Lightning risk assessment model for transmission lines with lift-based improved analytic hierarchy process

Yuesheng Zheng1 | Zhiwei Wu1 | Shengwen Shu1 | Jun Xu2 | Chaoying Fang2 | Wenbing Xie2

1 College of Electrical Engineering and Automation, Fuzhou University, Fuzhou, People’s Republic of China
2 Electric Power Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou, People’s Republic of China

Correspondence
Shengwen Shu, College of Electrical Engineering and Automation, Fuzhou University, Xueyuan Road No. 2, Fuzhou 350108, P.R. China.
Email: shushengwen@fzu.edu.cn

Funding information
Joint Project of Science and Technology Innovation Strategy Research of Fujian Province, Grant/Award Number: 2020R0172; Young and Middle-aged Teachers Education Scientific Research Projects of Fujian Province Education Department, Grant/Award Number: JAT190040

Abstract
This paper proposes a new assessment method for lightning trip-out of transmission lines that includes an improved analytic hierarchy process (IAHP) algorithm based on a lift in association rules. First, the k-means clustering algorithm is used to discretize the geomorphic information along transmission lines, and ten feature factors are obtained. Next, by analysing the correlation between each feature factor and a lightning trip-out event using the lift algorithm, the strong correlation factors are extracted. On this basis, an IAHP algorithm optimized by lift is proposed to establish an objective judgment matrix. Then, multiple indexes are combined as a criterion layer of the IAHP, and a lightning risk assessment model of transmission lines is established. Finally, the proposed method is validated by simulations using 28 transmission lines with 220 kV voltage grade of a provincial power grid of China. An assessment model is established using the lightning trip-out data in the period 2010–2014, and then this model is employed to estimate the lightning trip-out risk in the period 2015–2018. Based on the number of phase faults, the agreement between the estimation results and the actual values is 76%. The proposed model provides a new method for the lightning risk assessment of transmission lines.

1 INTRODUCTION

Among the disasters that affect the safe operation of the power grid, such as typhoons, ice disasters, and bird damage, lightning represents the greatest threat. According to the statistics of the State Grid Corporation of China (SGCC), in regions with frequent lightning strikes in China, such as Fujian, the power network suffers from about 100,000 lightning strikes every year. Also, the line fault tripping caused by lightning accounts for 30% of all disasters and even 60% in some provinces, so the lightning trip-out risk must not be ignored [1]. The overhead transmission lines are mostly distributed in mountainous areas with a high soil resistivity and local topography, and the lightning protection performance of the tower itself is restricted, which results in frequent lightning strikes. Risk assessment is a precondition for the governance of lightning trip-out events on transmission lines. However, before deploying the lightning protection measures, it is necessary to conduct certain lightning strike assessments.

At present, there are two main methods for lightning risk assessment, physical analysis and data mining. In the physical analysis, a large number of formulas are used to describe a quantitative relationship between the lightning trip-out rate and various influencing factors. For instance, the regular method proposed in the regulation DL/T 620 ‘Overvoltage Protection and Insulation Coordination for AC Electrical Installations’ [2], the classic electrical geometric model (EGM) [3], the Monte Carlo method [4], and the leader progression model (LPM) [5, 6]. All these methods are based on the physical modelling of the lightning strike mechanism. Besides, many optimizations and
improvement models have been proposed [7, 8], which lay a foundation for the research on risk assessment of lightning disaster.

Physical models for lightning shield analysis of towers represented by the EGM have been widely used in the engineering field, but it is difficult to consider the influence factors, such as topography and environment, comprehensively when these models are used. At the same time, with the continuous accumulation of lightning strike data of transmission lines, the advantages of data-driven models based on the analytic hierarchy process (AHP) in risk assessment have gradually emerged. These models have been practically used in the line lightning management, such as selection of lightning protection measures [9] and evaluation of trip-out risks [10, 11].

Although it is one of the most commonly used assessment methods, the AHP has two main shortcomings. The first shortcoming is that the considered influencing factors are not comprehensive. The essence of the AHP method lies in the mixed ordering of multiple factors used to conduct a comprehensive analysis, which can be difficult achieved using physical models. However, at present, most factors used in the criterion layer for lightning risk assessment of transmission lines are obtained from the history data of lightning strikes in the production management system. These factors include the unscheduled downtime, number of reclosing, the importance level of lines, the operating year of lines, and many other factors. The problem with traditional lightning risk assessment is that the feature factor selection is ambiguous, and the application of topography and transmission line property is ignored. Another shortcoming is that the weights are calculated subjectively, so the establishment of a judgment matrix depends on the group decision-making experience. Even when the nine-point scale of psychological limits is improved to the three-point scale, the core high-weight matrix is still highly dependent on expert experience, thus affecting model performance significantly.

