Resilience-Oriented Transmission Line Fragility Modeling and Real-Time Risk Assessment of Thunderstorms

Jie Bao, Xin Wang, Yihui Zheng, Feng Zhang, *Member, IEEE*, Xuyong Huang, Peng Sun, and Zuyi Li, *Senior Member, IEEE*

***Abstract*—Fragility modeling and real-time risk assessment can be widely applied to evaluate and enhances the resilience of the power system to High-Impact and Low Probability events. In previous studies, fragility modeling generally targets extreme weather conditions other than thunderstorm. This paper proposes a fragility model to describe the relationship between the duration of a thunderstorm and the probability of lightning related trip-out. The duration of thunderstorms, which can usually be forecasted from the meteorological department, together with the fragility function expression can help a power company to predict the possibility of lightning related trip-out. Furthermore, this paper proposes a real-time risk assessment model that can dynamically adjust the risk value based on the update of the location, peak current, and subsequent stroke of real-time thunderstorm. A case study conducted on the lightning related trip-out data in Southwest China demonstrates that the average risk of transmission line trip-out in high risk group is about ten times that in low risk group. It clearly demonstrates that real-time risk assessment can efficiently distinguish the trip-out risks of different real-time thunderstorms.**

***Index Terms*—Resilience-oriented fragility modeling, real-time risk assessment, transmission line, extreme weather, lightning.**

I. Introduction

EXTREME weather has posed a severe threat to the operation of legacy power systems. High-voltage overhead transmission lines are especially more susceptible to extreme weather due to their wide-area distribution and harsh natural environment for operation. In Southwest China, over 82% of transmission line trip-outs from 2014 to 2016 were related to extreme weather. More specifically, more than 67% of transmission line trip-outs were related to the thunderstorm. The lightning-related transmission line trip-out does not refer to the lightning flashover rate of an overhead transmission line network (flashovers/100 km-yr), but to the number of unscheduled interruptions of the transmission power system due to lightning flashover of the interconnected overhead lines. Besides, the lightning location system detected more than 17 million lightning strikes in Southwest China from 2014 to 2016, but there were only a few hundred records of lightning related trip-out during the three years. Therefore, such High-Impact and Low Probability (HILP) thunderstorm events in terms of causing trip-outs have a significant impact on the operational resilience of the transmission system. In order to improve the resilience of the transmission system, it is essential to build a fragility model to describe the relationship between the trip-out of the transmission line and the historical thunderstorm. Besides, in terms of the real-time thunderstorm, this paper proposes the real-time risk assessment to helps provide more accurate early warning information for the operation of the system and enhances the resilience of the system to HILP events.

A diverse range of weather-related reliability models has been proposed in the literature [1]. Ref. [2] evaluates composite power system reliability considering weather effects based on Markov cut-set method. Authors in [3] discuss the impacts of wind storms in all categories based on a Sequential Monte Carlo method enhanced by a temporal wind storm sampling strategy. On the other hand, research on the resilience of power systems under natural disasters is also widely studied [4]. Ref. [5] focuses predominantly on proactive resilience against natural disasters. However, these previous works pay less attention on how to build the fragility model. Authors in [6] present a simulation-based fragility model to capture the interaction between the hurricane and physical power infrastructure. In addition, a fragility model of individual components is built for mapping the real-time impact of severe wind events in [7], [8]. Ref. [9] proposes an integrated assessment model to assess the impact of seismic events on the electric power system. Moreover, authors in [10] present an optimal resilient operation strategy in dealing with geomagnetic storms. Whereas, the above methods are not aimed at the lightning related fragility model. Ref. [11] summarizes the impact of lightning intensity on failure rate based on a non-homogenous Poisson process. However, the fragility model

Manuscript received July 31, 2020; revised February 2, 2021; accepted March 10, 2021. Date of publication March 17, 2021; date of current version July 23, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61673268 and in part by the Key Project of NSFC under Grant 61533012. Paper no. TPWRD-01175-2020. *(Corresponding author: Xin Wang.)*

Jie Bao, Xin Wang, Yihui Zheng, and Feng Zhang are with the Department of Center of Electrical and Electronic Technology, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: leprince@sjtu.edu.cn; wangxin26@sjtu.edu.cn; zhengyihui@sjtu.edu.cn; fzhang@sjtu.edu.cn).

Xuyong Huang is with Electric Power Research Institute, Yunnan Power Grid Company Limited, Kunming 650217, China (e-mail: 15504024@qq.com).

