IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 14, 2020, accepted April 24, 2020, date of publication May 6, 2020, date of current version May 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2992528

A Hybrid Prediction Model for Damage Warning of Power Transmission Line Under Typhoon Disaster

HUI HOU[1], (Member, IEEE), SHIWEN YU[1], HAO WANG[2], (Member, IEEE), YAN XU[3], (Senior Member, IEEE), XIANG XIAO[4], YONG HUANG[4], AND XIXIU WU[1]

[1]Electrical Engineering Department, Wuhan University of Technology, Wuhan 430070, China
[2]Department of Data Science and Artificial Intelligence, Monash University, Melbourne, VIC 3800, Australia
[3]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798
[4]Electric Power Research Institute, Guangdong Power Grid Company Ltd., Guangzhou 510080, China

Corresponding authors: Shiwen Yu (yswwhut@163.com) and Xixiu Wu (wuxixiu@163.com)

This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant WUT: 191011005, and in part by the Guangdong Power Grid Company Ltd., through the Science and Technology Project under Grant GDKJXM20198441.

ABSTRACT To bolster the resilience of power systems against typhoon disasters, this paper develops a holistic framework of wind disaster warning for transmission lines. This paper proposes a hybrid prediction model to quantify the transmission line damage probability under typhoon disaster based on extreme value type I probability distribution, Monte Carlo method, and Random Forest. Specifically, this paper uses the extreme value type I probability distribution and the Monte Carlo method to simulate the random wind field, and predict the damage probability of transmission lines under each wind field using the Random Forest method. This paper takes typhoon ''Mangkhut'' in 2018 as a case study, and compare the performance of the hybrid model based on random wind field with the Random Forest method under predicted and measured wind field. The results demonstrate that the hybrid model can effectively utilize wind speed data to obtain a more reliable prediction and achieves the best synthetic similarity to the actual damage situations.

INDEX TERMS Typhoon, power system resilience, transmission line, extreme value type I probability distribution, Monte Carlo, Random Forest, random wind field.

I. INTRODUCTION

Typhoon disaster, as one of the most catastrophic natural disasters, brings severe hazards like storm and flooding and may cause unprecedented damages to the power grid in coastal areas, including tower tilt and falling, power transmission lines breaking, and even wide-scale power outage [1], [2]. Therefore, it is urgent to establish an effective wind disaster warning system for transmission line damage under typhoon to provide decision support for dispatch, operation, maintenance, and overhaul.

However, the damage mechanism of power grid under typhoon is complex and requires further study. For example, the failure of power equipment is caused by various factors including project construction quality, physical and chemical environment, and extreme weather. Some of the factors are difficult to quantify and thus bring challenges to the cause analysis and damage prediction modeling. Therefore, this paper aims to develop an effective method for damage warning of power transmission lines under typhoon disaster. In particular, this paper focuses on the transmission line damage caused by tower tilt and falling in this paper.

There has been a body of literature [2], [3] that explored the correlation between transmission line failure and the wind speed. Yang *et al.* in [3] utilized a tropical cyclone wind model and fragility curve to simulate the failure probability of transmission lines. An exponential function in [4] was constructed to describe the relationship between the transmission line segment failure rate and wind speed, and thus further established contingency probability models. In [5], fragility curve was used to illustrate the relationship between the failure probability and the wind speed. However, the wind speed probability distribution was ignored and the transmission line failure model under typhoon disaster was oversimplified. Other studies in [6]– [8] proposed failure probability models using the structural stress theory. Geng *et al.* in [6] computed

The associate editor coordinating the review of this manuscript and approving it for publication was Shunfeng Cheng.

the failure probability of transmission lines based on the stress intensity interference model. Yang *et al.* in [2] established the logarithmic normal distribution model of tower damage by using Yan Meng wind field model [7], [8] and structural stress theory. However, models using the structural stress theory have high computational overhead and thus cannot be used in large-scale studies. Moreover, many key information such as microtopographic information and operation time were not captured in the structural-stress-theory-based models.

Recent studies in [9]– [11] took multi-source information into consideration to establish transmission line failure models. An early warning model for transmission line galloping with input of weather conditions, the conductor type, cross-section and span of transmission line was constructed based on support vector machine and Adaboost [9]. Gao *et al.* in [10] established a transmission line failure prediction model based on the microclimate prediction and historical tripping data. But the microclimate prediction substantially increased the computational complexity. The study in [11] improved the wind load model of power tower with the Weibull distribution aging curve based on the batts model, but the analysis of microtopography is insufficient. However, the destruction of the typhoon will lead to dynamic changes of the power grid structure, which leads to the changes of the weighted betweenness and other indicators as well.

The transmission line failure models have also been inspired by the studies of outage prediction. For example, Nateghi *et al.* in [12] found that the most critical variables were the wind characteristics of the storms and the climatic and geographic characteristics of the service area in outage prediction. Guikema *et al.* in [13] used the Random Forest method and Bayesian mass-balance multiscale model to build a high-precision outage prediction model. However, it did not consider the influence of microtopography. Reference [14] constructed an ensemble decision tree regression consisting of decision tree, Random Forest and boosted gradient tree, and achieved a higher accuracy than other models in predicting outages.

The aforementioned studies have demonstrated that multi-source information and hybrid methodology are promising in predicting transmission line failure. In our previous work [15], a combined model based on the equipment operation information, meteorological information, and geographic information was established, but it only utilized measured maximum gust without considering the uncertainty of predicted wind speed. This paper develops a holistic framework for wind disaster warning of the transmission line under typhoon disaster. We set up a hybrid prediction model for transmission line damage probability calculation based on extreme value type I probability distribution, Monte Carlo method and Random Forest method. The former model in [15] only utilized certain maximum gust, which was not applicable for scenario of predicted gust that contained occurrence probability. Thus, in this paper, the extreme value type I probability distribution is employed to calculate the occurrence probability of a certain gust based on the hypothesis that the maximum gusts at one point follow extreme value type I probability distribution [16]. However, the predicted maximum gust at one point is time-varying. Therefore, the Monte Carlo method [17] is employed to carry out stochastic simulation.

This enables the comprehensiveness, efficiency and applicability in the warning of typhoon "Mangkhut" in Guangdong power grid, China. The main contributions of this paper are as follows:

1) This paper develops a hybrid method to predict the damage probability of transmission line under typhoon disaster, which can provide decision support for the disaster prevention and mitigation;

2) The extreme value type I probability distribution and Monte Carlo method are used to simulate the wind field, and the Random Forest is used to predict the damage probability of transmission lines.

The rest of the paper is organized as follows: Section II briefly introduces the technical framework for wind disaster warning of transmission line damage. Section III describes the construction of modules in the technical framework. In Section IV, results are demonstrated through comparison with single model Random Forest (RF) method under predicted wind field and measured wind field. We conclude the paper in Section V.

II. TECHNICAL FRAMEWORK

To prevent and control the transmission line damage caused by a typhoon, this paper constructs a technical framework of wind disaster warning for transmission lines. It consists of three functional modules: information acquisition module, information processing module, and damage warning and evaluation module. The framework is depicted in Fig. 1.

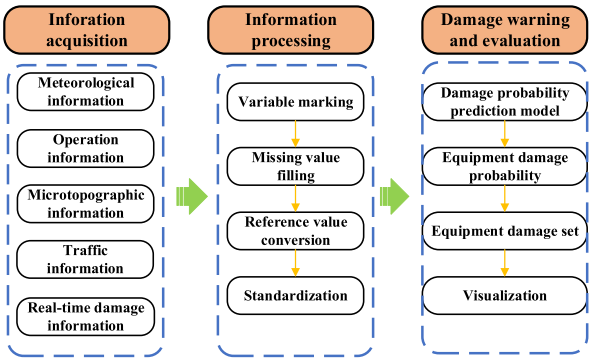


FIGURE 1. The technical framework of wind disaster warning of transmission line.

