

Article

# Operational Probabilistic Forecasting of Coastal Freak Waves by Using an Artificial Neural Network

Dong-Jiing Doong <sup>1</sup>, Shien-Tsung Chen <sup>1,\*</sup>, Ying-Chih Chen <sup>1</sup> and Cheng-Han Tsai <sup>2</sup>

<sup>1</sup> Department of Hydraulic and Ocean Engineering, National Cheng Kung University, Tainan City 701, Taiwan; doong@mail.ncku.edu.tw (D.-J.D.); venwhah@gmail.com (Y.-C.C.)

<sup>2</sup> Department of Marine Environmental Informatics, National Taiwan Ocean University, Keelung City 202, Taiwan; chtsai@mail.ntou.edu.tw

\* Correspondence: chen@gs.ncku.edu.tw

Received: 8 February 2020; Accepted: 27 February 2020; Published: 3 March 2020



**Abstract:** Coastal freak waves (CFWs) are unpredictable large waves that occur suddenly in coastal areas and have been reported to cause casualties worldwide. CFW forecasting is difficult because the complex mechanisms that cause CFWs are not well understood. This study proposes a probabilistic CFW forecasting model that is an advance on the basis of a previously proposed deterministic CFW forecasting model. This study also develops a probabilistic forecasting scheme to make an artificial neural network model achieve the probabilistic CFW forecasting. Eight wave and meteorological variables that are physically related to CFW occurrence were used as the inputs for the artificial neural network model. Two forecasting models were developed for these inputs. Model I adopted buoy observations, whereas Model II used wave model simulation data. CFW accidents in the coastal areas of northeast Taiwan were used to calibrate and validate the model. The probabilistic CFW forecasting model can perform predictions every 6 h with lead times of 12 and 24 h. The validation results demonstrated that Model I outperformed Model II regarding accuracy and recall. In 2018, the developed CFW forecasting models were investigated in operational mode in the Operational Forecast System of the Taiwan Central Weather Bureau. Comparing the probabilistic forecasting results with swell information and actual CFW occurrences demonstrated the effectiveness of the proposed probabilistic CFW forecasting model.

**Keywords:** coastal freak wave; probabilistic forecasting; artificial neural network

---

## 1. Introduction

A coastal freak wave (CFW) is a sudden, unpredictable large wave that occurs in coastal areas as a result of complex interactions between shoaling waves and coastal topography. CFWs can occur even when the sea is calm and in areas without previous notable CFWs. Although CFWs are uncommon, they have been reported in a variety of areas [1–6]. CFWs can cause numerous casualties when people are present in the area where the CFW occurs. Therefore, CFW forecasting and timely warning information can mitigate the potential harm of CFWs. Numerical models based on the wind wave energy spectrum are usually applied for forecasting wave characteristics. However, the complex mechanisms that cause CFWs may include the propagation of nearshore waves, wave structure interactions, and topographic bathymetry effects, rendering CFWs difficult to simulate using physical laws. Therefore, numerical models are currently unable to simulate and forecast CFWs. Nevertheless, some recent works on the characterization of large scale extreme events in geophysical flows shed a new light on the CFW genesis and forecasting [7]. Simulating and forecasting CFWs in a deterministically physical method may be attainable in the near future.

Artificial intelligence (AI) methods can be used for CFW simulation and forecasting. Various AI methods have been applied to hydrologic and ocean forecasting. In hydrologic forecasting, artificial neural networks [8–12], support vector machines [13–20], and fuzzy set theory [21–27] have been successfully used to perform rainfall and flood forecasting. In ocean forecasting, AI methods have been employed to forecast significant wave height. For example, Deo and Sridhar Naidu [28], Deo et al. [29], Tsai et al. [30], and Mandal and Prabakaran [31] applied artificial neural networks to predict wave height on the basis of wave height data and observed or simulated wind speed. Mahjoobi and Mosabbebi [32] and Elbisy [33] employed support vector machines to simulate significant wave height and wave characteristics. Moreover, Kazeminezhad et al. [34], Özger and Şen [35], and Mahjoobi et al. [36] used fuzzy logic models to simulate or forecast wave height, and Chang and Chien [37,38], Tsai et al. [39], Wei [40], and Chen [41] applied typhoon characteristics as the inputs of AI models to forecast wave parameters during typhoons.

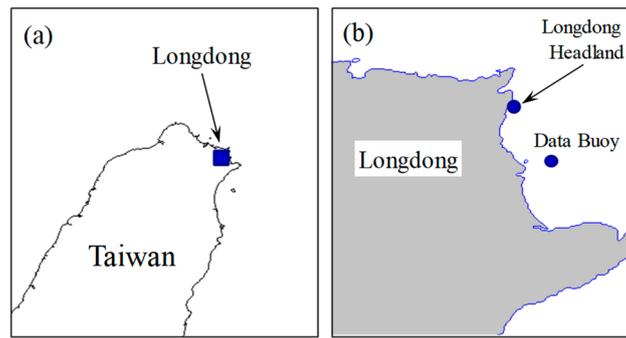
However, few studies have applied AI methods to CFW forecasting. Doong et al. [6] first proposed applying artificial neural networks to CFW forecasting. Their proposed model adopted various sea state characteristics to forecast CFW occurrence. The methodology is similar to a clustering process that deterministically predicts whether a CFW will or will not occur. The present study improved on the deterministic methodology used by Doong et al. [6] by developing a probabilistic CFW forecasting model, which is an advance to fulfill the probabilistic forecasting of CFW occurrence. This probabilistic forecasting model is more practical than the deterministic model, and agencies can use it to establish risk-based criteria for warnings and responses.

