

# Risk Management of Weather-related Failures in Distribution Systems Based on Interpretable Extra-trees

Ying Du, Yadong Liu, Yingjie Yan, Jian Fang, and Xiuchen Jiang

**Abstract**—Weather-related failures significantly challenge the reliability of distribution systems. To enhance the risk management of weather-related failures, an interpretable extra-trees based weather-related risk prediction model is proposed in this study. In the proposed model, the interpretability is successfully introduced to extra-trees by analyzing and processing the paths of decision trees in extra-trees. As a result, the interpretability of the proposed model is reflected in the following three respects: it can output the importance, contribution, and threshold of weather variables at high risk. The importance of weather variables can help in developing a long-term risk prevention plan. The contribution of weather variables provides targeted operation and maintenance advice for the next prediction period. The threshold of weather variables at high risk is critical in further preventing high risks. Compared with the black-box machine learning risk prediction models, the proposed model overcomes the application limitations. In addition to generating predicted risk levels, it can also provide more guidance information for the risk management of weather-related failures.

**Index Terms**—Extra-tree, machine learning, interpretability, weather-related failure, distribution system.

## I. INTRODUCTION

**W**EATHER-RELATED failures pose a significant challenge to the reliability of distribution systems. Service interruptions often occur under unfavorable weather conditions [1]. Because weather-related failures are random, time-varying, and destructive, they can generate huge economic losses in distribution systems [2]. In addition, cascading failures are often caused by initial weather-related failures [3]. Therefore, managing weather-related risks in advance and making reasonable decisions about crew arrangements, material reserves, and inspection plans are critical [4]. A risk management plan usually refers to the information produced by weather-related risk prediction models,

which is mainly the predicted risk level for the next prediction period. The aim of this study is to develop an interpretable machine learning (ML) weather-related risk prediction model to produce more guidance information for risk management, enabling utility companies to better withstand weather-related risks.

Many studies have focused on improving the performance of weather-related risk prediction. Poisson regression models [1], negative binomial regression models [5], [6], generalized linear mixed models [7], linear regression models [8], and exponential regression models [9] have been used to model the risks in previous studies. These statistical models require the distribution of data to meet certain assumptions that actual data are difficult to meet, leading to weak prediction performance as compared with some ML methods. For example, [1] shows that the performance of a Bayesian network is better than that of a Poisson regression model. In [10], a prediction model using ensemble learning outperforms the simple regression models. An artificial neural network is used in [11], and it is concluded that the performance is better than that of the statistical models presented in [8] and [9]. We propose a Bayesian neural-network-based risk prediction model in [12] that has an advanced risk prediction performance and can provide prediction confidence.

ML methods have always been a hot topic in risk prediction. However, ML methods currently in use do not provide thorough support for risk management because interpretability is sacrificed [13]. Interpretability is the degree to which a person can consistently predict the results of a model [14]. The lack of interpretability causes current ML methods to become black boxes [15], which means the methods are capable of outputting the predicted risk but are unable to figure out the role of each variable in the prediction. Black-box ML methods are not satisfactory for risk managers. This is because risk management plans expect more guidance information from the risk-prediction model. The predicted risk level produced by these models is only part of a useful reference. In addition, the action mechanism of variables, for example, the effect of the variable's values on the severity of a risk, is significant. Therefore, ML methods currently in use are unable to help risk managers formulate effective plans.

To solve these problems, developing an interpretable ML model for weather-related risk prediction is necessary. Gener-

Manuscript received: July 27, 2022; revised: August 24, 2022; accepted: October 8, 2022. Date of CrossCheck: October 8, 2022. Date of online publication: January 25, 2023.

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

Y. Du, Y. Liu (corresponding author), Y. Yan, and X. Jiang are with the Department of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: duyong@sjtu.edu.cn; lyd@sjtu.edu.cn; yanyingjie@sjtu.edu.cn; xcjiang@sjtu.edu.cn).

J. Fang is with the the Guangzhou Power Supply Bureau Co., Ltd. Electric Power Test Institute, Guangzhou 510000, China (e-mail: fjenglish@163.com).

DOI: 10.35833/MPCE.2022.000430



al ML models mine hidden rules from data, and, therefore, the source of knowledge is the data. Because the interpretable ML model is not exactly a black-box model, the model itself can provide valuable information [16]. In this manner, the interpretable ML model can explain or present understandable terms to humans [17], which means it can produce more interpretable guidance information for weather-related risk managers.

This study introduces an interpretable ML model for weather-related risk prediction and illustrates how interpretability can help in developing risk management plans.

An interpretable weather-related risk prediction model based on extra-trees is proposed. Extra-tree algorithm is an ensemble of decision trees in particular random patterns, which randomizes strongly in both attribute and cut-point choice while splitting a tree node [18]. This is suitable for prediction tasks because it pursues low variance in the trade-off between bias and variance [19]. Through analysis and processing of the path of decision trees in extra-trees, combined with using a trick in the training phase, extra-trees can be defined in an interpretable manner rather than as a black box. Therefore, the proposed model can provide the guidance in developing risk management plans. This guidance is mainly reflected in the following three aspects.

#### 1) *Long-term Weather-related Risk Management Plan (Long-term Plan)*

Regarding long-term plan for a region, effectively using limited investments to strengthen and update the weaknesses of power systems is critical. The task is how to determine which weather-induced risks are prioritized. The proposed model can derive the importance of weather variables, which represents the overall degree of impact of each weather variable on weather-related risk, helping in the formulation of a long-term plan.

#### 2) *Short-term Weather-related Risk Management Plan (Short-term Plan)*

The short-term plan is intended to guide ex-ante resource preparation and formulate preventive measures by clarifying the source of risks in the next prediction period. The proposed model can meet this requirement because of its interpretability. The proposed model can derive the contributions of weather variables, which indicates each weather variable's contribution to the predicted risk level in the next prediction period.

#### 3) *High-risk Prevention Plan*

High risks require more attentions when preparing risk management plans. The interpretability of the proposed model can derive the threshold of weather variables at high risk, which measures when weather variables promote the occurrence of high risks. Therefore, it can be used as a guide to developing a quantified high-risk prevention plan.

The main contribution of this study is obtaining and harnessing the valuable guidance information using the proposed model for weather-related risk management. This study represents a new perspective on weather-related risk management beyond merely pursuing prediction performance using previous black-box models, thus making the risk pre-

diction model more practical. The obtained information can be a useful reference for making long-term, short-term, and high-risk prevention plans for weather-related risk management. In addition, the proposed model offers better prediction performance than other ML models that can provide the same degree of interpretability. Therefore, the proposed model is an excellent choice for weather-related risk management, whether from the interpretability that brings sufficient guiding information or prediction performance.