Peng Sun is with Dispatching Control Center, Yunnan Power Grid Company Limited, Kunming 650217, China (e-mail: ynsp@163.com).

Zuyi Li is with Electrical and Computer Engineering Department, Illinois Institute of Technology in Chicago, Chicago, IL 60616 USA (e-mail: lizu@iit.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPWRD.2021.3066157

Digital Object Identifier 10.1109/TPWRD.2021.3066157

based on lightning intensity is mostly static and cannot be updated in real time because the lightning density is calculated in years. In this paper, the fragility model is based on the duration of the thunderstorm, which can usually be forecasted from the meteorological department. Accordingly, the fragility of the power system can be dynamically updated based on the forecasted duration of thunderstorms.

Furthermore, when the meteorological department cannot accurately predict the duration of a thunderstorm, the real-time lightning location data can still be used to assess the risk. Research to assess the risk due to lightning has been conducted for a long period [12]–[16]. Ref. [17] applies the whole span calculation method to evaluate the risk of lightning shielding flashover. Authors in [18] propose that the number of lightning strikes shield depends on the transmission-line geometry. On the other hand, the impact of the environmental factor is regarded as the most vital factor in lightning risk assessment [19]. Ref. [20] quantifies lightning risk by transmission line parameters and soil resistivity. In addition, authors in [21] discuss the correlation between the environmental factor and the lightning related trip-out based on entropy-weighted grey correlation analysis. However, none of the above literatures consider the impact of real-time lightning strike. In our previous work [22], real-time lightning movement track is applied to assess the medium and long-term risk of transmission line trip-out. Furthermore, most of the existing lightning risk assessment models are off-line models and cannot be dynamically adjusted based on the real-time thunderstorm. This paper proposes a novel real-time risk assessment model based on the location, peak current, and subsequent stroke of real-time thunderstorm.

To sum up, the main contributions of this paper are as follows.

- Previous studies of fragility modeling generally target extreme weather conditions other than thunderstorm. This paper proposes a fragility model to describe the relationship between the duration of the thunderstorm and the probability of lightning related trip-out based on a statistical analysis of actual trip-out records. The duration of thunderstorms, which can usually be forecasted from the meteorological department, together with the fragility function expression can help a power company to predict the possibility of lightning related trip-out. Meanwhile, the fragility function can be applied to research on the resilience of power systems under the background of the thunderstorm.
- This paper proposes a real-time risk assessment model that can dynamically adjust the risk value of transmission line trip-out based on the update of the location, peak current, and subsequent stroke of thunderstorm in real time. Previous studies are generally off-line study and consider only historical trip-out events, environment factors or lightning intensity in calculating the trip-out risk.
- This paper conducts a real-time risk assessment case study for the transmission system of Southwest China, which has 109835 towers and 29 digital lightning detection stations, and recorded over 17 million lightning strikes yet only a few hundred lightning related trip-outs during the three-year period from 2014 to 2016. The case study clearly demonstrates that the proposed real-time risk assessment model can efficiently assess the trip-out risk of different real-time thunderstorms.

The rest of this paper is organized as follows. In Section II, the fundamental theory of the proposed model is introduced, while Section III presents the implementation of the proposed fragility model. In Section IV, the real-time risk assessment model is introduced. And in Section V, the case study is performed in the entire transmission system of Southwest China. Section VI summarizes the conclusions of this paper.

II. Fundamental Theory

For the thunderstorms with different duration, the proportion of lightning related trip-out is the key index to a fragility model and real-time risk assessment. In this paper, the proportion of lightning related trip-out is calculated based on the Affinity Propagation (AP) algorithm and the two-layer feed-forward network. The former is an unsupervised clustering method that is applied to cluster the historical thunderstorm in off-line. Based on the clustering result of the former, the latter is a supervised classification method that can improve the speed of the model to meet the requirement of real-time risk assessment.

A. Affinity Propagation Clustering Algorithm

The Affinity Propagation is an unsupervised clustering algorithm based on the concept of "message passing" between points. For each point, the AP algorithm continuously updates the responsibility and availability values with the iterative process. Finally, the central points are generated, and the remaining points are assigned to the corresponding clusters [23]. The flowchart of the AP algorithm is shown in Fig. 1, where the damping factor is mainly used for convergence, the responsibility matrix describes the degree to which point k is suitable as the central of point i, and the availability matrix describes the suitability of point i to select point k as central point.