A back-propagation neural network (BPNN) was used to develop the probabilistic CFW forecasting model. This study proposed a novel probabilistic scheme that assigns the BPNN outputs to probability values indicating the physical tendency of a CFW to occur. Variables potentially related to CFW occurrence were selected as the model inputs on the basis of current knowledge of CFWs. The model output is a numerical value ranging from 0 to 1 that indicates the probability of CFW occurrence. Probabilistic CFW forecasting was conducted using validation data, and an operational system was developed and tested to examine the effectiveness of the developed CFW forecasting model. Validation results based on real CFW data confirmed the predictive ability of the proposed probabilistic CFW forecasting model. The CFW forecasting model was tested in operational mode in the Operational Forecast System of the Taiwan Central Weather Bureau (TCWB) during 2018. Comparing the forecasting results with swell information and actual CFWs during a typhoon period demonstrated the effectiveness of the proposed probabilistic CFW forecasting model.

## 2. Wave Data and Probabilistic Forecasting

### 2.1. CFW Data

Coastal fishing is a popular activity in Taiwan, and people fishing in coastal areas sometimes encounter CFWs, which have occasionally resulted in casualties. CFW accidents in Taiwan occur the most frequently in the Longdong headland of northeast Taiwan (Figure 1). During 2000–2016, 63 CFW accidents were reported in the Longdong headland and its neighboring coast. These CFW accidents, which were checked and validated by the TCWB, were used as CFW data in this study. A CFW accident refers to a CFW that results in casualties. When a CFW does not result in injury, it is not recorded as a CFW accident. Accordingly, the number of CFW events is considerably higher than that of CFW accidents.



**Figure 1.** Map of (a) Longdong headland and (b) the Taiwan Central Weather Bureau (TCWB) data buoy.

### 2.2. Buoy Observation and Wave Model Simulation Data

The TCWB operates a data buoy in the Longdong area to record meteorological and wave data at intervals of 1 h. The buoy is located at 121.92° E and 25.10° N, which is approximately 1 km from the coast and has a water depth of approximately 23 m (Figure 1). This data buoy measures meteorological variables, including wind speed, wind direction, air pressure, air temperature, and surface water temperature, and wave parameters, such as the wave height, period, and direction. A data quality control procedure proposed by Doong et al. [42] was adopted to ensure the quality of the hourly observations of meteorological and wave data.

In addition to data buoy observations, the TCWB uses the WaveWatch III (WWIII) model to simulate the marine meteorology in the western Pacific. The spatial resolution of TCWB’s WWIII model is 25 km for coastal ocean around Taiwan. A directional wave spectrum at each grid is available at hourly time steps. This study used an interpolation method to derive the directional wave spectrum at the buoy location. Therefore, the meteorological and wave variables from the WWIII model are also available for CFW prediction. The hourly meteorological and wave data from the data buoy and wave variables from the WWIII model were collected and organized for a period covering the studied CFW accidents.

### 2.3. Back-Propagation Neural Network

This study adopted a BPNN, which is the most commonly used type of artificial neural network. BPNNs feature a feedforward network structure with input, hidden, and output layers. BPNNs are trained using the back-propagation algorithm—a supervised learning method based on the method of steepest descent [43]. The model is briefly introduced as follows, and detailed descriptions of the model and its development are available in previous studies [44,45]:

Let  $X_i$  ( $i = 1, 2, \dots, m$ ) be the input variables in the input layer and  $\hat{Y}_k$  ( $k = 1, 2, \dots, p$ ) be the output variables in the output layer. A BPNN with  $n$  neurons in the hidden layer is formulated as follows:

$$\hat{Y}_k = \sum_{j=1}^n W_{jk} \cdot f\left(\sum_{i=1}^m W_{ij} \cdot X_i + b_j\right) + d_k \quad (1)$$

where  $W_{ij}$  is the weight of the connection between the  $i$ th input neuron and the  $j$ th hidden neuron ( $j = 1, 2, \dots, n$ ),  $b_j$  is the bias term of the  $j$ th hidden neuron,  $W_{jk}$  is the weight of the connection between the  $j$ th hidden neuron to the  $k$ th output neuron,  $d_k$  is the bias term of the  $k$ th output neuron, and  $f$  is the activation function of the hidden neurons. In the present study, the hyperbolic tangent sigmoid transfer function was adopted as the activation function.

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (2)$$

The output of the BPNN  $\hat{Y}_k$  is fit to the actual output  $Y_k$  by minimizing the objective function  $E$ :

$$E = \frac{1}{2} \sum_{k=1}^p (\hat{Y}_k - Y_k)^2 \tag{3}$$

During model training, the method of steepest descent is used to adjust the weights and biases of the network. Adjustments are repeated until an acceptable convergence of the weights and biases is reached.

### 2.4. Probabilistic Forecasting

Probability is the measure of the likelihood of an event. In this study, probability is regarded as the physical tendency of a CFW to occur. Probability values range from 0 to 1, where 0 implies impossibility and 1 implies certainty. The higher the probability of a CFW occurrence, the more likely it is that the CFW event will occur. Figure 2 illustrates the process of the probabilistic forecasting method used in this study. On the basis of the input of relevant variables (selection detailed in Section 3.1), the BPNN model outputs a probability value for the occurrence of a CFW. A high probability value (e.g., 0.91) indicates that a CFW is highly likely to occur. A low probability value (e.g., 0.18) indicates that a CFW is highly unlikely to occur.

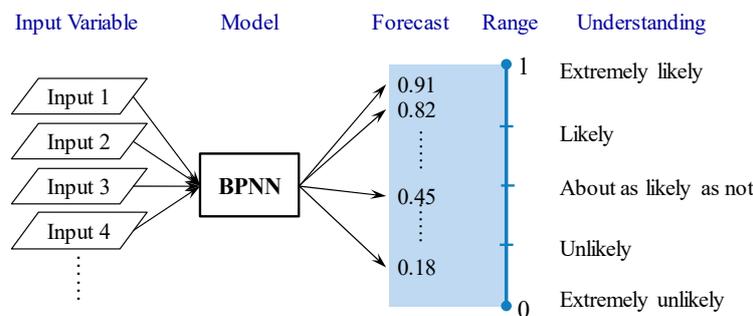


Figure 2. Probabilistic forecasting method.