III. CONSTRUCTION OF MODULES

A. CONSTRUCTION OF INFORMATION ACQUISITION MODULE

The information acquisition module is designed to acquire the information needed for damage warning of transmission lines, including meteorological information, equipment operation information, microtopography information, traffic

information, real-time damage information, etc. The Information in Fig. 1 are provided by electric power research institute, Guangdong power grid Co., Ltd. The wind speed data have missing values and thus cannot be used directly in the analysis. Therefore, this paper uses ArcGIS10.4.1 [18] to collect and process information. For instance, we use ArcGIS10.4.1 to perform inverse distance weight interpolation for gust and microtopography, and relevant data are extracted to the equipment coordinate. Note that ArcGIS10.4.1 is a geographic information system created by ESRI (Environmental Systems Research Institute, Inc.), of which 10.4.1 is the widely used version. Though the transmission system is composed of the transmission tower, insulator, fittings, guide line, and ground wire, this paper mainly considers the transmission line damage involving tower tilt and break.

B. CONSTRUCTION OF INFORMATION PROCESSING MODULE

The information processing module is used for variable marking, missing value filling, reference value conversion, and standardization. Variable marking is to flag damage status. Missing values are filled using the median. Then the reference value conversion is used to convert gust wind speed and tower design wind speed to 10 meters height.

According to the technical specification [16], the variation of wind speed along with the height obeys the exponential law shown as follows,

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

where V_z (m/s) is the wind speed at height z, V_1 (m/s) is the wind speed at height z_1, α is the surface roughness coefficient, and α can be selected using Table 1 according to [16].

TABLE 1. Surface roughness coefficient.

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 high houses	0.28-0.44

Meteorological stations are generally located in open plains that the surface roughness is generally class B.

Normalize all the data x as

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

where x^* is the standardized feature, x is the original feature, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the original feature.

C. CONSTRUCTION OF DAMAGE WARNING AND EVALUATION MODULE

The damage warning and evaluation module establishes the hybrid prediction model of transmission line damage probability based on the extreme value type I probability distribution, Monte Carlo method, and Random Forest method.

In the reliability study, the extreme value type I probability distribution is more conservative, which is helpful to make more reliable prevention and mitigation strategies [19]. It is demonstrated that amongst various machine learning techniques such as the Random Forest (RF) method, Adaboost, Classification and Regression Tree (CART), Gradient Boosted Regression Tree (GBRT), Support Vector Machine (SVM) and Logistic Regression (LR), the RF method often achieves better performance in predicting damage caused by typhoon [15]. The Monte Carlo method is one of the most accurate and effective method for solving structural reliability problems [20]. Therefore, the extreme value type I probability distribution, RF method and Monte Carlo method are selected to construct the hybrid prediction model. However, the traditional RF method was utilized alone in [15] and the wind speed was used only once, which ignores the occurrence probability of the wind speed. Therefore, their results are heuristic and inaccurate. This paper uses the random wind field generated by the Monte Carlo method and extreme value type I probability distribution to improve the prediction of RF. The hybrid model in this paper utilizes random wind samples from the extreme value type I probability distribution to get a convergent prediction that achieves higher accuracy.

To be more specific, the extreme value type I probability distribution and Monte Carlo method are used to simulate the wind field. The RF method is used to predict the damage probability of transmission line under each wind field. This hybrid method takes uncertainty of predicted wind speed into account and realizes weighted probability output based on each prediction of RF method.

1) EXTREME VALUE TYPE I PROBABILITY DISTRIBUTION

The extreme value type I distribution (also known as the Gumbel distribution) [21] is

$$F(x) = e^{-e^{-a(x-u)}}, \tag{3}$$

where a is the scale parameter, u is the location parameter, and $a > 0$, $-\infty < u < +\infty$. The probability density distribution function is

$$f(x) = a \cdot e^{-a(x-u)} \cdot e^{-e^{-a(x-u)}}. \tag{4}$$

The extreme value type I probability distribution and density function are depicted in Fig. 2 and Fig. 3, respectively.

There are three commonly used parameter estimation methods: the method of moments, the Gumbel's method and the maximum likelihood method. It is proved that when the sample size is small, Gumbel's method is less effective [22].

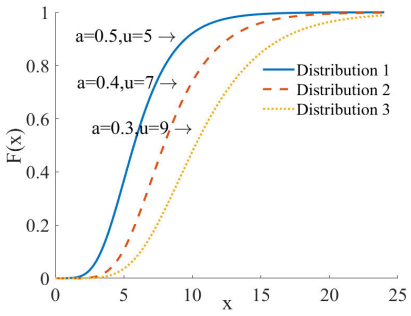


FIGURE 2. Extreme value type I probability distribution function.

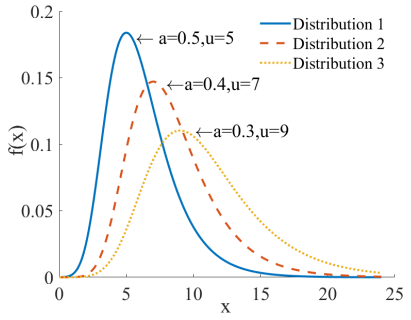


FIGURE 3. Extreme value type I probability density function.

The forecasting data of wind speed in this paper is provided as a matrix of 800 longitudes multiplying 600 latitudes and each longitude-latitude pair is a wind speed point. Since the wind speed series at each wind speed point is just 24, this paper doesn't use Gumbel's method. As for accuracy, the method of moments, Gumbel's method, and maximum likelihood method all have satisfactory parameter estimation results for the extreme type I distribution [22]. However, the maximum likelihood method and Gumbel's method are more computationally expensive than the method of moments. Given the advantages in both accuracy and calculation, this paper chooses the method of moments [23] to estimate scale parameters and position parameters.

The method of moments estimation is based on the law of large numbers, and the sample means converge to the distributional mean as the number of observations increase.

Independent random variables $X_1, X_2, \ldots$ chosen according to the probability distribution derived from the parameter value θ and m are a real valued function, if $k(\theta) = E_\theta m(X_1)$, then

$$\frac{1}{n}\sum_{i=1}^{n} m(X_i) \to k(\theta) \quad \text{as} \quad n \to \infty. \tag{5}$$

The method of moments results from the choices $m(x) = x^m$. Write

$$\mu_m = EX^m = k_m(\theta). \tag{6}$$

for the m-th moment.

The first moment (mathematical expectation) of the extreme value type I distribution is:

$$E(x) = \frac{y}{a} + u, \tag{7}$$

where $y \approx 0.57722$, the second moment (variance) is:

$$\sigma^2 = \frac{\pi^2}{6a^2}. \tag{8}$$

Hence, obtain

$$a = \frac{1.28255}{\sigma}. \tag{9}$$

$$u = E(x) - \frac{0.57722}{a}. \tag{10}$$

According to (3), the relationship between wind speed x_R and recurrence period R is:

$$x_R = u - \frac{1}{a}\ln\left[\ln\left(\frac{R}{R-1}\right)\right]. \tag{11}$$

The occurrence probability of the maximum value x_P once in N years is the guaranteed rate [16]:

$$P_N = P(x > x_P) = 1 - P(x \le x_P) = 1 - F(x_p), \tag{12}$$

where P_N is the guaranteed rate.

This paper employs the Kolmogorov-Smirnov test [24] to validate the accuracy of the method of moments for its effectiveness on samples with small sizes.

For random variable x, construct a hypothesis test:

$$H_0 : F(x) = F_0(x), \tag{13}$$

where $F(x)$ is a distribution function and $F_0(x)$ is the completely determined continuous distribution function.

Then construct the sample distribution function as

$$F_n(x) = \begin{cases} 0 & x < X_{(1)} \\ \dfrac{i}{n} & X_{(i)} < x < X_{(i+1)}, i = 1, 2, \ldots, n \\ 1 & x \ge X_{(n)} \end{cases} \tag{14}$$

where $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$.

Set the test statistic

$$D_n = \sup_{-\infty < x < +\infty} |F_n(x) - F_0(x)|. \tag{15}$$

The supremum of D_n can be found at

$$d_i = \max\left\{\left|F_0(X_{(i)}) - \frac{i-1}{n}\right|, \left|\frac{i}{n} - F_0(X_{(i)})\right|\right\}, \quad i = 1, 2, \ldots, n, \tag{16}$$

where d_i is the maximum variance between the sample distribution and the completely determined continuous distribution function of $X_{(i)}$.

Therefore, the test statistic is written as

$$D_n = \max\{d_1, d_2, \cdots, d_n\}. \tag{17}$$

If F_0 and F_n fit well, D_n should be small, otherwise D_n is large. Specifically, the test rule is set under a significance level α: if $D_n > D_{\alpha,n}$, reject H_0; otherwise, accept it.

