Article

Risk Assessment and Its Visualization of Power Tower under Typhoon Disaster Based on Machine Learning Algorithms

Hui Hou [1], Shiwen Yu [1],* , **Hongbin Wang [2], Yong Huang [3], Hao Wu [3], Yan Xu [4], Xianqiang Li [1] and Hao Geng [1]**

[1] School of Automation, Wuhan University of Technology, Wuhan 430070, China; husthou@126.com (H.H.); lxq@whut.edu.cn (X.L.); whutgenghao@163.com (H.G.)
[2] Guangzhou Power Supply Bureau Co., Ltd., Guangzhou 510620, China; forest_whb@tom.com
[3] GuangDong Power GRID Co., Ltd., Electric Power Research Institute, Guangzhou 510080, China; huangyong_wh@126.com (Y.H.); wh_dky@163.com (H.W.)
[4] Department, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore; eeyanxu@gmail.com
* Correspondence: yswwhut@163.com

check for updates

Abstract: For power system disaster prevention and mitigation, risk assessment and visualization under typhoon disaster have important scientific significance and engineering value. However, current studies have problems such as incomplete factors, strong subjectivity, complicated calculations, and so on. Therefore, a novel risk assessment and its visualization system consisting of a data layer, knowledge extraction layer, and visualization layer on power towers under typhoon disaster are proposed. On the data layer, a spatial multi-source heterogeneous information database is built based on equipment operation information, meteorological information, and geographic information. On the knowledge extraction layer, six intelligent risk prediction models are established based on machine learning algorithms by hyperparameter optimization. Then the relative optimal model is selected by comparing five evaluation indicators, and the combined model consisting of five relatively superior models is established by goodness of fit method with unequal weight. On the visualization layer, the predicted results are visualized with accuracy of 1 km × 1 km by ArcGIS 10.4. In results, the power tower damage risk assessment is carried out in a Chinese coastal city under the typhoon 'Mujigae'. By comparing predicted distribution and similarity indicator of the combined model with those of the other models, it is shown that the combined model is superior not only in quality but also in quantity.

Keywords: typhoon; power tower; risk assessment; visualization; machine learning; intelligent prediction model

1. Introduction

As one of the extreme disasters, typhoon has a tremendous impact on the power system. It may not only cause damage to power facilities, but also lead to large scale blackouts, which will seriously threaten people's production and life. Therefore, it is necessary and urgent to study the safety risk of power system under typhoon disaster.

At present, several studies evaluated the reliability of power grid under extreme weather conditions. Zhou et al. analyzed the resilience of distribution network under extreme weather, and provided the distribution network resilience index and assessment idea based on information entropy and multiple failure model [1]. Yin et al. pointed out that most of the faults in typhoon

weather were caused by failures of the power tower and pole, so it was necessary to pay attention to the power tower and pole [2]. Based on transmission line segmentation, the dynamic safety was analyzed by the optimization method by Song et al. [3]. However, the segmentation of the transmission line was not accurate enough. Although the prediction accuracy was improved by establishing a micro-meteorological model by Gao et al. [4], it increased the amount of calculation. Aiming at the complex structure of the distribution network, a health index for system operation status and importance indicator for user difference were proposed by Huang et al. [5], but with strong subjectivity for quantification. In the risk assessment of distribution network, the reclosing success rate was regarded as a dichotomous variable and fitted with logistic regression, and the overall risk assessment speed was improved by Yang et al. [6]. However, the accuracy of the whole model was lower than the traditional method. The relationship between wind speed (m/s) and transmission line failure rate was described with a vulnerability curve by Yang et al. [7], but the use of average wind speed (m/s) of 10 min is not accurate enough for predicting, and the wind speed at turning point of the vulnerability curve is subjective. It can be seen that current studies still have problems such as incomplete factors, strong subjectivity, and complicated calculations.

In order to visually display the risk severity and location after risk assessment of power system under typhoon disaster, it is necessary to visualize the risk. However, at present, the visualization technology is rarely used in the early warning of power system risk in China, except for the comprehensive disaster prevention and mitigation system and the grid icing automatic forecasting system of Fujian Electric Power Company [8,9], and the mountain fire prevention system for transmission lines of Hunan Power Company [10]. In view of the importance of visualization technology in disaster warnings, Wang et al. summarized the GIS-based monitoring and early warning technology for grid meteorological disasters [11]. In general, visualization technology has been applied in domestic engineering practice, but it is mainly used in emergency response after disasters, and still relatively rare in early warning research. There were relatively more visualization technologies applied in foreign early warning related research. Liu et al. employed a negative binomial regression model and showed the average number of power outages per square kilometer [12]. They also proposed the GLMM (generalized linear mixed model) to predict the number of power outages in the target area under hurricane and ice storm, and finally showed the predicted and actual number of power outages in each zip code area [13], but insufficient consideration of microtopography factors, such as altitude, aspect, slope, etc. Mensah et al. proposed a flexible evaluation framework and showed the proportion of users' power outages in each geographic grid in the target area [14], but the empirical function was used to replace the actual data, which reduced the credibility. Han et al. used GIS technology to deal with weather conditions and environmental conditions, and constructed statistical models to assess the risks within the geographic grid [15]. They also proposed GAM (generalized addictive model) to predict the number of power outages in the target area before typhoon comes, and the results showed that the GAM were more precise than the GLM (generalized linear model) by visual comparison [16]. It can be seen that although foreign early warning researches applied visualization technology more generally, the risk assessment methods also have problems such as incomplete factors and strong subjectivity.

In summary, the current risk assessment methods for power system security under typhoon disasters have problems such as incomplete factors, strong subjectivity, complex calculation, and less application of visualization techniques to power system risk warnings, and so on. Therefore, based on a machine learning algorithm, this paper presents a novel method that comprehensively considers spatial multi-source heterogeneous information such as equipment operation information, meteorological information, and geographic information. First, by using hyperparameter optimization and goodness of fit method with unequal weight, six single intelligent models, and a combined model for power tower damage predicting are established. Second, these models are applied to predict damage risk under a specific typhoon. Finally, predicted results are compared to actual damage and the best model is selected.

The combined model utilizes advantages of different learning strategies, so it is more objective and reliable than a single intelligent model. It is easy to calculate and can visually exhibit predicted results with an accuracy of 1 km × 1 km. Therefore, the research results can provide a theoretical basis and practical guidance for the decision making of disaster prevention and mitigation for the power sector.

2. Study Area

In recent years, the typhoon had huge impact on China's coastal power grids [17–19]. Taking the typhoon 'Mujigae', no. 22 of 2015, as an example, it was strengthened from the tropical depression that was originally located near the Philippines. It entered eastern part of the South China Sea around noon on 2 October 2015, and its intensity gradually strengthened. It landed on the coast of the Potou District of Zhanjiang City around 14:10 on 4 October 2015, and affected as far as the Pearl River Delta. At the time of landing, the maximum wind near the typhoon center was 15 class (about 50 m/s), and the lowest pressure in the center was 940 hPa. For the first time, the gust in Zhanjiang reached 18 class and set a historical record. 'Mujigae' was the strongest typhoon that landed on China since October 1949. It caused 256 trips to lines of 35 kV and above, and damaged 80 towers [20].

In order to reveal the connection between typhoon disaster and tower damage and provide support for the power sector in disaster prevention and mitigation, the tower damage data of main network under typhoons 'Rammasun' [17] and 'Hato' [19] are utilized to train intelligent models. Then the tower damage risk in a Chinese coastal city under typhoon 'Mujigae' is predicted and visualized. Finally, the predicted results of intelligent model and combined model are compared.

3. Framework for Risk Assessment and Its Visualization

This paper proposes a risk assessment and its visualization method of power tower under typhoon disasters based on machine learning algorithms. The overall framework is shown in Figure 1. It is divided into three layers: the data layer, knowledge extraction layer, and visualization layer.
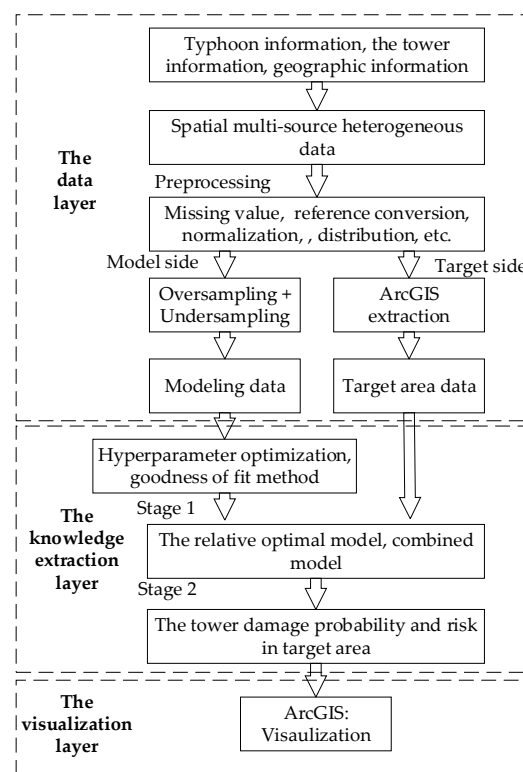

Figure 1. Overall structure of the method.