The AP clustering algorithm is selected to cluster thunderstorms as it has three significant advantages as follow over other clustering algorithms [24]:

- It can overcome the large differences in the density of different classes.
- It does not need to set the scope of the search field which is the difficulty for other clustering algorithms.
- It has a robust clustering result.

B. Two-Layer Feed-forward Network

A two-layer feed-forward network is a neural network for supervised classification. The neural network utilizes a set of data and corresponding data classification labels for training. For newly input data, a trained neural network can output their corresponding classification labels. For example, a neural network can recognize a type of real-time thunderstorm based on the lightning movement track or the peak current of the historical thunderstorms.

As shown in Fig. 2, there are two layers in the neural network, including the hidden layer and output layer. The hidden layer introduces the sigmoid function to handle nonlinear multi-classification. It can classify vectors arbitrarily well based

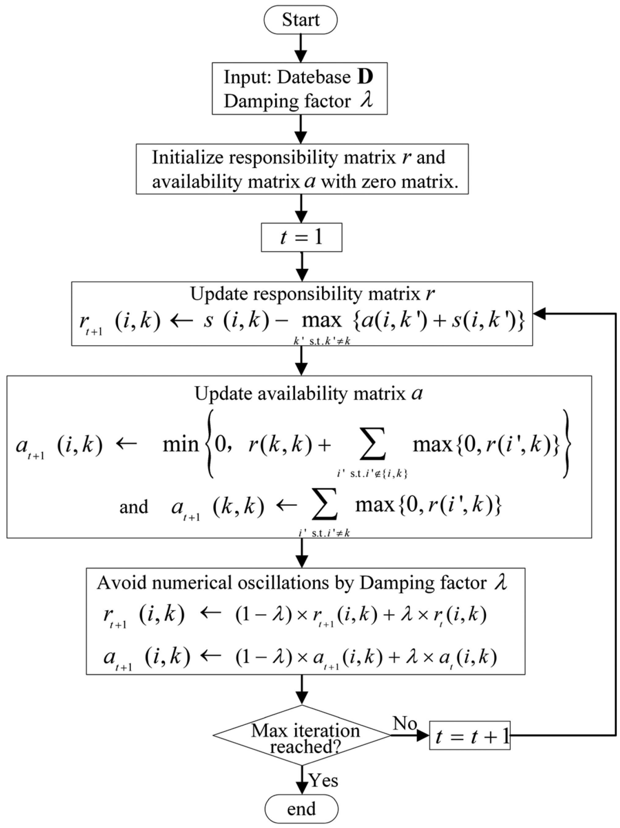


Fig. 1. The flowchart of AP clustering algorithm.

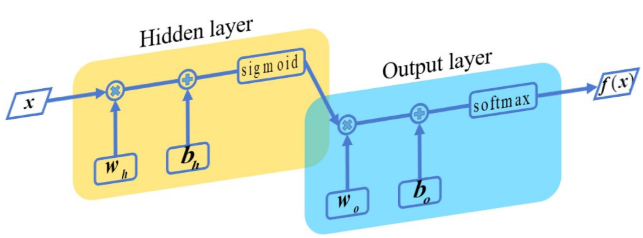


Fig. 2. Two-layer feed-forward network.

on enough neurons in its hidden layer. Moreover, the output layer introduces the softmax function to output one classification result with the largest proportion.

$$\mathrm{sigmoid}(x) = \frac{1}{1 + \mathrm{e}^{-(\boldsymbol{w}_h \cdot x + \boldsymbol{b}_h)}} \tag{1}$$

$$\mathrm{softmax}(\mathrm{x})_j = \frac{\mathrm{e}^{(w_o \cdot \mathrm{x_j} + b_o)}}{\sum_{\mathrm{i}}^{\mathrm{N}} \mathrm{e}^{(w_o \cdot \mathrm{x}_i + b_o)}}, \mathrm{j} = 1, 2, \ldots, \mathrm{N} \tag{2}$$

Furthermore, the weight $\boldsymbol{w}_h$ and bias $\boldsymbol{b}_h$ of the hidden layer are trained with scaled conjugate gradient backpropagation [25]. The weight $\boldsymbol{w}_o$ and bias $\boldsymbol{b}_o$ of the output layer are trained in the same way.

The feed-forward neural network has a simple structure and a wide range of applications. Its classification and pattern recognition capabilities are generally stronger than recurrent networks. Meanwhile, it has higher accuracy and fast calculation speed. In this paper, the feed-forward network can meet the requirements of real-time risk assessment for calculation accuracy and speed.