## 3. Model Development

### 3.1. Selection of Input Variables

A probabilistic CFW forecasting model was developed to forecast the likelihood of CFW occurrence in the subsequent time interval. Input variables should be physically related to CFW occurrence. Eight meteorological and wave variables, namely significant wave height, peak wave period, onshore wind speed, misalignment between wind and wave directions, kurtosis of sea surface elevation, Benjamin–Feir index (BFI), groupiness factor, and abnormality index, were selected on the basis of expert knowledge and a preliminary experiment. These variables were used in the BPNN model to predict the probability of CFW occurrence. With the exception of misalignment between wind and wave directions, maximum values during the 6 h period before CFW occurrence were used as the model input. For misalignment between wind and wave directions, the value during CFW occurrence was used. The variables are detailed as follows.

#### 3.1.1. Significant Wave Height

Severe sea states feature large waves and an increased probability of CFW occurrence. Significant wave height is the basic variable describing sea states and is defined as the mean of the highest third wave heights of all waves during a certain period.

### 3.1.2. Peak Wave Period

Chien et al. [1] and Tamura et al. [46] identified a strong relationship between swell and CFW occurrence. Swell has a period longer than that of waves caused by wind. It can transfer more energy and has more momentum compared with wind waves, and it can create a large spray of water when it breaks at the coast. Swell is represented by the peak wave period, which is equivalent to the period with the highest energy.

### 3.1.3. Onshore Wind Speed

Strong wind can generate large wind waves in coastal areas, especially when the wind direction is perpendicular to the coast. When the wind speed is high, large wind waves can increase the probability of CFW occurrence.

### 3.1.4. Misalignment between Wind and Wave Directions

As mentioned in the previous variable descriptions, wind waves and swell may produce CFWs. Gramstad and Trulsen [47] indicated that the combination of local wind waves and swell can increase the probability of CFWs by 5%–20%. When wind waves and swell propagate in similar directions, the probability of a CFW is higher. Therefore, the difference between local wind and swell directions is adopted as an input variable.

### 3.1.5. Kurtosis of Sea Surface Elevation

A CFW is an extreme outlier in a wave time series. Kurtosis measures the type of tail of a probability distribution. Mori et al. [48] discussed the relationship between the probability of CFWs and kurtosis, and they indicated that wave kurtosis is a suitable input for CFW forecasting.

### 3.1.6. Benjamin–Feir Index

When wave nonlinearity and instability increase, wave coalescence can produce CFWs. The BFI [49] is the ratio of wave steepness to spectral bandwidth, which quantitatively measures wave instability and nonlinearity. A higher BFI corresponds to a higher probability of CFW occurrence. Janssen [50] and Mori and Janssen [51] have provided details of BFI computation.

### 3.1.7. Groupiness Factor

Chien et al. [1] and Tsai et al. [2] claimed that a strong connection exists between wave groupiness and CFW occurrence. When a wave breaks at the coast, interaction with subsequent waves in the wave group may produce a CFW. The groupiness factor, proposed by Funke and Mansard [52], is adopted to quantify wave grouping. Sea surface elevation data obtained from buoy observations can be used to calculate the groupiness factor. However, the groupiness factor cannot be computed from the WWIII model outputs that do not include sea surface elevation data. For these outputs, the present study employed the peakedness factor [53], which is related to wave energy and groupiness. Wave energy is higher when the wave grouping phenomenon is more prominent. Thus, the peakedness factor replaced the groupiness factor when WWIII model data were used.

### 3.1.8. Abnormality Index

When sea states include abnormal waves, CFWs are likely to occur. The abnormality index [54,55] is the ratio of the wave height to the significant wave height. An extreme wave event is defined as a wave event with an abnormality index higher than 2. The abnormality index can be calculated from buoy observations but cannot be obtained from WWIII model data. Therefore, the abnormality index was not used for CFW forecasting based on WWIII model data.

### 3.2. Model Training

A total of 63 CFW accidents were recorded in the study area. For the training set, 40 of these waves were selected, and the remaining 23 formed the validation set. The training set was used for training the BPNN model to predict both CFW and non-CFW conditions. Although only 40 sea states leading to CFW accidents were available, numerous sea states without CFW occurrence (i.e., calm sea states) were existing. Therefore, 40 conditions without CFW occurrence were selected to represent these sea states for model training. To ensure that no CFWs occurred in these selected conditions, the following criteria were established for identifying calm sea states: significant wave height, peak wave period, and wind speed within the first quartile. That is, significant wave height less than 0.6 m, peak wave period less than 6.9 s, and wind speed less than 2.6 m/s were considered to be indicative of a calm sea state without CFWs. On the basis of these criteria, 40 conditions without CFW occurrence were randomly selected from the database. Finally, a data set of size 80 (40 instances with CFW and 40 instances without CFW) was established for training the BPNN model.

For BPNN model training, a supervised learning algorithm was used to assign inputs and outputs to the BPNN model during the training phase. The inputs were the variables detailed in Section 3.1 obtained from both buoy observations and WWIII model data, and the outputs were probability values. Typically, the outputs would be set as 1 and 0, indicating the occurrence and absence of a CFW, respectively. However, preliminary analysis revealed that this approach causes BPNN predictions to tend toward 1 and 0. In this circumstance, the BPNN model resembles a classification model that categorizes the occurrence or absence of a CFW rather than a regression model that predicts the probability of CFW occurrence. Therefore, the output values of the training set were modified to increase the flexibility and practicality of the probabilistic forecasting model. For the 40 CFWs, output values were randomly assigned in the interval of 0.70–0.99. Similarly, for the 40 conditions without CFWs, output values were randomly assigned between 0.01 and 0.30. Thus, a probability higher than 0.7 was regarded as a high probability of CFW occurrence, and a probability smaller than 0.3 was regarded as a low probability of CFW occurrence. The input data were normalized to the interval from 0 to 1 on the basis of the minima and maxima for the variables in the training data set.