The remainder of this study is organized as follows. Related works are described in Section II. Section III introduces the development of interpretable extra-trees. The interpretable extra-tree based weather-related risk prediction model is described in Section IV. Section V presents the weather-related risk management with help of proposed interpretability. Section VI concludes this paper.

## II. RELATED WORKS

Interpretability is critical for weather-related risk prediction. Decision makers in high-risk fields do not make decisions easily based on the prediction results of the used model without knowing the operating principle of the black-box model. When the interpretability of the used prediction model can be revealed, the interpretability will facilitate the practical application of the risk prediction model based on ML because it is more credible and can produce useful guiding information.

The research path of interpretable ML has two main directions: intrinsic and post-hoc interpretability. The models with intrinsic interpretability, which include many statistical models, have relatively simple structures and are already interpretable when they are designed. The relationship between the model variables and outputs can be easily explained for statistical models such as linear regressions. This is due to the availability of model parameters and their statistical significance. The coefficients of the linear regression model intuitively reflect the degree of influence of the variables on the predicted results. For black-box models such as extra-trees, this information is hidden inside the model structure. For intrinsic interpretable models, the model structure itself can explain why the model makes a certain prediction. When the model is too complex to interpret based on its structure, its interpretability can be tested using post-hoc methods. Compared with post-hoc interpretable models, intrinsic interpretable models are more intuitive and easier to understand for decision makers.

The relationship between prediction performance and model interpretability is illustrated in Fig. 1 [20]. The ML models (blue circles) perform better, whereas the traditional statistical models (red circles) are usually more interpretable. Thus, a trade-off exists between the prediction performance and interpretability. In our studied scenario, for weather-related prediction, the prediction model is expected to be more interpretable (preferably intrinsically interpretable) while maintaining an acceptable prediction performance. Therefore, we propose an interpretable extra-tree based prediction model to predict weather-related risks.

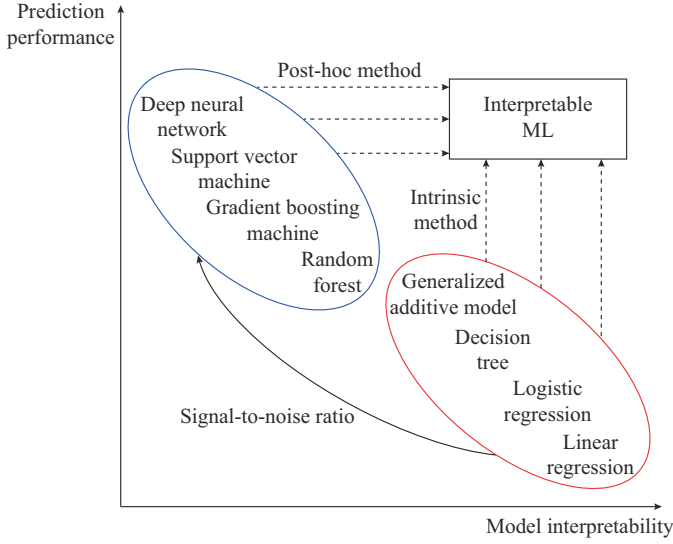


Fig. 1. Relationship between prediction performance and model interpretability.

### III. DEVELOPMENT OF INTERPRETABLE EXTRA-TREES

#### A. Extra-trees

Extra-tree algorithm [18] is one of the ensemble algorithms, the goal of which is to combine the predictions of several base estimators to improve the generality and robustness of the algorithm. The averaging method is used in extra-trees. Its base estimator is decision tree [21], which is named from using tree presentation. Each internal node of the tree corresponds to an attribute, and each leaf node corresponds to a class label. The pseudocode of the decision tree algorithm is shown in Algorithm 1.

---

**Algorithm 1:** decision tree algorithm (attribute list  $A$ )

---

1. Create a node  $N$
  2. If all samples are of the same class  $C$  then label  $N$  with  $C$ ; terminate
  3. Select  $a \in A$ , with the lowest Gini index; label  $N$  with  $a$
  4. For each value  $v$  of  $a$ :
    - 1) Grow a branch from  $N$  with the condition  $a = v$
    - 2) Let  $S_v$  be the subset of samples in  $S$  with  $a = v$
    - 3) If  $S_v$  is empty then attach a leaf labeled with the most common class in  $S$
  5. Repeat the above steps until leaf nodes are found
- 

The criterion like Gini index [22] can be used to select the attribute used in splitting the node, which represents the impurity or uncertainty of an attribute list. Gini index is more preferred due to its lower calculation complexity, which is a metric to measure how often a randomly chosen element would be incorrectly identified. It means that an attribute with a lower Gini index should be preferred. The formula is shown as:

$$Gini(D) = 1 - \sum_{i=1}^m p_i^2 \quad (1)$$

where  $m$  is the number of output labels; and  $p_i$  is the probability that a sample belongs to the  $i^{\text{th}}$  output label in the data set  $D$ . Compared with other tree-based ensemble algorithms, the randomness of extra-trees is greater in attribute selection,

as shown in Algorithm 2.

---

**Algorithm 2:** extra-tree splitting algorithm (attribute list  $A$ )

---

1. Select  $K$  attributes  $\{a_1, a_2, \dots, a_K\}$  in  $A$
  2. Let  $a_{i,\min}$  and  $a_{i,\max}$  denote the minimal and maximal values of  $a_i$  in  $A$   $\forall i = 1, 2, \dots, K$
  3. Draw a random cut-point  $a_i^*$  uniformly in  $[a_{i,\min}, a_{i,\max}]$
  4. Return the split  $s_i$   $[a < a_i^*]$ , where  $s_i$  represents a set smaller than  $a_i^*$  in  $A$
  5. Draw  $K$  splits  $\{S_1, S_2, \dots, S_K\}$
  6. Return a split  $s_*$  such that  $Score(s_*, S) = \max_{i=1,2,\dots,K} Score(s_i, S)$
- 

#### B. Interpretable Extra-trees

##### 1) Interpretable Decision Trees

Due to the difficulty in interpreting and fully understanding decision trees, the decision trees often become black boxes. This derives from the fact that the extra-tree model consists of a large number of deep trees, and each tree is split with strong randomness. However, when the fundamentals of the extra-tree model are thoroughly analyzed, the model can be better understood, and its interpretability can be expressed [23], [24]. First, we provide an explanation of the terms used.