III. Transmission Line Fragility Modeling of Thunderstorms

The transmission line fragility model describes the probability of transmission line trip-out under thunderstorms. This probability depends on the potential intensity of a hazard [7]. The fragility model in this paper is used to describe the relationship between the duration of the thunderstorm and the probability of trip-out. There are two approaches to deriving a fragility model, including:

- Statistical analysis of actual trip-out records.
- Trip-out analysis using simulation software.

Of the above approaches, the actual trip-out records are indispensable data. The simulation software can be used for better evaluation of the actual data on power outages and predicting lightning-related outages. Therefore, the fragility model described in this paper is based on a statistical analysis of actual trip-out records and simulation software.

In recent years, power companies have installed lightning location systems on a large scale. This system has accumulated a lot of historical lightning strike data and can provide real-time lightning location data. For the transmission system, a single lightning strike near the transmission line is almost impossible to cause a trip-out. However, a series of lightning strikes near the transmission line means that there is a certain possibility of causing a trip-out. The thunderstorm is the collection of a series of lightning strikes. Therefore, it is essential to study thunderstorms before building the transmission line fragility model. Based on our previous work [22], a Lightning Identification, Tracking, and Analysis (LITA) algorithm has been proposed to calculate the track of the thunderstorm.

For historical lightning strike data, it is integrated into the corresponding thunderstorm by the LITA algorithm. Moreover, existing thunderstorm tracking methods are mainly based on weather radar [35]. However, the accuracy of radar is generally poor. In comparison, the LITA algorithm tracks thunderstorms based on lightning strike points rather than thunder cloud. Therefore, the LITA algorithm can track the location of thunderstorms with sufficient accuracy. Accurate thunderstorm location is essential to build a fragile model between thunderstorm and the trip-out of transmission lines.

Generally, the duration of thunderstorm is different. In particular, there are four types of the thunderstorm, including single cell storm, multicell storm, multicell line storm, and supercell storm [27], which have different durations. Therefore, this paper utilizes duration to distinguish thunderstorms in fragility modeling.

As long as there is a lightning strike near a transmission line, the corresponding thunderstorm is included in the statistics of the fragility model. For the lightning strike point closest to the transmission line at the time of the trip-out, the corresponding thunderstorm is recorded as causing the transmission line trip-out.

Then, the proportion of trip-outs corresponding to a thunderstorm duration can be expressed as follows:

$$\mathrm{P}_t = \frac{O_t}{L_t} \tag{3}$$

where t represents the duration of thunderstorm. L_t is the number of thunderstorms whose duration is t.O_t represents the number of lightning related trip-outs corresponding to the thunderstorms whose duration is t. As the time unit of the LITA algorithm which tracks thunderstorms is every 5 minutes, the time unit of the duration of thunderstorms is 5 minutes.

Furthermore, a fragility function is fitted based on the proportion of trip-out at different time durations. In fitting, the least squares method treats the importance of the data in the sequence equally. Considering that the cumulative number of trip-outs corresponding to diverse durations is different, the actual impact on the transmission line is also disparate. In other words, the thunderstorm with a large number of trip-outs is more important in function fitting. Therefore, this paper uses the weighted least squares to reflect the importance, giving a larger weight coefficient to the duration with a larger number of trip-outs, and a smaller weight coefficient to the duration with a smaller number of trip-outs [28]. For most lightning related trip-outs, the fitting function based on weighted least squares can more effectively reflect the probability of trip-outs. The weight coefficient w is defined as follow:

$$\mathrm{w}_t = \frac{\mathrm{O}_t}{\mathrm{O}_{\mathrm{all}}} \tag{4}$$

where w_t is the weight coefficient for the thunderstorms whose duration is t. $\mathrm{O}_{\mathrm{all}}$ represents the total number of lightning related trip-out.

By applying the weight coefficient to the fragility curve fitting, the fragility function can describe the probability of trip-out more effectively.

This section quantitatively describes the probability of trip-out under different durations of the thunderstorm. However, when the meteorological department cannot accurately predict the duration of a thunderstorm, the real-time lightning location data can still be used to assess the risk, as introduced in the next section.