Instability introduced by the random selection of the 40 conditions without CFWs and the random assignment of output probability values was accounted for by performing the training process 10 times. That is, a 10-fold cross validation method was applied during model training. Moreover, because of the temporal arrangement of input and output variables, two BPNN models with forecasting lead times of 12 and 24 h were developed. The BPNN model based on buoy data was termed Model I, and that using WWIII model data was termed Model II. Both models had network architectures with five hidden nodes and the hyperbolic tangent sigmoid transfer function serving as the activation function.

### 3.3. Model Performance

The proposed BPNN model forecasts the probability of CFW occurrence in the subsequent 12 or 24 h period. Although the model forecasts are probability values, the practical result is information regarding the presence or absence of a CFW. Therefore, the model performance could be determined using the evaluation scheme typically used for classification models. When the BPNN model performs predictions, a probability greater than or equal to 0.7 is considered to predict a CFW occurrence and a probability less than or equal to 0.3 is considered to predict that no CFW will occur. Table 1 presents an example confusion matrix for the CFW forecasting results. The rows of the matrix represent the predicted class, and the columns represent the actual class. In Table 1, the model predicts A + B CFW events, where A represents the correctly forecasted cases and B represents the incorrectly forecasted cases. Similarly, C represents cases in which the model predicts no CFW but a CFW did occur, and D represents cases in which the model predicts no CFW occurrence and none occurred. Thus, the model performance could be assessed on the basis of the indices of accuracy and recall.

$$\text{Accuracy} = (A + D) / (A + B + C + D) \quad (4)$$

$$\text{Recall} = A / (A + C) \tag{5}$$

Accuracy is the ratio of correct predictions to the total number of predictions and is used to assess the overall correctness of prediction results. Recall is the ratio of correctly predicted CFW events to the total number of CFW events and is more suitable than accuracy for assessing the prediction of rare events such as CFWs.

Table 2 lists forecasting results of the BPNN models for the training data. Table 3 lists the forecasting performance of the models in terms of accuracy and recall. The results of Models I and II were similar. The 24 h forecasting results were slightly inferior to the 12 h forecasting results. Overall, the probabilistic CFW forecasting models were well trained, and model performance with the training data was satisfactory.

**Table 1.** Example confusion matrix for coastal freak wave (CFW) forecasting results.

		Actual	
		CFW	No CFW
Predicted	CFW	A	B
	No CFW	C	D

Note: A, B, C, and D are the numbers of CFW or non-CFW cases.

**Table 2.** Training results of the CFW forecasting models.

Lead Time		Model I		Model II	
		CFW	No CFW	CFW	No CFW
12 h	CFW	39	1	37	1
	No CFW	1	39	3	39
24 h	CFW	36	4	37	1
	No CFW	4	36	3	39

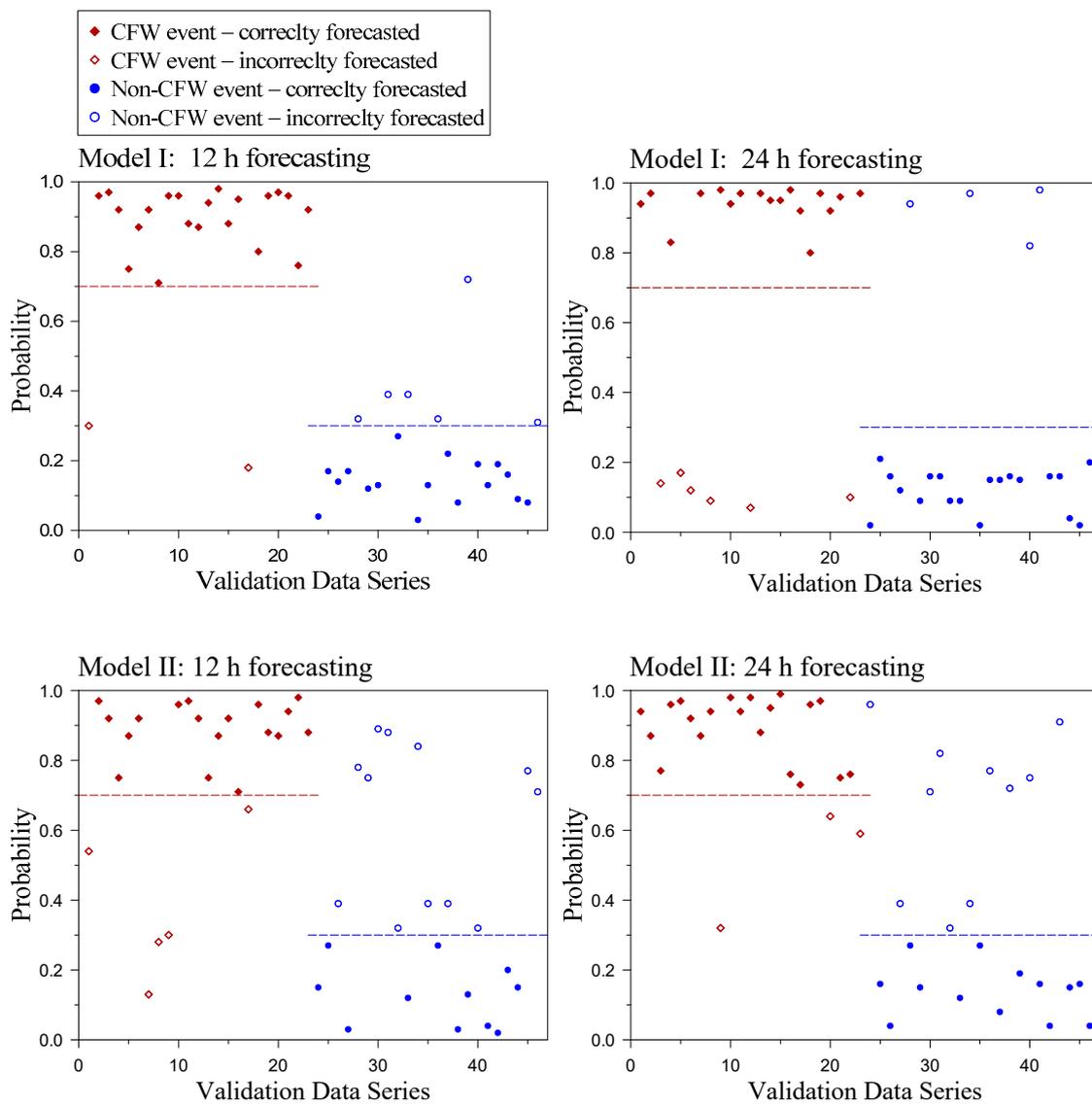
**Table 3.** Forecasting performance in terms of accuracy and recall (training data).