The first layer is the data layer, which first preprocesses the data, mainly including missing values filling, reference conversion, normalization, distribution, and so on. The model side mainly deals with historical disaster data, which is used to build intelligent models. The target side mainly processes the target area data for actual prediction. The data consists of spatial multi-source heterogeneous information such as equipment operation information, weather information, geographic information, etc.

The second layer is the knowledge extraction layer, which is divided into two stages: modeling and predicting. Firstly, based on machine learning algorithms, the power tower damage prediction model is established on the model side through hyperparameter optimization. Then, on one hand, the relative optimal model is selected by comparing the model evaluation indicators, on the other hand, the combined model is constructed by goodness of fit method with unequal weights. Secondly, an actual prediction on the target side is made.

The third layer is the visualization layer, which visualizes the predicted results of models in the second layer and provides a theoretical basis and practical guidance for the decision-making in disaster prevention and mitigation.

4. Construction of Power Tower Risk Assessment System

4.1. Data Layer

The spatial multi-source heterogeneous database contains equipment operational information including V' (design wind speed), meteorological information including V (maximum gusts), and geographic information such as H (altitude), A (slope direction), S (slope), P (slope position), U (underlying surface type), R (surface roughness), and so on.

4.1.1. Data Preprocessing

The missing value is filled with the median and V and V' are converted to 10 m high. According to China's current wind load specification [21], the variation of wind speed along the height can be calculated by exponential law

$$V_z = V_1 \left(\frac{z}{z_1} \right)^{\alpha} \tag{1}$$

where V_z (m/s) denotes the wind speed at height z, V_1 (m/s) denotes the wind speed at height z_1, α denotes the surface roughness coefficient, which can be chosen from Table 1 [21] that derives from relevant materials and specifications.

Table 1. Ground roughness factor.

Class	Surface Features	α
A	Offshore seas, islands, coasts, lakeshores, and desert areas	0.10–0.13
B	Fields, villages, jungles, hills, small and medium size towns, and suburbs of large cities where housing density is sparse	0.13–0.18
C	Urban areas with dense buildings	0.18–0.28
D	Urban area of a large city with dense buildings and tall houses	0.28–0.44

Meteorological stations are generally located in open plains, and the surface roughness is generally classified as Class B. V_1 in this study is V or V', and z_1 is the monitoring height or reference height of design wind. In this study, α is set as 0.16.

The data is normalized

$$X^* = (x - x_{\min})/(x_{\max} - x_{\min}) \tag{2}$$

where, X^* is the normalized feature; x is the original feature; $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the original feature, and they are set to be integer for convenient. The processed features are shown in Table 2.

Table 2. Feature interpretation table.

Feature Name	Meaning	Value Range
V_{10} (m/s)	Maximum gust at 10 m	0–70
V'_{10} (m/s)	Design wind at 10 m	0–50
H (m)	Altitude	−102–2483
A (°)	Aspect	−1–360
S (°)	Slope	0–90
P	Slope position	0–3
U	Underlying surface	70–79
R (m)	Surface roughness	0–30

Finally, conduct homogeneous and completeness analysis by use of *k*-means clustering in Python 3.6. Provide distribution of the data.

4.1.2. Processing of Model Side Data

On the model side, since the amount of damaged towers are much less than undamaged ones, the model training is faced with seriously unbalanced classes. Therefore, all the towers tend to be predicted as undamaged, thus reducing the credibility of the model. This problem can be solved by improving the sampling method, mainly including oversampling and undersampling. Oversampling is used to enlarge small class, and undersampling is used for partial sampling of large class.

In this paper, two sampling methods are synthesized. On one hand, the data of damaged towers are copied, on the other hand, the data of undamaged towers are randomly sampled with the same amount. Therefore, a space multi-source heterogeneous database on the model side is constructed. The class label is set to be dichotomous ($y = 0$ means the tower is not damaged, and $y = 1$ means the tower is damaged).

4.1.3. Processing of Target Side Data

On the target side, ArcGIS 10.4 is used to mesh and extract data from the target area through the following actions:

- Geographically mesh (rectangular grid constructed with latitude and longitude lines, hereinafter referred to as grid) the target area.
- Covert the maximum gust to 10 m high by method in Section 4.1.1 at each monitoring station under a typhoon, and the gust map is generated by the inverse distance weight interpolation method [22,23], and $V_{i,10}$ (the maximum gust at 10 m) in each grid is extracted, where i ($i = 1,2,$ ... , n) represents the grid serial number.
- Load the coordinates of the main network tower in the target area by ArcGIS 10.4, and extract N_i (the total number of towers), $V_{i,10}$ (design wind speed) from grid i.
- Extract geographic information within the grid, including H_i, A_i, S_i, P_i, U_i, R_i, and so on.

4.2. Knowledge Extraction Layer

4.2.1. Introduction to Algorithms

This layer is used to establish predicting models for damage probability prediction on the target side. Machine learning algorithms can be divided into individual learning and ensemble learning according to learning strategies. Individual learning uses a single model for predicting, while ensemble learning employs multiple models and calculates final results by combining predictions of each model.

Typical individual learning regression algorithms include LR (logistic regression) [24], SVR (support vector regression) [25], CART (classification and regression tree) [26,27], and so on. Typical ensemble learning regression algorithms include Adaboost iteration [28,29], GBRT (gradient regression tree) [30], RF (random forest) [31], and so on.

Here is brief introduction to the six algorithms:

LR is a classical method in statistical learning. The common binomial logistic regression model is defined by Equations (3) and (4)

$$P(Y = 1|x) = \frac{\exp(w \cdot x + b)}{1 + \exp(w \cdot x + b)} \tag{3}$$

$$P(Y = 0|x) = \frac{1}{1 + \exp(w \cdot x + b)} \tag{4}$$

where, $x \in \mathbf{R}^n$ is the input, $Y \in \{0,1\}$ is the output, $w \in \mathbf{R}^n$ and $b \in \mathbf{R}$ are parameters, $w \in \mathbf{R}^n$ is called the weight vector, b is the bias, and $w \cdot x$ is the inner product of w and x. It can not only predict class, but also provide the probability.

SVR transforms the problem into a high-dimensional feature space through nonlinear mapping and constructs a linear regression function in this space to replace the nonlinear function in original space [32]. A training set is established

$$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\} \tag{5}$$

where $x_i \in \mathbf{R}^n$, $y_i \in \mathbf{R}$, and SVR use the regression function

$$y = f(x) = w \cdot \varphi(x) + b \tag{6}$$

where $\varphi(x)$ is the nonlinear mapping of the input space to the high-dimensional feature space; $w \in \mathbf{R}^n$ is the weight vector, and $b \in \mathbf{R}$ is the bias, the way to obtain these 2 parameters is shown in Appendix B.

CART is a decision tree learning method that can be used for classification and regression. Assume that x and y are input and output variables, respectively, and y is a continuous variable, given a data set as shown in Equation (5). A regression tree divides the input space and specify an output value on each unit. Assuming that the input space has been divided into M units $R_1, R_2, \dots, R_M$ and there is a fixed output value c_m on each unit R_m, a regression tree is obtained

$$f(x) = \sum_{m=1}^{M} c_m I(x \in R_m) \tag{7}$$

Solving the optimal output on each unit with the squared error minimum criterion $\sum_{x_i \in R_m}^{M} (y_i - f(x_i))^2$. Detailed derivation of the algorithm is shown in Appendix B.

The Adaboost algorithm is a boosting algorithm, which aims to obtain a strong classifier from the weak classifier through iteration learning. The core of this algorithm is that each iteration uses updated data weights. Each iteration training gives greater weight to the data with large prediction errors in the previous training round, so that each weak classifier minimizes predicting error of some data. Finally, linear weights are used to combine weak classifiers. Detailed derivation of the algorithm is shown in Appendix B.

GBRT is an ensemble learning algorithm. Its basic classifier is boosting tree, and its learning method is gradient boosting. The negative gradient of the loss function is used as the approximation of the residual, and a regression tree is fitted for the residual [30]. Each training will reduce the residual in the previous training round. Finally, all the classifiers are added up. Detailed derivation of the algorithm is shown in Appendix B.

RF is an algorithm based on decision tree and Bagging (a parallel ensemble learning). Bagging is a bootstrap based sampling and performs a put back sampling on a given data set, thereby generating a series of data subsets. It trains a base learner on each subset, and finally combining the base learner. RF not only uses the data subset, but also randomly selects partial features to train the base decision tree [33]. When dividing the input space, traditional decision tree like CART selects an optimal feature

from the features (assuming there are d features) of current node. However, RF only randomly selects a subset including k features. Aforementioned dual perturbations ensure the differentiation between classifiers. Therefore, the final model is more adaptive in predicting. Parameter k controls the degree of randomness. When $k = d$, the base decision tree is the same as the traditional decision tree. Generally recommended as $k = \log_2 d$.

4.2.2. Single Intelligent Model

Six algorithms are used to establish the intelligent prediction model for tower damage probability. The main processes include original model evaluation, hyperparameter optimization, relative optimal model selection, full data fitting, actual prediction, and so on.