Calculate the rejection rate at level α of all wind speed points to validate the accuracy of the method of moments as

$$RR(\alpha) = \frac{num_{\alpha}(\text{rejected points})}{num(\text{wind speed points})}, \tag{18}$$

where $RR(\alpha)$ is the rejection rate at significance α, num_{α} (rejected points) is the number of rejected points under level α, num (wind speed points) is the account of wind speed points.

2) MONTE CARLO METHOD

The techniques used in Monte Carlo simulations can be basically classified into two categories known as sequential and non-sequential [17], [20] techniques. Because non-sequential methods are typically easier to implement and require much less CPU resource and memory as compared to sequential methods [25], non-sequential Monte Carlo method is used in this paper.

As for sample generation, this paper utilizes direct simulation [26] because the extreme value type I probability distribution is specific function.

A system state is a combination of all component states, and each component state can be sampled by the probability that the component appears in that state. When the components are independent, this can be achieved simply by sampling states of individual components to construct system states [25]. The tower destruction mainly due to the wind speed exceeding the design wind speed, so it is reasonable to neglect the interactions between towers and assume their damages are independent. And, since wind speed points are matched to towers as introduced in 4) for damage calculation, assuming that wind speed points are independent.

In this paper, each wind speed point is equivalent to a component of the system and the component state conforms to the extreme value type I probability distribution. All the wind speed points amount to the system. The set of generated wind speeds at all wind speed points, namely the whole wind field, represents the system state.

3) RANDOM FOREST METHOD

Random Forest [27] is a statistical learning method based on bagging [28], and its basic strategy is to adopt bootstrap sampling. Given a dataset $\mathbf{D}$, L sampling subsets containing l training samples are randomly selected. Then a base learner is trained based on each sampling subset, and finally all the base learners are combined. Bagging generally adopts simple voting method for the classification and simple average method for regression.

RF is an ensemble method consisting of the decision tree by bagging [29]. It not only uses bootstrap to select training samples randomly, but also selects features randomly when partitioning a decision tree. In general, traditional decision tree selects an optimal feature when splitting. In RF, for each node of the decision tree, a subset containing m features is randomly selected from the feature set of the node, and then an optimal feature is selected from each subset for partitioning. Therefore, the diversity of base learners of RF comes not only from sample perturbation, but also from feature perturbation. The generalization performance of the final integration can be improved by increasing the difference between individual learners [30].

4) TRANSMISSION LINE DAMAGE PROBABILITY HYBRID PREDICTION MODEL

In order to calculate the transmission line damage probability more accurately, this paper proposes a hybrid prediction model of transmission line damage probability.

Firstly, the model utilizes the extreme value type I probability distribution to simulate the predicted wind field distribution. The interval of wind speed point can reach 1km $\times$ 1km. Since the predicted wind speed is 10-minute-averaged wind speed at 10 meters height, the maximum gust wind speed at 10 meters height can be obtained by multiplying the gust coefficient. The gust coefficient is the parameter converted from average wind speed to instantaneous wind speed, which is generally defined as the ratio between the gust wind speed and the average wind speed at a time interval of 10min [31]. Assuming that the gust distribution at each wind speed point obeys the extreme value type I probability distribution. Based on (5)-(10), the forecast 24-hour wind speeds of each point are used for parameter estimation. Each wind speed point is fitted with an extreme value type I distribution, to realize the simulation of predicted wind field distribution.

Although the forecast wind speeds data adopted by the power sector are abundant, the engineering experience shows that the prediction data within 24 hours before the typhoon landfall achieve the best accuracy. Therefore, this paper uses the hourly forecast data within 24 hours before the typhoon landfall to fit the extreme value type I distribution. Then, according to the geographical distance, the maximum value type I distribution is matched for each tower, and each tower is taken as a new wind speed point. Assuming that each wind speed series is a vector of 24 elements, which contains hourly forecasting data of 24 hours before the typhoon makes landfall. Use (7)-(10) to calculate scale parameters and position parameters of each wind speed series.

The wind speed series is matched for each tower according to the distance restriction:

$$|X_1 - X_2| + |Y_1 - Y_2| \leq 0.01, \tag{19}$$

where (X_1, Y_1) is the geographic coordinate of a wind speed series, and (X_2, Y_2) is the geographic coordinate of a certain power tower. If (X_1, Y_1) and (X_2, Y_2) satisfy (19), the wind speed point is matched with this tower and the coordinate of the tower is taken as a new wind speed point. Going through all the wind speed points, a wind speed point satisfying the restriction will be obtained. This completes the selecting procedure of wind speed samples.

Secondly, the probability generation of random wind field was realized based on Monte Carlo method. At each wind speed point, M wind speed w_{ij} are generated randomly according to uniform distribution, where i ($i = 1, 2, \ldots, N$)

is the sequence of wind speed points, and j ($j = 1, 2, \ldots, \mathrm{M}$) is the sequence of random samples. At the same time, the probability of each sample is calculated by using the well-fitted extreme value type I probability distribution.

Finally, in each random wind field, use RF method to calculate the damage probability $f\left(x_{\mathrm{i}}|w_{\mathrm{i}} = w_{\mathrm{ij}}\right)$ of the tower, where x_{i} is the feature vector at the wind speed point i, and w_{i} is the wind speed component. According to the Monte Carlo method, the tower damage probability at the wind speed point is equivalent to the average of M predicted results. The tower damage probability at each wind speed point is P_i.

To obtain a stable and reliable prediction result, define:

$$P_i = \frac{1}{\mathrm{M}} \sum_{j=1}^{\mathrm{M}} P\left(w_{ij}\right) f\left(x_i \left| w_i = w_{ij}\right.\right). \tag{20}$$

The deviation of P_i at iteration M is defined as

$$\varepsilon_{\mathrm{M}} = \frac{1}{\mathrm{M}} \sum_{j=1}^{\mathrm{M}} P\left(w_{ij}\right) f\left(x_i \left| w_i = w_{ij}\right.\right) - \frac{1}{\mathrm{M\text{-}1}} \sum_{j=1}^{\mathrm{M\text{-}1}} P\left(w_{ij}\right) f\left(x_i \left| w_i = w_{ij}\right.\right), \tag{21}$$

where ε_{M} is deviation of P_i between the iterations M and $\mathrm{M} - 1$, and $\mathrm{M} > 1$.

The convergence restriction in this paper is defined as

$$|\varepsilon_{\mathrm{M}}| < 0.01, \tag{22}$$

and M is determined according to (22).

It shows that the hybrid prediction model of damage probability considers the uncertainty of wind speed prediction. Damage probability is evaluated based on the confidence of wind speed occurrence $P(w_{\mathrm{ij}})$, and Monte Carlo method is used for multiple simulations, to consider various possible wind fields. Therefore, the hybrid prediction model of damage probability is more effective than the single model RF method. In addition, since the wind speed point has been matched to the tower coordinates, the damage probability at the wind speed point represents the transmission line damage probability.

D. SIMILARITY ANALYSIS

In order to numerically compare the prediction results of different models, this paper introduces four similarity indicators: distribution similarity, magnitude similarity, cumulative similarity, and synthetic similarity defined as follows.

1) DISTRIBUTION SIMILARITY

Distribution similarity indicates how close the geographic distribution of predicted damage is to the actual damage. This paper uses equation (20) in [15] to calculate the distribution similarity between the predicted result and the actual damage situation. The larger the distribution similarity is, the closer the predicted damage distribution to the real situation.

2) MAGNITUDE SIMILARITY

The higher the damage probability is, the more likely the damage is to occur. To formulate an accurate disaster prevention and mitigation strategy, it is necessary to set a threshold for the damage probability and take precautions actions. In other words, when the predicted damage probability is greater than the threshold, the damage is considered to occur, and no damage is likely to occur when the predicted value is less than the threshold.

To analyze the similarity between the predicted damage magnitude under different probability thresholds and the actual damage magnitude, the magnitude similarity of the prediction model i under probability threshold P^j is defined as:

$$MS_i\left(P^j\right) = \frac{1}{\left|M - M_i\left(P^j\right)\right|}, \tag{23}$$

where P^j denotes the probability threshold of level j, M denotes the actual damage magnitude, $M_i\left(P^j\right)$ denotes the predicted damage magnitude of model i under P^j, and $MS_i\left(P^j\right)$ is the damage magnitude similarity of model i under P^j. The larger $MS_i\left(P^j\right)$ is, the closer the predicted magnitude to the actual situation.