In order to overcome the two shortcomings, this paper proposes a lightning risk assessment model for transmission lines with a lift-based improved analytic hierarchy process (IAHP). Aiming at overcoming the defect of subjectivity, each factor is subjected to the grid extraction and $k$-means clustering processing successively, and then the corresponding lift is calculated by the lift algorithm using the association rules to order factors at each layer. In order to address the problem of incomplete factors, a set of lightning disaster-pregnant factors is introduced. This set consists of six environmental feature factors, which are elevation, pollution level, pollution source, topography, underlying surface, and slope position, and four ontology feature factors of a transmission line, which are tower height, tower weight, tower structure, and horizontal span. In combination with five types of traditional assessment criteria for lightning tripping, an IAHP model with a total of 15 criteria is established. The proposed model is verified by the provincial power grid of China. The results show that the proposed model can identify a line with a high risk of lightning tripping more effectively than the traditional AHP model. Thus, the proposed model can provide a reference for the lightning risk assessment of transmission lines.

### Table 1: Reference value of ground resistance

| Soil resistivity (Ω·m) | ≤100 | 100–500 | 500–1000 | 1000–2000 | >2000 |
|------------------------|------|---------|----------|-----------|-------|
| Grounding resistance (Ω) | 10   | 15      | 20       | 25        | 30    |

According to the regulation DL/T 620 ‘Overvoltage Protection and Insulation Coordination for AC Electrical Installations’ [2], for lines with lightning conductors, the power frequency grounding resistance of a base tower must meet the values listed in Table 1. At an elevation of 1 km or less, the minimum number of suspended insulator strings required at a 220-kV voltage level is 13 pieces. The lightning withstand level mentioned in this paper refers to the one defined in DL/T 620. The lightning withstand level I1 of the 220 kV lines is 75–110 kA, and the probability of lightning current exceeding I1 is 6–14%.

In the model established, the grounding resistance of every tower and insulation level of a transmission line should meet the above-mentioned standard. This paper considers lines where protective gaps and arc horns are not widely used. These factors are not included in the proposed model as feature factors. The environment along a transmission line is complex, which makes the temperature and humidity data in the area change greatly day and night, resulting in limited effective data. In addition, the environmental features including humidity and temperature are incomplete, due to a small number of on-line monitoring devices. Therefore, the temperature and humidity factors are not considered.

The remainder of the article is organized as follows. The theoretical basis of risk assessment is provided in Section 2. The lightning risk assessment model based on the IAHP is introduced in Section 3. The case studies implemented into a provincial power grid of China are presented in Section 4. Finally, the conclusions are given in Section 5.

### 2 | THEORETICAL BASIS OF RISK ASSESSMENT

#### 2.1 | GIS gridding extraction

In order to extract the feature parameters of transmission towers accurately, a gridding method based on the geographic information system (GIS) is used, as shown in Figure 1. In Figure 1, blue pushpins and lines denote the transmission line and towers, respectively, and yellow pushpins and line respectively denote the GIS grid and points with a spatial resolution of 3 km that contain the space shuttle mapping data of the area, such as elevation and underlying surface.

The extraction process mainly consists of two main steps, which are explained in the following.

**Step 1:** Take the longitude and latitude of a target tower as a centre (point $T_1$ in Figure 1), and set a square area with a length slightly greater than 3 km (white dashed box in Figure 1) to ensure at least a grid point can be searched.
Step 2: Calculate the distance between each grid point and the target tower to obtain the located GIS grid of the tower, and then use the mapping data of this GIS grid as the target tower data.

The original and extracted GIS grids of towers along a 220-kV transmission line are presented as yellow squares in Figures 2(a,b), respectively. A rectangle region that has 886 GIS grids is needed to cover all 409 towers along the transmission line. After extraction, only 73 GIS grids are needed to obtain the mapping data used for lightning risk assessment of the transmission line, which accounts for 8% of the initial number of grids.

2.2 K-means clustering method based on fusion of contour coefficient and elbow method

The $k$-means method represents an unsupervised clustering algorithm proposed by Macqueen, J. [12]. In order to perform frequent pattern mining and analysis, continuous data of lightning disaster-pregnant environment indexes should be discretized. When there is no obvious classification intention or criterion, contour coefficient [13] and elbow method [14] have been generally used to evaluate the number of categories so as to determine the optimal number of clusters, which represents the value of parameter $k$. According to Ketchen, D. J. [15], using more than two processing methods can significantly improve classification accuracy and validity. Hence, this study adopts the contour coefficient and elbow method to evaluate the value of $k$ comprehensively.

The combined evaluation process can be divided into three steps as follows.