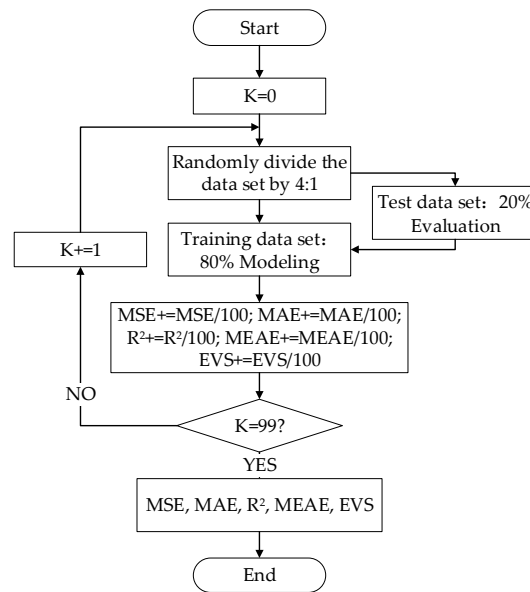The flowchart of original model evaluation is shown in Figure 2.

Figure 2. Model evaluation process.

As shown in Figure 2, firstly, the model side data is randomly divided into training set and test set according to a ratio of 4:1, and the model is built on training set by using default hyperparameters [34] in Python 3.6. Because this study uses regression methods, regression evaluation indicators are employed to evaluate the models. Five indicators including *MSE* (mean squared error), *MAE* (mean absolute error), R^2 (R–squared), *MEAE* (median absolute error), and *EVS* (explained variance score) are evaluated on test dataset. The cycle index K is set as 100, and the average value of each indicator is taken as the evaluation result of the original model.

The five indicators are displayed in Equations (8)–(12)

$$MSE = \frac{1}{N}\sum_{j=1}^{N}\left(y_j - f(x_j)\right)^2 \tag{8}$$

$$MAE = \frac{1}{N}\sum_{j=1}^{N}\left|y_j - f(x_j)\right| \tag{9}$$

$$R^2 = 1 - \frac{\sum_{j=1}^{N}\left(y_j - f(x_j)\right)^2}{\sum_{j=1}^{N}\left(y_j - \overline{y}\right)^2} \tag{10}$$

$$MEAE = median(|y_1 - f(x_1)|, \ldots, |y_N - f(x_N)|) \tag{11}$$

$$EVS = 1 - \frac{\mathrm{var}(y_i - f(x_i))}{\mathrm{var}(y_i)} \tag{12}$$

where j is the sequence number of test set data, N is the total number of test set data, y_j is the class label of data j, $f(x_j)$ is the predicted value of x_j, $\bar{y}$ is the mean of true class label on test set, *median* $(\cdot)$ is a function to obtain the median, var$(\cdot)$ is a function to obtain the variance.

Optimization of hyperparameters. Hyperparameters are parameters that need to be set in advance, not obtained through training. With the goal of minimizing $-R^2$, use the "hyperopt" in Python 3.6 to optimize part hyperparameters of each model and evaluate the model by method in Figure 2. The hyperparameters of each model are shown in Appendix A.

Selection of relatively optimal model. Determine the relative optimal model by comprehensively comparing five indicators.

For full data fitting, all the data is used to train models with optimized hyperparameters for target side predicting use.

According to risk assessment theory [35], risk is defined as the product of probability and severity. It is worth noting that this paper simply takes density of power towers in each grid as the severity index. Therefore, the grid risk is defined as Equation (13)

$$r_i = P_i N_i \tag{13}$$

where P_i is the damage probability of grid i, and N_i is the density of towers in grid i.

Actual predicting. On the target side, the spatial multi-source heterogeneous information is input into intelligent models to predict the damage probability. The risk value of the grid is calculated by Equation (13).

4.2.3. Combined Model

Combination forecasting is a method of predicting the same problem by using different methods, and it includes equal weighting combination and unequal weighting combination: equal weighting combination uses predictions of different methods have equal weight; the unequal weighting combination uses predictions of different methods have unequal weight. In this paper, different intelligent models are combined by the goodness of fit method with unequal weight. The corresponding combination model is

$$\hat{Y} = \sum_{k=1}^{L} W_k \hat{Y}_k \tag{14}$$

$$\sum_{k=1}^{L} W_k = 1 \tag{15}$$

where W_k is the weight of each model, which satisfies the constraint of Equation (15); $\hat{Y}_k$ is prediction model k, and L is the amount of models. According to the goodness of fit method, the weight of model k is

$$W_k = \frac{\sum\limits_{k=1}^{L} SE_k - SE_k}{\sum\limits_{k=1}^{L} SE_k} \frac{1}{L-1} \tag{16}$$

where SE_k is the standard error of prediction model k

$$SE_k = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(y_j - f(x_j))^2} \tag{17}$$

where j denotes the sequence number of test set data, N denotes the amount test set data, y_j denotes the class label of data x_j, and $f(x_j)$ is the predicted value of data x_j. When the prediction results are scattered, this method can assign the largest weight to model with the smallest prediction standard error, so that the prediction result can guarantee goodness of fit. Noting that Equation (17) is the square root of Equation (8), thus Equation (16) becomes

$$W_k = \frac{\sum\limits_{k=1}^{L}\sqrt{MSE_k} - \sqrt{MSE_k}}{\sum\limits_{k=1}^{L}\sqrt{MSE_k}}\frac{1}{L-1} \tag{18}$$

where MSE_k is the mean squared error of model k.

4.3. Visualization Layer

The visualization of damage probability or risk will make the decision-making process of power sector intuitive, simple, and specific in disaster prevention and mitigation, and is of great significance for promoting efficiency and reducing cost. Therefore, in this paper, the predicted damage results are visualized by use of z score method in ArcGIS 10.4, as shown in Equation (19)

$$z = \frac{x - \mu}{\sigma} \tag{19}$$

where x is the original data, which refers to damage probability or risk, μ is mean and σ is standard variance of all data. The damage probability interval (0,1) is divided into 10 equal classes by 0.1, and the risk is divided into 10 classes by natural discontinuous grading.

5. Materials and Methods

5.1. Processing of Model Side Data

Using the historical data of typhoons 'Rammasun' and 'Hato' to establish a spatial multi-source heterogeneous database on the model side. All the data are from the power sector.

The missing values are filled with the median. Because the features have different unit of measurement, they need to be normalized before homogeneity and completeness detection. Then the features are normalized by Equation (2).

Because the dataset is not large, the homogeneity and completeness are analyzed by means of 5-fold k-means clustering. The homogeneity score and completeness score are 1.0 respectively, which means all of clusters contain only data points of a single class and all the data points of a given class are elements of the same cluster.

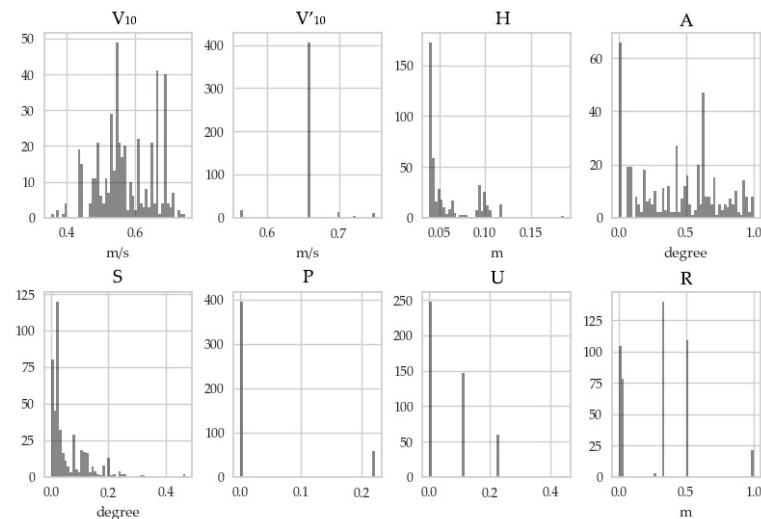The model side data distribution is depicted in Figure 3.

Figure 3. Distribution of the model side data.

Figure 3 illustrates distribution of the model side data. As seen from it, V_{10} shows a bimodal distribution, V_{10} and R illustrate unimodal distribution, H, A, S, P and U are consistent with exponential distribution.

The parallel coordinates visualization of features is shown in Figure 4. The abscissa represents normalized features, and the ordinate represents their value. $Y = 0$ and $Y = 1$ represent the undamaged state and the damaged state, respectively. Using parallel coordinate visualization, the relationship between each feature and Y can be initially diagnosed.
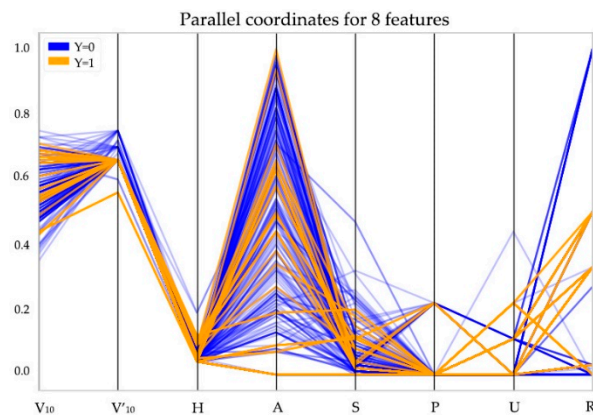
Figure 4. Parallel coordinates for eight features.