1) Path: for a sample input to a decision tree, the path refers to the combination of all inference rules that the sample passes through from the root node to the leaf node such as path 1, as illustrated in Fig. 2, where  $v_2 - v_1$  indicates the gain/loss from  $a$ ;  $v_4 - v_2$  indicates the gain/loss from  $c$ ; and  $v_8 - v_4$  indicates the gain/loss from  $b$ . Each sample passes through its corresponding path to reach the final leaf node. And  $a'$ ,  $b'$ ,  $c'$ ,  $b''$ ,  $c''$ ,  $b'''$ , and  $c'''$  indicate related thresholds in the node.

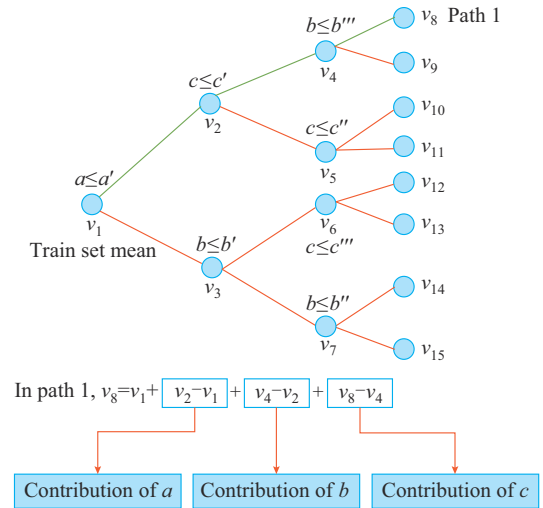


Fig. 2. Schematic of an interpretable decision tree.

2) Value: each node of the decision tree has a value represented by  $v_i$  ( $i = 1, 2, \dots, 15$ ), as shown in Fig. 3, which indicates the value of the predicted target.

3) Contribution: the contribution value is derived from the value of the current node minus that of the previous node and represents the contribution of the split attribute to the prediction path.

The paths of decision trees can be used to obtain more in-

formation. Each path of the tree is from the root of the tree and includes a series of decisions guarded by a particular attribute. The decision of paths either adds or subtracts from the value given in the parent node. All decision paths contribute to the final prediction results in decision trees. Therefore, the prediction process can be defined as the sum of the attribute contributions and the value of the root node, i.e., the mean value given by the topmost region that covers the entire training set. The prediction function can be written as (2).  $V_{root}$  is the value at the root of the node and  $contribution(x, k)$  is the contribution from the  $k^{th}$  attribute and the attribute vector  $x$ .

$$f(x) = V_{root} + \sum_{k=1}^K (contribution(x, k)) \quad (2)$$

Note that the contribution of each attribute is not a single predetermined value. It depends on the rest of the attribute vector which determines the decision path that traverses the tree. In this way, it also determines the contributions passing along the way.

Figure 2 shows how to analyze the decision paths in a decision tree. A sample that follows the path 1 in the prediction is exhibited with green color. Each decision of a node is made by an attribute, and each decision either adds or subtracts from the value provided by the parent node. According to the attributes which split the nodes in path 1, the prediction function of the exemplified sample can be written at the bottom of Fig. 3. It shows how the contributions of example attributes  $a$ ,  $b$ ,  $c$  are calculated. So far, the contribution of each attribute can be obtained in one prediction for decision tree. For extra-trees, the contribution of each attribute is the corresponding ensemble of the contained decision trees.

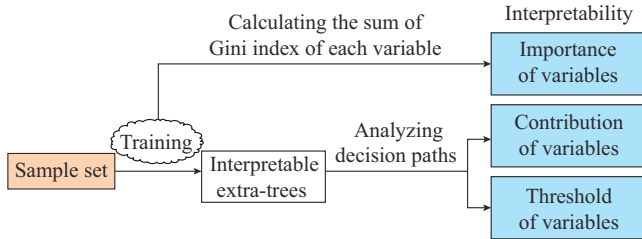


Fig. 3. Three aspects of interpretability in interpretable extra-trees.

## 2) Pseudocode of Interpretable Extra-trees

As the ensemble of decision trees, the interpretable definition of extra-trees is based on the interpretable extra-trees, the pseudocode for which is shown in Algorithm 3, where  $T$  is the number of trees in extra-trees.

---

### Algorithm 3: interpretable extra-trees

---

1. Obtain the value  $c_{root}$  at the root of the node in each tree of extra-trees
2. Calculate  $contribution(x, k)$
3. The prediction function in a decision tree can be written as:

$$f(x) = V_{root} + \sum_{k=1}^K (contribution(x, k))$$

4. The prediction of extra-trees is the average of the prediction of its trees:

$$F(x) = \frac{1}{T} \sum_{i=1}^T f_i(x)$$


---

## 3) Three Aspects of Interpretability

As Fig. 3 shows, the importance of variables, contribution

of variables, and threshold of variables can be produced when using interpretable extra-trees for risk prediction. Their meaning and specific production process are described as follows.

1) Importance of variables: the importance of variables measures the degree of impact of variables on the overall prediction, which can be obtained by calculating the Gini index in interpretable extra-trees. The importance ranking reflects the extent to which each variable determines the prediction. For a decision tree, it is necessary to find the attribute that can best distinguish the data set as the prioritized inference condition. The Gini index can be used to select the best attribute when dividing nodes. The smaller the Gini index of an attribute, the better its ability to divide nodes. The definition of the Gini index is given in (1). We first select a variable and then calculate the sum of the Gini index of all nodes split by this variable in each decision tree derived from the extra-trees. The variables can be ranked in the order of importance based on the sum of the Gini index of variables. The importance of variables is calculated during the training phase. Its source is the training data, i.e., the historical data, so it is a static index.

2) Contribution of variables: the importance of variables is used to evaluate those variables which are crucial for the overall prediction model, whereas the contribution of variables can provide more information when we are interested in a particular variable. The contribution of variables is a dynamic index of interpretability. It is used to evaluate the contribution of variables to a certain prediction. Therefore, the contribution of variables can be determined during each prediction period. The contribution value of variables on the prediction can be positive or negative. A positive or negative value indicates that the variable has a facilitating or hindering effect on the prediction, respectively. For example, as shown in Fig. 2, the contribution values of  $a$ ,  $b$ , and  $c$  can be positive or negative, representing their varying degrees of influence on the output.

3) Threshold of variables: managers may sometimes pay greater attention to specific classes of predicted outputs and their interpretability. For example, in the prediction of weather-related failure risks, risk managers are more vigilant against high risks. With respect to the prediction of the specific class, the threshold of variables can measure when variables will promote this prediction. For each variable, its contribution to the prediction of a specific class is related to its own value. In general, the larger the value of a variable, the larger its contribution to the high risk. Therefore, if we try to analyze the relationship between the contribution values of variables on the predicted output and their own values, the threshold of variables under different classes of prediction can be obtained, providing quantifiable guidance information.