Step 1: Set $k=2$ and use the elbow method to obtain the distortion value curve from $k=2$ to $k=10$, and then normalize the data on the vertical axis in the interval of $[0, 1]$. Confirm the corresponding range of $k$ values according to the elbow position.

Step 2: Calculate the average sample contour coefficients from $k=2$ to $k=10$, sequentially, and find several top $k$-values that have a coefficient value close to one.

Step 3: Take the intersection of $k$ values obtained in steps 1 and 2 as a final number of clusters.

2.3 Lift-based IAHP

The AHP method was first proposed by Saaty, T. L. [16]. Many improved and modified AHP methods have been proposed [17, 18]. In this paper, a lift-based IAHP is proposed to assess the lightning risk of transmission lines. The flowchart of the proposed method is presented in Figure 3.

The specific steps of the lift-based IAHP algorithm are as follows.

Step 1: Calculate lift value. Lift represents a correlation measure obtained by the support-confidence framework. Compared with the traditional association rule analysis, the lift value can reveal more internal connections between frequent patterns [19], and it can be calculated by:

$$
\text{lift}_{A \Rightarrow B} = \frac{P(B|A)}{P(B)} = \frac{\text{confidence}_{A \Rightarrow B}}{P(B)} \\
\text{confidence}_{A \Rightarrow B} = \frac{P(A \cap B)}{P(B)} = \frac{\sigma(A \cup B)}{\sigma(B)}
$$

where lift$_{A \Rightarrow B}$ denotes the lift value between A and B, confidence$_{A \Rightarrow B}$ denotes the confidence value between A and B, and $\sigma$ represents the number of incidents; A represents a feature factor, and B represents a lightning strike event. When lift = 1, there is no correlation between occurrences of events A and B. For the convenience of analysis, this study defines the lift value intervals as given in Table 2. When lift > 1, there is a positive
correlation between occurrences of events A and B. The specific range in Table 2 is divided according to experience. The lift value range in Table 2 only used in the model proposed in this paper.

Since positive and negative properties of lift are bounded by one, by subtracting one from a calculated lift value, each discrete label can be replaced, and it is referred to as ‘minus-one value (MO)’. After averaging the MO values, the absolute average value, which is referred to as the ‘mean-absolute-one value (MAO)’, is determined in order to characterize positive and negative correlations between the feature factors and a lightning trip-out event. The lift analysis can convert sample labels under the feature factors into specific values for comparison, thus making the scale selection process more objective.

Step 2: Construct a scale matrix for the criterion layer. For target \( T \) that has \( v \) influencing factors, the scale matrix \( W = (w_{ij})_{v \times v} \) is constructed according to the MAO value of each feature factor set, and it can be expressed as:

\[
W = \begin{pmatrix}
    w_{11} & w_{12} & \ldots & w_{1v} \\
    w_{21} & w_{22} & \ldots & w_{2v} \\
    \vdots & \vdots & \ddots & \vdots \\
    w_{v1} & w_{v2} & \ldots & w_{vv}
\end{pmatrix}
\]

(2)

where \( w_{ij} \) represents the scale of factor importance, and the specific rules are given by:

\[
w_{ij} = \begin{cases} 
0 & \text{MAO}_i > \text{MAO}_j \\
1 & \text{MAO}_i = \text{MAO}_j \\
2 & \text{MAO}_i < \text{MAO}_j 
\end{cases}
\]

(3)

Step 3: Calculate the criterion layer weight. By the two previous steps, the scale matrix \( W \) containing the ranking information of \( v \) influencing factors can be established, and then the scale matrix can be converted into a consistent judgment matrix by the range method [17], which is defined by:

\[
l_i = \sum_{j=1}^{v} w_{ij}
\]

\[
f(r_i, r_j) = \frac{(r_i - r_j)}{(r_{\text{max}} - r_{\text{min}})}
\]

(4)

where \( z_{ij} \) denotes the coefficient representing the relative importance of range elements, and it is equal to nine in this study. After the transformation, a consistent judgment matrix \( Z = (z_{ij})_{v \times v} \) can be obtained. Since matrix \( Z \) conforms to consistency, only the eigenvector corresponding to the maximum eigenvalue needs to be calculated, and the influence weight \( O_{vq} = \{O_{11}, O_{12}, O_{13}, \ldots, O_{1q}\} \) representing \( q \) influence factors on target \( T \) can be obtained after normalization.