It can be seen in Figure 4 that the damage event is prone to occur in areas with smaller V'_{10} and R, indicating that smaller design wind speed and surface roughness contribute to tower damage. The feature correlation is shown in Figure 5.
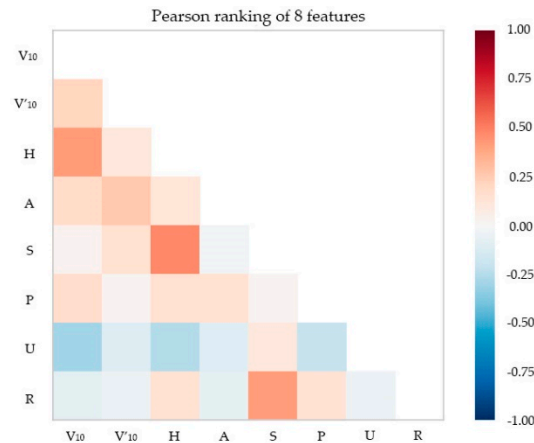
Figure 5. Correlation analysis of features.

It can be seen that some feature pairs such as H and V_{10}, R and S, H and S have strong positive correlations, and there is a strong negative correlation between U and V_{10}. Therefore, this study should not use linear model, and it can be foreseen that the prediction of LR is not accurate. According to the engineering experience, the higher the altitude, the greater the wind speed, because the air friction coefficient decreases with the increase of height. Generally, high altitude areas are mountainous areas, so the slope will increase accordingly, and the roughness will be greater than plain. Therefore, it is reasonable that H and V_{10}, R and S, H and S have strong positive correlations.

It is shown in Figure 5 that U is in negative correlation with H, and it has been shown that H is in positive correlation with V_{10}, so it is understandable that U and V_{10} have negative correlation.

Finally, the feature importance is analyzed by utilizing RF classification algorithm, as shown in Figure 6.
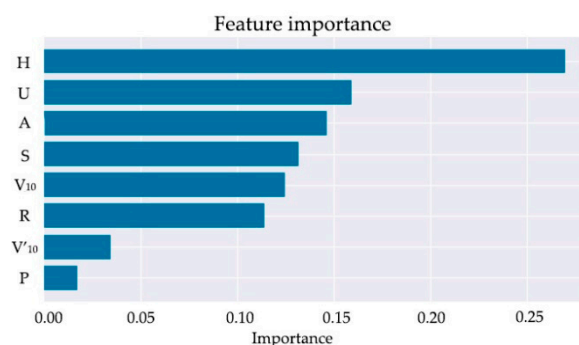

Figure 6. Feature importance ranking.

Figure 6 provides the importance ranking of eight features. H is the most important feature in this dataset and P is the least important feature. The importance of geographic information is more important than meteorological information and equipment operational information. Importance of DW and P are ranked at the last 2, because according to the entropy theory [36], the more diverse the data, the more useful information it provides, and the more important the feature is. As seen from Figure 3, V_{10} and P contain little differences in data, so these two features are relatively less important than the others. However, the feature importance ranking is only suitable for this dataset and should be evaluated again on a new dataset.

5.2. Processing of Target Side Data

The maximum gust of typhoon 'Mujigae' at 10 m and main network tower distribution are as shown in Figure 7.
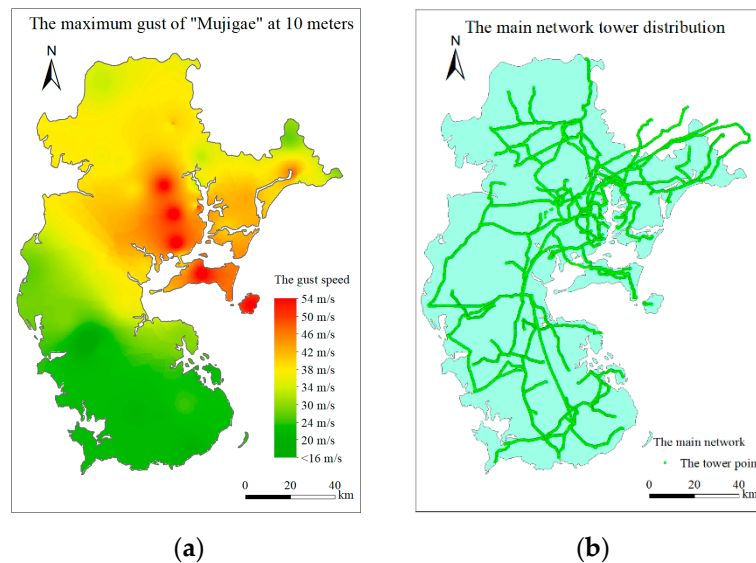
Figure 7. Distribution of maximum gust and main network tower. (**a**) The maximum gust of 'Mujigae' at 10 m. (**b**) The main network tower distribution.

Figure 7a shows that windspeed at the upper right (actually in the northeast) area is the highest, indicating that this area is close to the landing point of typhoon 'Mujigae'. The gust distribution reveals that the typhoon track is from the lower right (the southeast) area to the upper left (the northwest) area. Figure 7b presents the distribution of the main network tower.

The maximum gust, number of towers, and geographic information are extracted by 1 km × 1 km grid and normalized, and a multi-source heterogeneous information database on the target side is established.

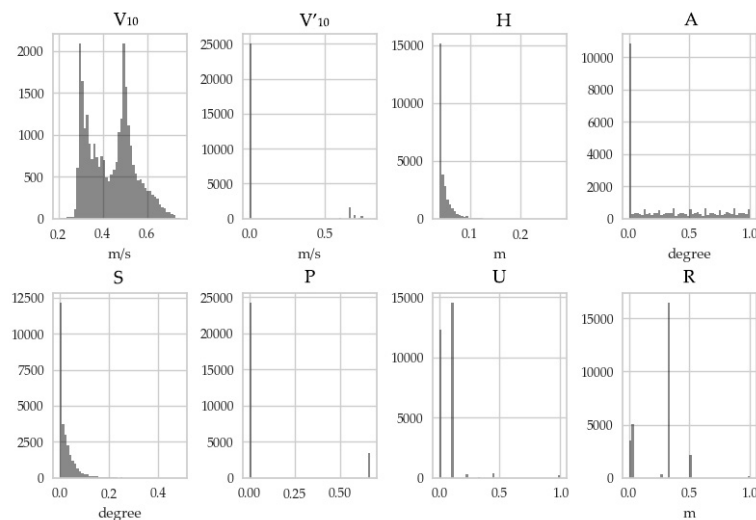The distribution of the target side data is shown in Figure 8.

Figure 8. Distribution of the target side data.

As observed in Figure 8, V_{10} and U shows a bimodal distribution, R illustrates unimodal distribution, V_{10}, H, A, S and P are consistent with exponential distribution. It is worth noting that most values of V_{10} are 0.0, because the design wind of grids with no power towers are set to be 0.0. It can be seen that the target side data distribution in Figure 8 is similar to the model side data distribution in Figure 3.

6. Results

6.1. Knowledge Extraction Layer

6.1.1. Single Intelligent Model

In this paper, six machine learning algorithms in Section 4.2.2 are used for modeling, and some hyperparameters are selected for optimization. The optimized hyperparameters are shown in Appendix A. Table 3 and Figure 9 show five indicators of six models before and after optimization.

Table 3. Evaluation results of six models before and after optimization.

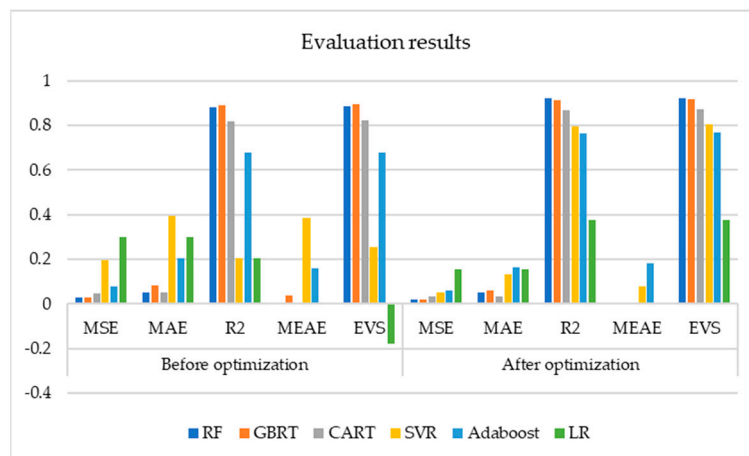| Model	Before Optimization					After Optimization				
	MSE	*MAE*	R^2	*MEAE*	*ESV*	*MSE*	*MAE*	R^2	*MEAE*	*ESV*
RF	0.0296	0.0529	0.880	0.000	0.886	0.0196	0.0503	0.920	0.007	0.922
GBRT	0.0267	0.0843	0.892	0.035	0.896	0.0213	0.058	0.914	0.003	0.916
CART	0.0451	0.0527	0.817	0.000	0.822	0.033	0.033	0.866	0.000	0.871
SVR	0.196	0.395	0.206	0.386	0.256	0.0502	0.131	0.796	0.076	0.803
Adaboost	0.0794	0.206	0.678	0.160	0.677	0.0577	0.165	0.766	0.182	0.769
LR	0.297	0.297	0.204	0.000	−0.180	0.154	0.154	0.375	0.000	0.375

Figure 9. Evaluation results of six models before and after optimization.

