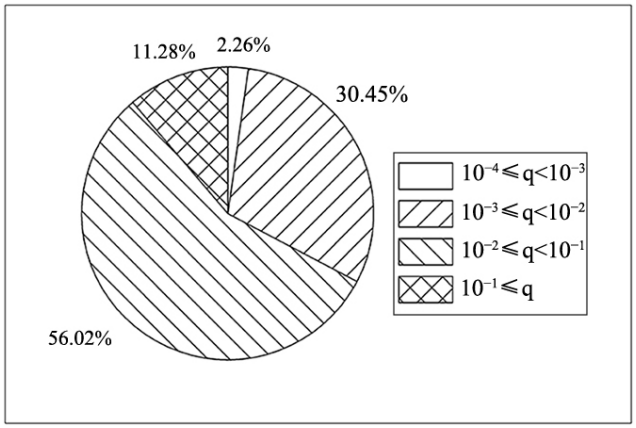

Figure 2. TIWTE dataset distribution.

The dataset captured 266 pieces of experimental data, in which all overtopping discharges were $q \geq 10^{-4}$. There were 149 experimental data with overtopping discharge ranging from $10^{-2} \leq q < 10^{-1}$, which accounted for the largest portion of the dataset, at 56.02%. There were six experimental data points with overtopping discharge ranging from $10^{-4} \leq q < 10^{-3}$, which accounted for the smallest portion of the dataset, at 2.26%.

2.3. Dataset Combination

Figure 3 shows the statistical results following the combination of the European and TIWTE single-slope wave overtopping discharge values into one dataset. Numbers 1 to 6 on the horizontal axis represent six overtopping discharge values between $q = 10^{-6}$ and $q = 10^{-0}$. TIWTE's dataset is a supplement to the European dataset in the area of overtopping discharge, which further expands the dataset and provides more training sets for training the neural network.

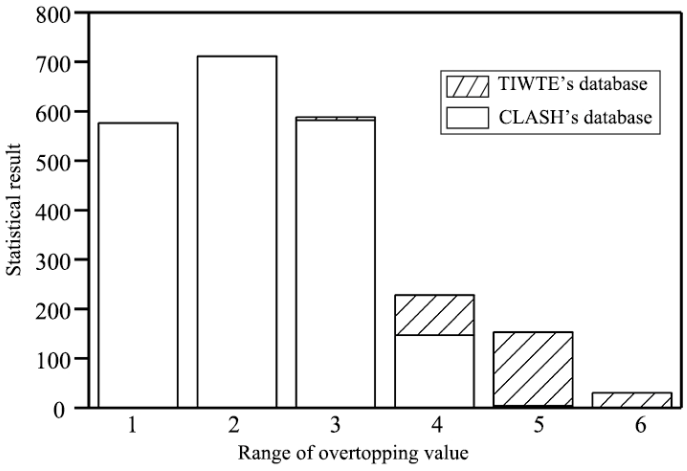

Figure 3. Statistical distribution of all datasets.

Figure 4 shows the distribution of the aggregated dataset. The graph shows that such a dataset is broader, has more data, and increases the percentage of overtopping discharge $q \geq 10^{-3}$ compared to the European dataset, which enriches the dataset for single-slope overtopping discharges.

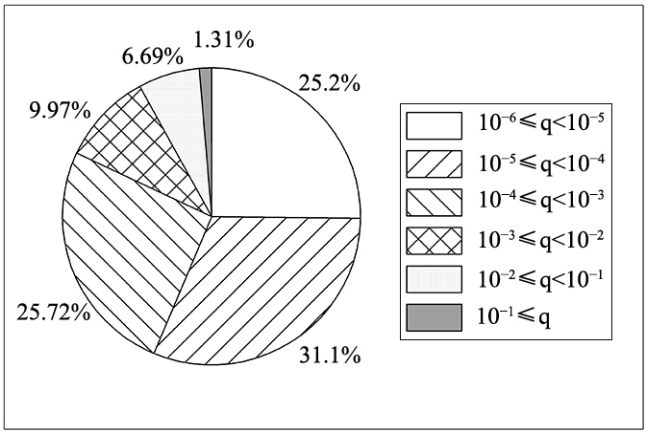

Figure 4. Distribution of all datasets.

3. Method

3.1. The Principles of the BP Algorithm

The BP (backpropagation) neural network algorithm is a multi-layer feedforward network trained according to the error backpropagation algorithm[28].. By optimizing the function weight, the corresponding error is propagated to form a three-layer network, which has an input layer, a hidden layer, and an output layer [29]. In order to develop a BP neural network, the first step is to determine the number of layers and number of neurons contained in each layer. The learning principle is through gradient descent and continuously adjusts to the weighting parameter X and threshold parameter b by applying the backpropagation algorithm to minimize the mean square error.

Suppose that $X = [x_1, x_2, \cdots, x_M]$ is used as the input space and $Y = [Y_1, Y_2, \cdots, Y_j]$ is used as the output space. The function expression for this network's forward propagation output is $\hat{Y} = f_{w,b}(X)(w, b \in R^n)$. The process of network training is to continuously adjust the value of w and b based on the training sample set's backpropagation algorithm and minimize the value of the defined loss parameter $Loss(Y, f_{w,b}(X))$.