Lead Time	Model I		Model II	
	Accuracy	Recall	Accuracy	Recall
12 h	98%	98%	95%	95%
24 h	90%	90%	93%	93%

#### 4. Model Validation and Operational Forecasting Results

##### 4.1. Model Validation

The trained probabilistic CFW forecasting models were validated using the validation data set. A total of 23 CFW events and 23 randomly selected non-CFW events (46 events in total) were used for model validation. Figure 3 presents the forecasting results of these 46 events for the two models and two lead times. Model I outperformed Model II because most of the CFW and non-CFW events were correctly predicted by Model I. The majority of CFW events were correctly predicted, but the prediction of non-CFW events was less accurate. Table 4 lists the accuracy and recall results for the validation data. For 12 h forecasting, Model I outperformed Model II. However, for 24 h forecasting, Models I and II demonstrated similar performance. The accuracy and recall rates were mostly higher than 72% (except for the accuracy of Model II for 12 h forecasting), indicating that the trained models were able to forecast CFW occurrence. Although the model validation is satisfactory, the proposed model could be improved to attain higher accuracy and recall rates as long as more CFW data are collected. According to the results for the training and validation data sets, Model I outperformed Model II for 12 h forecasting, and Models I and II demonstrated similar performance for 24 h forecasting.



**Figure 3.** Probabilistic forecasting of CFW occurrence for validation data.

**Table 4.** Forecasting performance in terms of accuracy and recall (validation data).

Lead Time	Model I		Model II	
	Accuracy	Recall	Accuracy	Recall
12 h	83%	91%	63%	78%
24 h	78%	74%	72%	78%

#### 4.2. Operational Probabilistic Forecasting

The proposed probabilistic CFW forecasting models were applied by the Operational Forecast System of the TCWB in experimental mode in 2018. Model I was the main model used for forecasting the probability of CFW occurrence, and Model II was a backup model used when real-time buoy observations were unavailable. The probabilistic CFW forecasting system provided probability forecasts of CFW occurrence at 6-h intervals, namely 00:00, 06:00, 12:00, and 18:00 CST. Figure 4 illustrates the forecasted probability values (rounded to the nearest integer) for August 2018. The forecasted probabilities for the lead times of 12 and 24 h were similar, but the probability yielded by 24 h forecasting was generally higher than that yielded by 12 h forecasting. The probabilistic CFW

forecasting system predicted low probability of CFW occurrence during dates close to the 5th of August and at the end of the month; high probability was predicted during dates close to the 12th and 23rd of August. Some data in Figure 4 are missing because the system failed to produce forecasts because of the unavailability of real-time input data and technical problems in terms of connecting to the database. Although calculations can be performed in offline mode, the real situation of the CFW forecasting system based on its operational test in experimental mode is presented in the current study.

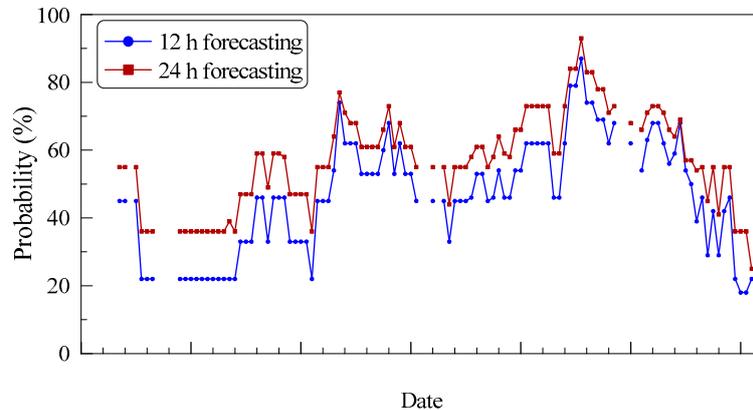


Figure 4. Probability of CFW occurrence during operational testing.

As described in Section 3, swell has a strong relationship with CFW occurrence [1,46], and the combination of swell and local wind waves may generate CFWs [47]. Therefore, the probabilistic CFW forecasting results were compared with relevant swell information. The TCWB regularly issues swell alerts when an observed or forecasted swell features a significant wave height exceeding 1.5 m and a period longer than 8 s. Thus, the swell alert data were used to assess the probabilistic CFW forecasting results. However, the TCWB does not issue swell alerts when typhoon or strong wind warnings have been issued. Therefore, swell alert data can only be compared with CFW forecasting results when typhoon or strong wind warnings have not been issued. Table 5 lists the swell alert times during 2018 and the 12 h forecasts of CFW probability for the 6 h alert period. The results indicate that the CFW forecasting system predicted high CFW probability when a swell watch was issued. As swells are related to the probability of CFW occurrence, the comparison result supports the effectiveness of the CFW forecasting system.

Table 5. Swell alert issuing time and forecasted CFW probability.