## IV. INTERPRETABLE EXTRA-TREE BASED WEATHER-RELATED RISK PREDICTION MODEL

An interpretable extra-tree based weather-related risk prediction model is developed in this study using an actual data set.



### A. Weather-related Failure Data Source and Analysis

The data used in this study were collected from a city in eastern China. The utility company recorded weather-related failures from January 2011 to October 2018 that included failure information by date, time, location, and type, and a simple weather description. To quantify the effects of the weather, we obtained quantified weather parameters from the meteorological bureau in the studied city.

Figure 4 shows the distribution of the weekly weather-related failure counts. The figure shows that no weather-related failures occur over the course of many days. In general, when we use a data-driven prediction model to predict weather-related risk, most samples belong to zero risk, which is common. For example, there are more than 1500 zero-failure days from 1998 to 2003 in the data set of [1]. Because the number of failure samples is more than that of other failure counts, there is an unbalanced phenomenon in weather-related risk prediction problems, which will bring a negative effect on prediction performance [25]. More attention will be paid to samples with zero risk in the prediction model, and the information on other failure counts is hard to be investigated [26]. This unbalanced phenomenon easily causes overfitting in the prediction algorithms, leading to a reduction in prediction performance.

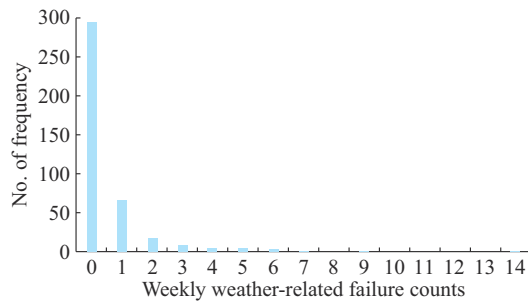


Fig. 4. Distribution of weekly weather-related failure counts.

Weather-related risk increases with the number of failures. The greater the number of failures in a given period, the higher the requirement for the coordination and preparation of manpower and material resources for power recovery. High weather-related failure counts indicate high risk, which imposes higher requirements on the risk management capabilities of utility companies. Therefore, accurately predicting the occurrence of high risk is critical. However, as Fig. 5 shows, very few high-risk samples are available. Limited high-risk information poses a significant challenge to the robustness of the prediction model. Table I presents an asset overview of the lines of the distribution system.

### B. Weather-related Risk-level Classification

In order to reasonably characterize the risk caused by weather-related failure counts, the failure counts are classified into three risk levels. The classification details are presented in Table II. The occurrence frequency percentages for the three risk levels are 72.66%, 23.15%, and 4.19%, respectively. Zero failure is naturally classified as one class due to its percentage.

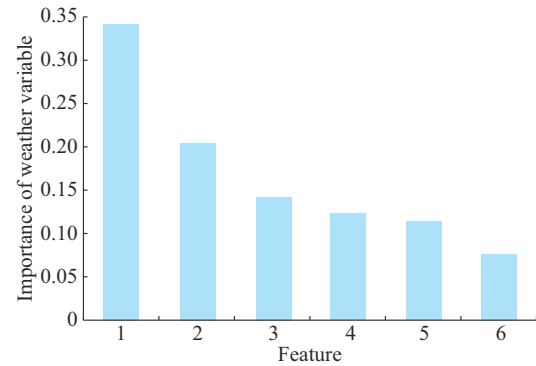


Fig. 5. Importance of weather variables in studied city.

TABLE I  
ASSET OVERVIEW OF LINES

| Type                    | Rich percentage (%) |
|-------------------------|---------------------|
| Overhead bare line      | 9.10                |
| Overhead insulated line | 35.40               |
| Cable line              | 55.50               |

TABLE II  
FAILURE-COUNT RISK-LEVEL CLASSIFICATION

| Failure level | Number of weather-related failures |
|---------------|------------------------------------|
| 0             | 0                                  |
| 1             | 1, 2, 3                            |
| 2             | [4, 14]                            |

The failure counts from 1 to 3 are classified as one class due to the high occurrence frequency, which can be thought of as the common risk level. However, when the failure counts are larger than 3, the failure occurrence frequency is reduced to 4.19%, which is beyond the 95% confidence level [27]. These failures have low occurrence frequency, but have high risk.

It is reasonable for utility companies to classify the risk level into common and rare because a well-designed power grid should perform well under both conditions. In addition, the classification of failure counts helps utility companies in conducting risk management because different operation and maintenance plans correspond to different risk levels.

### C. Weather Variables and Prediction Period

According to the investigation of historical failure data, weather-related failures are mainly caused by wind, rain, and thunder. There are various reasons specifically. For example, a strong wind will blow down trees, overhead lines, and equipment in a distribution system. Mild winds can also blow small objects up into the air such as plastic bags and branches, result in contacting with lines. Many failures occur in humid environments. When the rain is heavy, strong winds are also present, and when thunder occurs, the probability of failure increases. In general, several weather stations in cities produce different weather parameters. The degree of difference depends on the geography of the city under study. Due to the small area and single geographical en-

vironment in a city, the difference between different weather stations is typically small. In this study, a weather station located in the center of the city is used to obtain weather parameters. We chose the following six weather variables as attributes of the proposed model.

- 1) Feature 1: weekly average wind speed.
- 2) Feature 2: weekly maximum wind speed.
- 3) Feature 3: weekly average rainfall.
- 4) Feature 4: weekly maximum rainfall.
- 5) Feature 5: thunder days within a week.
- 6) Feature 6: weekly average humidity.

The prediction period should be reasonably determined. Daily and weekly predictions were used in previous studies. In [1] and [2], the daily weather-related failure counts are predicted. This means that the sample set consists of daily failure counts and related weather variables. The number of samples in the daily prediction is seven times that in the weekly prediction. However, because the area of the studied city is small, and the frequency of weather-related failure occurrence is not high (as shown in Fig. 5), the distribution of failure counts is scattered. It is difficult to build a robust prediction model [28]. Therefore, weekly prediction is chosen in this study, as in [2] and [28]. In addition, weather forecasts are typically accurate within one week.

#### D. Prediction Performance

##### 1) Evaluation Metrics

Due to the unbalanced nature of the data in terms of risk levels, evaluating the prediction performance based on accuracy is not reasonable, where accuracy is defined as the proportion of all correctly predicted samples to the total samples in the test set. We introduced the F1 score to evaluate the performance of risk prediction, which is suitable for evaluating ML methods under unbalanced sample sets [29]. A detailed definition of the F1 score can be found in [12]. The higher the F1 score, the better the performance of the prediction models under an unbalanced data set.