3.2. Details of the BP Algorithm Implementation

Since there are many parameters involved in wave overtopping processes, it is difficult to describe all of them. Given the structural features of single-slope breakwaters, 10 data were selected as input parameters for the network, including wave element and breakwater structural parameters, as shown in Figure 5: offshore effective wave height $H_{m0,t}$, average wave period $T_{m-1,0}$, offshore water depth h, toe submergence h_t, toe width B_t, slope tangent $cot\,\alpha$, armor rock surface roughness factor γ_f, crest height with respect to the static water level R_c, wall height with respect to the static water level A_c, and crest width G_c.

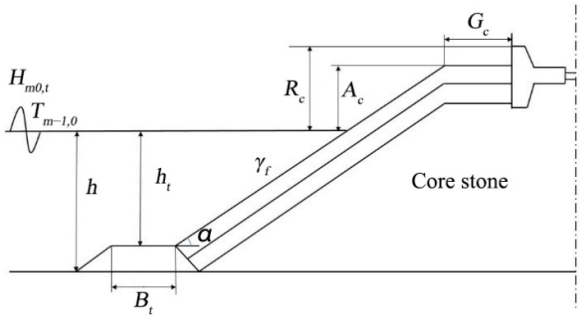

Figure 5. Parameters of straight slopes [27].

To eliminate the influence of the model scale and dimensions among the data and facilitate the training of the neural network, these data were non-dimensionalized. For the purpose of this research, the offshore effective wave height for each experimental record was scaled to $H'_{m0,t} = 1m$, and the scaling was recorded as λ. Meanwhile, other data were scaled in accordance with Froude scaling. Table 1 shows the maximum and minimum values of each parameter. After normalization of the input parameters for single-slope breakwaters, all parameters were within the designated range, which is under the coverage of the neural network elaborated in this paper.

Table 1. Distribution characteristics of input parameter data of the training set.

Parameters	Mean Value	Variance	Maximum Value	Minimum Value
h'	3.34	1.88	14.40	1.03
$H'_{mo,t}$	1	0	1	1
$T'_{m-1,0}$	4.05	0.98	15.57	2.55
h'_t	3.00	2.02	14.40	0.66
B'_t	0.92	1.47	10.18	0
$cot\,\alpha'$	0.46	0.06	0.66	0.38
γ'_f	1.84	0.68	5	1.30
R'_c	1.35	0.48	3.04	0
A'_c	1.32	0.60	3.75	0
G'_c	1.46	0.96	12.50	0
q^*	0.45	0.15	1	0

The number of hidden layer nodes has a great impact on the performance of neural networks. This paper used the step-by-step test method to determine the number of hidden layer nodes. That is, we gradually increased the number of hidden layer nodes from an initial value, compared the prediction performance of each network, and finally selected the corresponding number of nodes with the best performance as the number of hidden layer neuron nodes, as shown in Figure 6.

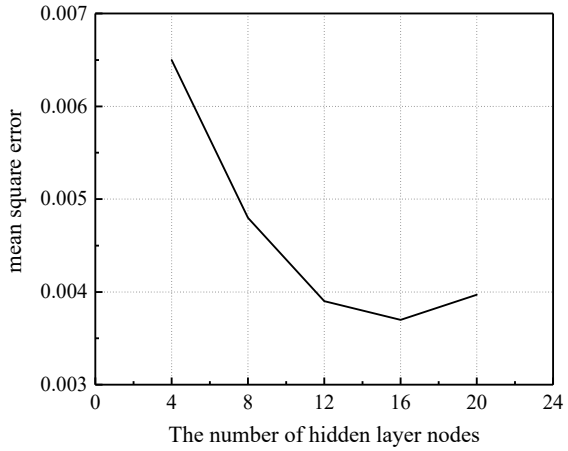

Figure 6. The influence of the number of hidden layer nodes on the network.

There were 16 hidden nodes and one output node; $tanh$ was the activation function, and the linear function $pure()$ was the transfer function. Following the establishment of the architecture of the neural network, the loss function was the mean square error, defined in Equation (1) as follows:

$$MSE(q_m, q_{nn}) = \frac{\sum_{i=1}^{n}(q_{mi} - q_{nni})^2}{n} \quad (1)$$

where q_{mi} is the experimental value of data i noted in the experimental record, and q_{nni} is the estimated value from the neural network. $trainlm()$ was selected as the training function, namely, the Levenberg–Marquardt algorithm. The developed network is illustrated in Figure 7.