3) CUMULATIVE SIMILARITY

To evaluate the robustness to probability threshold, the cumulative similarity is defined as:

$$CS_i = \sum_{j=1}^{l} MS_i\left(P^j\right), \tag{24}$$

where l is the number of probability threshold levels, and CS_i is the cumulative similarity of model i. The cumulative similarity reflects the ability of models to maintain the prediction accuracy under different probability thresholds. The larger CS_i is, the more robust the model to the probability threshold.

4) SYNTHETIC SIMILARITY

Damage prediction needs to focus on both the distribution and magnitude of damage so that disaster prevention and mitigation decisions can be accurate without causing underestimation or overestimation. Therefore, the synthetic similarity is defined as the arithmetic mean of the distribution similarity and cumulative similarity:

$$SS_i = \frac{DS_i + CS_i}{2}, \tag{25}$$

where SS_i denotes the synthetic similarity of model i, and DS_i denotes the distribution similarity of model i. The higher the SS value is, the closer the predicted damage distribution and magnitude to the actual damage situation.

IV. CASE STUDY FOR TYPHOON MANGKHUT

A. OVERVIEW OF TYPHOON "MANGKHUT"

At 5 pm on Sept. 16, 2018, tropical cyclone "Mangkhut" (strong typhoon) made landfall in Haiyan town, Taishan of

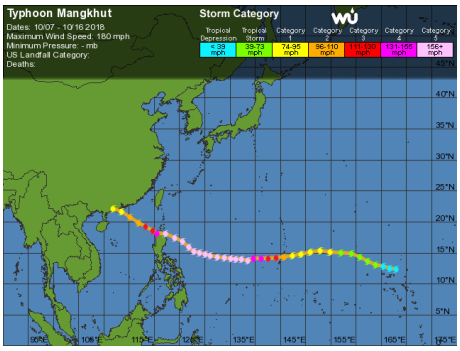


FIGURE 4. The path of typhoon "Mangkhut" [32].

Guangdong province. The maximum wind near the center was class 14 (45m/s, equivalent to 162km/h), and the minimum pressure at the center was 955hPa. Its moving path is shown in Fig. 4.

B. APPLICATION OF THE TECHNICAL FRAMEWORK

1) INFORMATION ACQUISITION MODULE

For the typhoon ''Mangkhut'', information acquired include hourly forecast gusts within 24 hours before typhoon landfall, the maximum gust measured at each of 1,340 monitoring points, and microtopographic information. In order to visually display the distribution of predicted gusts and measured gusts, Fig. 5 and Fig. 6 show the predicted and measured maximum gust distribution.

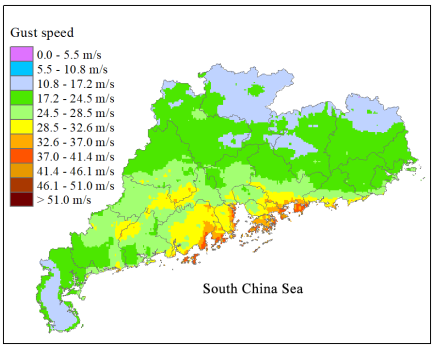


FIGURE 5. Predicted maximum gust of "Mangkhut".

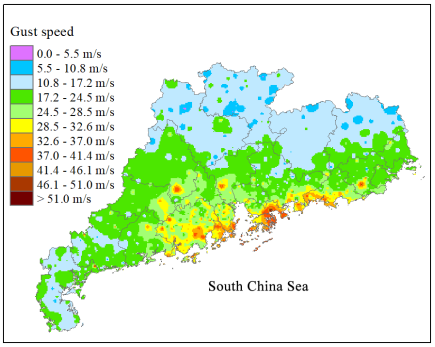


FIGURE 6. Measured maximum gust of "Mangkhut".

The predicted and measured maximum gust data come from the power grid department, among which the predicted maximum gust data are from the forecast gust series of 24 hours before landing because according to experience, the forecast data of 24 hours before landing are the most accurate. The measured maximum gusts were detected during the passage of the typhoon. Fig. 5 shows that the wind speed near the landing point is the maximum, and the wind speed is decreasing from the coastal area to the inland area. In Fig. 6, the maximum wind speed is also distributed near the landing point, and shows an attenuation trend from the coast to the inland. The gust distribution in both images are roughly the same along the coast. However, the attenuation trend in Fig. 6 is faster because each gust zone is narrower than that in Fig. 5, and the non-uniformity in the attenuation is more prominent. Fig. 7 shows the main network of 11 cities in coastal area of Guangdong province.

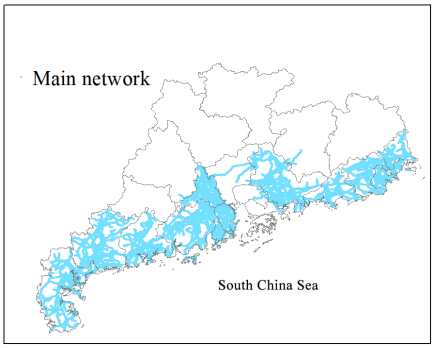


FIGURE 7. Main networks along coast of Guangdong.

In Fig.7, main networks of 11 cities including Chaozhou, Huizhou, Jiangmen, Jieyang, Maoming, Shantou, Shanwei, Yangjiang, Zhanjiang, Zhongshan and Zhuhai are presented except for those lacking operational information such as Guangzhou, Dongguan, Shenzhen, Hongkong.

Geographic information data with 1km$\times$1km resolution of Guangdong province, including altitude, aspect, slope, slope position, underlying surface, and surface roughness collected by ArcGIS10.4.1 are shown in Fig. 8.

Fig. 8(a) shows the characteristics of the topography of Guangdong province, which is high in the northwest and low in the southeast. The distribution characteristics of slope in Fig. 8(b) are consistent with altitude. Fig. 8(c) displays the aspect distribution. Fig. 8(d) exhibits the slope position distribution. Fig. 8(e) shows the underlying surface distribution. Roughness in Fig. 8(f) is large in mountainous areas with high altitude and in delta areas with dense urban buildings. The slope direction is -1, indicating ''flat without slope''. Definitions of slope position and underlying surface are given in Table 2 and Table 3.

The damage information of main and distribution networks under four typhoons including ''Rammasun'' and ''Kalmaegi'' in 2014, ''Mujigae'' in 2015, and ''Hato'' in 2017 collected by the power department will be shown in the following section.

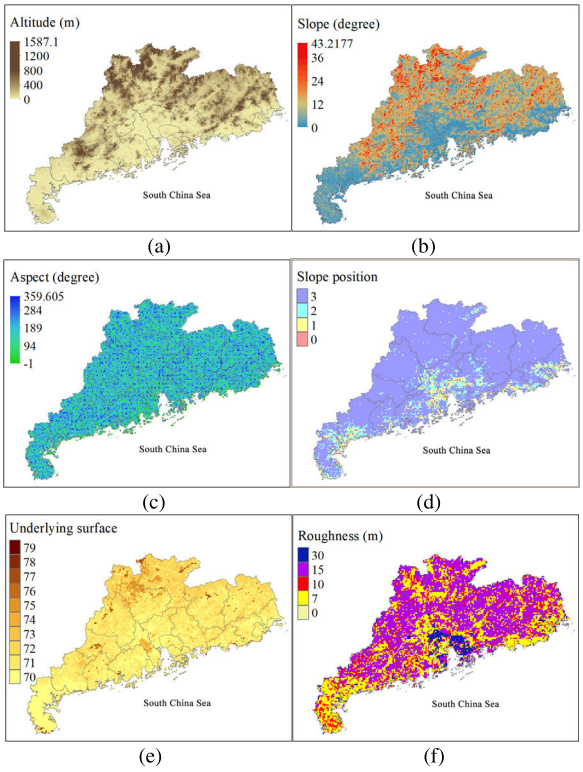


FIGURE 8. Geographical information distribution of Guangdong province: (a) Altitude, (b) Slope, (c) Aspect, (d) Slope position, (e) Underlying surface, (f) Surface roughness.

TABLE 2. Slope position interpretation table.