Step 4: Calculate the scheme layer weight. The scheme layer adopts the quantitative AHP to calculate the weight and converts the value of each sample into a judgment matrix to reflect the influence of each sample on the criterion layer objectively. For a certain influencing factor \( f \) of the criterion layer, a total of \( q \) scheme values are determined, \( J = \{J_1, J_2, J_3, \ldots, J_q\} \), where the scheme value refers to the lift value corresponding to the feature factor in a scheme. Then, a judgment matrix \( Y = (y_{ij})_{p \times q} \) can be formed similar to Equation (2), where \( y_{ij} \) denotes the numerical ratio of scheme values \( J_i \) and \( J_j \), and it is calculated by:

\[
y_{ij} = \begin{cases} 
J_j/J_i & J_j/J_i > 0 \\
0 & J_j = 0 \text{ or } J_i = 0 \\
J_i/J_j & J_i/J_j < 0
\end{cases}
\]

(5)

The consistency of matrix \( Y \) was proven in [17]. Therefore, the weight set \( O_{pq} = \{O_{11}, O_{12}, O_{13}, \ldots, O_{pq}\} \) consisting of \( q \) schemes of \( p \) factors can be obtained by normalizing the eigenvector corresponding to the maximum eigenvalue. After the above four steps, the weights of each of the layers can be established.

### 3 | PROPOSED ASSESSMENT MODEL

The lightning trip-out risk assessment model of transmission lines based on the optimized lift-based IAHP includes five steps, as shown in Figure 4.

1. Data pre-processing. Data quality is improved by the GIS grid information extraction algorithm. The continuous data are then clustered by the \( k \)-means algorithm using the fusion of contour coefficient and elbow rule. Lastly, a set of disaster-pregnant environmental feature factors is constructed.

2. The quantitative association rules are used to conduct frequent pattern mining, and then the lift algorithm is employed to analyse the correlation between the feature factor and lightning trip-out. Next, each discrete label is converted to the MO and MAO values based on the calculation result.
3. The feature factors are ranked based on the MAO values obtained in the previous step, and the three-scale method is employed to model the target-criterion layer. Lastly, the range method is used to transform the scale matrix to calculate the weight.

4. The sample labels are converted into MO values to establish the judgment matrix, and the weight vector is calculated by the quantitative AHP algorithm for the criterion-scheme layer modeling.

5. The traditional lightning assessment indexes are first obtained by the three-scale AHP method and then integrated with the disaster-pregnant environment indexes. A lightning trip-out risk assessment model with 15 factors is established, and the transmission lines are categorized according to their lightning trip-out risk values.

4 | CASE STUDY

4.1 | Data collection

The experimental data consisted of the ten-year data of a 220-kV transmission line production management system of Fujian province, China. This system included a total of 35,756 base towers. The data were collected in the province that had had frequent lightning strikes. In the period of 2010–2018, there were 4738 trip-out events, of which 3162 were caused by the lightning strike, which accounted for 66.7% of the total strike number. In recent years, the lightning trip-out events in this province have shown a clear upward trend. Namely, lightning strikes occur frequently, so lightning disasters have been in the focus of the operation and maintenance departments. The total number of lightning strikes in this province in 2018 was 579,083, with a flash density of 4.645 times per square kilometre, which represented an increase of 48.2% over the same period the year before, which was 3.135 times per square kilometre.

The lightning strike records of the 220-kV transmission line of Fujian province from 2010 to 2018 were used for model development. Four types of transmission-line ontology feature factors, namely, tower structure, tower height, tower weight, and horizontal span, were used. There were six environmental features: elevation, pollution level, topography, pollution source, slope position, and underlying surface. The ten environmental indexes used in the lightning risk assessment are given in Table 3. The traditional lightning assessment indexes included five common evaluation parameters of power grids: unscheduled downtime, number of operation years, lightning reclosing times, importance level, and lightning monetary loss. The data involved a total of 32,852 data samples, and there were small numbers of missing data samples and over-range data samples.
TABLE 3  Data used for model development

| Data classification       | Affecting factor | Data source                                                                 |
|---------------------------|-----------------|-----------------------------------------------------------------------------|
| Environmental feature factors | Elevation     | PMS, GIS system, OMAP software (combines the "Atlas of Polluted Areas in Power System", "Revision of the Pollution Atlas in Power System", "Current Land Use Classification", and "Series and basic requirements of digital topographic maps") |
|                           | Pollution level | GIS system combined with the grid extraction algorithm                      |
|                           | Topography      |                                                                             |
|                           | Pollution source|                                                                             |
|                           | Slope position  |                                                                             |
|                           | Underlying surface|                                                                             |
| Transmission line ontology feature factors | Tower form | PMS, GIS system combined with the grid extraction algorithm, "Rules of Nomenclature for Transmission Poles and Towers" |
|                           | Tower height    |                                                                             |
|                           | Tower weight    |                                                                             |
|                           | Horizontal span |                                                                             |

The common evaluation parameters of the traditional lightning assessment indexes were retrieved from the lightning trip-out records. The main record included the information on voltage level, line name, reclosing condition, trip-out time, recovery time, trip-out cause, fault tower number, and trip-out report, as presented in Table 3.