It can be seen from Table 3 and Figure 9 that the 5 indicators of each model have been greatly improved after optimizing. However, the five indicators of LR (the green column) are not ideal because it occupies the maximum of *MSE* and minimum of R^2 and *EVS*, indicating the necessity and correctness of the correlation analysis in Figure 5. Meanwhile, as shown in Figure 9, RF (the dark blue column) occupies the minimum of *MSE* and maximum of R^2 and *EVS* after optimization, so RF is relative optimal model. Therefore, the model ranking is: RF, GBRT, CART, SVR, Adaboost, LR.

After the full data fitting, the target side data is input into the model to predict the tower damage probability, and the risk is calculated by Equation (13).

6.1.2. Combined Model

According to R^2 in Table 3, LR is inferior to the other models even after optimization. Therefore, this paper excludes LR when constructing the combined model.

According to Equation (18) and Table 3, the weight of each model is shown in Table 4. The damage probability and risk of each grid are calculated by Model Equations (14) and (13).

Table 4. Weight of each model.

Model	RF	GBRT	CART	SVR	Adaboost
weight	0.212 441	0.210 846	0.201 265	0.189 891	0.185 557

6.2. Visualization Layer

6.2.1. Single Intelligent Model

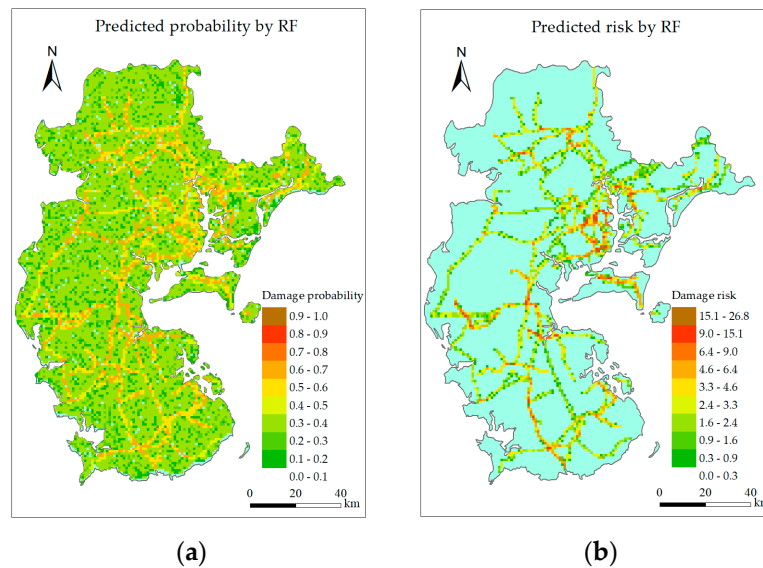Visualize the predicted damage probability and risk of RF, as shown in Figure 10.

(**a**)　　　　　　　　　　　　　　(**b**)

Figure 10. Prediction results of RF. (**a**) Predicted damage probability. (**b**) Predicted damage risk.

Figure 10a moderately illustrates the trend that damage probability is higher in the upper right area, and this trend consistence to the maximum gust distribution in Figure 7a. Therefore, the predicted damage probability is reasonable. Figure 10b shows that the highest risk grids are mainly distributed around the landing point, which is in line with the engineering experience. There are several discrete grids with high risk distributed at the lower left area.

The predicted damage probability and risk of GBRT are visualized, as shown in Figure 11.
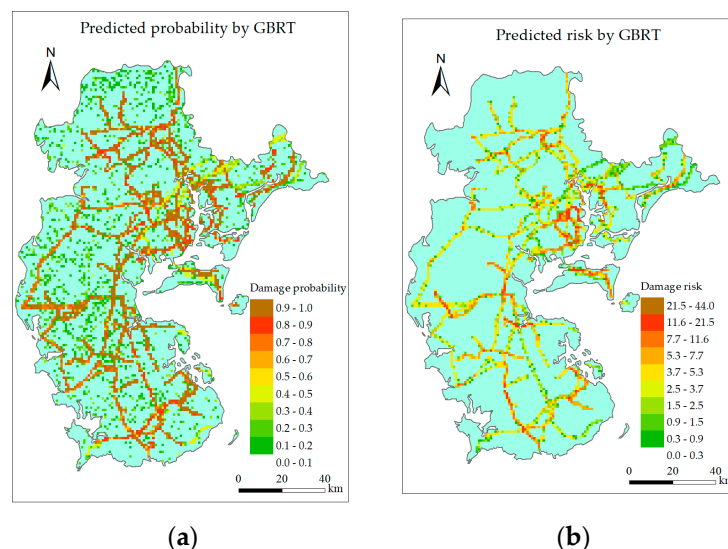

(**a**)　　　　　　　　　　　　　　(**b**)

Figure 11. Prediction results of GBRT. (**a**) Predicted damage probability. (**b**) Predicted damage risk.

Figure 11a shows that the grids with highest damage probability are distributed along the coastline and typhoon track, which conforms to engineering experience. The highest damage probability is mainly distributed at the upper right area, in line with gust distribution. Some of the highest risk grids in Figure 11b distribute mainly around landing point and the others spread around the lower left area, consistent with the gust distribution.

The predicted damage probability and risk of CART is visualized, as shown in Figure 12.
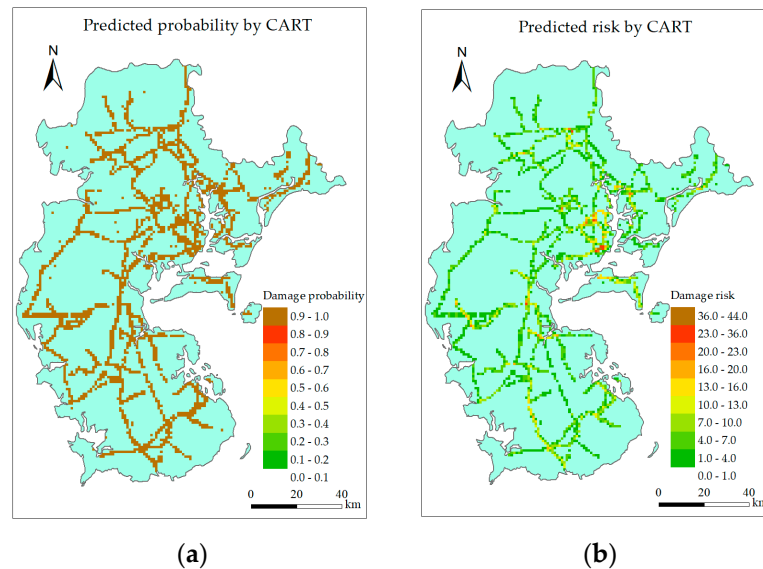

(**a**)   (**b**)

Figure 12. Prediction results of CART. (**a**) Predicted damage probability. (**b**) Predicted damage risk.

It can be seen from Figure 12a that all the grids contain towers are predicted to be high probability, so Figure 12b can only provide both density distribution of the power tower and predicted damage probability distribution.

The predicted damage probability and risk of SVR are visualized, as shown in Figure 13.
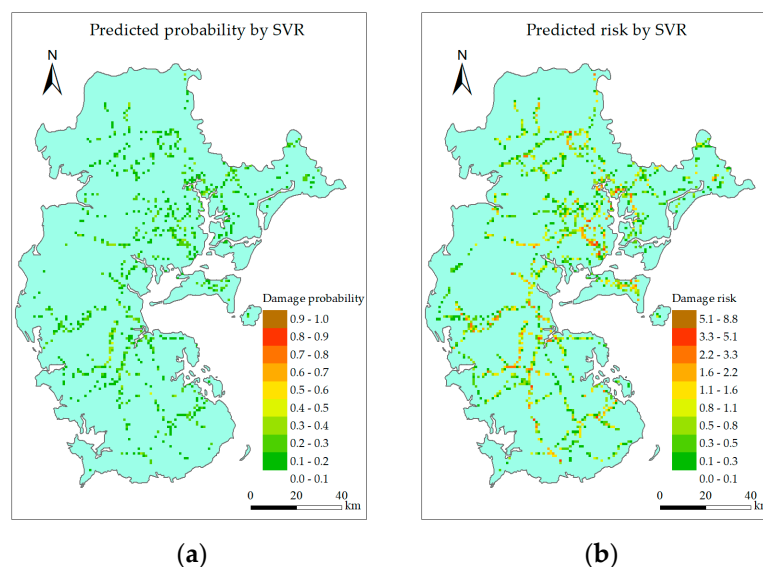

(**a**)   (**b**)

Figure 13. Prediction results of SVR. (**a**) Predicted damage probability. (**b**) Predicted damage risk.