Underestimating weather-related risk may cause utility companies to neglect its prevention, leading to the inability to cope with the risk. Overestimating the risk results in an increase in risk prevention costs, including waste of workforce as well as material and financial resources. To better reflect the model's prediction performance and the prediction propensity for risk, we define two evaluation metrics in this study: risk underestimation rate (RUR) and risk overestimation rate (ROR). The definitions of RUR and ROR are given in (3) and (4), respectively. In the test set, the numbers of samples with underestimated and overestimated risk are denoted as  $u$  and  $o$ , respectively, and the total number of samples is denoted as  $t$ .

$$RUR = \frac{u}{t} \times 100\% \quad (3)$$

$$ROR = \frac{o}{t} \times 100\% \quad (4)$$

##### 2) Experiments and Results

In our experiments, the weather-related risk data were divided into a training data set (from 2011 to 2016) and a test data set (from January 2017 to October 2018). The numbers

of training and test samples were 312 and 94, respectively. Previously, we introduced the interpretable way of the proposed model, which was based on an interpretation of the decision tree through an analysis of decision paths. Therefore, decision tree (DT) and random forest (RF) [30] are chosen as the contrast models because they also have the same interpretability. Table III lists the prediction performances of comparative models under optimal parameters.

TABLE III  
PREDICTION PERFORMANCES OF COMPARATIVE MODELS UNDER OPTIMAL PARAMETERS

| Model       | The maximum depth of trees | Number of trees | F1 score | RUR (%) | ROR (%) |
|-------------|----------------------------|-----------------|----------|---------|---------|
| Extra-trees | 20                         | 50              | 0.939    | 3.191   | 3.191   |
| RF          | 20                         | 200             | 0.918    | 4.255   | 4.255   |
| DT          | 3                          |                 | 0.877    | 4.255   | 7.446   |

In Table III, it can be observed that the extra-tree based prediction model has leading performance under all three different evaluation metrics. It has the highest F1 score, and the lowest RUR and ROR, which indicates that the proposed model has the best prediction performance under an unbalanced data set.

The solutions to the tasks with unbalanced data can be divided into two methods [31]. First, the distribution of samples is changed by using various sampling methods and training-set division methods. Second, modern algorithms can be developed or the existing ones can be upgraded. These are known as algorithm-centered methods, where "ensemble learning" is a typical representative. In our comparative experiment, the proposed model is based on the bagging method to achieve ensemble. In random forest, several attributes are selected randomly from all attributes and then the optimal split node is selected to build decision trees. However, in extra-trees, it draws the random cutting points by uniform selection of attributes. The randomness is stronger in the node split process of trees than random forest [18], which means more unstable base learners in ensemble learning. Under this condition, the performance of ensemble will be better and so the generalization ability. Therefore, the proposed model performs better with small and unbalanced data, which has been verified the lowest RUR and the highest F1 score of the proposed model.

The proposed model not only provides rich interpretability but also exhibits the best risk prediction performance among the models that have the same degree of interpretability. Thus, the proposed model is an excellent choice for utility companies in managing weather-related risk.

##### 3) Experiments on Robustness

The weather data we used are the monitoring data from weather stations, which are similar to many studies [1], [10]. Obviously, a difference exists between monitoring and forecasting weather data. To verify the robustness of the proposed model and to test the actual application effects of the model, this study performs robustness experiments within the range of  $\pm 10\%$  of the error between the forecasting and monitoring weather data, thus simulating actual situations.

The experimental results are presented in Table IV. It can be observed that the proposed model predicts an average F1 score of 0.908 with  $\pm 10\%$  of the error. The performance is more robust than the contrast models. Therefore, in practical applications, although the forecasting weather data will have errors compared with the monitoring data used for training, the model still has acceptable and relatively better prediction performance.

TABLE IV  
PREDICTION PERFORMANCES UNDER DIFFERENT ERRORS BETWEEN  
FORECASTING AND MONITORING WEATHER DATA

| Error (%) | Prediction value |       |       |
|-----------|------------------|-------|-------|
|           | Extra-trees      | RF    | DT    |
| 0         | 0.939            | 0.918 | 0.877 |
| +1        | 0.929            | 0.888 | 0.877 |
| -1        | 0.907            | 0.897 | 0.877 |
| +2        | 0.927            | 0.897 | 0.877 |
| -2        | 0.890            | 0.875 | 0.877 |
| +3        | 0.930            | 0.884 | 0.802 |
| -3        | 0.876            | 0.863 | 0.837 |
| +4        | 0.930            | 0.900 | 0.809 |
| -4        | 0.859            | 0.851 | 0.845 |
| +5        | 0.930            | 0.910 | 0.802 |
| -5        | 0.868            | 0.875 | 0.861 |
| +6        | 0.942            | 0.901 | 0.879 |
| -6        | 0.900            | 0.873 | 0.869 |
| +7        | 0.950            | 0.901 | 0.884 |
| -7        | 0.855            | 0.873 | 0.869 |
| +8        | 0.950            | 0.927 | 0.884 |
| -8        | 0.864            | 0.848 | 0.857 |
| +9        | 0.950            | 0.900 | 0.876 |
| -9        | 0.873            | 0.848 | 0.857 |
| +10       | 0.930            | 0.900 | 0.876 |
| -10       | 0.861            | 0.840 | 0.845 |
| Mean      | 0.908            | 0.884 | 0.859 |

## V. WEATHER-RELATED RISK MANAGEMENT WITH HELP OF PROPOSED INTERPRETABILITY

Compared with the previous ML risk prediction models that output only risk levels, the proposed model can further reveal the relationship between weather variables, which considerably helps in developing risk management plans. In this section, we describe in detail how the proposed interpretability helps weather-related risk management.

Interpretability includes the “importance of weather variables”, “contribution of weather variables”, and “threshold of weather variables”. As Fig. 6 shows, these three aspects of interpretability can help create weather-related risk management plans from different perspectives. The “importance of weather variables” can help create a long-term plan, which aims to escalate weak points in distribution systems on long-term scales. The “contribution of weather variables” can provide guidance information for making short-term plan, which aims to take timely measures to mitigate risks in the next prediction period. The “threshold of weather vari-

ables” can help develop high-risk prevention plans when the threshold is set for high risks.

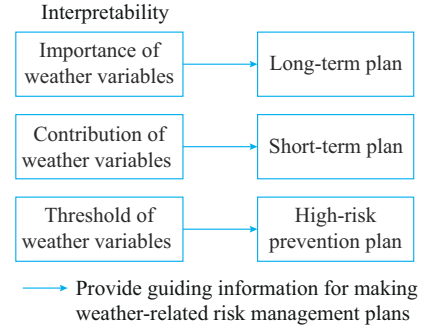


Fig. 6. Different guidance functions of three aspects of interpretability.