When different classification methods are used, there can appear differences in the number of samples falling into intervals, which affects the accuracy of analysis results. In order to avoid the appearance of such differences, sample labels should be corrected according to the relevant national standard.

For instance, in the collected data, in terms of pollution level at the tower, there was an error interference label that referred to the old standard GB/T 16434 entry system. Thus, it was necessary to use the latitude and longitude of the tower to re-determine the pollution level by the OMAP software in conjunction with the province’s atlas of polluted areas in the power system.

The pollution source conditions mainly include road dust, industrial gas, vehicle exhaust gases, household smoke, salt fog, and mountain air pollution. When there are multiple sources of environmental pollution, new labels should be determined. The tower topography can be mainly divided into mountainous areas, plains, gardens, fields, river networks, and forest areas. According to the national standards for land use classification and topographic maps of China [20–21], a small number of irregular labels can be merged. In terms of the tower structure, the classification was made according to the China Power Industry Standard [22]. The horizontal span, tower height, tower weight, and elevation were continuous data, so the discrete classification was required.

4.2  Discrete and lift analysis

The evaluation results of the tower height are shown in Figure 5. As presented in Figure 5, when \( k \) was equal to four, the distortion value of the elbow curve greatly decreased, and the elbow area was \((3, 6)\). The first five points whose contour coefficient values were close to one were in the following order \( k = \{10, 6, 4, 3, 2\} \). The intersection of the two methods’ results was at \( k = 4 \). When \( k = 4 \), the contour coefficient value was 0.6752, and the value of \( k \) was within the elbow area.

Similarly, the tower weight, horizontal span, and elevation data were discretized by the hybrid evaluation method. The obtained intervals are shown in Table 4.

After completing the discrete step, according to Equation (1) combined with the data of the tower account and the geomorphological information along the route, the lightning strike...
TABLE 5  Lift analysis of lightning disaster-pregnant factors

| Feature factor      | Average lift | MAO value |
|---------------------|--------------|-----------|
| Pollution source    | 4.687        | 3.687     |
| Tower structure     | 2.342        | 1.342     |
| Topography          | 1.924        | 0.924     |
| Pollution level     | 1.323        | 0.323     |
| Horizontal span     | 1.210        | 0.210     |
| Slope position      | 0.849        | 0.151     |
| Tower weight        | 0.893        | 0.107     |
| Tower height        | 1.095        | 0.095     |
| Elevation           | 1.090        | 0.090     |
| Underlying surface  | 1.063        | 0.063     |

FIGURE 6  Lift value of high-correlation feature factors

records of 28,220-kV transmission lines in the Fujian province in the period of 2010–2018 were analysed, and the relationship between the lightning disaster-pregnant factors and lightning trip-out event was determined. The MAO values of ten types of lightning disaster-pregnant factors are given in Table 5.

As given in Table 5, lightning strike trip-out had the greatest correlation with the four following factors: pollution source, tower structure, topography, and pollution level. The relationship between the high-correlation factors and lightning trip-out on the transmission lines is displayed in Figure 6.

The lift value can characterize the effect of a feature factor on a lightning trip-out event. The relationships between each label of the pollution source and the lightning trip-out of the transmission line and the corresponding lift values are presented in Figure 6.

As shown in Figure 6, there was a strong positive correlation between the food processing area and the lightning trip-out event, while the living area (with household smoke) had a strong negative correlation with the same event. This was because the household smoke came from densely populated living areas, and its lightning protection facilities were relatively complete and not prone to lightning trip-out. In contrast, the food processing area has an impact on the environment, so an analysis was required in conjunction with the pollution level at the tower.

In Figure 6, c1, c2, d1, and d2 represent different levels of environmental pollution, rising in turn from c1, c2, d1, and d2, where d2 represents the most severe pollution. The environmental pollution level was divided according to the equivalent salt density. The results showed that as the pollution level increased, the lift value of the lightning trip-out event significantly increased, showing a strong positive correlation. It should be noted that $U_{50}$ means that under the impulse voltage of $U$ the gap has a 50% probability of breakdown. In [23], it was shown that as pollution level increased, $U_{50}$ voltage value of a dirty insulator string reduced by 25% compared to that of a clean insulator, and it could be easily broken down, which increased the occurrence of lightning trip-outs. The food processing area mainly produced a large amount of wastewater and biogas, making the organic matter and salt content in the soil soar, and thus affecting the level of environmental pollution and soil resistivity, thereby reducing lightning withstand level. Moreover, multi-source air pollution, such as exhaust gas, salt fog, smoke, and dust, could cause the pollution and aging of insulators to increase.