Date (yyyy/mm/dd)	Issuing Time (hh:mm)	Forecasted CFW Probability (%)
2018/05/22	16:00	80
2018/07/10	14:00	91
2018/07/22	06:00	88
2018/09/14	07:00	81
2018/09/14	16:00	82
2018/09/15	02:00	90
2018/09/15	06:00	95
2018/09/15	14:00	96
2018/09/26	01:00	95
2018/09/28	03:00	96
2018/09/28	15:00	95
2018/09/28	22:00	97
2018/09/30	23:00	94
2018/10/04	23:00	93
2018/10/06	06:00	90
2018/10/28	22:00	88

### 4.3. CFW Accidents during the Period of Typhoon Jebi in 2018

Typhoons introduce large swells to coastal areas and may increase the probability of CFW occurrence. During the afternoon of 2 September 2018, the weather in Taiwan was sunny and calm, but swells with a wave weight of 1.5 m were observed on the east coast. At that time, Typhoon Jebi, a category 4 typhoon according to the Saffir–Simpson hurricane wind scale, was located 1300 km east of Taiwan (Figure 5). However, four CFW accidents were reported on the east coast, with six fatalities. Three accidents occurred between 15:00 and 17:00 approximately 70–80 km south of Longdong, and one occurred at 20:00 approximately 25 km northwest of Longdong. Figure 6 shows the probabilistic forecasting of CFW occurrence in the Longdong headland during the period of Typhoon Jebi. The CFW forecasting model predicted high probability of CFW occurrence on September 2nd and 5th. At the times when the mentioned CFW accidents occurred, the forecasted probability was 86%. Although the probabilistic CFW forecasting model was developed for CFW occurrence in the Longdong headland, it can serve as a reference for CFW occurrence throughout the east coast of Taiwan. The casualties caused by CFWs during Typhoon Jebi demonstrate the necessity of the CFW forecasting model.

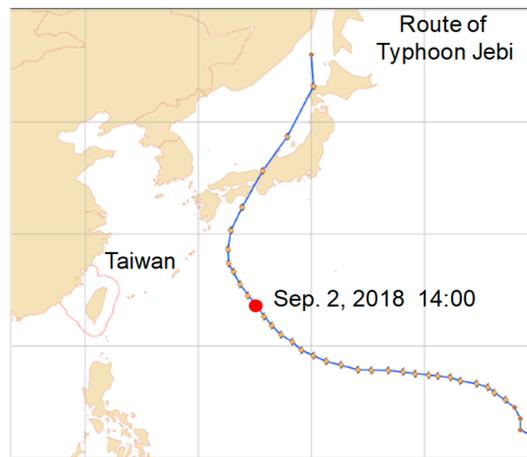


Figure 5. Route of Typhoon Jebi in 2018.

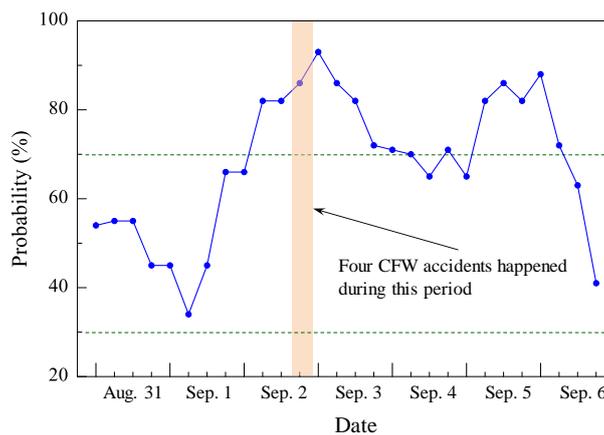


Figure 6. Probabilistic CFW forecasting during the period of Typhoon Jebi.

## 5. Conclusions

Although CFWs are uncommon, they have been reported worldwide. In Taiwan, CFWs have caused many accidents and casualties, mostly on the northeast coast. Forecasting the probability of CFW occurrence and disseminating warning information can mitigate the potential harm caused by CFWs. However, the CFW is difficult to simulate using physical wave models. Therefore, the present

study used an AI method to forecast the occurrence of CFWs. A BPNN, which is a classical artificial neural network, was adopted to provide a probabilistic CFW forecasting model in this study.

Records of 63 CFW accidents that occurred during 2000–2016 in the Longdong headland and its neighboring coast were collected; this is the area in which CFW accidents occur most frequently in Taiwan. CFW accident data were checked and confirmed by the TCWB and were used to train and validate the forecasting model. Non-CFW data were randomly selected from calm sea states for model training and validation. A scheme was designed to assign the probability of CFW occurrence for model training. Model inputs were wave and meteorological variables with physical relationships to CFWs, namely significant wave height, peak wave period, onshore wind speed, misalignment between wind and wave directions, kurtosis of sea surface elevation, Benjamin–Feir index, groupiness factor, and abnormality index.

The input data were obtained from two sources, namely buoy observations and a wave model simulation. Therefore, two forecasting models, Models I and II, were trained using these data sources to perform probabilistic CFW forecasting. The models yield predictions at 6 h intervals with lead times of 12 and 24 h. The results of Models I and II were similar during training, but Model I outperformed Model II for the validation set in terms of accuracy and recall. Therefore, for the operational forecasting system, which was tested during 2018, Model I was used as the main model, and Model II was a backup model. The probabilistic CFW forecasting results were closely correlated with swell, which is physically related to CFWs, suggesting that the proposed model is suitable for forecasting CFW occurrence. Moreover, four CFW accidents occurred on the east coast of Taiwan on 2 September 2018, when Typhoon Jebi was located in the Pacific Ocean far away from Taiwan. The CFW forecasting model predicted a probability of 86% for CFW occurrence during the time that these accidents happened. The CFW accidents during the Typhoon Jebi period indicate the necessity of the CFW forecasting model. The probabilistic CFW forecasting model is undergoing tests and improvements and will be employed in the Operational Forecast System of the TCWB for CFW warnings.

**Author Contributions:** Conceptualization, D.-J.D., S.-T.C. and C.-H.T.; Data curation, Y.-C.C.; Formal analysis, D.-J.D., S.-T.C. and Y.-C.C.; Methodology, D.-J.D., S.-T.C. and C.-H.T.; Writing—original draft, D.-J.D.; Writing—review & editing, D.-J.D. and S.-T.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Central Weather Bureau and the Ministry of Science and Technology (grant no. MOST 106-2628-E-006-008-MY3) of Taiwan.

