Received 1 February 2024, accepted 8 April 2024, date of publication 12 April 2024, date of current version 22 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3387968

RESEARCH ARTICLE

Considering the Methods of Lightning Protection and Early Warning for Power Transmission Lines Based on Lightning Data Analysis

HONG HUO, DEZHI WANG, HAO CHEN, CHUNLEI ZHAO, AND QI CHENG
State Grid Dandong Electric Power Supply Company, Dandong 118000, China

Corresponding author: Hong Huo (huohong_2023@126.com)

This work was supported by the Science and Technology Project funded by State Grid Liaoning Electric Power Company Ltd., through the project titled "Research on Lightning Risk Assessment and Lightning Protection Configuration Optimization Technology for Transmission Lines Based on Improved Trip Rate Calculation Method," under Grant 2023YF-31.

ABSTRACT The lightning discharges generated by lightning strikes can reach several hundred kiloamperes, causing significant electromagnetic, mechanical, and thermal effects. These effects induce extremely high overvoltages in power transmission systems, leading to the tripping of transmission lines. Therefore, the study of lightning protection and early warning methods is crucial for safeguarding transmission lines. However, previous research in the realm of active lightning protection has been limited, focusing either solely on assessing the risk of lightning-induced tripping on the power grid side or merely predicting the patterns of lightning activity on the disaster side. This paper proposes a novel method for lightning protection and early warning of power transmission lines based on the analysis of lightning data. It encompasses both the risk of lightning-induced tripping of transmission lines on the grid side and the hazards of real-time lightning activities on the disaster side. In the warning phase, the method extensively mines historical lightning location data and, based on predictions of thunderstorm cloud movement trajectories, determines if thunderstorm clouds are nearing tightly packed transmission corridors. It then assesses whether these clouds present a high-risk storm or are approaching transmission lines with low lightning protection capabilities, subsequently integrating these findings to issue a continuous lightning-induced trip warning. In a case study conducted in a certain part of northern China, the proposed continuous lightning strike trip warning method achieved an accuracy rate of 80% for transmission corridors of particular interest to power grid companies' dispatch departments, enhancing the efficiency of warnings and reducing false alarms and missed reports.

INDEX TERMS Data-driven, lightning movement trajectory prediction, high-risk thunderstorm assessment, lightning-induced trip alarm warning.

I. INTRODUCTION

Lightning strikes impacting transmission lines are known to provoke disturbances within the power system [1].Such discharges can hit the equipment directly or hit transmission lines originating overvoltage surges sufficiently high to cause severe fails on equipment insulation and consequently to unexpected outages [2].In extreme scenarios, these incidents can inflict substantial financial losses on power utilities and businesses reliant on electricity [3]. The development and implementation of a robust lightning protection system, which includes line lightning rods, grounding systems, surge protectors, and overhead ground wires, have undergone continuous refinement and gained widespread acceptance in the global electric power industry [4], [5], [6], [7]. The persistent endeavor to diminish the frequency of lightning strikes on transmission lines and reduce the incidence of lightning-induced outages is a key objective for scientists, engineers, and practitioners in the electric power sector worldwide. This effort is crucial for ensuring the safety and stability of the power system's operation [8], [9], [10].

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

Numerous studies have embarked on risk analyses concerning lightning strike-induced trip-outs in transmission lines. In particular, Document [11] utilized ground flash density metrics to identify segments of transmission lines vulnerable to lightning. This study integrated these metrics with parameters characterizing line lightning faults to compute the frequency of lightning strike-induced trip-outs. The findings confirmed the utility of ground flash density in facilitating a differentiated assessment of lightning protection efficacy across various line segments. Meanwhile, Document [12] embarked on a quest to elucidate the underlying causes of lightning strike-induced trip-outs in transmission lines. This endeavor entailed a comprehensive multi-dimensional statistical analysis, taking into account variables such as the timing of lightning strikes, geographical regions, voltage levels of the affected lines, and the nature of the trip-outs. This analysis successfully pinpointed vulnerabilities in line lightning protection systems. Focusing on a specific case, Document [13] examined a 220kV multi-circuit transmission line mounted on a single tower to analyze lightning faults. The study revealed a significant correlation between ground flash density, lightning current intensity, and the incidence of line lightning strike trip-outs, thereby elucidating the direct causes of concurrent trip-outs attributable to lightning strikes. In an effort to address the limitations inherent in the lightning hazard distribution map methodology and the quantitative calculation of lightning strikes, Document [14] introduced a novel, differentiated, and multilevel lightning strike risk assessment method. This method was predicated on ground flash density values, adjusted for a variety of influencing factors. However, it is noteworthy that these methodologies predominantly cater to lightning protection assessments from the power grid perspective, omitting the influence of real-time lightning activities on the disaster side on the risk of trip-outs. Furthermore, prior research on lightning strike trip-outs has predominantly been qualitative in nature, broadly categorizing them as lightning-induced without a quantitative differentiation of the thunderstorm intensities concurrent with the lightning strike trip-outs.

In the realm of researching lightning activities on the disaster side, Chinese power grid lightning protection professionals have made substantial contributions. Document [15] leveraged a Backpropagation (BP) neural network to forecast future lightning current amplitudes. This was achieved by utilizing meteorological radar data, including echo intensity, echo top height, and vertically integrated liquid water content. Significantly, the study incorporated the lateral distance between thunderclouds and transmission lines into a trip-out probability distribution model, enabling the issuance of lightning strike warnings based on predetermined probability thresholds. Moreover, Document [16] integrated data from atmospheric electric field meters and lightning location systems. By using variations in electric field strength and the proximity of lightning strikes as criteria, the study successfully implemented a tiered lightning warning system specifically for the Guangzhou power grid. In a similar vein, Document [17] employed forecasted trajectories of thunderclouds to predict future lightning strike zones. It employed a grid-based method to directly calculate the imminent probability of lightning strikes on transmission lines, thereby facilitating the prediction and early warning of lightning strike-induced trip-outs. Document [18] introduced an innovative approach, utilizing information gleaned from Lightning Location Systems (LLS) and image recognition technology. This methodology enabled dynamic prediction of lightning strike trip-outs, demonstrating noteworthy efficacy. Finally, Document [19] focused on computing the near-future probability of lightning strike trip-outs based on the proximity of ongoing lightning activities. The resultant data served as a foundational element in completing lightning strike warning tasks, representing another step forward in this critical area of electrical engineering research.

This paper concurrently addresses the risk of lightning-induced tripping in power transmission lines and the hazards posed by real-time lightning activities, aiming to enhance the efficiency of early warning systems while reducing the occurrences of false alarms and missed reports. Initially, the study analyzes a vast array of lightning location data and employs lightning identification, tracking, and analysis algorithms to predict the movement trajectories of thunderstorm clouds. A high-risk thunderstorm assessment model is proposed based on hypothesis testing.

During the warning phase, the research firstly calculates the predicted positions of the thunderstorm clouds' movement trajectories, and then assesses whether these clouds pose a high risk or are approaching transmission lines with low lightning protection capabilities. The evaluation process incorporates a comprehensive set of data, including the current position and dynamics of the thunderstorm clouds, the specific layout of transmission lines, the current status of lightning protection facilities, and historical lightning strike records. This facilitates the accurate identification of potential lightning strike risk areas and the timely issuance of continuous lightning strike tripping alerts.

Finally, the proposed warning method is validated through case studies in a certain part of northern China. For the tightly-knit transmission corridors that are of significant interest to power grid companies, the accuracy rate of the proposed continuous lightning strike tripping alert system reached 80%, demonstrating its potential in enhancing the safety and reliability of power grids.

The innovative aspects of this paper are primarily manifested in two key areas:

1.Dual Risk Integration: This research marks the first instance in the field of lightning protection studies where both the risk of lightning-induced tripping in power grid transmission lines and the hazard of real-time lightning activities are concurrently considered. This dual-perspective integration offers a novel viewpoint and methodology for research in electrical grid lightning protection and early warning systems. Unlike previous studies that typically focus on a single aspect of risk, this paper simultaneously analyzes the conditions of transmission lines on the grid side and the patterns of lightning activities on the disaster side. This approach allows for a more comprehensive assessment of the impact of lightning on power grids, thereby significantly enhancing the accuracy and practicality of early warning systems.

2.Comprehensive Data Analysis and High-Risk Thunderstorm Assessment Model: Innovatively utilizing big data analysis techniques, this study significantly improves the accuracy of predicting thunderstorm cloud movement trajectories by analyzing extensive lightning location data and integrating lightning identification, tracking, and analysis algorithms. Moreover, based on hypothesis testing methods, a high-risk thunderstorm assessment model is proposed. This model comprehensively considers the characteristics of the thunderstorm clouds themselves, along with various factors such as geography and meteorology, resulting in more precise prediction outcomes.