Regarding the tower structure, gantry towers have a strong positive correlation. In [24], it has been shown that different tower shapes may have a certain influence on SSFR and BFR performances of transmission lines.

Regarding the tower topography, the first five types of labels had a high correlation with lightning trip-out. In Figure 6, it can be seen that the peaks, hillsides, fields, bamboo areas, and gardens all show strong positive correlations with the lightning trip-out.

Due to the high terrain and soil resistivity and poor soil quality, the peaks and hillsides made it difficult to lay the ground grid and hinder the effectiveness of lightning protection facilities. Namely, fields and gardens were mostly in open areas, and there were no buildings or lightning protection facilities around them. Thus, pole towers in them were equivalent to small-scale bulges, which could easily accumulate electric charges and cause lightning strikes. Namely, the Moso bamboo grows mostly in hilly areas in moist environments, where the soil has high electrical conductivity and air humidity, which reduces the strike distance between the leader part of lightning and the pole tower, and thus can easily induce lightning strikes.

4.3  Risk assessment results of lightning trip-out

First, using the MAO values of each feature factor that are presented in Table 5, a three-scale matrix was established for a total of ten lightning disaster-pregnant factors, from the most to the least important (but still really important), as shown in Table 6. The range method was used to convert the matrix to a consistency matrix, and the eigenvector corresponding to its maximum eigenvalue was calculated. After data normalization,
the subordinate weight vector $W_E = \{w_{10}, w_{11}, w_{12}, \ldots, w_{19}\}$ was obtained; the obtained weight values are presented in Table 6.

The meaning of each of the weight vectors presented in Table 6 is described below:

- $w_{10}$—Types of pollution sources identified according to the sewage discharge, chemical oxygen demand, annual ammonia nitrogen emission, sulphur dioxide index, annual emissions of smoke and dust, and scale of livestock and poultry farming;
- $w_{11}$—Optimal tower structure selected according to the terrain, construction technology, and economic constraints;
- $w_{12}$—Vector referring to the topography of the area occupied by the corridor of overhead transmission lines;
- $w_{13}$—Environmental pollution level based on equivalent salt density;
- $w_{14}$—Vector referring to the distance between the midpoints of two adjacent spans; it is determined according to construction cost, tower usage conditions, conductor arrangement type, and terrain characteristics;
- $w_{15}$—Vector referring to the landform location where the slope is located and it is defined by the vertical position of a vertical section of the terrain slope;
- $w_{16}$—Vector referring to the tower weight, and it is related to the tower structure and transmission distance;
- $w_{17}$—Vector referring to the height of the tower, and it is determined by the voltage level of the line, terrain, and tower foundation;
- $w_{18}$—Vector referring to the height of the tower erection point counted from the geoid in the passing area of the transmission line corridor;
- $w_{19}$—Vector referring to the land type in contact with the ground in the lower atmosphere, including cities, river systems, farmland, grasslands, and other categories;

The traditional lightning assessment indexes differ significantly from the lightning disaster-pregnant indexes for lightning trip-out, and the former ones are more decisive. Therefore, based on operating experience and opinions of experts regarding the State Grid, the first layer of the lift-based IAHP model was ranked as follows: the traditional lightning assessment indexes were more important than the lightning disaster-pregnant indexes. The importance of traditional criteria was ranked as follows: unscheduled downtime was the most important, importance level of lines was the second-highest, lightning loss was the third-highest, while the operation years of lines was the last. The lightning loss could be further sorted as follows. The lightning monetary loss was more important than the lightning reclosing times. The first-layer weight vector $W_C = \{w_{20}, w_{21}\}$ that was calculated according to the ranking represented the traditional criterion $w_1$ and the disaster-pregnant criterion $w_0$. The traditional criterion $w_0$ subordinated vector $W_L = \{w_{22}, w_{23}, w_{24}, w_{25}\}$ and the lightning loss $w_{22}$ subordinated vector of $W_C = \{w_{30}, w_{31}\}$ are given in Table 7.

According to the regional lift analysis results, weight vectors $W_{C1}$, $W_{C2}$, and $W_{C3}$ of the target-criterion layer denoted sets of fixed values. The weight vector would change only if the assessment area changed so that the results of the lift analysis between the influencing factors changed accordingly. Then, the quantitative AHP algorithm was used to calculate the weight of the criterion-scheme layer. Afterward, the weights of each layer were calculated to establish the model.