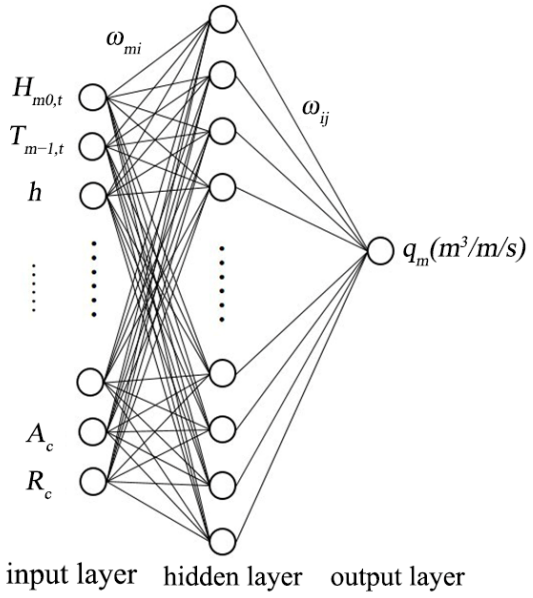

Figure 7. ANN model for estimating overtopping.

3.3. Ensemble Learning

3.3.1. Bootstrap Resampling

Bootstrap resampling is a random selection of N data, out of the original N data, and the probability of being sampled for each experimental record is $1/N$. The probability of not being sampled for each experimental record is $(1-1/N)^N$. When the N is increased infinitely, the probability will be $\lim_{N\to\infty}\left(1-\frac{1}{N}\right)^N = 1/e$ (equals approximately 0.37). Therefore, for an extensive sample database, approximately 63% of the sample data are selected for model training; the other 37% of the sample data are not selected and can be used as a verification set to test the state and convergence of the model in the training process.

3.3.2. Ensemble Learning Model

The BP neural network is based on data-driven supervised learning. Therefore, the performance of the BP model is highly affected by the data used for training. Any change in the selected training data will bring huge uncertainty to the model output [30,31]. Ensemble learning accomplishes the learning task by developing multiple learners, which is known to generally achieve a better learning result than a single learner. The common architecture is to produce one group of single learners and then incorporate these learners based on a certain strategy. A total of 500 groups of training subsets were selected based on the bootstrap sampling principle; then 500 groups of neural network sub-learners were generated based on this training subset. Subsequently, a comprehensive estimation model for overtopping discharge was established based on the combination strategy of the average technique, i.e., $q_{nn} = \frac{1}{500}\sum_{j=1}^{500} q_{nn_j}$.

3.4. Mean Impact Value

The mean impact value (MIV) was proposed by Dombi [32] and is one of the indicators showing the change in the weighting matrix in the neural network. We can analyze the importance of factors affecting overtopping by calculating the MIV value, which represents the extent of influence on overtopping. A positive value means a positive correlation; a negative value means a negative correlation. The detailed calculation process of the MIV is summarized as follows.

Suppose that matrix $X_{m\times n}$ is the original data sample, where each row of X represents a sample, and each column represents a feature (influencing factor). When adding noise to the jth feature value of each sample, that is, increasing or decreasing its T% times, two data samples with noise, X_1 and X_2, are obtained, as shown in Equation (2):

$$X_{m\times n}=\begin{bmatrix} x_{11} & x_{12}\cdots x_{1j}\cdots x_{1n} \\ x_{21} & x_{22}\cdots x_{2j}\cdots x_{2n} \\ \vdots & \vdots \quad \vdots \\ x_{m1} & x_{m2}\cdots x_{mj}\cdots x_{mn} \end{bmatrix} \quad X_1, X_2=\begin{bmatrix} x_{11} & x_{12}\cdots x_{1j}(1\pm t\%)\cdots x_{1n} \\ x_{21} & x_{22}\cdots x_{2j}(1\pm t\%)\cdots x_{2n} \\ \vdots & \vdots \quad\quad \vdots \\ x_{m1} & x_{m2}\cdots x_{mj}(1\pm t\%)\cdots x_{mn} \end{bmatrix} \tag{2}$$

Let X_1 and X_2 be input into the established BP neural network for simulation. Two column vectors, $Y_j^{(1)}$ and $Y_j^{(2)}$, will be obtained, which represents the predictive value after adding noise to the jth feature, as shown in Equation (3):

$$Y_j^{(1)}=\begin{bmatrix} y_{1j}^{(1)} & y_{2j}^{(1)} & \cdots & y_{mj}^{(1)} \end{bmatrix}^T \quad Y_j^{(2)}=\begin{bmatrix} y_{1j}^{(2)} & y_{2j}^{(2)} & \cdots & y_{mj}^{(2)} \end{bmatrix}^T \tag{3}$$

The jth feature's *MIV* value can be calculated using Equation (4):

$$MIV_i=\sum_{j=1}^{m}\left(Y_i^{(1)}-Y_i^{(2)}\right)/m \tag{4}$$

where m is the amount of data in the sample. Similarly, other features' *MIV* values can be calculated the same way.

3.5. Evaluation Index

Our work needs a statistical measure to evaluate the performance of the prediction model. The correlation coefficient (R) is a statistical index used to reflect the close degree of correlation between variables. A correlation coefficient closer to 1 or −1 represents a strong correlation; a correlation coefficient closer to 0 indicates a weak correlation. It is defined by Equation (5):

$$R=\frac{\sum q_m q_{nn}-\frac{\sum q_m q_{nn}}{N}}{\sqrt{\left(\sum q_m^2-\frac{(\sum q_m)^2}{N}\right)\left(\sum q_{nn}^2-\frac{(\sum q_{nn})^2}{N}\right)}} \tag{5}$$

where q_m is the experimental value, q_{nn} is the estimated value, and N is the number of samples.