II. THE RESEARCH FOUNDATION OF THE EARLY WARNING METHOD FOR CONTINUOUS LIGHTNING STRIKE TRIPPING IN TIGHTLY SPACED TRANSMISSION CORRIDORS

The focus of this paper is on the continuous tripping of tightly spaced transmission corridors, which is a key concern for the dispatch departments of power grid companies. Moreover, the foundation of the lightning strike tripping early warning research presented in this paper is based on analyzing the movement patterns of thunderclouds from extensive historical lightning location data, which is then used to predict the real-time movement positions of thunderclouds.

A. CONTINUOUS TRIPPING IN TIGHTLY SPACED TRANSMISSION CORRIDORS

Tightly spaced transmission corridors refer to sections of power transmission lines clustered in narrow areas with multiple lines arranged closely together. Following the definition proposed in reference [20], tightly spaced transmission corridors are defined as: 1) comprising three or more transmission lines; 2) with neighboring lines having a center-to-center distance not exceeding 600 meters; 3) each transmission line having at least 15 consecutive towers within the corridor. Senior experts in the field of lightning protection note that the lightning protection for transmission corridors with closely arranged lines requires further in-depth study. Tightly spaced transmission lines often originate from the same power source or supply electricity to the same load. Continuous or simultaneous lightning-induced tripping can lead to severe consequences, such as disruption of power generation or load supply, significantly threatening the reliability of power supply. Given these concerns, the focus of this paper is on the early warning of continuous lightning strike tripping in tightly spaced transmission corridors. Specifically, this study investigates the likelihood of subsequent lightning-induced trippings within 1 hour following an initial lightning-induced trip in such corridors.Based on historical data and statistical analysis, it has been determined that lightning activities are typically concentrated within an hour. Consequently, the selection of a one-hour time window for this study is grounded in the observed characteristics of lightning activities. This duration is chosen to encompass the majority of the time frame in which lightning strike events occur.

B. LIGHTNING DETECTION, TRACKING, AND ANALYSIS ALGORITHMS

In order to monitor and analyze the movement trajectories of thunderstorm clouds and understand the patterns of thunderstorm cloud motion, it is imperative to classify historical lightning location data into their respective thunderstorm clouds. In this paper, the authors have proposed a lightning identification, tracking, and analysis algorithm, as described in [21], which accomplishes the aforementioned classification task. The basic approach is outlined as follows:

1) Identification. The algorithm employs the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method [22], [23], [24] to perform density clustering of lightning strike points every 5 minutes within the study area, based on their geographical coordinates. Lightning points classified in the same cluster are generally geographically close and are marked as belonging to the same thundercloud. The DBSCAN clustering algorithm is widely used in analyzing lightning location data and can identify clusters of various shapes. In tracking thunderclouds, it is crucial to consider both the horizontal scale of the thunderstorm cell and its movement distance within 5 minutes. Typically, the horizontal range of a thunderstorm cell varies from a few kilometers to 20 kilometers, with an average movement speed of about 16.4 m/s. Hence, a thunderstorm can move approximately 5 km within 5 minutes. Based on the horizontal scale and movement distance of the storm, the search neighborhood for the DBSCAN algorithm is set to 25 km. Additionally, following the parameter settings obtained through multiple experiments in [25], the minimum number of lightning strikes contained within a thundercloud in 5 minutes is set to 10. Furthermore, the mean latitude and longitude of the lightning points within the same thundercloud are taken as the centroid of the thundercloud. The choice of 5 minutes as the basic time unit for analyzing thunderclouds is made considering that the lightning range of a thundercloud typically does not expand further after 5 minutes of continuous activity.

2) Tracking. Based on the concept of Moving Clusters in Spatio-Temporal (MCST) data [26], in the temporal dimension, thundercloud centroids that are close in time and geographic location are connected to form a complete thundercloud trajectory. The criterion for judging proximity is set such that two centroids are within 25 km of each other. At the same time, the merging and splitting of thunderclouds are also considered.

3) Analysis. Dense lightning strike points are allocated to different thunderclouds. By connecting the centroids of the same thundercloud, the spatio-temporal movement trajectory of the thundercloud is obtained, which can be applied

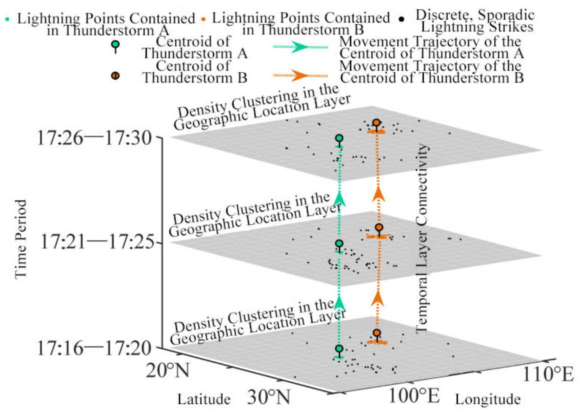

FIGURE 1. **Lightning detection, tracking, and analysis algorithm schematic diagram.**

in the thundercloud movement prediction model discussed in the following section. Additionally, discrete, sporadic lightning strikes, which account for less than 30% of all lightning strikes, are of less concern in this study focusing on continuous lightning strike tripping. The impact of these sporadic strikes will be further analyzed in the case study section of this paper. Since sporadic lightning strikes exhibit less regularity, the installation of atmospheric electric field instruments on towers is suggested for their prevention.

Taking the lightning data from 17:16 to 17:30 on a specific day in 2022 as an example, the results of the lightning detection, tracking, and analysis are illustrated in Figure 1.

C. THUNDERCLOUD MOVEMENT TRAJECTORY PREDICTION MODEL

For the early warning of continuous lightning strike tripping, accurately predicting the next movement position of a thundercloud can significantly enhance the efficiency of the warning system. Building upon the obtained thundercloud centroid movement trajectories, this paper employs Extreme Gradient Boosting (XGBoost) to predict the movement trajectory of thundercloud centroids [27]. XGBoost is an ensemble machine learning method based on decision trees, widely used for solving predictive problems. In machine learning competitions, XGBoost has become the method of choice for many winning teams.

Existing methods for predicting thundercloud trajectory, including linear fitting [28], nonlinear fitting methods [29], and the Holt two-parameter linear exponential smoothing method [30], all utilize real-time lightning location data to extrapolate the movement trajectory of thunderclouds. However, the patterns of historical thundercloud movement trajectories have yet to be integrated into these trajectory prediction models. Since thundercloud movement trajectories are typically nonlinear, a method with strong nonlinear approximation capabilities is required. Therefore, this paper introduces a machine learning approach (XGBoost) to learn the patterns of historical thundercloud movement trajectories. The XGBoost method continually decomposes trees to learn and fit the residuals of predictions, exhibiting a strong capability for nonlinear approximation.

Compared to other machine learning predictive methods, including Gradient Boosting Decision Tree (GBDT) and Light Gradient Boosting Machine (LightGBM), XGBoost (Extreme Gradient Boosting) has the following four advantages in solving regression prediction problems [31], [32]:

1) XGBoost performs a second-order Taylor expansion of the objective function, which aids in faster and more accurate gradient descent, whereas GBDT only uses first-order derivatives in optimization.

2) XGBoost employs pre-sorting before iteration and traverses to select the optimal splitting point. In contrast, although LightGBM uses a histogram-based algorithm, which requires less memory and has lower complexity for data splitting, it struggles to find the most accurate data split points.

3) XGBoost, similar to the Random Forest algorithm, supports column and row sampling. This approach can reduce computational load and the risk of overfitting.

4) XGBoost introduces regularization in the objective function. It regularizes the structure of trees to reduce the model's variance and introduces shrinkage in each step to avoid overfitting. This reduces the influence of individual trees on the outcome and enhances the model's generalization ability.

The present study selects the XGBoost method as the predictive approach for thunderstorm movement trajectories. Additionally, traditional neural network methods and deep learning approaches are included for comparison, further demonstrating the advantages of XGBoost in predicting thundercloud movement trajectories.

The following section introduces the model and parameters of XGBoost. For a dataset D with n samples and each sample having m features, XGBoost predicts the output $\hat{y}_i$ using K additive functions $f_k(x_i)$.

$$\boldsymbol{D} = \{(\boldsymbol{x}_i, y_i)\} \left(|\boldsymbol{D}| = n, \boldsymbol{x}_i \in \mathbf{R}^m, y_i \in \mathbf{R}\right) \tag{1}$$

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathcal{F} \tag{2}$$

In the formula, $\mathcal{F}$ represents the space of decision trees.