Part of the risk assessment data of lightning trip-out for 28,220-kV mountain transmission lines in some areas of Fujian province are given in Table 8. The number of phase faults referred to the number of phase faults during the lightning trip, which represents the number of faults a line has experienced. For instance, a three-phase fault caused by lightning strikes was recorded as three, and a single-phase fault was recorded as one, and finally the numbers summed as the number of phase faults of a line. Most of 28 lines were double-circuit lines with similar attributes, so only some typical lines are shown in Table 8. The assessment was based on the operating data from 2010 to 2014, and the verification analysis was conducted using the number of phase faults during the lightning trip-out from 2015 to 2018. The assessment results are shown in Figure 7, where it can be seen that the JC, ZK II, and ZJ II lines did not have unscheduled downtime for a period between 2010 and 2014, and the operation years of lines were short.

The traditional standard model could not identify potential risks. However, in a ZJ II line #60 pole tower BC phase lightning trip failure in 2015, a reclosing failure occurred. The unscheduled shutdown lasted 92 min, and the accident was severe. For
TABLE 8  Part of the assessment area data

| Names of line | Year of operation | Main tower Structure | Pollution levels along the line | Topography along the line | Number of phase fault |
|---------------|------------------|----------------------|-------------------------------|---------------------------|----------------------|
| DZ II         | 1989             | Cat head             | c1                            | Mountain                   | 9                    |
| ZJ II         | 2012             | Cat head             | c1                            | Mountain                   | 7                    |
| ST II         | 2006             | Cat head             | c1                            | Peaks and mountain         | 5                    |
| MZ II         | 2001             | Drum                 | c1                            | Mountain                   | 5                    |
| JC            | 2009             | Cat head             | c1 and c2                     | Mountain                   | 5                    |
| SM II         | 2011             | Drum                 | c1                            | Mountain                   | 4                    |
| ZK II         | 2009             | Drum                 | c2                            | Mountain                   | 4                    |
| ZL I          | 2010             | Drum                 | c2                            | Mountain                   | 4                    |
| PX II         | 2010             | Umbrella and shofar  | c1                            | Mountain                   | 3                    |
| SM I          | 2011             | Drum                 | c1                            | Mountain                   | 3                    |
| TT II         | 2009             | Cat head             | c1                            | Mountain                   | 2                    |
| HP            | 2008             | Drum                 | d2                            | Mountain                   | 2                    |
| NT I          | 2004             | Cat head             | d1                            | Hills and mountain         | 1                    |
| QL I          | 1988             | Umbrella             | d1                            | Hills and mountain         | 1                    |

FIGURE 7  Risk ranking of lightning trip-out lines

The JC and ZK II lines, there were multiple two-phase and three-phase trip-out accidents between 2016 and 2018, respectively. This is because the amplitude of the lightning current was much higher than the lightning withstand level mentioned in DL/T 620, causing a lightning overvoltage higher than the withstand voltage of insulators.

The red part on the left side in Figure 7 represents the actual number of phase faults during the lightning strikes in the assessment area in the period 2015–2018, and the right side represents the assessment result obtained by the proposed lift-based IAHP model. From the perspective of the phase fault number, the lift-based IAHP model concentrated mostly on the lightning trip-out risk of a small number of lines and could identify high-frequency lightning trip-out lines that could not be recognized by the traditional lightning assessment models, such as ZJ II, JC line, and ZK II.

The number of lightning trip faults of each of the lines was calculated, and the top 35% of the lines with the most number of faults were selected, which corresponded to 10 out of 28 lines. That is, among the top 10 lines with the most frequent lightning trip-out in the assessment area from 2015 to 2018, 49 trip-outs occurred. Among them, 37 were in line with the assessment results of the lift-based IAHP model in the period of 2010–2014, and the percentage of phase fault matches was 76%. In contrast, the model adopting the traditional lightning assessment indexes achieved an agreement of 57%. Accordingly, if the first 18% of the lines were marked, the lift-based IAHP model would match 61%, while the traditional model would match only 42%. Thus, compared to the traditional model, the phase matching obtained by the proposed model was 19% higher. This optimization effect refers only to the model and data range studied in this paper.

5  CONCLUSION

This paper proposes an improved data processing method that includes grid extraction, continuous data clustering, and missing-data recovery, to correct the data that is not statistically significant. The lift algorithm based on the association rule is used to analyse the correlation between the lightning trip-out and lightning disaster-pregnant factors. The case study is conducted to verify the proposed model. The results show that
pollution source, pollution level, topography, and tower structure have a strong correlation with the lightning trip-out. Within the voltage level and assessment area given in this study, a risk assessment model of transmission line lightning trip-out using the lift-based IAHP algorithm is proposed. The actual operating data of Fujian province from 2010 to 2014 are used to train the model. The results show that the proposed model is flexible, and a series of correlations between environmental and ontology feature factors and lightning trip can be objectively obtained. The variable weight matrix is formed, which adds new factors to the lightning risk assessment model. The results show that the proposed model can achieve 19% better assessment results than the traditional method in the studied data range, which verifies the effectiveness of the proposed model.