Figure 13a illustrates a general low probability distribution and only exhibit the moderate tendency that damage probability at the upper right area is higher than the lower left area. The high-risk

grids are mainly distributed around landing point, which is consistent with the gust distribution and engineering experience.

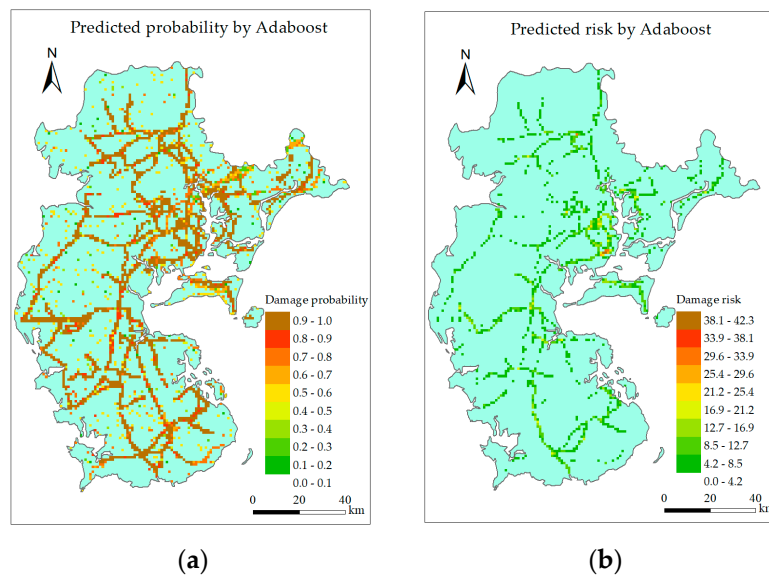The predicted damage probability and risk of Adaboost are visualized, as shown in Figure 14.

Figure 14. Prediction results of Adaboost. (**a**) Predicted damage probability. (**b**) Predicted damage risk.

Figure 14a,b are almost the same as Figure 12a,b, but the risk is lower than CART at the lower left area.

The predicted damage probability and risk of LR are visualized, as shown in Figure 15.
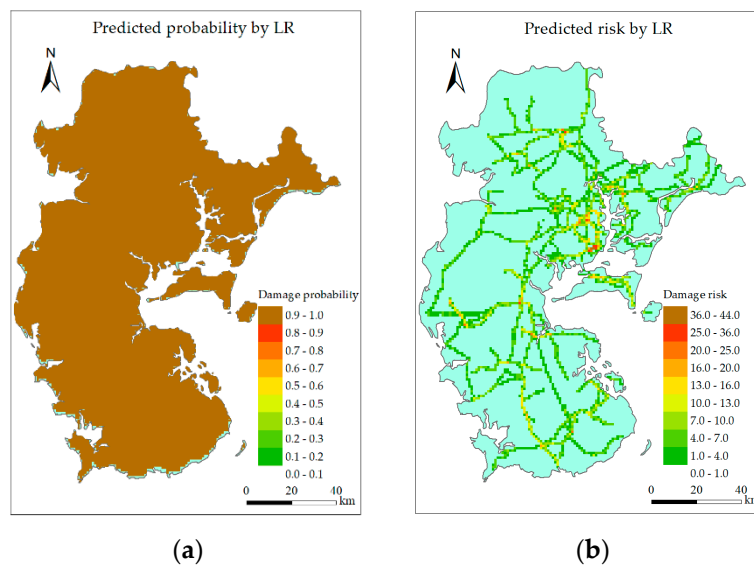

Figure 15. Prediction results of LR. (**a**) Predicted damage probability. (**b**) Predicted damage risk.

Figure 15a proves that LR cannot even identify features related to the power tower because its probability distribution is not similar to the power tower distribution in Figure 7b. Therefore, Figure 15b can only provide information as Figure 12b dose. Because the predicted probability including both true and false predicted results, it is hard to determine how much Figure 12b or Figure 15b reflect the true prediction results. However, according to indicators of the contrast between CART and LR in Table 3, Figure 12b should present more true prediction than Figure 15b.

6.2.2. Combined Model

The predicted damage probability and risk of the combined model are visualized, as shown in Figure 16.
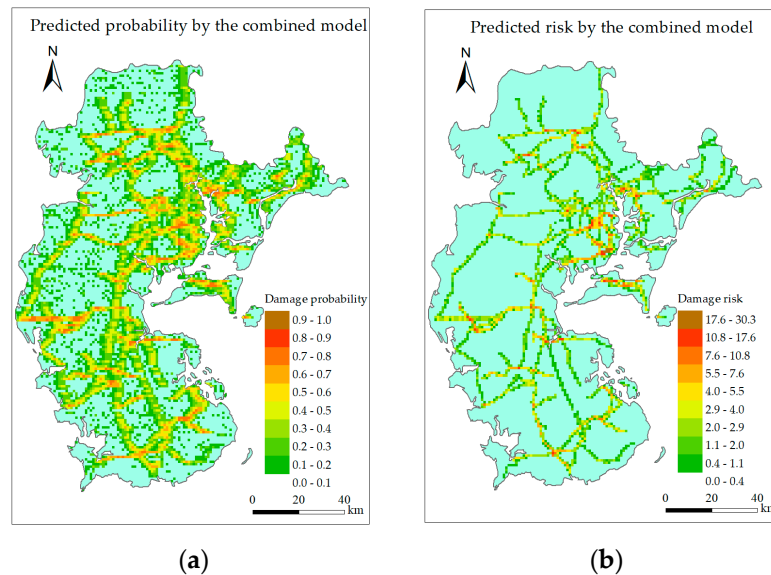

Figure 16. Prediction results of the combined model. (**a**) Predicted damage probability. (**b**) Predicted damage risk.

Figure 16a shows a clear trend that damage probability around landing point and along typhoon track are higher than most other places, which is most consistent with gust distribution in Figure 7a. High-risk grids in Figure 16b are mainly clustered around landing point and along typhoon track, which is also in line with engineering experience.

6.2.3. Discussion
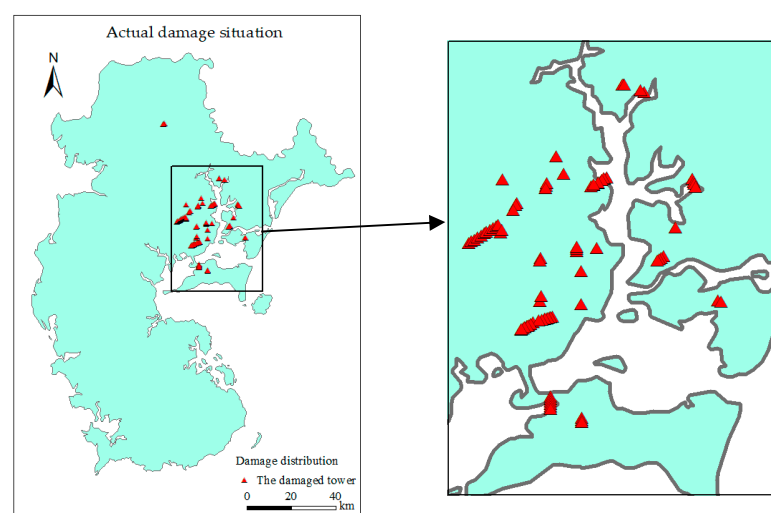
Figure 17 exhibits the actual damage situation.

Figure 17. Actual damage situation.

As shown in Figure 17, actual damage occurred mainly around the landing point, and a few on the typhoon track. Furthermore, the damage around the landing point mainly distributed at the left bank of the bay because water surface will have an acceleration effect on typhoons.

As seen from Figure 10 to Figure 17, only RF, GBRT, and the combined model can provide better prediction. Furthermore, only predicted risk of the combined model is most similar to actual situation. On one hand, predicted risk of the combined model are mainly distributed around the landing point and along the typhoon track as GBRT and RF do, on the other hand, the combined model has lower predicted risk than GBRT and RF at the lower left area, which is better consistent with the actual situation. As shown in Figures 10b and 11b, RF and GBRT overestimated risk at the lower left area. Comparing Figures 16b and 17, it can be seen that the combined model reduces the predicting error at the lower left area and at the same time maintains the accuracy around the landing point.

Therefore, the combined model is superior to the others. For contrast convenience, predicted risk of the combined model around landing point is depicted in detail in Figure 18.
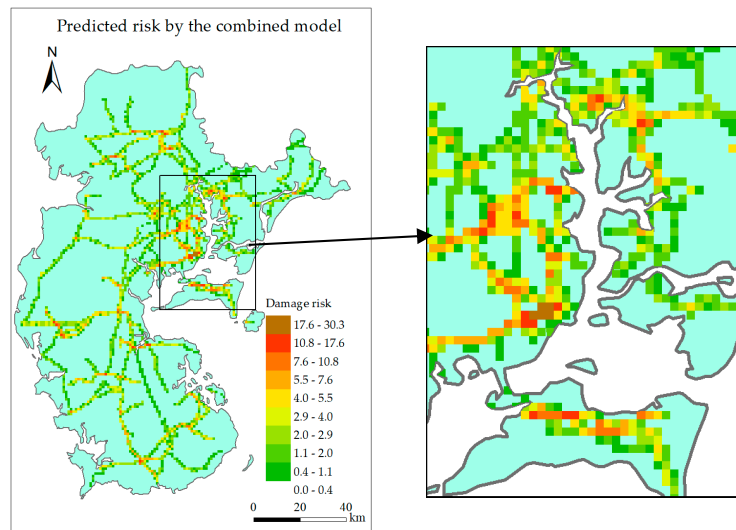

Figure 18. Predicted damage risk of the combined model.