### A. Interpretation 1: Importance of Weather Variables

The importance of weather variables reflects the influence of each weather variable on the severity of risk, which helps inform the development of long-term plan for utility companies. The calculated importance of the weather variables in the studied city is shown in Fig. 5. The maximum wind speed has the most critical effect on the risk level, followed by the average wind speed and number of thunder days within a week. Therefore, when developing the long-term plan in the studied city, attention should be paid to reinforcing the weaknesses in distribution systems related to the failures caused by wind and thunder. These include gradually increasing the cable rate and replacing old insulators in conjunction with outage plans.

### B. Interpretation 2: Contribution of Weather Variables

For the development of short-term plans, the contribution of weather variables, which is produced dynamically during each prediction, is useful. The interpretation can provide dynamic guidance information by creating a targeted operation and maintenance plan for the next prediction period.

As previously stated, a positive contribution value indicates that the weather variable has a facilitating effect on the predicted risk, whereas a negative value indicates a hindering effect on the predicted risk. Under the attempts to prevent weather-related failure, more attention should be given to weather variables that contribute to the emergence of risk. Thus, decision makers can develop more specific prevention strategies based on the different contribution values of the variables. For example, in Fig. 7(a), the average and maximum wind speeds have the leading positive contribution values, which are actually the causes of risk. Based on this interpretability, utility companies can take precautionary operation and maintenance measures against the wind such as checking the objects that could be blown by the wind and endanger the safety of lines.

It is worth mentioning that when the model misjudges the risk as risk-free, the contribution of weather variables can still help in risk prevention. Errors inevitably occur in prediction due to the randomness of weather-related failures. In this case, a negative contribution value of the weather variable would indicate that the weather variable hinders the

risk-free prediction results. Therefore, it is instructive to focus on weather variables with large negative contribution values for ex-ante operation and maintenance decisions when the risk is underestimated as risk-free, and they are most likely to be the causes of the risk. This means that when the prediction model indicates risk-free, utility companies have avenues to further stifle the risk by taking reasonable risk prevention measures with respect to weather variables with large negative contribution values. As an example, in Fig. 7(b), the actual risk level is level 1 but is incorrectly predicted as risk-free. The contribution of the maximum wind speed has the largest negative value, which is consistent with the cause of failure. In this case, wind-prevention measures may prevent failures.

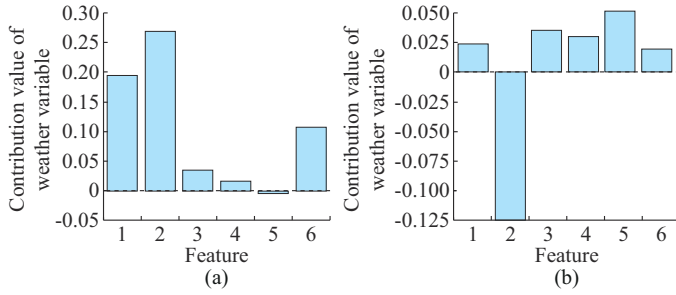


Fig. 7. Examples of application analysis of contribution values of weather variables. (a) Sample occurred from 2018-06-04 to 2018-10-06. (b) Sample occurred from 2017-10-23 to 2017-10-26.

Statistically, when the largest negative contribution value corresponds to the true cause of failure, and when we assume that taking measures in advance can effectively avoid failure, the *RUR* can be further reduced from 3.191% to 2.128%. Therefore, this interpretability provides a targeted prevention direction for the situations underestimated as risk-free, further improving the risk-management capabilities of distribution systems.

For the validity of the interpretability, it is crucial to realize that the results are not artifacts of one particular realization of an extra-tree model but that they convey actual information held by the data [24]. Therefore, we propose a method for robustness analysis of variable contributions. We remove this instance from the original data set to allow us to perform tests with an unseen instance. We generate 100 extra-tree models with each model built using an independent randomly generated training set, where the number of samples is 2/3 of the original training data set. For each generated model, we collect the contribution values of weather features for this instance, which is shown in Fig. 8. These results are qualitatively compared with those obtained in Fig. 7(a). Hence, we can conclude that weather variable contributions computed for an unseen instance provide reliable information.

### C. Interpretation 3: Threshold of Weather Variables

In risk management, more attention should be given to high risk (level 2) because its lower frequency and insufficient learning samples make the prediction difficult. Simultaneously, it has a more severe effect on distribution systems.

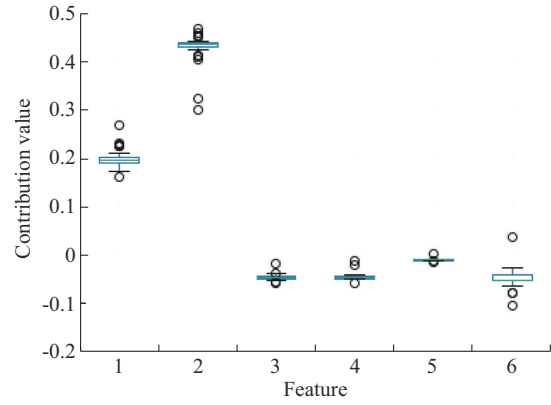


Fig. 8. Boxplot of contribution values of weather features for an instance.

When the contribution value of weather variables to high risk is positive, the occurrence probability of high risk increases. When the contribution value changes from negative to positive, the transition point can be used as a warning threshold. The contribution is related to the values of weather variables. Therefore, we can analyze the patterns of contributions of weather variables at high risk and then obtain the threshold of weather variables that trigger high risk, thereby providing quantitative reference information for high-risk management.

When the value of a single weather variable exceeds the captured threshold, the weather variable begins to make a positive contribution to the occurrence of high risk. If the positive contribution accumulates to a certain level, a high risk will occur, and only a single variable exceeding the threshold will be insufficient to warn of high risk. However, when each weather variable exceeds the captured threshold simultaneously, it can be used as a quantitative early-warning signal of high risk. When all weather variables begin to contribute positively to high risk, multiple weather factors facilitate the occurrence of high risk, indicating a complex weather situation. The probability of high risk is greatly increased. The theoretical reasons are as follows.

After we endow the extra-tree prediction model with intrinsic interpretability, the process of one prediction can be expressed by (2). Therefore, the proposed model can be expressed by the following equation in our application.

$$f(x) = V_{root} + c_1 \cdot feature_1 + c_2 \cdot feature_2 + c_3 \cdot feature_3 + c_4 \cdot feature_4 + c_5 \cdot feature_5 + c_6 \cdot feature_6 \quad (5)$$

where  $c_1$ - $c_6$  represent the contribution values of weather features  $feature_1$ - $feature_6$  to a certain predicted risk level.