In the following work, the voltage level and data range will be expanded to verify the superiority of the IAHP model over the traditional model. And some weights of factors in the model will be adjusted. The economic losses of factories and enterprises caused by lightning trips will be considered in lightning monetary loss, and the order of lightning monetary loss in the traditional criteria will be further discussed.

**REFERENCES**

1. Gu, S., et al.: Statistical and mining analysis of lightning detection data in power grid. High Volt. Eng. 42(11), 3383–3391 (2016)
2. DL/T 620: Overvoltage protection and insulation coordination for AC electrical installations (2016)
3. Brown, G.W., Whitehead, E.R.: Field and analytical studies of transmission line shielding II. IEEE Trans. Power App. Syst. PAS-88(5), 617–626 (1969).
4. Anderson, J.G.: Transmission Line Reference Book 345 kV and Above, 2nd ed. Electric Power Research Institute Press, Palo Alto (1982)
5. Dellera, L., et al.: Lightning stroke simulation by means of the leader progression model part I description of the model and evaluation of exposure of free-standing structures. IEEE Trans. Power Deliv. 5(4), 2009–2022 (1990).
6. Dellera, L., et al.: Lightning stroke simulation by means of the leader progression model part II exposure and shielding failure evaluation of overhead lines with assessment of application graphs. IEEE Trans. Power Deliv. 5(4), 2023–2029 (1990).
7. Silveira, F.H., Visacro, S.: Lightning performance of transmission lines: Impact of current waveform and front time on flashover occurrence. IEEE Trans. Power Deliv. 34(6), 2145–2151 (2019).
8. Dassios, Z.G., Mikropoulos, P.N., Tsiovlis, T.E.: Closed-form expressions for the estimation of the minimum backflashover current of overhead transmission lines. IEEE Trans. Power Deliv. 36(2), 522–532 (2021).
9. Wei, T., et al.: Study on the technicality and economy of lightning protection scheme for Multi-terminal VSC-DC distribution system. In: Asia Conference on Power and Electrical Engineering (ACPEE), Chengdu, China, pp. 1557–1562 (2020)
10. Xiao, B., et al.: Application research of lightning monitoring data in regional lightning disaster risk assessment. In: International Conference on Meteorology Observations (ICMO), Chengdu, China, pp. 1–7 (2019)
11. Du, Y., et al.: Risk assessment model of power network based on equipment condition and entropy combined weight method. In: IEEE Electrical Insulation Conference (EIC), Baltimore, USA, pp. 78–82 (2017)
12. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, UK, pp. 281–297 (1967)
13. Rousseeuw, P.: Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math 20(1), 53–65 (1987).
14. Goutte, C.: On clustering fMRI time series. NeuroImage 9(3), 298–310 (1999).
15. Ketchen, D.J., Shook, C.L.: The application of cluster analysis in strategic management research: An analysis and critique. Strateg. Manag. J. 17(6), 441–458 (1996).
16. Saaty, T.L.: Hierarchies and priorities—eigenvalue analysis. PhD thesis, University of Pennsylvania (1975)
17. Zhao, C.: Research on the Whole-level Evaluation and Harness System of Lightning Disaster Risk of Power System and Its Application. PhD thesis, Wuhan University (2012)
18. Zhao, C., Chen, J., Gu, S., et al.: Research on differentiated lightning protection comprehensive management for the 500-kV power network in the area near the three Gorges project. IEEE Trans. Power Deliv. 27(1), 337–352 (2012)
19. Han, J., Micheline, K., Pei, J.: DataMining: Concepts and Techniques, 3rd ed. China Machine Press, Beijing (2007)
20. GB/T 21010: Current Land Use Classification (2017)
21. GB/T 18315: Series and Basic Requirements of Digital Topographic Maps (2001)
22. DL/T 1252: Rules of Nomenclature for Transmission Poles and Towers (2013)
23. Pei, X., et al.: Experimental study on the flashover characteristics of polluted insulators under short-tail lightning impulse waveform. In: IEEE International Conference on High Voltage Engineering and Application (ICHVE), Athens, Greece, pp. 1–4 (2019)
24. Volpov, E., Linder, R.: Application of local environmental data to the lightning performance assessment of the IEEE transmission-line network. IEEE Trans. Power Deliv. 32(6), 2546–2554 (2017)

How to cite this article: Zheng Y, Wu Z, Shu S, Xu J, Fang C, Xie W. Lightning risk assessment model for transmission lines with lift-based improved analytic hierarchy process. IET Gener. Transm. Distrib. 2021;1–10 https://doi.org/10.1049/gtd2.12227. ORCID
Shengwen Shu https://orcid.org/0000-0003-0181-9563