To learn the set of additive functions used in the model, XGBoost minimizes a regularized objective $\mathcal{I}$.

$$\mathcal{I} = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{3}$$

$$\Omega(f_k) = \gamma T + \frac{1}{2} \|\omega\| \lambda^2 \tag{4}$$

In the formula: l is a differentiable convex loss function used to measure the difference between the actual value y_i and the predicted value $\hat{y}_i$; Ω is an additional regularization term used to penalize the complexity of the model, thus avoiding overfitting; γ and λ are penalty coefficients, which can be adjusted according to requirements in practical applications; ω represents the weights of the leaves of the decision tree.

The core concept of XGBoost is the continuous splitting of features to add trees. The addition of trees serves to introduce new additive functions $f_t(x_i)$ to fit the residuals from the previous prediction.

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i) \tag{5}$$

In the formula: $\hat{y}_i^{(t)}$ is the predicted value for instance i in the t round of iteration; $\hat{y}_i^{(t-1)}$ is the predicted value from the previous round of iteration.

After substituting into equation (5), the objective function can be rewritten as shown in equation (6). Then, the objective function is derived using a fast optimization second-order approximation, as shown in equation (7).

$$\mathcal{I}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{6}$$

$$\mathcal{I}^{(t)} \simeq \sum_{i=1}^{n} \left[l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \tag{7}$$

$$g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) \tag{8}$$

$$h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) \tag{9}$$

In the formula: g_i represents the first-oeder gradient of the loss function; h_i represents the second-order gradient of the loss function.

To find the optimal objective function for solving regression prediction problems and avoiding overfitting, appropriate parameters must be set for XGBoost. In regression prediction problems, key parameters of XGBoost are the maximum depth, learning rate, and the estimated number of trees, denoted as n More specifically, the maximum depth refers to the depth of the tree; the deeper the tree, the more complex the model, and the higher the risk of overfitting. The learning rate is the shrinkage step used in the updating process. The estimated number of trees is the count of decision trees. The latter two parameters are interrelated: the higher the learning rate, the fewer decision trees are required. This paper employs Bayesian optimization to automatically tune these three parameters. Bayesian optimization is a strategy for automatic parameter adjustment with fewer iterations and faster iteration speed. Ultimately, the real-time movement trajectory of thunderclouds is predicted using XGBoost, which has been trained on historical thundercloud movement trajectories and optimized in terms of parameters.

The input for the thundercloud movement trajectory prediction model consists of four centroids corresponding to four continuous time periods of a particular thundercloud. The output of the model is the centroid of the thundercloud for the next time period, specifically for the next 1 to 5 minutes. During the training phase of the algorithm, if a thundercloud exists for 97 minutes, it will have a total of 20 centroids, with the last centroid corresponding only to the last 2 minutes of lightning strikes. This thundercloud can generate 16 sets of training data; the first set contains 5 centroids corresponding to the initial 25 minutes of the thundercloud, and the last set contains 5 centroids corresponding to the final 22 minutes before the end of the thundercloud. Furthermore, under the condition that the input of the prediction model remains unchanged, the output of the model can also be the centroid of the thundercloud for the second upcoming time period, i.e., the next 6 to 10 minutes, or for the third upcoming time period, i.e., the next 11 to 15 minutes. As shown in Table 1, the locations of the centroids for the first, second, and third upcoming time periods are predicted at the current moment.

TABLE 1. Comparison of predictive performance across various methods.

Method	Average Absolute Error between the Actual and Predicted Position of the Centroid/km		
	The First Centroid	The Second Centroid	The Third Centroid
Generalized Regression Neural Network	17.46	19.65	23.04
Nonlinear Fitting	14.13	22.19	31.32
Linear Fitting	14.06	21.93	30.49
Long Short-Term Memory Network	12.57	14.41	16.80
Holt's Double Parameter Linear Exponential Smoothing	10.68	14.13	17.96
XGBoost	9.84	12.15	14.53

Table 1 compares the performance of various existing thunderstorm cloud trajectory prediction methods, including neural network and deep learning approaches, using the absolute error in the distance between the actual and predicted positions of the cloud centroid as the performance evaluation metric. For training, data from 87,029 sets corresponding to thunderstorm clouds in a certain part of northern China, for the years 2020-2021 were used. For testing, data from 43,272 sets corresponding to the 2022 thunderstorm clouds were utilized. Table 1 shows the average values of the absolute error in distance for the test group. The results demonstrate that XGBoost can consistently improve the precision in predicting the centroid position of thunderstorm clouds.

Particularly, this paper is concerned with the prediction of thunderstorm cloud movement trajectories in relation to transmission lines. For elongated transmission lines, there is a need to further improve the predictive performance of thunderstorm cloud movement trajectories. To this end, the paper introduces five meteorological factors that may influence the movement of thunderstorm clouds into the prediction model: temperature, dew point, wind direction, wind speed, and total cloud cover.The potential interactions or correlations among these meteorological factors are crucial for accurately predicting the movement and behavior of thunderstorm clouds. Firstly, the correlation between temperature and dew point is instrumental in determining atmospheric humidity levels. A combination of high temperatures and high dew points typically signifies a higher moisture content in the air, which is a critical condition for the development of thunderstorm clouds. Secondly, the close relationship between wind direction and wind speed plays a pivotal role in the movement and evolution of thunderstorm clouds. Wind direction influences the path of cloud movement, while wind speed affects the velocity and dispersion of the cloud. Additionally, the interaction between total cloud cover and temperature can impact ground-level temperatures, which in turn influences the thermodynamic conditions of thunderstorm clouds. The effects of wind speed and direction on total cloud cover are also significant. Strong winds might disperse cloud cover, whereas certain wind directions can contribute to the formation and accumulation of clouds, facilitating the genesis of thunderstorm clouds. Lastly, the combination of dew point, temperature, and cloud cover can lead to an increase in cloud cover, creating conducive conditions for thunderstorm activity. Considering these factors collectively allows for a more comprehensive understanding of thunderstorm cloud behavior, thereby enhancing the accuracy of predictions. However, as the length of the input sequence increases, the span of dependencies between sequences becomes longer, making the gradient descent-based parameter optimization calculations more complex and significantly increasing training time. Therefore, it is necessary to select meteorological factors highly correlated with the movement of thunderstorm clouds as input data.The distance correlation (dCor) method is utilized to calculate the distance correlation coefficients between each meteorological factor and the centroid of thunderstorm cloud segments [33], [34]. Considering that the centroid's position coordinates are two-dimensional, the dCor method can analyze the correlation between meteorological factors and the movement centroid of thunderstorm clouds without being limited by dimensions. Numerical fitting results indicate that dCor can effectively detect nonlinear or non-monotonic relationships between variables, especially when the correlation between random variables is nonlinear.

Based on the thunderstorm and meteorological data from a certain part of northern China, for the years 2020–2021, the distance correlation coefficients between temperature, dew point, wind direction, wind speed, total cloud cover, and the centroid of thunderstorm cloud segments were found to be 0.206, 0.591, 0.112, 0.099, and 0.296, respectively. In this study, only the three meteorological factors with higher distance correlation coefficients (temperature, dew point, and total cloud cover) were selected along with the centroid position data as inputs for the XGBoost model. This predictive approach is referred to as XGBoost-dCor.Furthermore, to verify the effectiveness of the selection of meteorological factors, a prediction using all five meteorological factors (temperature, dew point, total cloud cover, wind direction, and wind speed) combined with the centroid position data was made using XGBoost and is denoted as XGBoost-All. Additionally, a prediction method using only the centroid position data as the input for XGBoost is referred to as XGBoost.

Predicting the movement trajectories of thunderstorm clouds with respect to transmission lines is essential to gain more predictive information before lightning-induced trip-outs occur. Consequently, in further testing the performance of various methods, the predictive output has been adjusted to target the centroids of thunderstorm cloud segments that cause lightning strike trip-outs on transmission lines.

TABLE 2. Comparison of the performance of various methods in predicting the centroid of thunderstorm cloud segments causing lightning strike trip-outs in transmission lines.

Method	Average Absolute Error between the Actual and Predicted Position of the Centroid/km		
	The First Centroid	The Second Centroid	The Third Centroid
Generalized Regression Neural Network	13.25	17.88	18.30
Nonlinear Fitting	11.47	19.86	31.31
Linear Fitting	11.25	19.42	30.21
Long Short-Term Memory Network	11.24	11.31	16.26
Holt's Double Parameter Linear Exponential Smoothing	8.60	13.98	16.01
XGBoost	8.68	10.38	13.75
XGBoost-ALL	8.59	10.14	13.44
XGBoost-dCor	8.04	9.36	12.61

As indicated in Table 2, 10GBoost-dCor has further improved the prediction accuracy. Therefore, this study opts for XGBoost-dCor to predict the movement trajectories of thunderstorm clouds.

In summary, the prediction of thunderstorm cloud movement trajectories discussed in this section can be utilized to predict the distance between the centroid of thunderstorm clouds and transmission lines. This will serve as the basis for triggering continuous lightning strike trip-out warnings in closely spaced transmission corridors.