IV. Transmission Line Real-Time Risk Assessment of Thunderstorms

This section describes the steps of transmission line real-time risk assessment under thunderstorms. Real-time risk assessment is a dynamic process. Based on the real-time update of the lightning location system, the risk value is continuously updated during different periods of thunderstorm development. The lightning location system can provide the time, location, number of subsequent strokes, peak current, and polarity of each lightning strike [29].

Based on our previous work [22], a thunderstorm can be represented by a series of central points. A central point includes all lightning strike points of a thunderstorm within 5 minutes. More specifically, the center point $\mathrm{i}$of a thunderstorm can be described as follow:

$$lo_i = \frac{1}{\mathrm{N}_i}\sum_{j=1}^{\mathrm{N}_i} lo_j \tag{5}$$

$$la_i = \frac{1}{\mathrm{N}_i}\sum_{j=1}^{\mathrm{N}_i} la_j \tag{6}$$

$$Pp_i = \frac{1}{\mathrm{N}_i}\sum_{j=1}^{\mathrm{N}_i} Pp_j \tag{7}$$

$$Pn_i = \frac{1}{\mathrm{N}_i}\sum_{j=1}^{\mathrm{N}_i} Pn_j \tag{8}$$

$$Re_i = \frac{1}{\mathrm{N}_i}\sum_{j=1}^{\mathrm{N}_i} Re_j \tag{9}$$

where j is the index of the lightning strike point of which the center point is i. N_i is the number of lightning strike points to which center point i belongs. lo_j and la_j represents the longitude and latitude of point j, respectively. Pp_j and Pn_j represents the positive and negative peak current of point j, respectively. Re_j is the number of subsequent strokes of point j.

Then, the central point of the thunderstorm is regarded as the basic unit of risk assessment. A central point represents all lightning strike data of a thunderstorm in 5 minutes. Lightning strike data include the location, the peak current, and the number of subsequent strokes. The duration of a central point is 5 minutes. A thunderstorm can be represented by multiple consecutive central points. Risk assessment is mainly to compare the similarity between historical thunderstorms and real-time thunderstorm, and then define the real-time thunderstorm risk based on the trip-out proportion of historical thunderstorms. In order to improve the accuracy of real-time risk assessment and reduce the false alarm rate, it is essential to cluster historical thunderstorms based on the AP algorithm. In real-time risk assessment, only historical thunderstorms of the same type as real-time thunderstorms are selected for analysis instead of all historical thunderstorms.

In this paper, with the development of the real-time thunderstorm, the risk assessment model dynamically adjusts the risk value of transmission line trip-out. However, the AP algorithm is slow when processing massive historical thunderstorms and cannot meet the calculation speed requirement of real-time risk assessment. Therefore, the risk assessment model builds the AP clustering result database $\mathrm{f}_{AP-H}(\mathrm{t})$ of historical thunderstorms at different time points in off-line.

$$\mathrm{f}_{AP-H}(\mathrm{t}) = \mathrm{APcluster}\begin{bmatrix} \mathrm{Pp}_1, Pn_1, Re_1 \\ \vdots \\ \mathrm{Pp}_s, Pn_s, Re_s \\ \vdots \\ \mathrm{Pp}_{\mathrm{N_t}}, Pn_{\mathrm{N_t}}, Re_{\mathrm{N_t}} \end{bmatrix} \tag{10}$$

$$\mathrm{Pp}_s = \frac{1}{\mathrm{D_t}} \sum_{i=1}^{\mathrm{D_t}} Pp_i^s \tag{11}$$

$$Pn_s = \frac{1}{\mathrm{D_t}} \sum_{i=1}^{\mathrm{D_t}} \mathrm{Pn}_i^s \tag{12}$$

$$Re_s = \frac{1}{\mathrm{D_t}} \sum_{i=1}^{\mathrm{D_t}} \mathrm{Re}_i^s \tag{13}$$

where t is the time point of the AP clustering. s is the index of historical thunderstorms. $\mathrm{N_t}$ represents the number of historical thunderstorms whose duration are over t. $\mathrm{D_t}$ is the number of central points when time point is t. Pp_s and Pn_srepresents the average value of positive and negative peak current of the historical thunderstorm s, respectively. Pp_i^s and Pn_i^srepresents the positive and negative peak current of the central point i to which the historical thunderstorm sbelongs, respectively. Re_s is the number of the subsequent stroke of the historical thunderstorm s. Re_i^s is the number of the subsequent stroke of the central point i to which the historical thunderstorm sbelongs.