As observed from Figure 10 to Figure 17, predicted damage probability cannot effectively reflect actual damage situation. Because a grid with high damage probability may has low density of power towers, so that the risk will be small; but a grid with relative low damage probability and high density of power towers will have big risk. Therefore, risk is more appropriate to be used in engineering than damage probability.

In order to compare the combined model with the original models quantitively, this paper defines the cosine similarity indicator

$$S(\mathbf{a}, \mathbf{b}) = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \times \|\mathbf{b}\|} = \frac{\sum\limits_{i=1}^{I} a_i b_i}{\sqrt{\sum\limits_{i=1}^{I} a_i^2} \sqrt{\sum\limits_{i=1}^{I} b_i^2}} \tag{20}$$

where $S(\mathbf{a},\mathbf{b})$ denotes the similarity between vector $\mathbf{a}$ and vector $\mathbf{b}$, $\langle \mathbf{a}, \mathbf{b} \rangle$ denotes the inner product of $\mathbf{a}$ and $\mathbf{b}$, $\| \cdot \|$ denotes Frobenius-2 norm of a vector, a_i and b_i are elements in vector $\mathbf{a}$ and $\mathbf{b}$ respectively, $i = 1,2, \ldots ,I$ denotes the sequence numbers of the elements.

First, extract actual damage information from target area to grids and construct an actual risk vector with shape $N \times 1$, where N denotes the amount of grids. Each other model predicts a risk vector with shape $N \times 1$. The vector element is then normalized, and similarity between actual risk vector and predicted risk vectors is calculated. The results are shown in Table 5.

Table 5. Similarity index of each model.

Model	RF	GBRT	CART	SVR	Adaboost	LR	Combined Model
S	0.202 999	0.197 771	0.190 429	0.125 620	0.202 631	0.203 927	0.210 829

As observed in Table 5, the similarity of the combined model is the biggest, indicating that its predicted risk is more similar to the actual damage situation than the others. Therefore, the combined model is superior to the others.

However, as illustrated in Figure 17, there are still some high-risk grids distributed at the lower left area, where the windspeed is not high. This can be explained in two respects: data quantity and data quality. Limited to information collected by the power sector, this paper merely uses information of two historical typhoons to train models, and only eight features are employed in the dataset. Even though oversampling is adopted to enlarge the dataset, it seems to be not enough. Furthermore, due to the low coverage of relevant monitoring equipment at that time, the information collected is inaccurate enough. As a result, there is uncertainty existing in prediction. However, this problem is sure to be solved in the near future because the power sector is addressing the difficulties that hinder information collection.

7. Conclusions

This paper comprehensively considers spatial multi-source heterogeneous information such as meteorological information, equipment operation information, and geographic information, and proposes a method for estimating the tower damage risk under typhoon disaster by using machine learning algorithms and visualization techniques. This method is divided into three layers by function: the data layer, knowledge extraction layer, and visualization layer.

The intelligent models are trained with the historical data of typhoon 'Rammasun' and 'Hato'. By hyperparameter optimization and comparison of MSE, MAE, R^2, $MEAE$, and EVS, it shows that the relative optimal model is RF. After excluding LR, the combined model is established by using the goodness of fit method with unequal weight. Finally, the risk assessment of the main network tower of a Chinese coastal city under typhoon 'Mujigae' is carried out with the accuracy of 1 km × 1 km.

(1) Different machine learning algorithms have different adaptability to data, and LR is not suitable for this study according to Table 3 and Figure 15.

(2) Both the single intelligent model and combined model can identify the high-risk grids, but the combined model can reduce the predicting error at the low gust area while maintain the predicting accuracy around landing point and along typhoon track.

(3) RF, GBRT, and the combined model perform better than other models, but predicted risk distribution of the combined model is the most similar to the actual situation.

(4) The similarity indicator of predicted risk of the combined model is 0.210 829, which is the biggest among all the models, so the combined model is the optimal model.

(5) This study verifies the feasibility and scientificity of the presented method and can provide support for the power sector in disaster prevention and mitigation.

(6) The uncertainty existing in the model should be tackled in the future for higher predicting accuracy, and it is necessary to extend the evaluation object to other power equipment such as transmission lines and transformers, and further raise from the equipment level to the system level.

Author Contributions: The methodology was proposed by H.H., the funding acquisition work was done by Y.H., the investigation was done by H.W. (Hao Wu), the formal analysis work was done by H.W. (Hongbin Wang) and X.L., the validation work was done by Y.X., the data curation work was finished by H.G., and the original draft was finished by S.Y.

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

Appendix A

Table A1. Model hyperparameters and their optimization combination.

Model	Hyperparameters for Optimization	Range	Optimized Hyperparameters
RF	criterion; n_estimators.	mse, mae; 1–100.	mae; 69.
GBRT	learning_rate; loss; n_estimators.	0.01–1.0; ls, lad, huber, quantile; 1–200.	0.203 578; huber; 190.
CART	criterion; presort; splitter.	mse, mae, friedman_mse; 0/1; best, random.	mae; 1; best.
SVR	C; coef0; degree; epsilon; gamma; kernel; shrinking.	0–5; 0–10; 0–10; 0–10; 0–10; linear, poly, rbf, sigmoid; 0/1.	3.894 502; 7.335 600; 3; 0.076 037; 5.350 016; rbf; 0.
Adaboost	learning_rate; loss; n_estimators.	0.01–1; linear, square, exponential; 1–100.	0.176 719; linear; 58
LR	C; penalty.	0–10; l1,l2.	3.765 847; l1.

Appendix B

Appendix B.1. SVR

Set up training set

$$T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\} \tag{B1}$$

where $x_i \in \mathbf{R}^n$, $y_i \in \mathbf{R}$, and *SVR* uses regression function

$$y = f(x) = w \cdot \varphi(x) + b \tag{B2}$$

where $\varphi(x)$ is the nonlinear mapping of the input space to the high-dimensional feature space; $w \in \mathbf{R}^n$ is the weight vector, and $b \in \mathbf{R}$ is the bias, and these two parameters are solved by minimizing the structural risk

$$\min \varphi(w, \xi, \xi^*) = \frac{1}{2} \|w\|^2 + c \sum_{i=1}^{N} (\xi_i + \xi_i^*) \tag{B3}$$

$$s.t \begin{cases} y_i - w \cdot \varphi(x_i) - b \leq \varepsilon + \xi_i, \xi_i \geq 0 \\ w \cdot \varphi(x_i) + b - y_i \leq \varepsilon + \xi_i^*, \xi_i^* \leq 0 \end{cases} \tag{B4}$$

where c is a constant; ξ_i and ξ_i^* are slack variables, and the expression of Equation (B2) can be obtained by solving Equations (B3) and (B4).

Appendix B.2. CART

Assume that x and y are input and output variables, respectively, and y is a continuous variable, given a data set as shown in Equation (B1). A regression tree corresponds to a division of the input

space and an output value on each unit. Assuming that the input space has been divided into M units $R_1, R_2, \ldots, R_M$ and there is a fixed output value c_m on each unit R_m, a regression tree is obtained

$$f(x) = \sum_{m=1}^{M} c_m I(x \in R_m) \tag{B5}$$

Solving the optimal output on each unit with the squared error minimum criterion $\sum_{x_i \in R_m}^{M} (y_i - f(x_i))^2$, the optimal value $\hat{c}_m$ of c_m on unit R_m is the mean of the output y_i of all input instances x_i on R_m

$$\hat{c}_m = ave(y_i | x_i \in R_m) \tag{B6}$$

when dividing the input space, assume that $x^{(j)}$ (the variable j) and s (the value of variable j) are the splitting variable and the splitting point, and define two regions

$$R_1(j,s) = \left\{ x \middle| x^{(j)} \le s \right\}, \ R_2(j,s) = \left\{ x \middle| x^{(j)} > s \right\} \tag{B7}$$

then solve j and s

$$\min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right] \tag{B8}$$

Iterate through all the input variables to find the optimal segmentation variable, then form a pair (j,s) that divides the input space into two regions. Repeat the above process for each subregion until the stop condition is met, and finally generate a regression tree.