**Acknowledgments:** This research was supported by the Central Weather Bureau and the Ministry of Science and Technology (grant no. MOST 106-2628-E-006-008-MY3) of Taiwan. Buoy data were measured and processed by the Coastal Ocean Monitoring Center of National Cheng Kung University. The authors express their gratitude for all support.

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

## References

1. Chien, H.; Kao, C.C.; Chuang, L.Z. On the characteristics of observed coastal freak waves. *Coast. Eng. J.* **2002**, *44*, 301–319. [[CrossRef](#)]
2. Tsai, C.H.; Su, M.Y.; Huang, S.J. Observations and conditions for occurrence of dangerous coastal waves. *Ocean Eng.* **2004**, *31*, 745–760. [[CrossRef](#)]
3. Didenkulova, I.I.; Slunyaev, A.V.; Pelinovsky, E.N.; Kharif, C. Freak waves in 2005. *Nat. Hazards Earth Syst.* **2006**, *6*, 1007–1015. [[CrossRef](#)]
4. Didenkulova, I.; Anderson, C. Freak waves of different types in the coastal zone of the Baltic Sea. *Nat. Hazards Earth Syst.* **2010**, *10*, 2021–2029. [[CrossRef](#)]
5. Nikolkina, I.; Didenkulova, I. Rogue waves in 2006–2010. *Nat. Hazards Earth Syst.* **2011**, *11*, 2913–2924. [[CrossRef](#)]
6. Doong, D.J.; Peng, J.P.; Chen, Y.C. Development of a warning model for coastal freak wave occurrences using an artificial neural network. *Ocean Eng.* **2018**, *169*, 270–280. [[CrossRef](#)]

7. Feraco, F.; Marino, R.; Pumir, A.; Primavera, L.; Mininni, P.D.; Pouquet, A.; Rosenberg, D. Vertical drafts and mixing in stratified turbulence: Sharp transition with Froude number. *Europhys. Lett.* **2018**, *123*, 44002. [[CrossRef](#)]
8. French, M.N.; Krajewski, W.F.; Cuykendall, R.R. Rainfall forecasting in space and time using a neural network. *J. Hydrol.* **1992**, *137*, 1–31. [[CrossRef](#)]
9. Chen, S.T.; Yu, P.S.; Liu, B.W. Comparison of neural network architectures and inputs for radar rainfall adjustment for typhoon events. *J. Hydrol.* **2011**, *405*, 150–160. [[CrossRef](#)]
10. Lin, G.F.; Jhong, B.C.; Chang, C.C. Development of an effective data-driven model for hourly typhoon rainfall forecasting. *J. Hydrol.* **2013**, *495*, 52–63. [[CrossRef](#)]
11. Jhong, Y.D.; Chen, C.S.; Lin, H.P.; Chen, S.T. Physical hybrid neural network model to forecast typhoon floods. *Water* **2018**, *10*, 632. [[CrossRef](#)]
12. Dariane, A.B.; Azimi, S. Streamflow forecasting by combining neural networks and fuzzy models using advanced methods of input variable selection. *J. Hydroinform.* **2018**, *20*, 520–532. [[CrossRef](#)]
13. Bray, M.; Han, D. Identification of support vector machines for runoff modeling. *J. Hydroinform.* **2004**, *6*, 265–280. [[CrossRef](#)]
14. Yu, P.S.; Chen, S.T.; Chang, I.F. Support vector regression for real-time flood stage forecasting. *J. Hydrol.* **2006**, *328*, 704–716. [[CrossRef](#)]
15. Han, D.; Chan, L.; Zhu, N. Flood forecasting using support vector machines. *J. Hydrol.* **2007**, *9*, 267–276. [[CrossRef](#)]
16. Chen, S.T.; Yu, P.S. Real-time probabilistic forecasting of flood stages. *J. Hydrol.* **2007**, *340*, 63–77. [[CrossRef](#)]
17. Chen, S.T.; Yu, P.S. Pruning of support vector networks on flood forecasting. *J. Hydrol.* **2007**, *347*, 67–78. [[CrossRef](#)]
18. Chen, S.T.; Yu, P.S.; Tang, Y.H. Statistical downscaling of daily precipitation using support vector machines and multivariate analysis. *J. Hydrol.* **2010**, *385*, 13–22. [[CrossRef](#)]
19. Chen, S.T. Multiclass support vector classification to estimate typhoon rainfall distribution. *Disaster Adv.* **2013**, *6*, 110–121.
20. Chen, S.T. Mining informative hydrologic data by using support vector machines and elucidating mined data according to information entropy. *Entropy* **2015**, *17*, 1023–1041. [[CrossRef](#)]
21. Yu, P.S.; Chen, C.J.; Chen, S.J. Application of gray and fuzzy methods for rainfall forecasting. *J. Hydrol. Eng.* **2000**, *5*, 339–345. [[CrossRef](#)]
22. Yu, P.S.; Chen, S.T.; Wu, C.C.; Lin, S.C. Comparison of grey and phase-space rainfall forecasting models using fuzzy decision method. *Hydrol. Sci. J.* **2004**, *49*, 655–672. [[CrossRef](#)]
23. Yu, P.S.; Chen, S.T.; Chen, C.J.; Yang, T.C. The potential of fuzzy multi-objective model for rainfall forecasting from typhoons. *Nat. Hazards* **2005**, *34*, 131–150. [[CrossRef](#)]
24. Yu, P.S.; Chen, S.T. Updating real-time flood forecasting using a fuzzy rule-based model. *Hydrol. Sci. J.* **2005**, *50*, 265–278. [[CrossRef](#)]
25. Chen, C.S.; Jhong, Y.D.; Wu, T.Y.; Chen, S.T. Typhoon event-based evolutionary fuzzy inference model for flood stage forecasting. *J. Hydrol.* **2013**, *490*, 134–143. [[CrossRef](#)]
26. Lohani, A.K.; Goel, N.K.; Bhatia, K.K.S. Improving real time flood forecasting using fuzzy inference system. *J. Hydrol.* **2014**, *509*, 25–41. [[CrossRef](#)]
27. Chen, C.S.; Jhong, Y.D.; Wu, W.Z.; Chen, S.T. Fuzzy time series for real-time flood forecasting. *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 645–656. [[CrossRef](#)]
28. Deo, M.C.; Sridhar Naidu, C. Real time wave forecasting using neural networks. *Ocean Eng.* **1998**, *26*, 191–203. [[CrossRef](#)]
29. Deo, M.C.; Jha, A.; Chaphekar, A.S.; Ravikant, K. Neural networks for wave forecasting. *Ocean Eng.* **2001**, *28*, 889–898. [[CrossRef](#)]
30. Tsai, C.P.; Lin, C.; Shen, J.N. Neural network for wave forecasting among multi-stations. *Ocean Eng.* **2002**, *29*, 1683–1695. [[CrossRef](#)]
31. Mandal, S.; Prabakaran, N. Ocean wave forecasting using recurrent neural networks. *Ocean Eng.* **2006**, *33*, 1401–1410. [[CrossRef](#)]
32. Mahjoobi, J.; Mosabbeq, E.A. Prediction of significant wave height using regressive support vector machines. *Ocean Eng.* **2009**, *36*, 339–347. [[CrossRef](#)]