Slope position	0	1	2	3
Meaning	Other	Uphill	Mesoslope	Down hill

TABLE 3. Underlying surface interpretation table.

Value	Underlay surface type
70	Paddy soil, lakes, reservoirs
71	Latosolic red soil
72	Red soil, coastal saline soil
73	Yellow soil, acid sulfate
74	Tidal soil, coastal sand soil
75	Calcareous soil
77	Purple soil
78	Rocky soil
79	Skeleton soil

2) INFORMATION PROCESSING MODULE

First, the damage status of the tower is marked by classification variables. In this paper, the damaged state is denoted by 1, and the undamaged state is denoted by 0. Literature [33] proves that the median can obtain a good confidence interval in the missing value interpolation of multi-source data, and it is easy to operate, so this paper uses the median to fill in the missing value. All wind speeds are converted to the gust.

TABLE 4. Features interpretation table.

Feature name	Unit	Meaning
V_{10}	m/s	Maximum gusts at 10 meters
V'_{10}	m/s	Design wind speed at 10 meters
H	m	Altitude
A	degree	Aspect
S	degree	Slope
P		Slope position
U		Underlying surface
R	m	Surface roughness
T	year	Operation time

Convert wind speeds by using (1), and (2) is used for standardization. Table 4 shows the variables and their definitions. After processed, the damage information under four typhoons are used to train the model. For the large amount of training data in Section III, we list some of them in Table 14 of Appendix.

3) DAMAGE WARNING AND EVALUATION MODULE

First, train the RF model by use of the training dataset constructed by processed damage information under four typhoons. The training dataset is composed by cases divided into features and damage status. The features are shown in Table 4, and the damage status of the tower is marked by classification variables, which is 1 if the tower is damaged and otherwise 0. The cases representing damage status 0 in the dataset are randomly selected from historical record of these four typhoons. Then utilize the training dataset to fit a RF model [15].

Then, this paper generates wind speed randomly at each wind point, and calculate its guaranteed rate by the extreme value type I probability distribution. Then this paper puts the generated wind speed into the RF method to predict the damage probability. At last, this paper multiplies it with the guaranteed rate to get the predicted damage probability. Repeat above process until the probability converges.

We take a power tower at $(116.9°N, 23.643°E)$ as an example. Its feature vector denoted as $[V'_{10}, H, A, S, P, U, R, T]$ is shown in Fig. 9.

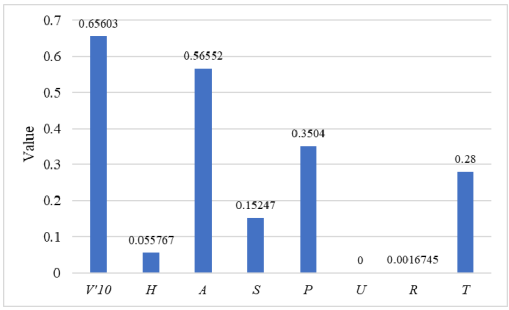


FIGURE 9. Features of the power tower at $(116.9°N, 23.643°E)$.

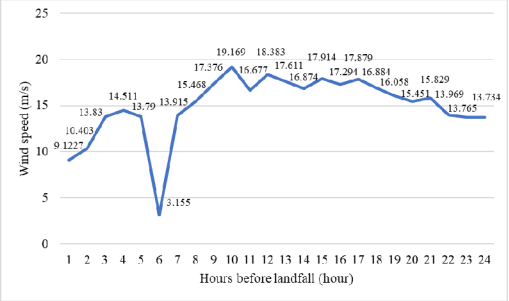


FIGURE 10. The 24-hour forecast wind speed series at (116.9°*N*, 23.643°*E*).

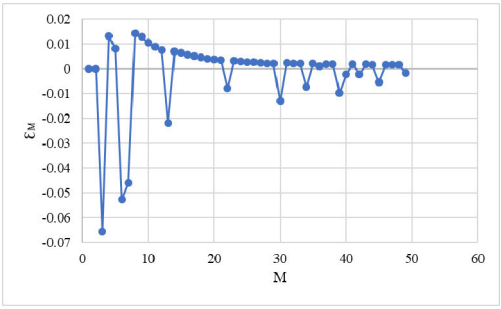


FIGURE 11. The deviation of average value of *P* as iteration times increase.

The 24-hour forecasting wind speed series in the coordinate of the illustrated power tower is shown in Fig. 10.

Scale parameter and location parameter of extreme value type I probability distribution at this tower are $a = 0.52072$ and $u = 14.269$. To evaluate the estimation accuracy of the method of moments, test the goodness of fit under 5 significance levels at this wind speed point using (13)-(17), and the result is shown in Table 5.

TABLE 5. The Hypothesis Test at (116.9° *N*, 23.643° *E*).

a	u	D_n	α	$D_{24,\alpha}$	Result
0.52072	14.269	0.1684	0.2	0.21205	Accept
			0.1	0.24242	Accept
			0.05	0.26931	Accept
			0.02	0.30104	Accept
			0.01	0.32286	Accept

As shown in Table 5, there is $D_n < D_{24,\alpha}$ at each value of α. Therefore, accept the hypothesis that wind speed samples at (116.9°N, 23.643°E) obey the extreme value type I probability distribution.

To calculate the damage probability, this paper first generates a wind speed using Monte Carlo method to get $w = 15.0024$ m/s. Considering the occurrence possibility of w, calculate its guaranteed rate $P(w) = 0.49471$. Then this paper uses RF to calculate the damage probability regarding the generated wind speed and obtain $f\ (x|w = 15.0024) = 0.5192$. However, this result should be amended by $P(w)$, i.e., $P = P(w)^* f\ (x|w = 15.0024) = 0.4878$. In order to get a stable result, repeat the above process until (22) is satisfied. As shown in Fig. 11, the deviation of average value of P is reduced to 0.01 after 30 iterations, and it further decreases as the iteration times increase. This paper repeated the above process for all the power towers and found that when the number of iterations exceeds 50, deviations of all the average values are smaller than 0.01. Therefore, this paper sets iteration times M = 50, and the result shows $P = 0.4409$ in Fig. 11.

To evaluate the accuracy of the method of moments for all wind speed points, calculate the rejection rate as shown in Table 6 by (18). Table 6 shows that when $\alpha = 0.2$,

TABLE 6. The rejection rate of all wind speed points.

α	$D_{24,\alpha}$	$RR(\alpha)$
0.2	0.21205	0.0915
0.1	0.24242	0.0372
0.05	0.26931	0.0152
0.02	0.30104	0.0029
0.01	0.32286	1.548e-4

the rejection rate $RR\,(\alpha) = 0.0915$. Therefore, all the wind speed points fit well with the extreme value type I probability distribution by using the method of moments, achieving an accuracy of 90.85%. Therefore, the method of moments is effective in parameter estimation for wind speed.

For purpose of damage calculation, we use the single model RF method described in Section III to respectively evaluate the damage situation of transmission lines in case of predicted wind field and measured wind field.

First, the evaluation results in case of predicted wind field are shown in Fig. 12.

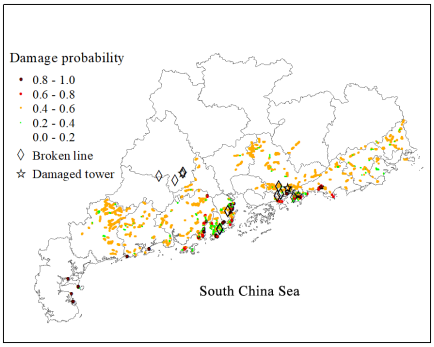


FIGURE 12. Damage prediction by the single model RF method under predicted wind field.

In Fig. 12, the diamond icons indicate the location of the broken line accidents, and the pentagrams indicate the location of damaged towers. The results show that transmission lines with high damage probability are mainly located in coastal areas, especially in areas with gust above 28.5 m/s in Fig. 5. Furthermore, the predicted damage distribution is consistent with the actual damage situation, the tower

damage is still influenced by modeling related features, so the predicted location is close to the actual damaged location. However, some areas in the upper left containing actual damage do not have predicted result due to the lack of operation information including design wind speed and operation time. The prediction of damage probability is high where no damage occurs.

Specifically, the four indicators including distribution similarity, the maximum magnitude similarity, cumulative similarity and synthetic similarity are shown in Table 7.