Appendix B.3. Adaboost

Taking the classification problem as an example, given a dichotomous training data set as shown in Equation (B1). Where $x_i \in \chi \subseteq \mathbf{R}^n$, $y_i \in \gamma = \{-1, +1\}$, x and γ are input and output spaces respectively. First, the weight distribution of the training data is initialized

$$D_1 = (w_{11}, \ldots, w_{1i}, \ldots, w_{1N}), \ w_{1i} = \frac{1}{N} \tag{B9}$$

where $i = 1, 2, \ldots, N$. Assuming that $m = 1, 2, \ldots, M$, basic classifiers are obtained by learning from training data set with weight distribution D_m

$$G_m(x) : \chi \rightarrow \{-1, +1\} \tag{B10}$$

Assign a coefficient to each classifier

$$\alpha_m = \frac{1}{2} \ln \frac{1 - e_m}{e_m} \tag{B11}$$

where e_m is the classification error rate. Then update the weight distribution of the training set

$$D_{m+1} = (w_{m+1,1}, \ldots, w_{m+1,i}, \ldots, w_{m+1,N}) \tag{B12}$$

Finally, construct a linear combination of basic classifiers

$$f(x) = \sum_{m=1}^{M} \alpha_m G_m(x) \tag{B13}$$

Appendix B.4. GBRT

Set the training data set as shown in Equation (B1), where $x_i \in \chi \subseteq \mathbf{R}^n$, $y_i \in \gamma \subseteq \mathbf{R}$, the loss function is $L(y, f(x))$. First, initialize

$$f_0(x) = \underset{c}{\arg\min} \sum_{i=1}^{N} L(y_i, c) \tag{B14}$$

For $m = 1, 2, \ldots, M$, $i = 1, 2, \ldots, N$, calculate

$$r_{mi} = -\left[\frac{\partial L(y, f(x_i))}{\partial f(x_i)}\right]_{f(x) = f_{m-1}(x)} \tag{B15}$$

Fit a regression tree for r_{mi} to get the leaf node area of tree m. For $j = 1, 2, \ldots, J$, calculate

$$c_{mj} = \underset{c}{\arg\min} \sum_{x_i \in R_{mj}}^{N} L(y_i, f_{m-1}(x_i) + c) \tag{B16}$$

Update $f_m(x) = f_{m-1}(x) + \sum_{j=1}^{J} c_{mj} I(x \in R_{mj})$, get the regression tree

$$\hat{f}(x) = f_M(x) = \sum_{m=1}^{M} \sum_{j=1}^{J} c_{mj} I(x \in R_{mj}) \tag{B17}$$

References

1. Zhou, X.M.; Ge, S.Y.; Li, T.; Liu, H. Assessing and Boosting Resilience of Distribution System Under Extreme Weather. *Chin. Soc. Electron. Eng.* **2018**, *38*, 505–513. [CrossRef]
2. Yin, C.X.; Tang, W.Q.; Wen, L.F.; Lan, J.J.; Lin, J.H.; Li, S.Y.; Li, N. A new method for reliability evaluation of distribution network considering the influence of typhoon. *Power Syst. Prot. Control* **2018**, *46*, 138–143. [CrossRef]
3. Song, X.Z.; Wang, Z.; Xin, H.H.; Gan, D.F. Risk-based dynamic security assessment under typhoon weather for power transmission system. In Proceedings of the IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Kowloon, China, 8–11 December 2013.
4. Gao, W.; Zhou, R.; Zhao, D. Heuristic failure prediction model of transmission line under natural disasters. *IET Gener. Transm. Distrib.* **2017**, *11*, 935–942. [CrossRef]
5. Huang, Z.; Wang, F.; Tan, Y.H.; Dong, X.Z.; Wu, Z.R.; Chen, C. Operational risk assessment based on health and importance indexes for distribution network. *Electron. Power Autom. Equip.* **2016**, *36*, 136–141. [CrossRef]
6. Yang, L. Operational Statistic Risk Assessment for Power Systems Under Extreme Weather Conditions. In Proceedings of the Advances in Power System Control, Operation & Management (APSCOM 2015), 10th International Conference, Hong Kong, China, 8–12 November 2015.
7. Yang, Y.H.; Xin, Y.L.; Zhou, J.J.; Tang, W.H.; Li, B. Failure probability estimation of transmission lines during typhoon based on tropical cyclone wind model and component vulnerability model. In Proceedings of the 2017 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Bangalore, India, 8–10 November 2017.
8. Han, L. *Construction and Application of Emergency Management System for Power Network Disaster Prevention and Mitigation*, 1rd ed.; National Defense Industry Press: Beijing, China, 2009.
9. Knowing Icing Early: Hunan Power Grid Disaster Prevention and Mitigation Technology Is at the International Leading Level. Available online: http://hn.rednet.cn/c/2015/01/15/3577523.htm (accessed on 8 June 2018).
10. Wu, C.P.; Lu, J.Z.; Liu, Y.; Yang, L. Technology of field fire prevention near transmission lines and its application in Hunan power grid. *Hunan Electron. Power* **2014**, *34*, 28–30.

11. Wang, J.; Shi, H.Y.; Fang, J.L. Discussion on GIS-based monitoring and early warning system for meteorological disasters in power grids. *Autom. Appl.* **2017**, *5*, 92–93.

12. Liu, H.B.; Rachel, A.D.; David, V.R.; Jery, S. Negative binomial regression of electric power outages in hurricanes. *J. Infrastruct. Syst.* **2005**, *11*, 258–267. [CrossRef]

13. Liu, H.B.; Rachel, A.D.; Tatiyana, V.A. Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms. *Reliab. Eng. Syst. Saf.* **2008**, *93*, 897–912. [CrossRef]

14. Mensah, A.F.; Duenas-Osorio, L. Efficient Resilience Assessment Framework for Electric Power Systems Affected by Hurricane Events. *J. Struct. Eng.* **2016**, *142*, C4015013. [CrossRef]

15. Han, S.; Guikema, S.D.; Quiring, S.M.; Lee, K.; Rosowsky, D.; Davidson, R.A. Estimating the spatial distribution of power outages during hurricanes in the Gulf Coast region. *Reliab. Eng. Syst. Saf.* **2009**, *94*, 199–210. [CrossRef]

16. Han, S.; Guikema, S.D.; Quiring, S.M. Improving the Predictive Accuracy of Hurricane Power Outage Forecasts Using Generalized Additive Models. *Risk Anal.* **2009**, *29*, 1443–1453. [CrossRef] [PubMed]

17. Sun, J.B.; Xin, T.; Wang, Y.W. Experience of Guangdong Power Grid Resisting Super Typhoon "Rammasun" and Introspection. *Guangdong Electron. Power* **2014**, *27*, 80–83. [CrossRef]

18. Lyu, H.; Chen, Y.L.; Tang, Y.; Wang, L.H.; He, F.S.; Xie, W.P. Reason Analysis on Disconnector Damage After Typhoon "Rainbow". *Guangdong Electron. Power* **2015**, *28*, 138–143. [CrossRef]

19. Zhang, Z.Q. Application of double-strength main material in the reinforcement project of power tower in Zhuhai under typhoon "Hato". *Mech. Electron. Inf.* **2017**, 51–53. [CrossRef]

20. *Summary Report on the Impact of Typhoon "Mujigae" on Transmission Lines*; Electric Power Research Institute, Guangdong Power Grid Co., Ltd.: Guangdong, China, 2015.

21. *Technical Specifications for Windproof Design of Distribution Lines of China Southern Power Grid Corporation*; Q/CSG 1201012-2016; China Southern Power Grid Co., Ltd.: Guangzhou, China, 2016.

22. Qu, C.X.; Jiang, Y.; Wu, Y.P.; Zou, X.N.; Qiao, Y.L. Study on the development of a choropleth atlas on cancer mortality using the inverse distance weight interpolation in the 1990's. *Chin. J. Epidemiol.* **2006**, *27*, 230–233.

23. Li, X.; Cheng, G.D.; Lu, L. Comparison of spatial interpolation methods. *Adv. Earth Sci.* **2000**, *15*, 260–265.

24. Li, H. *Statistical Learning Method*, 1st ed.; Tsinghua University Press: Beijing, China, 2012.

25. Vapnik, V. *Statistic Learning Theory*, 1st ed.; New York Wiley: New York, NY, USA, 1998.

26. Leo, B.; Jerom, H.F.; Richard, A.O.; Charles, J.S. *Classification and Regression Trees*, 1st ed.; Chapman & Hall: New York, NY, USA, 1984.

27. Zhao, Z.Y. *Python Machine Learning Algorithm*, 1st ed.; Publishing House of Electronics Industry: Beijing, China, 2017.

28. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to Boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [CrossRef]

29. Freund, Y.; Schapire, R.E. Experiments with a new Boosting algorithm. In Proceedings of the 13th Conference on Machine Learning, Bari, Italy, 3–6 July 1996.

30. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]

31. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

32. Wang, L.D.; Li, Z.Y.; Wen, J.Y. Short-term load fast forecasting based on support vector regression. *RELAY* **2005**, *33*, 17–20, 49. [CrossRef]

33. Zhou, Z.H. *Machine Learning*, 1st ed.; Tsinghua University Press: Beijing, China, 2016.

34. Scikit-Learn. Available online: http://sci-kit-learn.org/stable/index.html (accessed on 18 June 2018).

35. Li, W.Y. *Risk Assessment of Power System (Models, Methods and Applications)*, 1st ed.; Science Press: Beijing, China, 2006.

36. Wei, Z.J.; Ma, H.L.; Tang, D.L. Trend Assessment of Typhoon Disasters Based on the Improved Entropy Method. *J. Catastrophol.* **2017**, *32*, 7–11. [CrossRef]