III. METHOD FOR EARLY WARNING OF CONTINUOUS LIGHTNING STRIKE TRIP EVENTS IN COMPACT TRANSMISSION CORRIDORS

To prevent the hazards of lightning to transmission lines, lightning protection management strategies primarily consider the following factors:

1) Lightning activity, which is the direct cause of lightning-induced tripping in transmission lines. The magnitude of lightning over-voltage can reach several hundred thousand volts, a level generally beyond the endurance of power equipment insulation.

2) The structural characteristics of transmission lines are the inherent factors leading to lightning-induced tripping. Lines with poor lightning protection performance are more susceptible to lightning strikes. Intrinsic factors affecting line lightning protection include the protection angle of the lightning rod, grounding resistance, tower height, etc [35].

3) Topographical and geomorphological characteristics are environmental factors that lead to lightning-induced tripping in transmission lines [36]. Side flashes are the main cause of lightning-induced tripping in high-voltage lines, and are significantly influenced by topography and geomorphology. Environmental factors affecting line lightning protection include topographical factors, climatic factors, and the altitude of towers.

In light of the aforementioned factors, this article introduces a high-risk thunderstorm assessment model to determine the risk of lightning-induced tripping due to real-time lightning activity. Additionally, the lightning protection assessment model for transmission lines presented in literature [21] can be utilized to evaluate the risk of lightning-induced tripping relevant to the structural characteristics of the transmission lines and their topographical and geomorphological features.

A. CONTINUOUS HIGH-RISK THUNDERSTORM ASSESSMENT MODEL

The lightning positioning system records the location, time, lightning current value, and number of return strokes for each ground flash. This paper is concerned with the records of lightning activities that cause tripping. According to the case studies in the Electric Power Company's lightning protection manual, lightning activity records are sought within a 5-minute window before and after the tripping record time and within a 5-kilometer radius around the faulted line. Considering that the lightning strike points found in this manner are generally not unique, this paper opts to use segments of thunderstorm clouds to correspond to a particular tripping record. A thunderstorm cloud segment refers to dividing a thunderstorm cloud into multiple segments on the time dimension, where each segment contains all the lightning strike points within 5 minutes of that thunderstorm cloud. If a thunderstorm cloud disappears before 5 minutes, it is also recorded as a segment.

In researching lightning-induced tripping, the initial step involves identifying all the lightning strike points contained within multiple thunderstorm cloud segments corresponding to the time of the tripping record. Subsequently, in terms of geographical location, and in conjunction with the results of traveling wave ranging, the nearest lightning strike point around the approximate fault location is identified from the aforementioned strike points. The thunderstorm cloud segment to which this lightning strike point belongs is then designated as the "thunderstorm cloud segment responsible for the tripping."

The goal of the High-Risk Thunderstorm Assessment Model is to identify high-risk thunderstorm cloud segments from all thunderstorm cloud segments, that is, the segments most likely to cause tripping. The basic data that the lightning positioning system can provide includes the location, time, lightning current value, and number of return strokes for each lightning strike point. Accordingly, there are five parameters that can be identified for a thunderstorm cloud: 1) the number of lightning strikes contained within the thunderstorm cloud; 2) the average (absolute) value of the lightning current of the strike points contained within the thunderstorm cloud; 3) the average number of return strokes of the strike points contained within the thunderstorm cloud; 4) the maximum (absolute) value of the lightning current of the strike points contained within the thunderstorm cloud; 5) the maximum number of return strokes of the strike points contained within the thunderstorm cloud. For this purpose, a hypothesis testing method is introduced to determine whether there are significant differences in each parameter between all thunderstorm cloud segments and high-risk thunderstorm cloud segments.

TABLE 3. Mean values of each parameter and the p-value from hypothesis testing.

Parameter Name and Unit	Average Value for All Thunderstorm Cloud Segments	Average Value for Thunderstorm Cloud Segments Causing Tripping	p-value
Number of Lightning Strikes	56.60	102.40	3.62×10^{-15}
Average Lightning Current/kA	34.74	32.63	0.0033
Average Number of Return Strokes	1.49	1.64	0.001
Maximum Lightning Current/kA	114.24	137.32	6.47×10^{-7}
Maximum Number of Return Strokes	6.17	8.34	2.50×10^{-20}

Taking the thunderstorm clouds in a certain part of northern China, from 2020 to 2021 as an example, the mean values of the aforementioned five parameters for all thunderstorm cloud segments and those causing tripping are shown in Table 3. In Table 3, the p-value is the probability of hypothesis, an important indicator in inferential statistics, and serves as crucial evidence for judging the correctness of the original hypothesis.

Considering that the distribution of each parameter does not conform to a normal distribution or other common distributions, this paper employs non-parametric tests to conduct statistical inference on the various parameters of thunderstorm clouds.

The two-sample Kolmogorov-Smirnov (KS) test is often used to assess the difference between the cumulative distributions of two sample datasets [37]. Under a certain significance level α, the null hypothesis posits that there is no significant difference in the distributions of the two samples. The decision of whether there is a significant difference between the two samples is made based on the p-value obtained from the KS test. Specifically, for the parameters of thunderstorm clouds in a certain part of northern China, from 2020 to 2021, at a significance level of $\alpha = 0.001$, the p-value for the mean number of lightning strikes of the two samples is much less than 0.001, leading to the rejection of the null hypothesis. This implies, at a 0.001 error probability, that: first, there is a significant difference in the number of lightning strikes between all thunderstorm cloud segments and those causing tripping; second, there is a significant difference in the maximum lightning current between all thunderstorm cloud segments and those causing tripping; third, there is a significant difference in the maximum number of return strokes between all thunderstorm cloud segments and those causing tripping. Similarly, at an error probability of 0.001, there is no significant difference in the

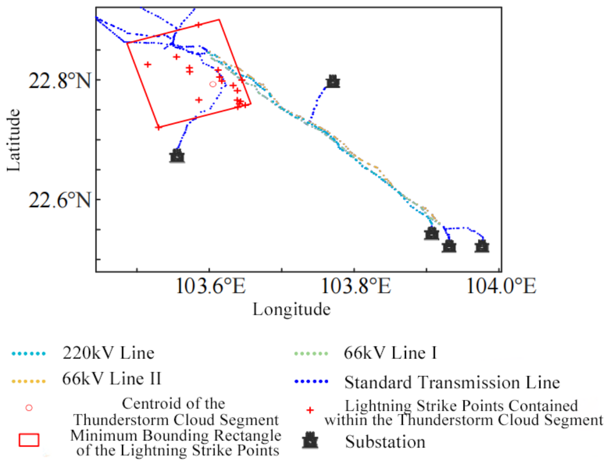

FIGURE 2. Schematic diagram of the minimum bounding rectangle for a thunderstorm cloud segment.

average lightning current between all thunderstorm cloud segments and those causing tripping, nor is there a significant difference in the average number of return strokes.

In summary, this paper selects the number of lightning strikes, the maximum lightning current, and the maximum number of return strokes as the candidates for the basis of high-risk thunderstorm assessment. Specifically, for the compact transmission corridors of interest in this study, a minimum bounding rectangle is used to determine whether a thunderstorm cloud segment is near a compact transmission corridor. The minimum bounding rectangle is defined as the smallest rectangle encompassing all the lightning strike points of a thunderstorm cloud segment, describing the potential impact range of the thunderstorm cloud segment. The solution for the minimum bounding rectangle is based on the approach provided in literature [38], [39], which primarily involves calculating the convex hull of the current set of lightning strike points (the convex hull being the convex polygon formed by connecting the outermost points) and then calculating the area of the rectangle corresponding to any two adjacent points on the hull. The specific steps are: 1) Take the line connecting two adjacent points on the convex hull as one side of the rectangle. 2) Find the point on the convex hull farthest from the obtained edge and draw a parallel line through this point to form the second side of the rectangle. 3) Project the points on the convex hull onto the obtained edge and find the two projection points that are the farthest apart; draw a line through these two points to form the other two sides of the rectangle. 4) Iterate over all adjacent pairs of points on the convex hull, i.e., repeat steps 1)–3), and take the rectangle with the smallest corresponding area as the minimum bounding rectangle. As shown in Figure 2, taking the thunderstorm cloud segment from June 5, 2020, 06:55:06 to 07:00:06 as an example, the minimum bounding rectangle of this thunderstorm cloud includes a compact transmission corridor (with four transmission lines).

Considering that some lightning strike points far from the centroid can enlarge the area of the minimum bounding rectangle, potentially leading to a certain degree of false alarms (i.e., even though the compact transmission corridor is inside the minimum bounding rectangle, there may be no nearby lightning strike points), this paper also takes into account the minimum distance from the centroid of the lightning strike point set to the compact transmission corridor. This metric reflects the average distance from the lightning strike points to the compact transmission corridor. If this metric is too large, it is not considered that the thunderstorm cloud segment impacts the compact transmission corridor. The paper uses a one-sided confidence interval to calculate the upper limit of this minimum distance under a certain confidence level.