TABLE 7. Indicators of single model RF method under predicted gust.

Distribution similarity	The maximum magnitude similarity	Cumulative similarity	Synthetic similarity
0.00985	0.0222	0.0471	0.0285

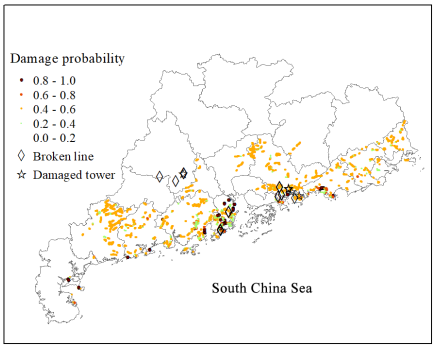


FIGURE 13. Damage prediction by single model RF method under measured wind field.

Second, in case of measured wind field, Fig. 13 shows damage probability evaluation results calculated by the single model RF method. In Fig. 13, under the actual gust, the predicted transmission line damage is mainly distributed in coastal areas with gust above 28.5 m/s, and in consistence with the actual damage distribution. The evaluation results cover almost all the actual damaged areas, indicating that the damage evaluation results under the actual gust are effective. However, the predicted damage probability is still high where no damage occurs in practice.

TABLE 8. Indicators of single model RF method under measured gust.

Distribution similarity	The maximum magnitude similarity	Cumulative similarity	Synthetic similarity
0.0152	0.0156	0.0416	0.0284

The four indicators including distribution similarity, the maximum magnitude similarity, cumulative similarity and synthetic similarity are shown in Table 8.

At last, this paper uses the hybrid prediction model to predict damage probability of transmission lines according to Section III. The predicted results are shown in Fig. 14.

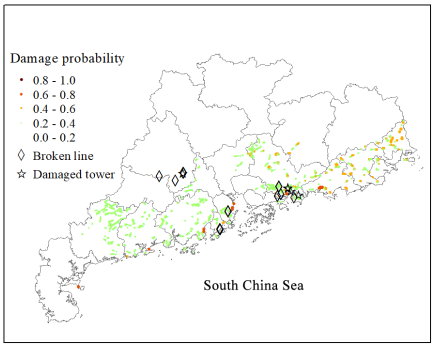


FIGURE 14. Damage prediction by hybrid prediction model under predicted wind field.

In Fig. 14, the damage probability of coastal areas is relatively high. Except for maintaining the consistence with actual damage distribution, the hybrid prediction model avoids or reduces damage probability in areas without actual damage. Therefore, areas with highest risk will be identified efficiently than Fig. 12 and Fig. 13, indicating that the prediction result of the hybrid prediction model is better. Compared with the predicted results of the single model RF method, the predicted distribution of the hybrid prediction model of damage probability improves the probability accuracy.

The four indicators including distribution similarity, the maximum magnitude similarity, cumulative similarity and synthetic similarity are shown in Table 9.

TABLE 9. Indicators of the hybrid prediction model.

Distribution similarity	The maximum magnitude similarity	Cumulative similarity	Synthetic similarity
0.0107	0.1667	0.347	0.179

The hybrid prediction model uses the extreme value type I probability distribution and Monte Carlo method to simulate the random wind field. Under the premise of giving the occurrence probability of wind field, the damage probability of transmission lines under each wind field is calculated by RF method, and the average damage effect is finally calculated. Compared with the hybrid prediction model, the single model RF method only uses the maximum gust field during the prediction period, and does not consider the occurrence possibility of maximum gust at each wind speed point, so it may lead to application limitations in actual use. In addition, in the actual situation, the measured wind speed data are abundant, and only using the maximum gust data will inevitably lead to the loss of effective information. Therefore, the hybrid prediction model of damage probability has more practical application value than the single model RF method, especially in the early-warning and prevention and control of medium-term and long-term typhoon disaster damage in the power grid. Abundant historical typhoon monitoring data can be used to fit the extreme value type I distribution, and Monte Carlo method and RF method can help to predict the risk of power grid damage in a specific time scale.

The Monte Carlo method calculates the probability of system states by combining components' states, but the number of samples will increase sharply as the number of components increases. This paper does not need to calculate the risk indicators at the system level, so the prediction problem is simplified to the component level, which greatly reduces the number of sampling. For example, sampling 50 times is effective for the prediction. The computation time is 3333.11 seconds with computer configuration of Intel i5-4210U and 4-gigabyte of memory, which is efficient to issue an alert at least 23 hours in advance.

The comparisons between four indicators of three methods are carried out in the following section.

4) SIMILARITY ANALYSIS

For comparing convenience, distribution similarity indicators of the three predicted results are shown in Table 10.

TABLE 10. Comparison of distribution similarity.

Prediction method	Single model RF method		Hybrid prediction model
	Predicted gust	Measured gust	
Distribution similarity	0.00985	0.0152	0.0107

Table 10 shows that the predicted damage distribution of single model RF method under the measured gust is the closest to the actual damage situation, which is 0.0152. However, it is only 0.00985 under the predicted wind speed, indicating that the uncertainty of the predicted wind speed has a great impact on the prediction results. After using the hybrid prediction model, the distribution similarity is improved to 0.0107, which is 8.63% higher than the result of single model RF method under predicted gust. It is demonstrated that the extreme value type I probability distribution and Monte Carlo method are effective to simulate typhoon wind field, and the uncertainty of typhoon can be counted to improve the accuracy of prediction results.

According to the statistics, Table 11 shows the damage magnitude of the three predicted results under various probability thresholds.

The magnitude similarities of different probability thresholds are shown in Fig. 15, which indicates that when the threshold value is 0.7, the magnitude similarity of the hybrid prediction model reaches to the highest. The magnitude similarity of single model RF method under predicted and measured gust reach to the highest at 0.9, but far lower than the hybrid prediction model, indicating that the predicted damage magnitude of hybrid prediction model is more accurate.

In order to facilitate the comparison, power towers related to broken lines are included into the actual damage magnitude. The actual damage magnitude is 16, and the cumulative similarities of the models are shown in Table 12. It can be seen that the cumulative similarity of single model RF method under predicted gust is 0.0471. In the case of measured gust, the cumulative similarity of single model RF method is only 0.0416. The cumulative similarity 0.347 of the hybrid prediction model is the highest, and higher than the two methods by 636.73% and 734.13%, respectively. It shows that the hybrid prediction model of damage probability is more effective and can reduce the prediction error.

TABLE 11. Damage statistics of different probability thresholds.

Probability threshold	Single model RF method		Hybrid prediction model
	Predicted gust	Measured gust	
0.1	9348	9152	7531
0.2	8894	8751	6878
0.3	8329	8213	3173
0.4	6966	6630	602
0.5	4300	4104	62
0.6	268	232	48
0.7	148	124	10
0.8	95	104	0
0.9	61	80	0

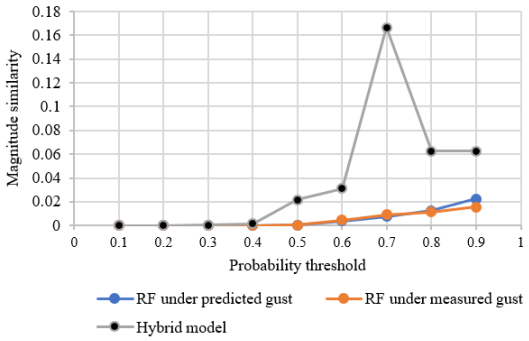


FIGURE 15. Magnitude similarity of different probability thresholds.

TABLE 12. Comparison of cumulative similarity.

Prediction method	Single model RF method		Hybrid prediction model
	Predicted gust	Measured gust	
Cumulative similarity	0.0471	0.0416	0.347

This paper compares the probability threshold for 3 methods simultaneously, intending to verify that there is a threshold that can be identified from all thresholds in the hybrid prediction model, which is critical to make a decision under typhoon disasters. Although there are also most suitable thresholds in the other two methods if they are studied individually, their predicted magnitudes at 0.9 are neither obvious nor close to the actual damage magnitude. Therefore, the hybrid prediction model is state-of-the-art model among the three methods.

The synthetic similarities are shown in Table 13. It can be seen that the synthetic similarity of the hybrid prediction model is the best (0.179), higher than the previous two methods by 528.07% and 530.28%, respectively. Therefore, the model proposed in this paper is scientific and effective, and can be applied to engineering practice.