4. Result

4.1. The Performance of the Network Model

The estimated overtopping discharge (q_{nn}) was produced after all data were input into the network and then compared with the experimental value (q_m). Figure 8 indicates that the network boosted the performance, obtaining high accuracy and a strong estimation. The solid slash in the middle of this graph is the 45-degree hypothetical line, and the space between the broken slash on both sides of the hypothetical line is the 10 times error tolerance. It is worth noting that overtopping is related to many factors; thus, the error for different calculation formulae is generally controlled under 10 times, i.e., within one order of magnitude. Further, there may still be a five times error even after repeated tests of the overtopping discharge in the laboratory. As shown in Figure 8, the vast majority of the network-estimated overtopping discharges are within a tenfold error; it is rare to see any estimated overtopping discharges beyond the tenfold error. The correlation coefficient

$R = 0.92$ also shows a high correlation between the network-estimated overtopping discharge and the experimental value.

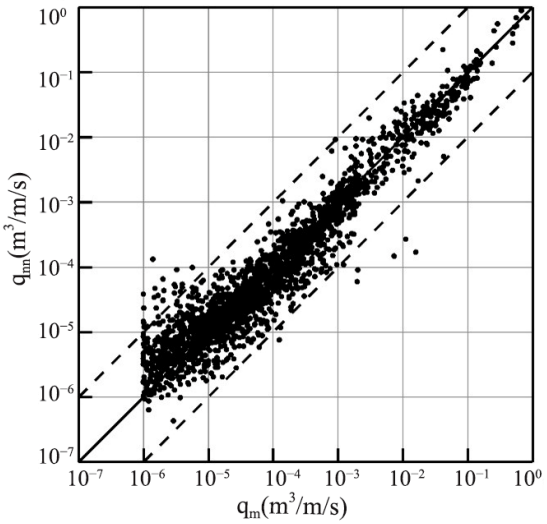

Figure 8. Comparison of the estimated value of the ANN model with the experimental value of all datasets.

4.2. Comparison with the Chinese Standard Formula

The Code of Hydrology for Port and Waterway Engineering [5] includes a formula for calculating the slope breakwater wall overtopping discharge under certain circumstances. The wall overtopping discharge calculation formula for a breakwater without a seawall is as follows:

$$Q = AK_A \frac{H_{1/3}^2}{T_P} \left(\frac{H_C}{H_{1/3}}\right)^{-1.7} \left[\frac{1.5}{\sqrt{m}} + th\left(\frac{d}{H_{1/3}} - 2.8\right)^2\right] ln \sqrt{\frac{gT_p^2 m}{2\pi H_{1/3}}} \quad (6)$$

Wall overtopping discharge calculation formula for breakwater with seawall:

$$Q = 0.07^{H_c'/H_{1/3}} exp\left(0.5 - \frac{b_1}{2H_{1/3}}\right) BK_A \frac{H_{1/3}^2}{T_p} \left(\frac{0.3}{\sqrt{m}} + th\left(\frac{d}{H_{1/3}} - 2.8\right)^2\right) ln \sqrt{\frac{gT_p^2 m}{2\pi H_{1/3}}} \quad (7)$$

where d represents water depth in front of the structure (m), $H_{1/3}$ is the effective wave height (m), L_{po} is the deep water wave length calculated based on peak spectral period (m), m is the slope gradient coefficient and the slope is $1{:}\,m$, b_1 is the berm width (m), H_c' is the crest height with respect to static water level (m), i is the underwater gradient.

Based on the application range of each parameter contained in Equations (6) and (7), which can be seen in the reference [5], the database established previously was screened. Independent screening refers to the quantity of remaining experimental records after the screening of data satisfying the current parameter. Association screening refers to the quantity of remaining experimental records after the screening of data satisfying each of these parameters one by one from left to right. It can be seen in Figure 9 that the experimental record with the largest portion was screened out due to the limited application range of the berm width b_1 contained in Equations (6) and (7); the experimental record with the smallest portion was screened out due to the limited application range of the deep water wavelength L_{po} contained in Equations (6) and (7). Given the limited application range of each parameter, the experimental records screened out of the dataset, from large to small portions, were the berm width b_1, effective wave height $H_{1/3}$, armor rock

type Ka, slope gradient coefficient m, offshore water depth, and deep water wave length calculated based on the peak spectral period L_{po}.

In the process involved in the association screening of data satisfying each of these parameters one by one from left to right, the changes in the remaining experimental records are displayed, as shown in Figure 9. There were only 47 experimental records left in the dataset following the screening of data satisfying all parameters contained in Equations (6) and (7).