TABLE 4. Distribution test of the minimum distance from thunderstorm cloud centroids to compact transmission corridors.

Distribution Test	p-value
Normal Distribution	0.1011
Poisson Distribution	3.19
Uniform Distribution	1.76
Rayleigh Distribution	7.34

A statistical analysis was conducted on the minimum distance from the centroids of thunderstorm clouds that caused tripping in compact transmission corridors in a certain part of northern China, from 2020 to 2021. Various distribution tests were performed on these distances, and the results are presented in Table 4.

At a significance level of $\alpha = 0.05$, the null hypothesis assumes that the distribution of the minimum distance follows a certain distribution. Following the distribution tests, the p-values for the Poisson, uniform, and Rayleigh distribution tests are all less than 0.05, meaning the null hypothesis is rejected, and the distribution of the minimum distance does not follow the Poisson, uniform, or Rayleigh distributions. However, the p-value for the normal distribution test is 0.1011, greater than 0.05, implying acceptance of the null hypothesis that the distribution of the minimum distance follows a normal distribution. Under the condition that the distribution of the minimum distance follows a normal distribution, the one-sided confidence upper limit at a confidence level of 0.95 is

$$0.95 \leq P\{X < d\} = P\left\{\frac{X-\mu}{\sigma} < \frac{d-\mu}{\sigma}\right\}$$
$$= \Phi\left(\frac{d-\mu}{\sigma}\right) \geq 0.95 = \Phi(1.6449) \quad (10)$$

After substituting in the mean and standard deviation, the solution yields d=39.4km. This means that at a confidence level of 0.95, the one-sided confidence upper limit for the minimum distance from the centroids of thunderstorm clouds that caused tripping in compact transmission corridors to the compact transmission corridors is 39.4km.

In summary, if the minimum bounding rectangle of a thunderstorm cloud segment includes a compact transmission corridor, and the predicted position of the next centroid of this thunderstorm cloud segment is less than 39.4km away from

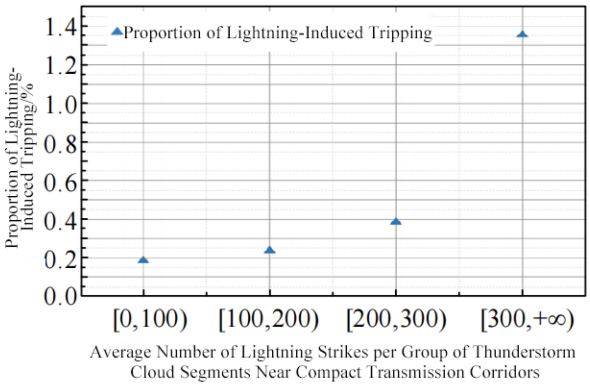

FIGURE 3. **Proportion of lightning-induced tripping corresponding to different average lightning strike count ranges.**

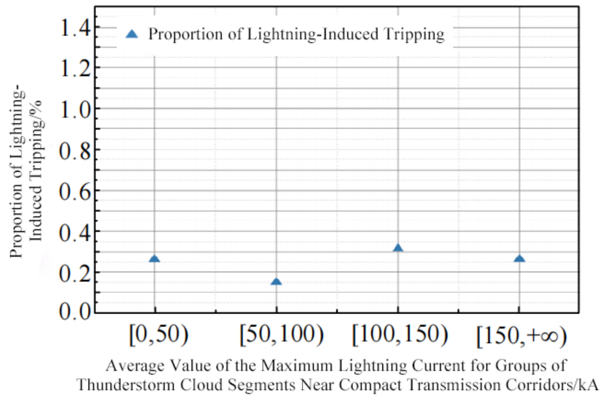

FIGURE 4. **Proportion of lightning-induced tripping corresponding to different ranges of average maximum lightning current values.**

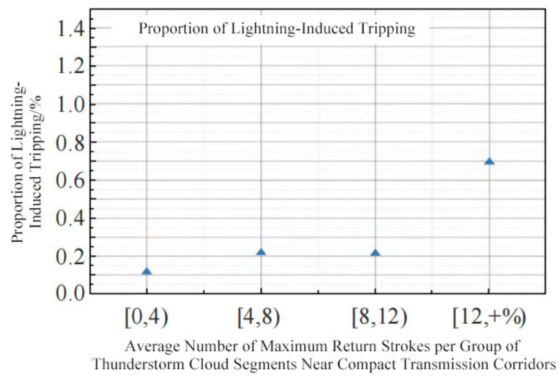

FIGURE 5. **Proportion of lightning-induced tripping corresponding to different ranges of average maximum number of return strokes.**

the compact transmission corridor, it is then designated as a ''thunderstorm cloud segment near the compact transmission corridor.''

In this paper, the input for the early warning system consists of groups of four consecutive thunderstorm cloud segments, and the output is the warning status of that thunderstorm cloud for the next time period (5 minutes). Initially, all groups of thunderstorm cloud segments in a certain part of northern China, from 2020 to 2021 were statistically analyzed. If the fourth thunderstorm cloud segment in a group is identified as being near the compact transmission corridor, the average number of lightning strikes for the four thunderstorm cloud segments in that group is calculated, and the group is labeled as a ''thunderstorm cloud segment group near the compact transmission corridor.'' If a group of thunderstorm cloud segments is identified as being near the compact transmission corridor and the subsequent thunderstorm cloud segment causes a lightning-induced trip, then that group is also marked as a ''thunderstorm cloud segment group causing lightning-induced tripping.'' The proportion of lightning-induced tripping shown in Figure 3 refers to the percentage of groups identified as causing lightning-induced tripping within the same average number of lightning strikes range among the groups that are near the compact transmission corridor.

Similar to the average number of lightning strikes for groups of thunderstorm cloud segments, the average values of the maximum lightning current and the maximum number of return strokes for these groups can also be calculated to determine the corresponding proportions of lightning-induced tripping. The results are respectively illustrated in Figures 4 and 5.

By comparing the proportions of lightning-induced tripping in each range in Figures 3, 4, and 5 (indicated by blue triangles in the graphs), it can be concluded that groups of thunderstorm cloud segments with an average number of lightning strikes exceeding 300 have the highest proportion of causing tripping in compact transmission corridors. Therefore, this paper selects this criterion for early warnings. Specifically, if the fourth thunderstorm cloud segment in a group is near a compact transmission corridor and the average number of lightning strikes for the group exceeds 300, it triggers an early warning for the thunderstorm cloud in the subsequent time period.

B. TRANSMISSION LINE LIGHTNING PROTECTION ASSESSMENT MODEL

Previous assessments of lightning-induced tripping risk typically qualitatively categorized such events as lightning-induced tripping, without quantitatively distinguishing the intensity of the thunderstorm at the time of the tripping. Literature [21] introduced the concept of real-time lightning intensity and quantitatively analyzed this intensity based on the proximity of lightning strikes, the magnitude of the lightning current, and the number of times a thunderstorm crosses a transmission line. It also considered the impact of lightning intensity and tower characteristics on the lightning protection performance of transmission lines during tripping events. The authors of this paper refer to the transmission line lightning protection assessment model proposed in the referenced literature [21].

For towers with records of lightning-induced tripping, the model defines their lightning protection performance based on the real-time lightning intensity at the time of tripping.

TABLE 5. Structural characteristics and their value ranges, as well as topographical and geomorphological characteristics and their value ranges.

	Feature	Feature Value
Structural Characteristics	Lightning Rod Protection Angle/(°)	5,10,15,20,25,30
	Design Value of Grounding Resistance/Ω	[1,30]
	Voltage Level/kV	110,220
	Number of Circuits	[1,6]
	Tower Height/m	[10,116]
	Tower Weight/kg	[1200,185883]
	Tower Type	Straight Tower, Tension Tower
Topographical and Geomorphological Characteristics	Terrain	Slope Orientation, Slope Position, Slope Surface, Slope Curvature, Soil Type
	Geomorphology	Mountainous Area, Low Mountain Area, Hill Area, Plain, Paddy Field, River, Road
	Tower Altitude/m	[87,4620]
	Climate	High Lightning Area, Micro-meteorological Area, Canyon Area
	Crossed Lines	Transmission Line, Communication Line

For towers without tripping records, the model searches for the most similar tower with a tripping record based on structural and topographical features (as shown in Table 5). The risk level of the most similar tower with a tripping record is then assigned to the tower without a tripping record. In summary, this approach allows for the determination of the risk level for all towers, and consequently, the assessment of the lightning protection risk level for the transmission line.