TABLE 13. Comparison of synthetic similarity.

Prediction method	Single model RF method		Hybrid prediction model
	Predicted gust	Predicted gust	
Synthetic similarity	0.0285	0.0284	0.179

TABLE 14. Partial damage information under four typhoons.

Sequence	V_{10}	V'_{10}	H	A	S	P	U	R	T	State
1	0.652	0.656	0.051	0.839	0.174	0.222	0.111	0.500	0.150	0
2	0.518	0.656	0.044	0.195	0.080	0.000	0.000	0.500	0.150	0
3	0.560	0.656	0.041	0.501	0.010	0.000	0.000	0.000	0.150	0
4	0.530	0.656	0.042	0.154	0.022	0.000	0.111	0.333	0.150	0
5	0.584	0.656	0.042	0.377	0.026	0.000	0.111	0.333	0.150	0
6	0.659	0.656	0.050	0.234	0.134	0.222	0.000	0.500	0.150	0
7	0.565	0.656	0.041	0.790	0.021	0.222	0.111	0.000	0.150	0
8	0.612	0.656	0.041	0.969	0.013	0.000	0.000	0.000	0.150	0
9	0.539	0.656	0.065	0.934	0.121	0.000	0.000	0.500	0.150	0
10	0.627	0.656	0.040	0.000	0.000	0.000	0.000	0.033	0.150	0
11	0.475	0.656	0.058	0.649	0.078	0.000	0.111	0.500	0.150	0
12	0.565	0.656	0.041	0.790	0.021	0.222	0.111	0.000	0.150	0
13	0.627	0.656	0.041	0.790	0.021	0.222	0.111	0.000	0.150	0
14	0.505	0.656	0.041	0.076	0.012	0.000	0.111	0.000	0.150	0
15	0.465	0.656	0.062	0.892	0.207	0.000	0.00	0.500	0.150	0
16	0.518	0.656	0.041	0.076	0.012	0.000	0.111	0.000	0.150	0
17	0.650	0.656	0.040	0.000	0.000	0.000	0.000	0.033	0.150	0
18	0.491	0.656	0.041	0.127	0.011	0.000	0.000	0.000	0.150	0
19	0.585	0.656	0.040	0.575	0.012	0.000	0.000	0.333	0.150	0
20	0.565	0.656	0.041	0.790	0.021	0.222	0.111	0.000.	0.150	0
21	0.490	0.656	0.043	0.296	0.038	0.222	0.000	1.000	0.150	0
22	0.505	0.656	0.067	0.232	0.256	0.000	0.111	0.500	0.150	0
23	0.523	0.656	0.041	0.479	0.018	0.000	0.000	0.000	0.150	0
24	0.584	0.656	0.049	0.740	0.039	0.000	0.111	0.033	0.150	0
25	0.686	0.656	0.054	0.324	0.186	0.000	0.111	0.500	0.150	0
26	0.553	0.656	0.041	0.626	0.022	0.000	0.222	0.000	0.150	1
27	0.553	0.656	0.041	0.626	0.022	0.000	0.222	0.000	0.150	1
28	0.543	0.656	0.041	0.626	0.022	0.000	0.222	0.000	0.150	1
29	0.543	0.656	0.041	0.626	0.022	0.000	0.222	0.000	0.150	1
30	0.543	0.656	0.041	0.626	0.022	0.000	0.222	0.000	0.150	1
31	0.543	0.656	0.041	0.626	0.022	0.000	0.222	0.000	0.150	1
32	0.553	0.656	0.041	0.626	0.022	0.000	0.222	0.000	0.150	1
33	0.606	0.656	0.039	0.000	0.000	0.000	0.111	0.333	0.150	1
34	0.569	0.656	0.094	0.268	0.076	0.000	0.111	0.500	0.150	1
35	0.569	0.656	0.062	0.339	0.184	0.000	0.000	0.500	0.150	1
36	0.573	0.656	0.045	0.936	0.020	0.000	0.000	0.500	0.150	1
37	0.652	0.656	0.051	0.070	0.124	0.000	0.222	0.500	0.150	1
38	0.607	0.656	0.044	0.712	0.043	0.000	0.000	0.333	0.150	1
39	0.533	0.656	0.118	0.194	0.203	0.000	0.111	0.500	0.150	1
40	0.533	0.656	0.118	0.194	0.203	0.000	0.111	0.500	0.150	1
41	0.493	0.656	0.063	0.646	0.146	0.000	0.111	0.333	0.150	1
42	0.545	0.656	0.044	0.090	0.109	0.000	0.111	0.500	0.150	1
43	0.545	0.656	0.044	0.090	0.109	0.000	0.111	0.500	0.150	1
44	0.545	0.656	0.044	0.090	0.109	0.000	0.111	0.500	0.150	1
45	0.652	0.656	0.051	0.070	0.124	0.000	0.222	0.500	0.150	1
46	0.445	0.562	0.039	0.000	0.000	0.000	0.111	0.333	0.150	1
47	0.445	0.562	0.039	0.000	0.000	0.000	0.111	0.333	0.150	1
48	0.436	0.562	0.039	0.000	0.000	0.000	0.111	0.333	0.150	1
49	0.434	0.656	0.039	0.000	0.000	0.000	0.111	0.333	0.150	1
50	0.434	0.656	0.039	0.000	0.000	0.000	0.111	0.333	0.150	1

V. CONCLUSION

This paper establishes a technical framework for transmission line damage warning under typhoon disasters. In the damage warning and evaluation module, a hybrid prediction model of transmission line damage probability based on the extreme value type I distribution, Monte Carlo method and RF method is proposed. Its application in typhoon "Mangkhut" is introduced by a case study.

The comparison between the predicted results and the actual damage distribution verifies the scientific nature and practicability of the proposed hybrid prediction model of transmission line damage probability. The hybrid prediction model can promote distribution similarity by 8.63% than single model RF method under predicted wind speed; the hybrid prediction model has the best cumulative similarity; the hybrid prediction model has the highest synthetic similarity. Therefore, the hybrid prediction model of transmission line damage probability can reduce the prediction error while ensuring the prediction accuracy, which is more practical for further risk indicator calculation.

According to the statistics of the tower damage reasons, the potential factors such as the construction quality, corrosion of the equipment caused by salt in the paddy field, and accidental meteorology could not be ignored. However, how to predict and quantify the potential risks caused by these factors are the direction of future efforts.

APPENDIX

See Table 14.

REFERENCES

[1] A. Kwasinski, F. Andrade, M. J. Castro-Sitiriche, and E. O'Neill-Carrillo, "Hurricane maria effects on puerto rico electric power infrastructure," *IEEE Power Energy Technol. Syst. J.*, vol. 6, no. 1, pp. 85–94, Mar. 2019.

[2] Y. Yang, W. Tang, Y. Liu, Y. Xin, and Q. Wu, "Quantitative resilience assessment for power transmission systems under typhoon weather," *IEEE Access*, vol. 6, pp. 40747–40756, 2018.

[3] Y. Yang, Y. Xin, J. Zhou, "Failure probability estimation of transmission lines during typhoon based on tropical cyclone wind model and component vulnerability model," in *Proc. IEEE PES Asia–Pacific Power Energy Eng. Conf. (APPEEC)*, Chengdu, China, Apr. 2017, pp. 1–6.

[4] X. Song, Z. Wang, H. Xin, "Risk-based dynamic security assessment under typhoon weather for power transmission system," in *Proc. IEEE Asia–pacific Power Energy Eng. Conf. (APPEEC)*, Kowloon, Hong Kong, Dec. 2014, pp. 1–6.

[5] M. Panteli, C. Pickering, S. Wilkinson, R. Dawson, and P. Mancarella, "Power system resilience to extreme weather: Fragility modeling, probabilistic impact assessment, and adaptation measures," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3747–3757, Sep. 2017.

[6] H. Geng, Y. Huang, S. Yu, "Research on early warning method of overhead transmission line damage caused by typhoon disaster," in *Proc. 9th Int. Conf. Ambient Syst., Netw. Technol. (ANT)*, Porto, Portugal, vol. 52, May 2018, pp. 1170–1175.