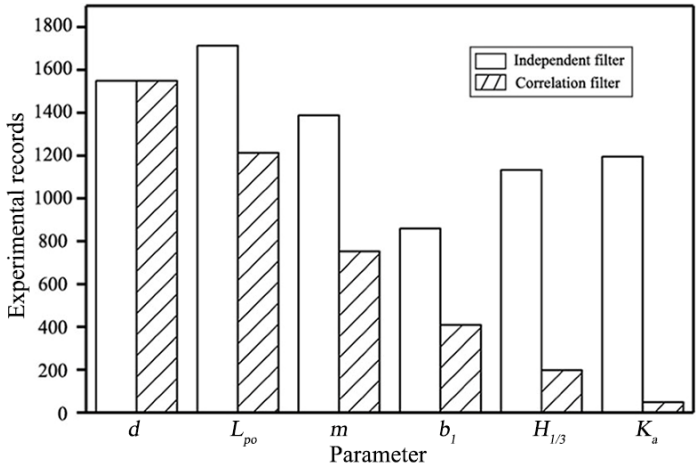

Figure 9. Result of all datasets filtered by the Equations (6) and (7).

Equations (6) and (7) were applied to these post-screening experimental records for calculating the overtopping discharge and comparing it with the neural network estimation value and test value. The results of the comparison of the correlation coefficient are shown in Table 2.

Table 2. Comparison of the evaluation index between our model and the Chinese standard formula.

Method	R
Chinese standard formula	0.61
This study	0.84

The estimation value proves that there was a high correlation between the network-estimated overtopping discharge and the tested value, with the correlation coefficient $R = 0.84$. It is rare to see any substantial estimation error, and the estimation result is reliable. The correlation between the overtopping discharge calculated based on the Chinese standard formula and the test value was low, $R = 0.61$. There was also a substantial estimation error for the overtopping discharge $q < 10^{-4}$, as shown in Figure 10. Therefore, the accuracy and application range of the neural-network-based overtopping discharge estimation model are better than those of the Chinese standard formula.

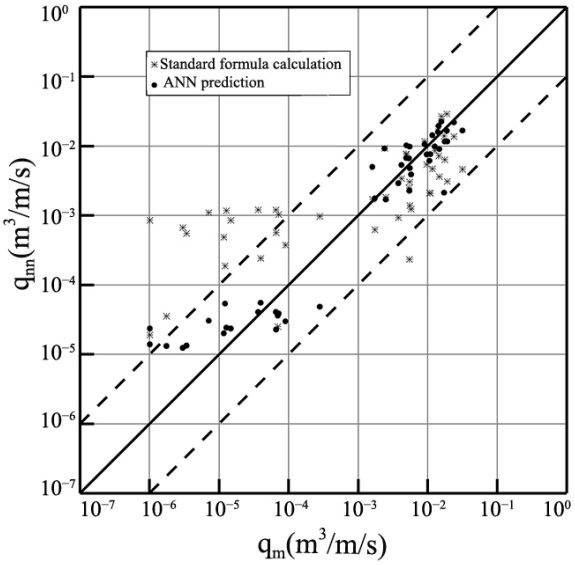

Figure 10. Comparing the estimated value of the ANN model, the Chinese standard formula, and the experiment.

4.3. Comparison with the Van Der Meer Formula

In EurOtop, Van der Meer proposed [27] a calculation formula for the wave overtopping discharge from the single-slope breakwater, and it is commonly accepted by most European countries:

$$\frac{q}{\sqrt{g \cdot H_{m0}^3}} = \frac{0.023}{\sqrt{tan\,\alpha}} \gamma_b \cdot \xi_{m-1,0} \cdot exp\left[-\left(2.7 \frac{R_c}{\xi_{m-1,0} \cdot H_{m0} \cdot \gamma_b \cdot \gamma_f \cdot \gamma_\beta \cdot \gamma_v}\right)^{1.3}\right] \quad (8)$$

Calculation of the maximum value per formula below：

$$\frac{q}{\sqrt{g \cdot H_{m0}^3}} = 0.09 \cdot exp\left[-\left(1.5 \frac{R_c}{H_{m0} \cdot \gamma_f \cdot \gamma_\beta \cdot \gamma_v}\right)^{1.3}\right] \quad (9)$$

where α is the angle between the slope and horizontal plane, γ_b is the berm reduction factor, $\xi_{m-1,0}$ is the wave breaking parameter, γ_f is the influence coefficient of different armor rock types, γ_β is the influence coefficient of wave incidence angle and $\gamma_\beta = 1$ under positive incidence, γ_v is the influence coefficient of seawall.

Similarly, based on the application range of each parameter contained in Equation (8), which can be seen in the reference [27], the database established previously was screened. As shown in Figure 11, the independent screening will squeeze out experimental records accounting for the largest portion of the dataset due to the limited application range of the seawall influence coefficient γ_v and eliminate experimental records accounting for the smallest portion of the dataset due to the limited application range of the crest height with respect to the static water level. Given the limited application range of each parameter, the experimental records screened out of the dataset, from large to small portions, were the seawall influence coefficient γ_v, rock surface roughness factor γ_f, slope gradient coefficient $s_{m-1,0}$, wave breaking parameter $\xi_{m-1,0}$, wave steepness $s_{m-1,0}$, and the wall height with respect to the static water level R_c/H_{m0}.