The early warning system in this paper selects transmission lines ranked in the lowest 10% in terms of lightning protection performance as low lightning protection performance transmission lines. If the fourth thunderstorm cloud segment in a group is identified as a segment near a compact transmission corridor, and this compact transmission corridor contains low lightning protection performance transmission lines, then an early warning is triggered for the thunderstorm cloud in the subsequent time period.

C. CONTINUOUS LIGHTNING STRIKE TRIP WARNING PROCESS FOR COMPACT TRANSMISSION CORRIDORS

Integrating the models from the two previous sections, the process involves firstly determining if a thunderstorm cloud segment is near a compact transmission corridor, and then assessing whether the compact transmission corridor contains low lightning protection performance transmission lines. Concurrently, it evaluates whether the current thunderstorm cloud belongs to a high-risk category. These two conditions are independent of each other; satisfying either one can trigger a continuous lightning strike trip warning for the compact transmission corridor. The flowchart for the continuous lightning strike trip warning process in compact transmission corridors is illustrated in Figure 6.

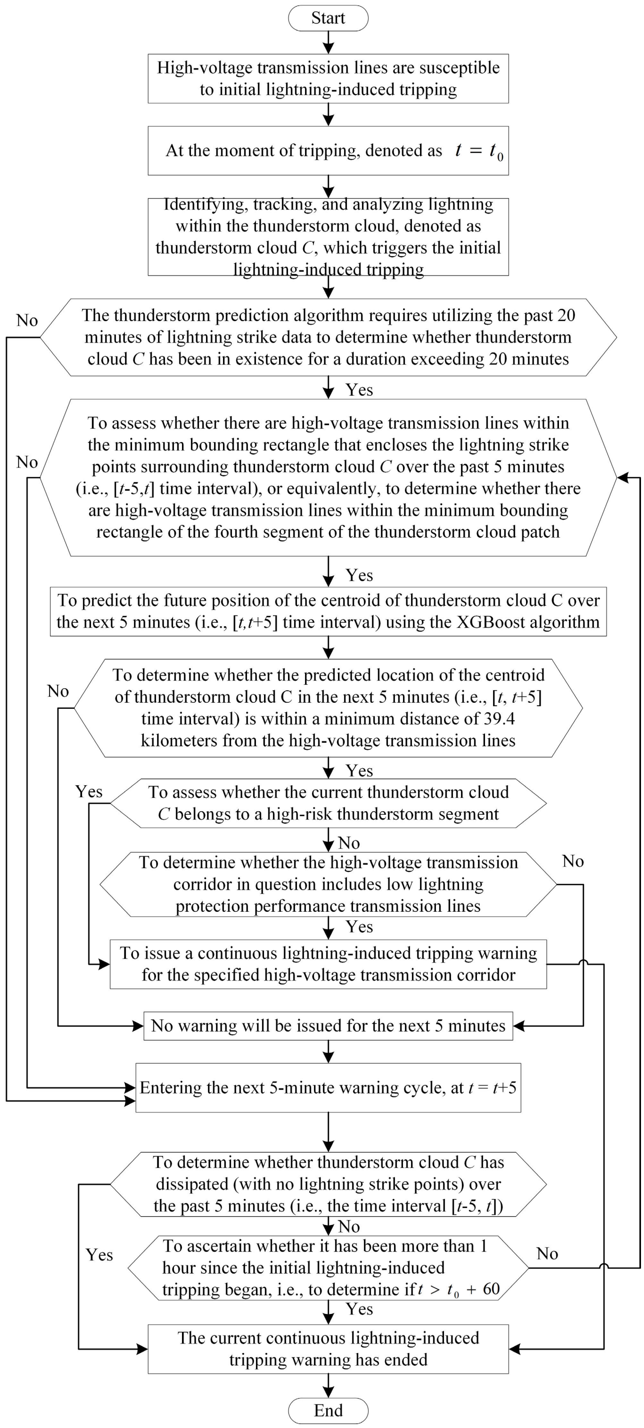

FIGURE 6. Continuous lightning-induced tripping warning process for high-voltage transmission corridor.

IV. NUMERICAL CASE VALIDATION

The model proposed in this paper has been validated through a case study conducted in a specific area in the northern regions of China. The original lightning activity data was sourced from lightning detection stations deployed by the electric power company in the area, with their actual detection range covering the regions in northern China intended for monitoring. Although there were lightning-induced power outages in the transmission system of this area, the thunderstorms causing these outages may not be confined to this region alone. Consequently, we utilized lightning location data from this region of China to track the thunderstorms, ensuring the comprehensiveness of our analysis. Based on the lightning detection stations deployed in a northern region, ground lightning data were acquired, totaling over 17 million ground lightning records from 2020 to 2022. The data from 2020 to 2021 were utilized for model training, while the 2022 data were employed for case study validation. All structural characteristics of the transmission towers, the local topography, climate, and line crossings around the towers, as well as the outage records of the towers, were provided by the local power grid company. Elevation data for the towers were obtained from the Shuttle Radar Topography Mission (SRTM), and soil type data for the areas around the towers were sourced from the GlobeLand land cover product, both with a spatial accuracy of 30 meters. Other topographical data for the towers were derived through interpolation based on SRTM data. Python and Matlab served as the development environments for this study.

In 2022, there were a total of 145 instances of lightning-induced power outages within the tightly integrated transmission corridors in a certain part of northern China. Among these, 8 cases involved consecutive power outages occurring within 1 hour of each other. For consecutive lightning-induced power outages occurring within the same transmission corridor, if a warning is triggered within 1 hour after the first outage and before the second outage, it is considered a successful warning. Typically, the conditions for triggering consecutive lightning-induced power outage warnings are satisfied immediately after the first outage.

Furthermore, the warning method presented in this paper is based on real-time lightning location data. The time required for the lightning detection station to process and detect lightning strikes at ground points is less than 1 millisecond, and the lightning monitoring system can locate lightning strikes within seconds. However, considering that there is some delay in transmitting real-time lightning location data and power outage information to the data center, the warning time indicated in this paper is uniformly delayed by 5 seconds.

Using the warning method proposed in this paper, the subsequent power outage warning situations for these 8 events are summarized in Table 6. Among them, 2 events did not trigger warnings. One of these events was caused by isolated and sporadic lightning strikes. The other event involved consecutive power outages occurring within the first 20 minutes of the initiation of the thunderstorm cloud. Due to the short duration of this event, the input data for the warning in this paper consisted of consecutive segments of at least 4 thunderstorm cloud slices, meaning that predictions could only be made at least 20 minutes after the onset of the thunderstorm. This limitation is inherent in the warning method. Compared to meteorological radar systems that provide large-scale warnings before the occurrence of thunderstorms, the use of lightning location system data for thunderstorm trajectory prediction can enhance the accuracy of predictions. However, the prerequisite is that the thunderstorm has been ongoing for some time, with at least 20 minutes of lightning location data available for prediction. For these two missed consecutive power outage events, the installation of atmospheric electric field instruments on the transmission lines is suggested for early warning purposes.

A. EARLY WARNING PERFORMANCE VERIFICATION METHOD

Utilizing performance evaluation criteria inspired by meteorology and previous assessments of meteorological disaster warnings for transmission lines, the performance of the consecutive lightning-induced power outage warning method is assessed, as depicted in Table 7. In this context, N_{TP} signifies the instances in which a warning was issued and consecutive power outages indeed transpired. N_{FP} denotes the occurrences when a warning was issued, but consecutive power outages failed to manifest. N_{FN} accounts for scenarios in which a warning was not issued, yet consecutive power outages did occur. Finally, N_{TN} corresponds to situations where no warning was issued, and consecutive power outages did not take place.

Based on the statistical values presented in Table 7, this paper employs four evaluation metrics to assess the performance of the consecutive lightning-induced power outage warning method. These specific evaluation metrics include classification accuracy, warning accuracy, missed detection rate, and false alarm rate. The details of each evaluation metric are as follows.

1) CLASSIFICATION ACCURACY, DENOTED AS A_{CR}

In this context, both scenarios of "issuing a warning, and consecutive power outages actually occurring" and "not issuing a warning, and consecutive power outages not occurring" are considered as "correct classifications." A_{CR} is the proportion of the number of correct classifications to all possible cases and is expressed by the following formula:

$$A_{CR} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{FP} + N_{FN} + N_{TN}} \tag{11}$$

2) WARNING ACCURACY, DENOTED AS A_{WR}

Consider the scenario where a warning was issued, and consecutive power outages actually occurred as "Warning Accuracy." A_{WR} is the proportion of the number of accurate

TABLE 6. Consecutive lightning-induced power outage event warning status.