Considering that the values of  $feature_1$ - $feature_6$  are all greater than or equal to 0 and that the trained value of  $V_{root}$  is also greater than 0, if the contribution value of each weather variable to high risk is also greater than 0 at this time, multiple weather factors will simultaneously contribute positively to the occurrence of high risk, and the occurrence probability of high risk will be high. Figure 9 shows the relationship between the contribution value of weather variables to high risk and weather variable values. It is known that as the value of a weather variable increases, its contribution to high risk also increases. Therefore, we use Lowess interpola-



tion to depict this volatile linear trend [32]. We could then capture the corresponding threshold of the contribution value from negative to positive. The threshold values of weather variables at high risk are listed in Table V.

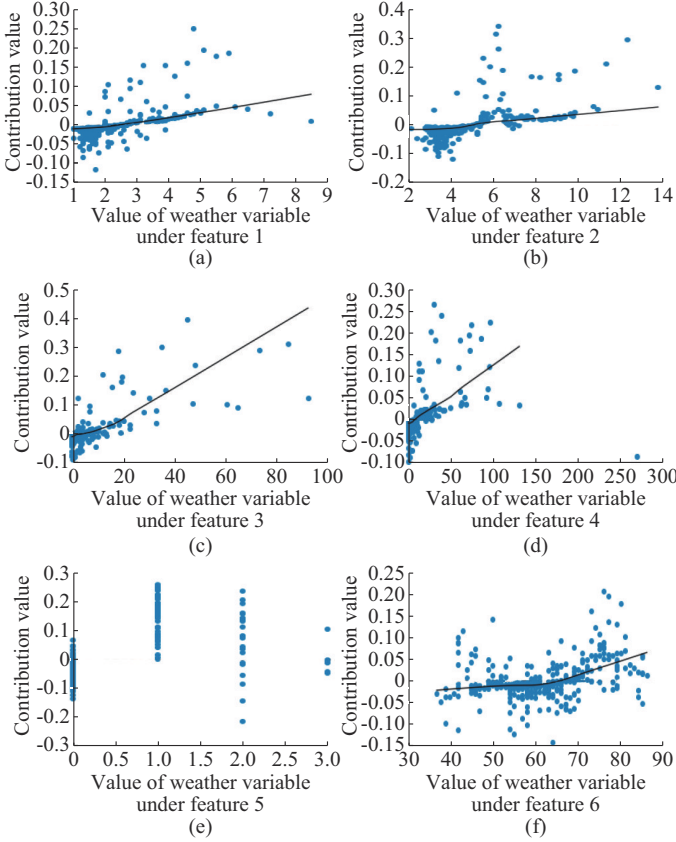


Fig. 9. Relationship between contribution value of weather variables to high risk and values of weather variable. (a) Feature 1. (b) Feature 2. (c) Feature 3. (d) Feature 4. (e) Feature 5. (f) Feature 6.

TABLE V  
THRESHOLD VALUES OF WEATHER VARIABLES AT HIGH RISK

| Feature | Threshold value | Precision_single |
|---------|-----------------|------------------|
| 1       | 2.3 m/s         | 0.737 m/s        |
| 2       | 5.1 m/s         | 0.842 m/s        |
| 3       | 5.2 mm          | 0.842 mm         |
| 4       | 8 mm            | 0.895 mm         |
| 5       | 1 day           | 0.895 day        |
| 6       | 64 %rh          | 0.642 %rh        |

We test the effectiveness of the proposed threshold in the high-risk data set. We chose the precision metric to verify the effectiveness of the proposed threshold. The precision metric can measure the proportion of samples that meet the threshold criteria as high-risk samples, i.e., the probability of high-risk occurrence when exceeding the threshold. The precision metrics calculated by the threshold of a weather variable and the threshold of all weather variables as criteria of high risk are named as precision\_single and precision\_all (100%), respectively. We can find that for just a single weather variable meeting the threshold, there is no necessari-

ly high risk. When the threshold of all weather variables is exceeded at the same time, the probability of the high-risk occurrence is 100%. Therefore, the proposed threshold can serve as a quantifiable early warning signal for high risk, guiding utility companies in making high-risk prevention arrangements.

## VI. CONCLUSION

Predicting weather-related failure risk can provide useful guidance information for utility companies to develop ex-ante risk prevention plans. The interpretable extra-tree based weather-related risk prediction model is proposed, which has interpretability with three aspects to provide effective advice for risk prevention. Specifically, the importance of weather variables helps in making long-term operation and maintenance plans. The weather variables support the development of specific risk prevention plans prior to the next prediction period. The threshold of weather variables at high risk yields a quantitative high-risk prevention plan. The proposed model overcomes the limitations of black-box ML models, making the risk prediction model more practical, further improving the weather-related risk management capabilities of utility companies. In comparison with ML models that can provide the same degree of interpretability, the proposed model has the best weather-related risk prediction performance. In addition, the proposed model can provide a way to guide decisions on other prediction issues in power systems.

## REFERENCES

- [1] D. H. Vu, K. M. Muttaqi, A. P. Agalgaonkar *et al.*, "Recurring multi-layer moving window approach to forecast day-ahead and week-ahead load demand considering weather conditions," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 6, pp. 1552-1562, Nov. 2022.
- [2] H. Li, L. A. Treinish, and J. R. M. Hosking, "A statistical model for risk management of electric outage forecasts," *IBM Journal of Research and Development*, vol. 54, no. 3, pp. 1-11, May 2010.
- [3] X. Wei, J. Zhao, T. Huang *et al.*, "A novel cascading faults graph based transmission network vulnerability assessment method," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2995-3000, May 2018.
- [4] J. He, D. W. Wanik, B. M. Hartman *et al.*, "Nonparametric tree-based predictive modeling of storm outages on an electric distribution network," *Risk Analysis*, vol. 37, no. 3, pp. 441-458, Mar. 2017.
- [5] H. Liu, R. A. Davidson, D. V. Rosowsky *et al.*, "Negative binomial regression of electric power outages in hurricanes," *Journal of Infrastructure Systems*, vol. 11, no. 4, pp. 258-267, Dec. 2005.
- [6] S. R. Han, S. D. Guikema, S. M. Quiring *et al.*, "Estimating the spatial distribution of power outages during hurricanes in the Gulf coast region," *Reliability Engineering & System Safety*, vol. 94, no. 2, pp. 199-210, Feb. 2009.
- [7] H. Liu, R. A. Davidson, and T. V. Apanasovich, "Spatial generalized linear mixed models of electric power outages due to hurricanes and ICE storms," *Reliability Engineering & System Safety*, vol. 93, no. 6, pp. 897-912, Mar. 2007.
- [8] P. Kankanala, A. Pahwa, and S. Das, "Regression models for outages due to wind and lightning on overhead distribution feeders," in *Proceedings of 2011 IEEE PES General Meeting*, Detroit, USA, Jul. 2011, pp. 1-4.
- [9] P. Kankanala, A. Pahwa, and S. Das, "Exponential regression models for wind and lightning caused outages on overhead distribution feeders," in *Proceedings of 2011 North American Power Symposium*, Boston, USA, Aug. 2011, pp. 1-4.
- [10] P. Kankanala, S. Das, and A. Pahwa, "AdaBoost: an ensemble learning approach for estimating weather-related outages in distribution systems," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 359-