In the process of association screening of data satisfying each of these parameters in sequence from left to right, the changes in the remaining experimental records are displayed, as shown in Figure 11. There were only 142 experimental records left in the dataset following the screening of data satisfying all the parameters contained in Equation (8).

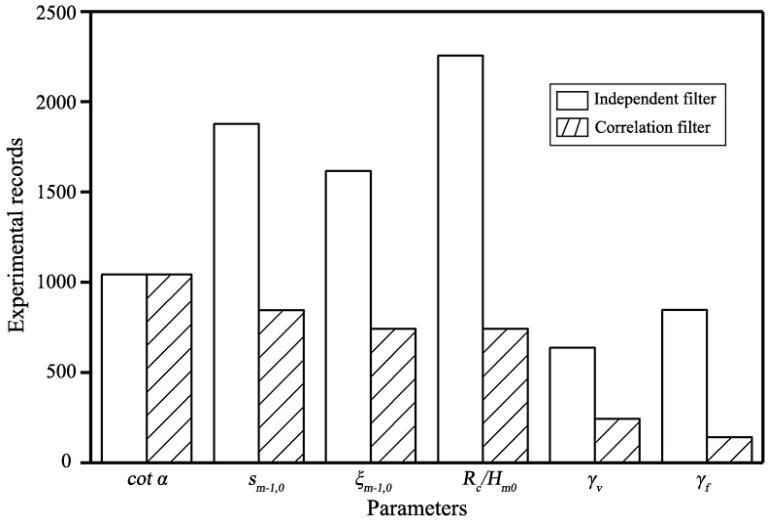

Figure 11. Result of all datasets filtered by Equation (8).

The Van der Meer formula was applied to these post-screening experimental records for calculating the overtopping discharge and comparing it with the neural network estimation value and test value. The results of the comparison of correlation coefficients are shown in Table 3.

Table 3. Comparison of the evaluation index between our model and Van der Meer's formula.

Method	R
Van der Meer's formula	0.64
This study	0.90

The estimation value shows that there was a high correlation between the network-estimated overtopping discharge and the test value, with the correlation coefficient $R = 0.90$. It is infrequent to see any substantial estimation error, and the estimation result is reliable. The correlation between the overtopping discharge calculated based on Equation (8) and the test value was low, $R = 0.64$. There are some large estimation deviations in some data, and the estimation results are scattered, as shown in Figure 12. Therefore, the accuracy and application range of the neural network based overtopping discharge estimation model are better than those of Equation (8).

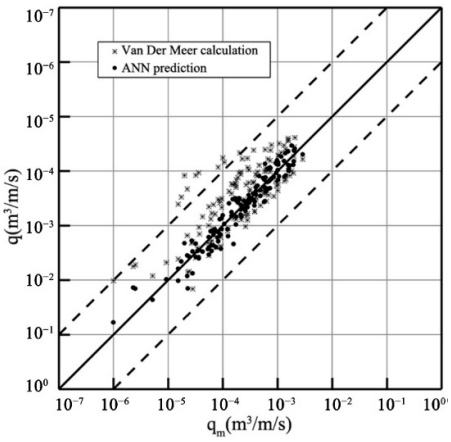

Figure 12. Comparison of the estimated value of the ANN model, the European Van der Meer formula, and the experiment.

4.4. Sensitivity Analysis Based on the MIV

The MIV algorithm was used to analyze the correlation and sensitivity of the BP neural network's input parameters. MIV mediation rates of 10%, 15%, and 20% were selected to evaluate the network's input characteristic parameters. Table 4 shows the detailed assessment result.

Table 4. MIV values of input characteristic parameters under different mediation rates.

No.	Characteristics Parameter	10% Mediation Rate	15% Mediation Rate	20% Mediation Rate
1	h	8.26×10^{-4}	0.0012	0.0016
2	$H_{m0,t}$	0.0058	0.0088	0.0118
3	$T_{m-1,0}$	0.0035	0.0053	0.0072
4	h_t	-5.24×10^{-4}	-7.84×10^{-4}	−0.001
5	B_t	1.15×10^{-4}	1.74×10^{-4}	2.34×10^{-4}
6	$\cot\alpha$	−0.0028	−0.0043	−0.0059
7	γ_f	4.14×10^{-4}	0.0011	0.0019
8	R_c	−0.0024	−0.0036	−0.0049
9	A_c	−0.0019	−0.0028	−0.0038
10	G_c	−0.0019	−0.0028	−0.0038

Figure 13 shows the MIV value under different mediation rates, which was produced based on the MIV value changes noted in Table 4.

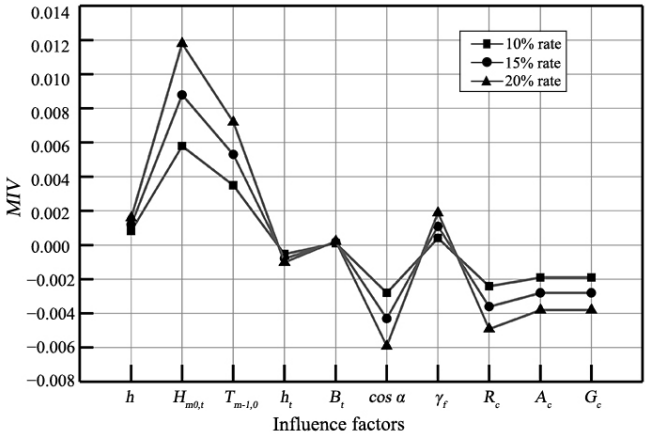

Figure 13. Curve of the MIV value changes in overtopping influence factors.