Event Number	Transmission Line	Fault Location	Fault Time	Channel Number	Subsequent Lightning-Induced Power Outage Warnings
1	220kV Kongdong Line	116 Tower	2022-04-19 18:58:02	50	—
	220kV Fengtie Line	60 Tower	2022-04-19 19:04:56	50	18:58:07 Triggered Warning
	220kV Shangcao Line	60 Tower	2022-04-19 19:04:56	50	18:58:07 Triggered Warning
2	220kV Shitie Line	194 Tower	2022-04-19 19:38:04	30	—
	220kV Shuijiu Line	173 Tower	2022-04-19 19:40:56	30	19:38:09 Triggered Warning
3	66kV Baofeng B Line	128 Tower	2022-06-06 22:21:04	54	—
	66kV Baofeng A Line	131 Tower	2022-06-06 22:28:59	54	22:21:09 Triggered Warning
4	66kV Baogang 1 Line	150 Tower	2022-07-24 21:25:03	54	—
	66kV Baogang 2 Line	148 Tower	2022-07-24 21:30:57	54	21:25:08 Triggered Warning
5	66kV Caosai Line	28 Tower	2022-08-10 22:40:57	58	—
	66kV Aigu Line	33 Tower	2022-08-10 22:55:34	58	22:41:02 Triggered Warning
	66kV Baofeng A Line	34 Tower	2022-08-10 22:55:34	58	22:41:02 Triggered Warning
6	66kV Baoxi B Line	77 Tower	2022-08-19 06:46:00	53	—
	66kV Baoxi A Line	43 Tower	2022-08-19 07:20:38	53	No Warning
7	66kV Baoxi B Line Danding Branch Line	19 Tower	2022-09-17 00:35:00	33	—
	66kV Baoxi B Line Yuanjiagou Branch Line	234 Tower	2022-09-17 00:43:00	33	No Warning
8	66kV Caosai Line Dongxing Branch Line	39 Tower	2022-10-01 13:44:59	40	—
	66kV Caosai Line Zhangjiagou Branch Line	120 Tower	2022-10-01 14:19:58	40	14:45:04 Triggered Warning

warnings to the total number of actual consecutive power outage occurrences, expressed by the following formula:

$$A_{WR} = \frac{N_{TP}}{N_{TP} + N_{FN}} \tag{12}$$

TABLE 7. Performance metrics for early warning.

Warning Situation	Actual Occurrence of Consecutive Power Outages	Actual Non-Occurrence of Consecutive Power Outages
Issued Warning	N_{TP}	N_{FP}
Not Issued Warning	N_{FN}	N_{TN}

TABLE 8. Performance of the early warning method in this paper.

Warning Situation	Actual Occurrence of Consecutive Power Outages	Actual Non-Occurrence of Consecutive Power Outages
Issued Warning	8	7
Not Issued Warning	2	128
Classification Accuracy A_{CR}:93.79%;Warning Accuracy A_{WR}:80%; Missed Detection Rate M_{WR}:20%;False Alarm Rate F_{WR}:46.67%		

3) MISSED DETECTION RATE, DENOTED AS M_{WR}

Consider the scenario where no warning was issued, but consecutive power outages actually occurred as "Missed Detection." M_{WR} is the proportion of the number of missed detections to the total number of actual consecutive power outage occurrences, expressed by the following formula:

$$M_{WR} = \frac{N_{FN}}{N_{TP} + N_{FN}} \tag{13}$$

4) FALSE ALARM RATE, DENOTED AS F_{WR}

Consider the scenario where a warning was issued, but consecutive power outages did not actually occur as "False Alarm." F_{WR} is the proportion of the number of false alarms to the total number of actual warnings issued, expressed by the following formula:

$$F_{WR} = \frac{N_{FP}}{N_{TP} + N_{FP}} \tag{14}$$

B. COMPARATIVE ANALYSIS OF EARLY WARNING PERFORMANCE

To validate the effectiveness of the early warning method proposed in this paper, a comparative analysis was conducted with scenarios that considered only the high-risk thunderstorm assessment model and scenarios that considered only the lightning protection assessment model for transmission lines. The performance of these three methods is presented in Tables 8 to 10. It's worth noting that, as shown in Table 6, the 8 consecutive lightning-induced power outage events corresponded to a total of 18 lightning-induced power outages, including 8 initial outages and 10 subsequent outages. In cases where two subsequent outages occurred simultaneously after triggering a warning, it is counted as issuing 2 warnings and is recorded as 2 actual occurrences of consecutive power outages.

The continuous power outage warning in closely monitored transmission channels is the primary focus of this paper. Therefore, it is desirable to have a lower missed detection rate, and given the relatively low total number of issued warnings, a moderate false alarm rate can be acceptable.

TABLE 9. Performance of the early warning method in this paper when considering only the high-risk thunderstorm assessment model.

Warning Situation	Actual Occurrence of Consecutive Power Outages	Actual Non-Occurrence of Consecutive Power Outages
Issued Warning	2	0
Not Issued Warning	8	135
Classification Accuracy A_{CR}:94.48%;Warning Accuracy A_{WR}:20%; Missed Detection Rate M_{WR}80%;False Alarm Rate F_{WR}:0		

TABLE 10. Performance of the early warning method in this paper when considering only the lightning protection assessment model for transmission lines.

Warning Situation	Actual Occurrence of Consecutive Power Outages	Actual Non-Occurrence of Consecutive Power Outages
Issued Warning	6	7
Not Issued Warning	4	128
Classification Accuracy A_{CR}:92.41%;Warning Accuracy A_{WR}:60%; Missed Detection Rate M_{WR}40%;False Alarm Rate F_{WR}:53.85%		

All three scenarios exhibit high classification accuracy (all exceeding 90%). The early warning method proposed in this paper has the lowest false alarm rate among them. Most importantly, compared to scenarios considering only high-risk thunderstorms or only low lightning protection performance transmission lines, the proposed early warning method has the lowest missed detection rate, while achieving an 80% warning accuracy. The high-risk thunderstorm assessment model and the lightning protection assessment model for transmission lines complement each other, effectively improving the accuracy of the warnings. After excluding consecutive power outages caused by discrete occasional lightning strikes and accounting for the inherent limitations of lightning prediction, the proposed early warning method issues warnings for 100% of the continuous power outages in closely monitored transmission channels.

V. CONCLUSION

This paper, driven by lightning location data, proposed a continuous lightning-induced power outage early warning method for closely monitored transmission channels, and the following conclusions were drawn:

1) The paper effectively harnessed the data value of the lightning location system, utilizing a vast amount of historical data for model training, and integrating real-time lightning location data to issue lightning-induced power outage warnings.

2) The introduction of a high-risk thunderstorm model and the simultaneous consideration of both high-risk thunderstorms and low lightning protection performance transmission lines in the early warning method enhanced the efficiency of warnings, reducing the occurrence of false alarms and missed detection events.

3) For closely monitored transmission channels of interest to power grid companies, the accuracy of the proposed continuous lightning-induced power outage warning reached 80%, demonstrating practical value in engineering applications and providing assistance to power grid companies in making relevant control decisions.

The currently proposed method is subject to constraints arising from geographical and meteorological factors, leading to potential significant variances in the system's effectiveness across different geographical areas. The performance of the early warning system may be influenced by the distinct characteristics of the terrain, prevalent weather patterns, and the frequency of lightning events in various regions. Concurrently, the system's reliance on predicting the movement of thunderstorm clouds is contingent upon meteorological conditions, whose variability and unpredictability could result in inaccuracies in forecasts, leading to false alarms or missed warnings. Furthermore, the quality and completeness of lightning location data, historical lightning strike records, and information regarding the layout and status of power transmission lines are crucial for the system's accuracy. Insufficiencies or inaccuracies in these data sets could impair the system's ability to accurately predict and assess risks. Therefore, these factors must be taken into account in practical applications to ensure the system's efficacy and accuracy.

As power grid companies start installing atmospheric electric field instruments on transmission lines, obtaining pre-warning data before ground flashes occur, this can compensate for the inherent limitations of lightning location systems used for early warning. In the future, it has the potential to further reduce the missed detection rate and improve the accuracy of early warnings.

REFERENCES

[1] E. A. Mahdiraji, M. S. Amiri, and S. M. Shariatmadar, "Analysis of lightning strikes on the transmission line by considering the frequency-dependent model," *Quantum J. Eng., Sci. Technol.*, vol. 2, no. 6, pp. 12–36, 2021.

[2] I. Petrovic, S. Nikolovski, H. R. Baghaee, and H. Glavaš, "Determining impact of lightning strike location on failures in transmission network elements using fuzzy decision-making," *IEEE Syst. J.*, vol. 14, no. 2, pp. 2665–2675, Jun. 2020.

[3] T. Debao, F. Yuhua, and Z. Xuehui, "Statistical analysis of lightning disaster accidents in China from 2012 to 2017," *Sci. Technol. Rev.*, vol. 36, no. 5, pp. 42–47, 2012.

[4] D. Wu and J. Wang, "Lightning protection of 10-kV distribution lines by multiple breakpoints arc-extinguishing lightning protection gap," *IEEE Trans. Plasma Sci.*, vol. 48, no. 2, pp. 531–536, Feb. 2020.