33. Elbisy, M.S. Sea wave parameters prediction by support vector machine using a genetic algorithm. *J. Coast. Res.* **2013**, *31*, 892–899. [[CrossRef](#)]
34. Kazeminezhad, M.H.; Etemad-Shahidi, A.; Mousavi, S.J. Application of fuzzy inference system in the prediction of wave parameters. *Ocean Eng.* **2005**, *32*, 1709–1725. [[CrossRef](#)]
35. Özger, M.; Şen, Z. Prediction of wave parameters by using fuzzy logic approach. *Ocean Eng.* **2007**, *34*, 460–469. [[CrossRef](#)]
36. Mahjoobi, J.; Etemad-Shahidi, A.; Kazeminezhad, M.H. Hindcasting of wave parameters using different soft computing methods. *Appl. Ocean Res.* **2008**, *30*, 28–36. [[CrossRef](#)]
37. Chang, H.K.; Chien, W.A. A fuzzy–neural hybrid system of simulating typhoon waves. *Coast. Eng.* **2006**, *53*, 737–748. [[CrossRef](#)]
38. Chang, H.K.; Chien, W.A. Neural network with multi-trend simulating transfer function for forecasting typhoon wave. *Adv. Eng. Softw.* **2006**, *37*, 184–194. [[CrossRef](#)]
39. Tsai, C.C.; Wei, C.C.; Hou, T.H.; Hsu, T.W. Artificial neural network for forecasting wave heights along a ship’s route during hurricanes. *J. Waterw. Port Coast.* **2017**, *144*, 04017042. [[CrossRef](#)]
40. Wei, C.C. Nearshore wave predictions using data mining techniques during typhoons: a case study near Taiwan’s northeastern coast. *Energies* **2017**, *11*, 11. [[CrossRef](#)]
41. Chen, S.T. Probabilistic forecasting of coastal wave height during typhoon warning period using machine learning methods. *J. Hydroinform.* **2019**, *21*, 343–358. [[CrossRef](#)]
42. Doong, D.J.; Chen, C.C.; Kao, C.C.; Lee, B.C.; Yeh, S.P. Data quality check procedures of an operational coastal ocean monitoring network. *Ocean Eng.* **2007**, *34*, 234–246. [[CrossRef](#)]
43. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
44. Haykin, S. *Neural Networks: A Comprehensive Foundation*; MacMillan: New York, NY, USA, 1994.
45. Hagan, M.T.; Demuth, H.B.; Beale, M.H. *Neural Network Design*; PWS Publishing: Boston, MA, USA, 1996.
46. Tamura, H.; Waseda, T.; Miyazawa, Y. Freakish sea state and swell-windsea coupling: Numerical study of the Suwa-Marui incident. *Geophys. Res. Lett.* **2009**, *36*, L01607. [[CrossRef](#)]
47. Gramstad, O.; Trulsen, K. Can swell increase the number of freak waves in a wind sea? *J. Fluid Mech.* **2010**, *650*, 57–79. [[CrossRef](#)]
48. Mori, N.; Janssen, P.A. On kurtosis and occurrence probability of freak waves. *J. Phys. Oceanogr.* **2006**, *36*, 1471–1483. [[CrossRef](#)]
49. Benjamin, T.B.; Feir, J.E. The disintegration of wave trains on deep water Part 1. Theory. *J. Fluid Mech.* **1967**, *27*, 417–430. [[CrossRef](#)]
50. Janssen, P.A. Nonlinear four-wave interactions and freak waves. *J. Phys. Oceanogr.* **2003**, *33*, 863–884. [[CrossRef](#)]
51. Mori, N.; Onorato, M.; Janssen, P.A. On the estimation of the kurtosis in directional sea states for freak wave forecasting. *J. Phys. Oceanogr.* **2011**, *41*, 1484–1497. [[CrossRef](#)]
52. Funke, E.R.; Mansard, E.P.D. On the synthesis of realistic sea states. *Coast. Eng. Proc.* **1980**, *1*, 2974–2991.
53. Goda, Y. *Random Seas and Design of Maritime Structures*; World Scientific Publishing Company: Singapore, 2000.
54. Dean, R. Freak waves: A possible explanation. In *Water Wave Kinematics*; Torum, A., Gudmestad, O., Eds.; Kluwer: Dordrecht, The Netherlands, 1990; pp. 609–612.
55. Kharif, C.; Pelinovsky, E.; Slunyaev, A. *Rogue Waves in the Ocean*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008.