It can be seen in Figure 13 that the influence of overtopping discharge impact factors under different mediation rates on the overtopping weight factors was minimal. The weight factors of the offshore effective wave height, average wave period, slope tangent, crest height with respect to the static water level, wall height with respect to the static water level, and berm width were large. Generally speaking, a higher wave element (i.e., a large wave height and wave period) will produce a larger overtopping discharge. A large slope gradient (an extreme case being a vertical seawall) will incur a small overtopping discharge. A large crest height with respect to the static water level, a large wall height with respect to the static water level, and a broad berm width will result in a small overtopping discharge.

5. Conclusions

This study collated a dataset associated with the overtopping discharge of single-slope breakwaters based on TIWTE's experimental records, which supplied data with $q \geq 10^{-4}\ m^3/m/s$ to the CLASH dataset and further enriched it. This allowed us to expand the new dataset and make it more widely applicable. Furthermore, this paper provided a novel prediction model, MIV-BP, for estimating overtopping discharge and discussing the impact factors of overtopping. The prediction results of the model are basically within the range of the 10 times error, and there are few large deviations. Compared to the Chinese standard formula and the European Van der Meer formula, the model's coefficient of correlation increased from 0.61 to 0.84 and from 0.64 to 0.90 in the studied dataset. The prediction model with R = 0.92 on all of the datasets was recognized as the better model. Furthermore, the results of the MIV evaluation show that overtopping discharge can be effectively decreased from a design perspective by reducing the wave height and wave period and by increasing the slope tangent and the crest height with respect to the static water level. The influence of the following factors on overtopping discharge is minimal: offshore water depth, rock surface roughness factor, and toe width. The above deduction provides a basis and reference for effectively controlling overtopping discharge and can be used to optimize the structural design of breakwaters.

Author Contributions: Conceptualization, H.C.; methodology, S.C.; writing—original draft preparation, S.C.; supervision, H.C. and C.P.; data collection, C.P. and Y.W.; writing—review and editing, S.C., H.C. and Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by China National Key R&D Program(2022YFE0104500), the National Natural Science Foundation of China (no. 52001149, no. 52039005, no. 51861165102) and the Research Funds for the Central Universities (no. TKS20210102, no. TKS20220301, no. TKS20220601).

Data Availability Statement: EurOtop dataset are available on website at http://www.overtopping-manual.com/assets/downloads(accessed on 18 October 2021).

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

References

1. Fan, H.; Zhou, Y. A review of studies on overtopping discharge and flow of seawall. *Port Waterway Eng*. **2008**, *8*, 14–19.
2. Chen, S.G.; Wang, Z.M.; Zhang, C.; Zheng, J. Experimental study on wave overtopping of a vertical seawall on coral reefs in large wave flume. *Chin. Sci Bull*. **2019**, *64*, 3049–3058.
3. Wang, H.; Zhou, J. Estimating of irregular wave overtopping quantities in single sloping. *Hydro-Sci. Eng*. **1996,** *1*, 58–63.
4. Zhang, J.; Zhou, J.; Wang, H. *Research Report on Overtopping Model Test*; Nanjing Hydraulic Research Institute: Nanjing, China, 1992.
5. *Hydrologic Regulations for Ports and Waterways (A-JTS145-2015)*; China Communications Publishing: Beijing, China, 2015.
6. Van der Meer, J.; Janssen, J. *Wave Run-Up and Wave Overtopping at Dikes and Revetments*; Delft Hydraulics: Delft, The Netherlands, 1994.
7. Van der Meer, J.W. Wave run-up and overtopping. In *Dikes and Revetments*; Routledge: London, UK, 2017; pp. 145–160.
8. Goda, Y. *Random Seas and Design of Maritime Structures*; World Scientific Publishing Company: Singapore, 2010.
9. Yu, Y. Comparison of Methods for Calculating Mean Wave Overtopping on Sloping Seawall. *Port Eng. Technol.* **2011**, *48*, 1–4.
10. Dong, S.; Abolfathi, S.; Tan, Z.; ; Pearson, J. Enhancing climate resilience of vertical seawall with retrofitting-A physical modelling study. *Appl. Ocean Res.* **2020**, *103*, 102331.
11. Dong, S.; Abolfathi, S.; Salauddin, M.; Pearson, J.M. Spatial distribution of wave-by-wave overtopping behind vertical seawall with recurve retrofitting. *Ocean Eng*. **2021**, *238*, 109674.
12. Dong, S.; Salauddin, M.; Abolfathi, S.; Jonathan M. Wave Impact Loads on Vertical Seawalls: Effects of the Geometrical Properties of Recurve Retrofitting. *Water* **2021**, *13*, 2849.
13. Salauddin, M.; O'Sullivan, J.; Abolfathi, S.; Dong, S. Distribution of individual wave overtopping volumes on a sloping structure with a permeable foreshore. *Coast. Eng. Proceed*. **2020**, 54. https://doi.org/10.9753/icce.v36v.papers.54.
14. Salauddin, M.; O'Sullivan, J.J.; Abolfathi, S.; Pearson, J.M. Eco-engineering of seawalls—an opportunity for enhanced climate resilience from increased topographic complexity. *Front. Mar. Sci*. **2021**, *8*, 674630.
15. Gao, J.L; Ma, X.Z; Zang, J; Dong, G.H; Ma, X.J; Zhu, Y.Z; Zhou, L. Numerical investigation of harbor oscillations induced by focused transient wave groups. *Coast. Eng*. **2020**, 158: 103670.