[5] D. Wu, Z. Ji, and J. Wang, "Simulation and experimental analysis of multi-chamber arc-quenching arresters (MCAA) for 10 kV transmission lines," *Energies*, vol. 14, no. 19, p. 6185, Sep. 2021.

[6] G. Sheng, Y. Qian, and L. Luo, "Key technologies and application prospects for operation and maintenance of power equipment in new power systems," *High Voltage Eng.*, vol. 47, no. 9, pp. 3072–3084, 2021.

[7] A. Ozdemir and S. Ilhan, "Experimental performance analysis of conventional and non-conventional lightning protection systems-preliminary results," *Electr. Power Syst. Res.*, vol. 216, Mar. 2023, Art. no. 109080.

[8] J. Zhao, C. Yu, and H. Ren, "Analysis of lightning strike fault cases in special forms of 1000kV ultra-high voltage transmission lines," *Electr. Porcelain Surge Arresters*, vol. 18, no. 1, pp. 107–112, 2022.

[9] H. Zhang, "Design and analysis of lightning protection and grounding for overhead transmission lines," *Electr. Technol. Econ.*, no. 4, pp. 69–71, 2021.

[10] Z. Lei and H. Liu, "Electromagnetic compatibility technology of power system automation equipment," *Electron. World*, no. 16, pp. 208–209, 2021.

[11] X. Tong, J. Chen, and Q. Chen, "Research on lightning parameter characteristics and application in Hainan power grid," *Electr. Porcelain Surge Arresters*, no. 1, pp. 134–138, 2015.

[12] Z. Gao, H. Pan, and F. Zhou, "Multidimensional analysis of lightning-induced trip-outs in Yunnan power grid transmission lines," *Yunnan Electr. Power Technol.*, vol. 46, no. 3, pp. 20–22, 2018.

[13] X. Peng, L. Jin, and R. Wang, "Analysis and prevention measures of lightning-induced simultaneous trip faults in 220kV same-tower lines," *Electr. Porcelain Surge Arresters*, no. 4, pp. 65–71, 2018.

[14] H. Deng, S. Li, and D. Dai, "Risk assessment of lightning strikes on overhead transmission lines based on hierarchical and differentiated approach," *Power Syst. Protection Control*, vol. 44, no. 4, pp. 69–75, 2016.

[15] L. Fang, S. Fang, and X. Xiong, "Probability analysis of indirect lightning faults in transmission lines and lightning forecasting methods," *Guangdong Electr. Power*, vol. 27, no. 3, pp. 95–100, 2014.

[16] Y. Liu, Y. Tian, and J. Xiong, "Research on lightning forecasting methods for Guangzhou power grid," *Shaanxi Electr. Power*, vol. 43, no. 7, pp. 88–91, 2015.

[17] B. Zhang, "Real-time prediction of lightning-induced trip probability for transmission lines based on lightning positioning system," Ph.D. dissertation, Dept. Power Syst. Automat., Southwest Jiaotong Univ., Chengdu, China, 2014.

[18] Z. Zhao, J. Shi, and Q. Yuan, "A new method for predicting lightning trip probability of transmission lines," *Autom. Electric Power Syst.*, vol. 39, no. 3, pp. 51–58, 2015.

[19] W. Wang, "Research on warning methods for wind-induced discharge and lightning trip risk of transmission lines under severe convective weather," Ph.D. dissertation, Dept. Elect. Eng., Chongqing Univ., Chongqing, China, 2017.

[20] S. Liu, X. Yang, and J. Yang, "Research on solutions to transmission capacity limitations of transmission corridors," *Northeast Electr. Power Technol.*, vol. 38, no. 5, pp. 51–54, 2017.

[21] H. Liu, "Research on distribution line lightning protection scheme evaluation system based on improved lightning trip rate calculation method," Ph.D. dissertation, South China Univ. Technol., Guangzhou, China, 2020.

[22] Y. Zhuang, Y. Xiao, and X. Chen, "Application of spatio-temporal trajectory clustering algorithm in lightning risk assessment," *Guangdong Meteorol.*, vol. 45, no. 3, pp. 95–98, 2023.

[23] M. Hahsler, M. Piekenbrock, and D. Doran, "DBSCAN: Fast density-based clustering with R," *J. Stat. Softw.*, vol. 91, no. 1, pp. 1–30, 2019.

[24] N. Gholizadeh, H. Saadatfar, and N. Hanafi, "K-DBSCAN: An improved DBSCAN algorithm for big data," *J. Supercomput.*, vol. 77, no. 6, pp. 6214–6235, Jun. 2021.

[25] P. Gao, H. Tian, and J. Li, "Thunderstorm mining and research based on improved DBSCAN algorithm," *High Voltage App.*, vol. 55, no. 4, pp. 169–177, 2019.

[26] C. Jiang, "Research on spatio-temporal trajectory clustering of mobile objects considering multidimensional semantics," Ph.D. dissertation, Dept. Geomatics Sci. Technol., Univ. Electron. Sci. Technol. China, Chengdu, China, 2022.

[27] H. Liu, "Analysis of thunderstorm characteristics in South China and potential forecast of XGBoost algorithm under multiple factors," Ph.D. dissertation, Dept. Geomatics Eng. GNSS Meteorol., Guilin Univ. Technol., Guangxi, China, 2021.

[28] S. Du, X. Liu, and H. Sun, "Study on the relationship between dual-polarization radar parameters and lightning activity in a typical thunderstorm process in South China," *J. Tropical Meteorol.*, vol. 37, no. 3, pp. 427–438, 2021.

[29] W. Yan, C. Xiao, and X. Wu, "Calculation of lightning trip rate for 10 kV distribution lines based on lightning activity characteristics," *High Voltage Eng.*, vol. 47, no. 3, pp. 1118–1127, 2021.

[30] W. Qin, "Analysis of severe convective weather processes in Guangdong province and research on lightning nowcasting warning," Ph.D. dissertation, Dept. Lightning Sci. Technol., Lightning Detection Warning, Nanjing Univ. Inf. Sci., Nanjing, China, 2017.

[31] S. Ramraj, N. Uzir, R. Sunil, and S. Banerjee, "Experimenting XGBoost algorithm for prediction and classification of different datasets," *Int. J. Control Theory Appl.*, vol. 9, no. 40, pp. 651–662, 2016.

[32] S. Dhaliwal, A.-A. Nahid, and R. Abbas, "Effective intrusion detection system using XGBoost," *Information*, vol. 9, no. 7, p. 149, Jun. 2018.

[33] L. Zhang, L. Kong, and H. Chen, "Hierarchical clustering method based on distance correlation coefficient," *Mathematica Numerica Sinica*, vol. 41, no. 3, pp. 320–334, 2019.

[34] C. Shen, S. Panda, and J. T. Vogelstein, "The chi-square test of distance correlation," *J. Comput. Graph. Statist.*, vol. 31, no. 1, pp. 254–262, Jan. 2022.

[35] S. Zeng, K. Liu, and Y. Han, "Analysis of lightning faults and lightning protection measures for overhead distribution lines," *Develop. Innov. Machinery Electr. Products*, vol. 36, no. 6, pp. 98–100, 2023.

[36] J. Zhang, Y. Zhang, and C. Liu, "Study on atmospheric electric field characteristics of lightning activity in complex terrain regions of Sichuan province," *Plateau Mountain Meteorol. Res.*, vol. 41, no. 4, pp. 56–60, 2021.

[37] Z. Zheng, Y. Yan, and J. Cai, "Non-specific TVLA method based on double-sample KS test," *J. Commun.*, vol. 44, no. 5, pp. 137–147, 2023.

[38] DSP Tian. *Minimum Enclosing Rectangle*. Accessed: May, 8, 2021. [Online]. Available: https://www.cnblogs.com/tiandsp/p/4044282.html

[39] A. Yadav and D. Yadav, "Wavelet tree based hybrid geo-textual indexing technique for geographical search," *Indian J. Sci. Technol.*, vol. 8, no. 33, pp. 624–632, Dec. 2015.

HONG HUO is currently the Leader of the Transmission and Maintenance Innovation Workshop, State Grid Dandong Electric Power Supply Company, and a Grid Craftsman with State Grid Liaoning Electric Power Company Ltd.

DEZHI WANG received the degree from Shenyang Engineering University. He has been engaged in the maintenance and inspection of transmission lines.

HAO CHEN has been involved in transmission line maintenance work for many years, accumulating extensive on-site experience.

CHUNLEI ZHAO has been consistently engaged in the maintenance and inspection of transmission lines, accumulating rich on-site work experience.

QI CHENG has been involved in the management of maintenance and inspection of transmission lines, providing guidance and participating in numerous national and provincial-level transmission projects.