For example, $\mathrm{f}_{AP-H}(25)$ represents the AP clustering result for the first 25 minutes of historical thunderstorms whose duration are over 25 minutes. In other words, only the data of the first 25 minutes of the above historical thunderstorm participated in the AP clustering. The above historical thunderstorm is clustered based on the average of the peak current and the number of the subsequent stroke.

Then, a two-layer feed-forward neural network is applied, which has fast calculation speed and very high accuracy for classification problems. More specifically, for the time point t, the two-layer feed-forward neural network $\mathrm{Net(t)}$ is trained based on the historical thunderstorm and the clustering result $\mathrm{f}_{AP-H}(\mathrm{t})$ of the AP algorithm. In real-time risk assessment, the peak current and the number of the subsequent strokes of real-time thunderstorm is input to the trained neural network. Then, the neural network outputs the classification result of real-time thunderstorm.

Moreover, this paper selects historical thunderstorms that meet the following conditions for real-time risk assessment:

- Belonging to the same type as the real-time thunderstorm which is currently assessed.
- Being close to the geographical location of the real-time thunderstorm which is currently assessed.

Furthermore, the AP algorithm is applied again to calculate the clustering results of historical thunderstorms and real-time thunderstorm based on their latitude and longitude data. In this step, the number of historical thunderstorms that meet the above conditions is much less than the number of all historical thunderstorms. Therefore, the calculation time can meet the requirements of real-time risk assessment.

$$\mathrm{f}_{\mathrm{AP}-R}^{t} = \mathrm{APcluster}$$

$$\begin{bmatrix} \mathrm{lo}_1^1, \mathrm{la}_1^1, \cdots \mathrm{lo}_\mathrm{i}^1, \mathrm{la}_\mathrm{i}^1, \cdots \mathrm{lo}_{\mathrm{D_t}}^1, \mathrm{la}_{\mathrm{D_t}}^1 \\ \vdots \ddots \vdots \ddots \vdots \\ \mathrm{lo}_1^\mathrm{h}, \mathrm{la}_1^\mathrm{h}, \cdots \mathrm{lo}_\mathrm{i}^\mathrm{h}, \mathrm{la}_\mathrm{i}^\mathrm{h}, \cdots \mathrm{lo}_{\mathrm{D_t}}^\mathrm{h}, \mathrm{la}_{\mathrm{D_t}}^\mathrm{h} \\ \vdots \ddots \vdots \ddots \vdots \\ \mathrm{lo}_1^{\mathrm{N_C}}, \mathrm{la}_1^{\mathrm{N_C}}, \cdots \mathrm{lo}_\mathrm{i}^{\mathrm{N_C}}, \mathrm{la}_\mathrm{i}^{\mathrm{N_C}}, \cdots \mathrm{lo}_{\mathrm{D_t}}^{\mathrm{N_C}}, \mathrm{la}_{\mathrm{D_t}}^{\mathrm{N_C}} \\ \mathrm{lo}_1^\mathrm{C}, \mathrm{la}_1^\mathrm{C}, \cdots \mathrm{lo}_\mathrm{i}^\mathrm{C}, \mathrm{la}_\mathrm{i}^\mathrm{C}, \cdots \mathrm{lo}_{\mathrm{D_t}}^\mathrm{C}, \mathrm{la}_{\mathrm{D_t}}^\mathrm{C} \end{bmatrix} \tag{14}$$

where t represents is the time point of clustering. $\mathrm{f}_{\mathrm{AP}-R}^{t}$ represents the clustering result of real-time thunderstorm and historical thunderstorms which meet conditions. h is the index of the historical thunderstorms which meet conditions. C represents the real-time thunderstorm which currently assessed. $\mathrm{N_C}$ is the number of historical thunderstorm which meets conditions. $\mathrm{D_t}$ is the number of central points when time point is t. lo_i and la_irepresents the longitude and the latitude of central point i, respectively.