16. Losada, I.J.; Lara, J.L.; Guanche, R.; Gonzalez-Ondina, J.M. Numerical analysis of wave overtopping of rubble mound breakwaters. *Coast. Eng*. **2008**, *55*, 47–62.
17. Guan, D. *Numerical Simulation of Seawall Breakwater over Regular and Irregular Waves*; South China University of Technology: Dalian, China, 2016.
18. Ye, X. *SPH Simulation of Wave Crossing Process of Slope Breakwater*; Dalian University of Technology: Dalian, China, 2011.
19. Ren, B.; Wen, H.; Dong, P.; Wang, Y. Improved SPH simulation of wave motions and turbulent flows through porous media. *Coast. Eng*. **2016**, *107*, 14–27.
20. Dasineh, M.; Ghaderi, A.; Bagherzadeh, M.; Ahmadi, M.; Kuriqi, A. Prediction of hydraulic jumps on a triangular bed roughness using numerical modeling and soft computing methods. *Mathematics* **2021**, *9*, 3135.
21. Bagherzadeh, M.; Mousavi, F.; Manafpour, M.; Mirzaee, R.; Hoseini, K. Numerical simulation and application of soft computing in estimating vertical drop energy dissipation with horizontal serrated edge. *Water Suppl*. **2022**, *22*, 4676–4689.
22. Abolfathi, S.; Dong, S.; Borzooei, S.; Pearson, J. Application of smoothed particle hydrodynamics in evaluating the performance of coastal retrofit structures. *Coast. Eng. Proceed*. **2018**, *1*, 109.
23. Daneshfaraz, R.; Aminvash, E.; Ghaderi, A.; Abraham, J.; Bagherzadeh, M. SVM performance for predicting the effect of horizontal screen diameters on the hydraulic parameters of a vertical drop. *Appl. Sci*. **2021**, *11*, 4238.
24. Daneshfaraz, R.; Bagherzadeh, M.; Esmaeeli, R.; Norouzi, R.; Abraham, J. Study of the performance of support vector machine for predicting vertical drop hydraulic parameters in the presence of dual horizontal screens. *Water Suppl*. **2021**, *21*, 217–231.
25. van Gent, M.R.; van den Boogaard, H.F.; Pozueta, B.; Medina, J.R. Neural network modelling of wave overtopping at coastal structures. *Coast. Eng*. **2007**, *54*, 586–593.
26. Formentin, S.M.; Zanuttigh, B.; van der Meer, J.W. A neural network tool for predicting wave reflection, overtopping and transmission. *Coast. Eng. J*. **2017**, *59*, 1750006.
27. Van der Meer, J.W.; Allsop, W.; Bruce, T.; De Rouck, J.; Kortenhaus, A.; Pullen, T; Zanutitigh, B. Manual on Wave EurOtop, Overtopping of Sea Defences and Related Structures. 2018. Available online: www.overtopping-manual.com (accessed on 18 October 2021).
28. Fang, W.F. Research on the mechanism and application of BP neural network algorithm. *Sci. Technol. Innovat. Herald*. **2020**, *17*, 150–151.
29. Chang, L. *Research on Grid Resource Scheduling Based on BP Algorithm*; Harbin University of Science and Technology: Harbin, China, 2007.
30. Ghiasi, B.; Noori, R.; Sheikhian, H.; Zeynolabedin, A.; Sun, Y.; Jun, C.; Hamouda, M.; Bateni, S.M.; Abolfathi, S. Uncertainty quantification of granular computing-neural network model for prediction of pollutant longitudinal dispersion coefficient in aquatic streams. *Sci. Reports* **2022**, *12*, 4610.
31. Noori, R.; Ghiasi, B.; Salehi, S.; Esmaeili Bidhendi, M.; Raeisi, A.; Partani, S.; Meysami, R.; Mahdian, M.; Hosseinzadeh, M.; Abolfathi, S. An Efficient Data Driven-Based Model for Prediction of the Total Sediment Load in Rivers. *Hydrology* **2022**, *9*, 36.
32. Dombi, G.W.; Nandi, P.; Saxe, J.M.; Ledgerwood, A.M.; Lucas, C.E. Prediction of rib fracture injury outcome by an artificial neural network. *J. Trauma Acute Care Surg*. **1995**, *39*, 915–921.