- 367, Jan. 2014.
- [11] P. Kankanala, A. Pahwa, and S. Das, "Estimation of overhead distribution system outages caused by wind and lightning using an artificial neural network," in *Proceedings of International Conference on Power System Operation & Planning (ICPSOP)*, Juja, Kenya, Jan. 2012, pp. 1-6.
  - [12] Y. Du, Y. Liu, X. Wang *et al.*, "Predicting weather-related failure risk in distribution systems using Bayesian neural network," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 350-360, Aug. 2020.
  - [13] I. Bratko, "Machine learning: between accuracy and interpretability," *Learning, Networks and Statistics*, vol. 382, pp. 163-177, Jan. 1997.
  - [14] B. Kim and R. Khanna, "Examples are not enough, learn to criticize criticism for interpretability," in *Proceedings of the 30th International Conference on Neural Information Processing Systems*, Barcelona, Spain, Dec. 2016, pp. 2288-2296.
  - [15] D. V. Carvalho, E. M. Pereira, and J. S. Cardoso, "Machine learning interpretability: a survey on methods and metrics," *Electronics*, vol. 8, no. 8, pp. 832-838, Jul. 2019.
  - [16] C. Molnar. (Aug. 2019). Interpretable machine learning: a guide for making black box models explainable. [Online]. Available: <https://christophm.github.io/interpretable-ml-book>
  - [17] F. Doshi-Velez and B. Kim. (2017, Mar.). Towards a rigorous science of interpretable machine learning. [Online]. Available: <https://arxiv.org/abs/1702.08608>
  - [18] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3-42, Mar. 2006.
  - [19] C. Desir, C. Petitjean, L. Heutte *et al.*, "Classification of endomicroscopic images of the lung based on random subwindows and extra-trees," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2677-2683, Sept. 2012.
  - [20] A. Zhang. (2021, Dec.). Explainable artificial intelligence. [Online]. Available: <http://statsoft.org/>
  - [21] R. J. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81-106, Mar. 1986.
  - [22] R. I. Lerman and S. Yitzhaki, "A note on the calculation and interpretation of the Gini index," *Economics Letters*, vol. 15, no. 3, pp. 363-368, Feb. 1984.
  - [23] G. Tam. (Sept. 2017). Interpreting decision trees and random forests. [Online]. Available: <https://engineering.pivotal.io/post/interpreting-decision-trees-and-random-forests/>
  - [24] A. Palczewska, J. Palczewski, R. M. Robinson *et al.*, "Interpreting random forest models using a feature contribution method," in *Proceedings of 2013 IEEE 14th International Conference on Information Reuse & Integration (IRI)*, San Francisco, USA, Aug. 2013, pp. 112-119.
  - [25] Y. Du, Y. Liu, Q. Shao *et al.*, "Single line-to-ground faulted line detection of distribution systems with resonant grounding based on feature fusion framework," *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1766-1775, Aug. 2019.
  - [26] Y. Sun, A. K. Wong, and S. K. Mohamed, "Classification of imbalanced data: a review," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 23, no. 4, pp. 687-719, Oct. 2009.
  - [27] B. Ci, "Confidence intervals," *Lancet*, vol. 1, no. 8531, pp. 494-497, Jan. 1987.
  - [28] G. Wang, T. Xu, T. Tang *et al.*, "A Bayesian network model for prediction of weather-related failures in railway turnout systems," *Expert Systems with Applications*, vol. 69, pp. 247-256, Oct. 2016.
  - [29] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5-32, Oct. 2001.
  - [30] C. Bouveyron, B. Hammer, and T. Villmann, "Recent developments in clustering algorithms" in *Proceedings of the 20th European Symposium on Artificial Neural Networks*, Bruges, Belgium, Apr. 2012, pp. 447-458.
  - [31] H. Kaur, H. S. Pannu, and A. K. Malhi, "A systematic review on imbalanced data challenges in machine learning," *ACM Computing Surveys (CSUR)*, vol. 52, no. 4, pp. 1-36, Jul. 2020.
  - [32] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 829-836, Apr. 1979.
- Ying Du** received the B.E. degree in electrical engineering from the Central South University, Changsha, China, in 2017. Currently, she is a Ph.D. degree candidate in Shanghai Jiao Tong University, Shanghai, China. Her research interests include energy security under more and more severe climate change, including weather-related risk prediction, resilience assessment, disaster evolution mechanism, disaster damage assessment, fault detection and power recovery after disasters in distribution systems.
- Yadong Liu** received the B.E. degree in electronic & information from China University of Geosciences, Wuhan, China, in 2004. He received the M.D. and Ph.D. degrees of electrical engineering from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2008 and 2012, respectively. Currently, he is an Associate Professor in School of Electronic Information and Electrical Engineering of SJTU. His research interests include fault location and diagnosis of power equipment.
- Yingjie Yan** received the B.E. and Ph.D. degrees from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2018. He is currently with the Department of Electrical Engineering, SJTU. His research interests include online monitoring of electrical equipment and transmission lines in distribution network.
- Jian Fang** is a Senior Engineer with the Guangzhou Power Supply Bureau Co., Ltd. Electric Power Test Institute, Guangzhou, China. His main research interests include distribution network operation management, distribution network technology supervision, and distribution network new technology promotion.
- Xiuchen Jiang** received the B.E. degree from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 1987, the M.D. degree of electrical engineering from Tsinghua University, Beijing, China, in 1992, and the Ph.D. degree of electrical engineering from SJTU, in 2001. Currently, he is a Full Professor at SJTU, acting as the Director of Electrical Power Department at SJTU since 2002. His main research interests include electrical measuring technology and electrical apparatus automation.