Based on real-time clustering of AP algorithm, the risk value of real-time thunderstorm is set identical to the trip-out proportion of the same type of historical thunderstorm.

$$\mathrm{Risk(C)} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{N}_C} \mathrm{Z_i(C)}}{\sum_{\mathrm{i}=1}^{\mathrm{N}_C} \mathrm{R_i(C)}} \tag{15}$$

$$\mathrm{R_i(C)} = \begin{cases} 0, \text{Thunderstorm } \mathrm{i} \text{ and } \mathrm{C}\text{not belong to same type} \\ 1, \text{Thunderstorm } \mathrm{i} \text{ and } \mathrm{C}\text{belong to same type} \end{cases} \tag{16}$$

$$\mathrm{Z_i(C)} = \begin{cases} 0 \times \mathrm{R_i(C)}, \text{Thunderstorm i has not caused trip - out after duration t} \\ 1 \times \mathrm{R_i(C)}, \text{Thunderstorm i has caused trip - out after duration t} \end{cases} \tag{17}$$

where $\mathrm{Risk(C)}$ represents the risk value of real-time thunderstorm C. $\sum_{\mathrm{i}=1}^{\mathrm{N}_C} \mathrm{R_i(C)}$ represents the number of historical thunderstorms which are the same type as real-time thunderstorm C. $\sum_{\mathrm{i}=1}^{\mathrm{N}_C} \mathrm{Z_i(C)}$ represents the number of trip-outs of historical thunderstorms which are the same type as real-time thunderstorm C.

The flowchart of transmission line real-time risk assessment under thunderstorm is shown as Fig. 3. As the time unit of the LITA algorithm which tracks thunderstorms is every 5 minutes, the time unit of the real-time risk assessment is 5 minutes.

Note that the above real-time risk assessment model uses the AP clustering algorithm twice. On the one hand, it is to reduce the amount of calculation for each clustering to meet the speed requirements of real-time risk assessment. On the other hand, after two times of clustering, for those historical thunderstorms that are classified into the same type as the real-time thunderstorms, they are similar to real-time thunderstorms in location, peak current, and subsequent stroke. Their corresponding proportion of lightning related trip-out can more accurately describe the risk of real-time thunderstorm which is assessed.

V. Case Study for the Entire Transmission Network of Southwest China

Southwest China is a major energy base that has an extra-long transmission network. As shown in Fig. 4, there were 1031 transmission lines (110 kV and above) by the end of 2019.

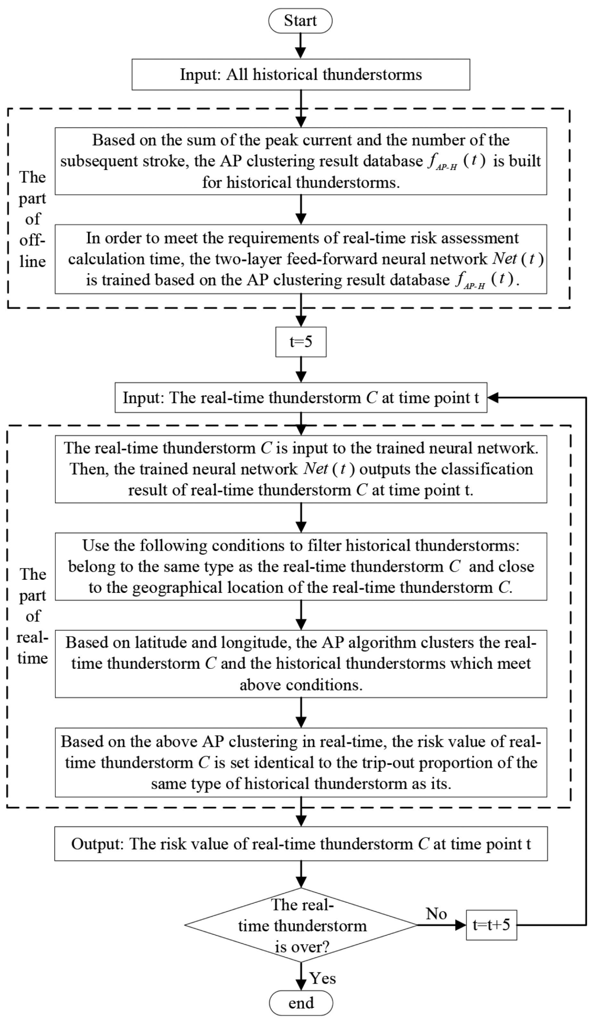


Fig. 3. The flowchart of transmission line real-time risk assessment under thunderstorm.