[7] Y. Meng, M. Matsui, and K. Hibi, "An analytical model for simulation of the wind field in a typhoon boundary layer," *J. Wind Eng. Ind. Aerodyn.*, vol. 56, nos. 2–3, pp. 291–310, May 1995.

[8] Y. Meng, M. Matsui, and K. Hibi, "A numerical study of the wind field in a typhoon boundary layer," *J. Wind Eng. Ind. Aerodyn.*, vols. 67–68, pp. 437–448, Apr. 1997.

[9] J. Wang, X. Xiong, N. Zhou, Z. Li, and W. Wang, "Early warning method for transmission line galloping based on SVM and AdaBoost bi-level classifiers," *IET Gener., Transmiss. Distrib.*, vol. 10, no. 14, pp. 3499–3507, Nov. 2016.

[10] W. Gao, R. Zhou, and D. Zhao, "Heuristic failure prediction model of transmission line under natural disasters," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 4, pp. 935–942, Mar. 2017.

[11] Y. Wang, Z. Yin, and L. Li, "Risk Assessment Method for Distribution Network Considering Operation State in the Scene of Typhoon Disaster," *Proc. CSU-EPSA*, vol. 30, no. 12, vol. 21, no. 3, pp. 60–65, Dec. 2018.

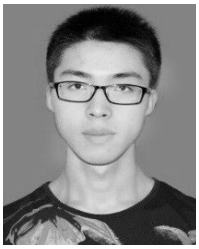
[12] R. Nateghi, S. D. Guikema, and S. M. Quiring, "Forecasting hurricane-induced power outage durations," *Natural Hazards*, vol. 74, no. 3, pp. 1795–1811, Dec. 2014.

[13] S. D. Guikema, R. Nateghi, S. M. Quiring, A. Staid, A. C. Reilly, and M. Gao, "Predicting hurricane power outages to support storm response planning," *IEEE Access*, vol. 2, pp. 1364–1373, 2014.

[14] D. W. Wanik, E. N. Anagnostou, B. M. Hartman, M. E. B. Frediani, and M. Astitha, "Storm outage modeling for an electric distribution network in northeastern USA," *Natural Hazards*, vol. 79, no. 2, pp. 1359–1384, Nov. 2015.
[15] H. Hou, S. Yu, H. Wang, Y. Huang, H. Wu, Y. Xu, X. Li, and H. Geng, "Risk assessment and its visualization of power tower under typhoon disaster based on machine learning algorithms," *Energies*, vol. 12, no. 2, p. 205, 2019.
[16] *Technical Specification for Wind Protection Design of Distribution Lines of China Southern Power Grid Company Limited, China Southern Power Grid Company Limited* Standard Q/CSG 1201012-2016, 2016.
[17] W. Li, "Basis of risk assessment methods," in *Risk Assessment Power Systerm*, 1st ed. Beijing, China: Science Press, 2006, pp. 69–71.
[18] ESRI. *ArcGIS 10.4.1 for Desktop Quick Start Guide*. Accessed: Dec. 19, 2019. [Online]. Available: https://desktop.arcgis.com/en/quick-start-guides/10.4/arcgis-desktop-quick-start-guide.htm
[19] B. R. Ellingwood and P. B. Tekie, "Wind load statistics for probability-based structural design," *J. Struct. Eng.*, vol. 125, no. 4, pp. 453–463, Apr. 1999.
[20] R. Billinton, W. Li. *Reliability assessment of electric power systems using Monte Carlo methods*. New York, NY, USA: Plenum, 1994.
[21] B.-H. Lee, D.-J. Ahn, H.-G. Kim, and Y.-C. Ha, "An estimation of the extreme wind speed using the korea wind map," *Renew. Energy*, vol. 42, pp. 4–10, Jun. 2012.
[22] Z. Duan and D. Zhou, "A comparative study on parameter estimate method for extremal value distribution," *J. Harbin Inst. Technol.*, vol. 36, no. 12, pp. 1605–1609, Dec., 2004.
[23] M. DeGroot and M. Schervish, "Estimation," in *Probability and Statistics*, 4th ed. Boston, MA, USA: Addison-Wesley, 2012, pp. 430–431.
[24] T. Shimokawa and M. Liao, "Goodness-of-fit tests for type-I extreme-value and 2-parameter weibull distributions," *IEEE Trans. Rel.*, vol. 48, no. 1, pp. 79–86, Mar. 1999.
[25] H. Lei and C. Singh, "Non-sequential Monte Carlo simulation for cyber-induced dependent failures in composite power system reliability evaluation," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1064–1072, 2017.
[26] F. Sharipov, *Rarefied fas Dynamics: Fundamentals for Research and Practice*, 1st ed. Weinheim, Germany: Wiley-VCH Verlag GmbH, 2016, p. 74.
[27] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45., no. 1, p. 5, 2001.
[28] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.
[29] H. Li, "Decision tree," in *Statistical Learning Method*, 1st ed. Beijing, China: Tsinghua Univ. Press, 2012, pp. 55–75.
[30] Z. Zhou, "Ensemble learning," in *Machine Learning*, 1st ed. Beijing, China: Tsinghua Univ. Press, 2016, pp. 171–196.
[31] *Electric Power Research Institute*, Guangdong Power Grid Co, Guangzhou, China.
[32] *Hurricane and Tropical Cyclones—Super Typhoon Mangkhut*. [Online]. Available: https://www.wunderground.com/hurricane/western-pacific/2018/super-typhoon-mangkhut
[33] E. Casleton, D. Osthus, and K. Van Buren, "Imputation for multisource data with comparison and assessment techniques," *Appl. Stochastic Models Bus. Ind.*, vol. 34, no. 1, pp. 44–60, Jan. 2018.

HUI HOU (Member, IEEE) was born in Wuhan, Hubei, China, in May 1981. She received the B.S. degree from Wuhan University, Wuhan, in 2003, and the Ph.D. degree from the Huazhong University of Science and Technology, Wuhan, in 2009.

She is currently an Associate Professor with the School of Automation, Wuhan University of Technology. Her research interests include power system risk assessment, and protection and security automatic control.



SHIWEN YU was born in Linyi, Shandong, China, in February 1995. He received the B.S. degree from the Wuhan University of Technology, Wuhan, in 2017, where he is currently pursuing the M.S. degree in electrical engineering.

His research interests include power system risk assessment and machine learning.

HAO WANG (Member, IEEE) is currently a Lecturer (Assistant Professor) with the Department of Data Science and Artificial Intelligence, Faculty of Information Technology, Monash University, Australia. He has been a Postdoctoral Research Fellow with Stanford University. His main research interests include optimization, machine learning, and data analytics of power and energy systems.

Dr. Wang was a recipient of the Washington Research Foundation Innovation Fellowship at the University of Washington, Seattle.

YAN XU (Senior Member, IEEE) received the B.S. and M.S. degrees from the South China University of Technology, Guangzhou, China, in 2008 and 2011, respectively, and the Ph.D. degree from The University of Newcastle, Australia, in 2013.

He is currently the Nanyang Assistant Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He was previously with The University of Sydney, Australia. His research interests include power system stability and control, microgrid and multienergy networks, and data-analytics for smart grid applications.

XIANG XIAO was born in Tianmen, Hubei, China, in November 1964.

He started work in July 1984. He is currently a Senior Engineer and the Deputy Chief Engineer of Guangdong Power Grid Company Ltd. He is also the President and the Deputy Secretary of the Electric Power Research Institute. He is also the Chairman of Guangdong Electric Power Research Institute Energy Technology Company Ltd., and a Secretary of the General Party Branch. His research interests include power system risk assessment and machine learning.

YONG HUANG was born in Wuhan, Hubei, China, in February 1986. He received the B.S. and M.S. degrees from Wuhan University, Wuhan, in 2009 and 2012, respectively.

He is currently a Senior Engineer and the Vice Minister of monitoring headquarters of Guangdong Power Grid Company Ltd. His research interest includes power system risk assessment.

XIXIU WU was born in Huanggang, Hubei, China, in June 1976. She received the Ph.D. degree in electrical engineering from the School of Electrical and Electronics Engineering, Huazhong University of Science and Technology (HUST), Wuhan, China, in 2005.

She is currently an Associate Professor with the School of Automation, Wuhan University of Technology. Her research interests include electromagnetic filed simulation and the EMC phenomena in high voltage of power systems.