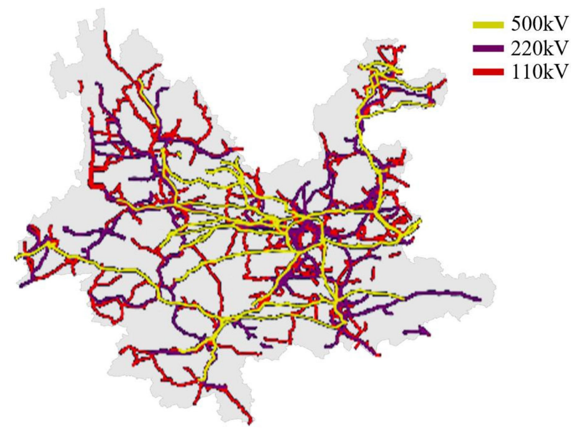


Fig. 4. The transmission network of Southwest China.

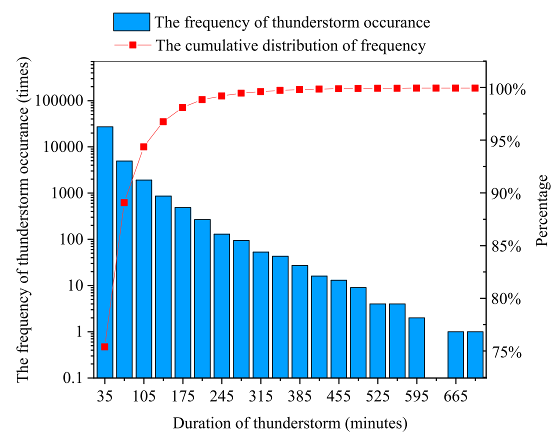


Fig. 5. The frequency of thunderstorm occurrence and cumulative distribution (The scale of y-axis is the logarithm base 10).

More specifically, there are 109835 towers in Southwest China. Moreover, there are 29 digital lightning detection stations [30] and all the lightning location systems in southern China had been interconnected, which can detect the border regions of each province and improve the accuracy of detection. More specifically, the location precision is within 500 meters, and the lightning current amplitude error is less than 15% [29]. The above performance of the lightning location system meets the requirement of IEC 62858:2015 [31].

In summary, an extra-long transmission network and an accurate lightning location system can provide a reliable basis for fragility modeling and real-time risk assessment.

A. *Transmission Line Fragility Modeling Under Thunderstorm*

The fragility function describes the probability of lightning related trip-outs. It can be derived from statistical analysis of actual trip-outs. In this paper, a lightning related trip-out that is cleared by a successful reclose is also counted as a trip-out.

Based on over 17 million lightning location data in Southwest China from 2014 to 2016, the frequency of thunderstorm occurrence and its cumulative distribution is shown as Fig. 5. To analyze the relationship between thunderstorm and lightning related trip-out, the thunderstorms are classified according to their durations. Since the durations of different thunderstorms are scattered, it is more convenient for the purpose of statistical analysis to group thunderstorms with similar durations into the same bin. Regarding the number of bins, too many bins will make it difficult to see a clear relationship between the frequency of trip-outs and the duration of thunderstorms, while too few bins cannot distinguish different types of thunderstorms well. In the case study of this paper, the durations of the thunderstorm are split into 20 bins. Considering that the life cycle of the single cell storm is generally about 30 minutes [26], appropriately extending to 35 minutes can cover the duration of most single cell storms. Therefore, the first bin includes the frequency of thunderstorms within 35 minutes, and the second bin includes the frequency of thunderstorms between 35 and 70 minutes, and so on. As shown in Fig. 5, more than 75% of the duration of thunderstorms is less than 35 minutes. But the longest duration of thunderstorm is 700 minutes.

Moreover, based on hundreds of trip-out records in Southwest China from 2014 to 2016, the frequency of lightning related trip-out and cumulative distribution are shown as Fig. 6. It can be

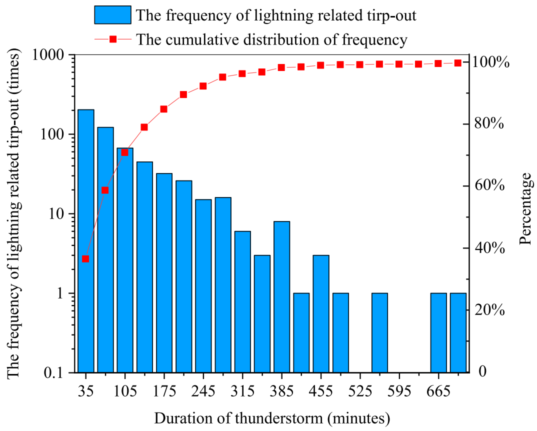


Fig. 6. The frequency of lightning related trip-out and cumulative distribution (The scale of y-axis is the logarithm base 10).

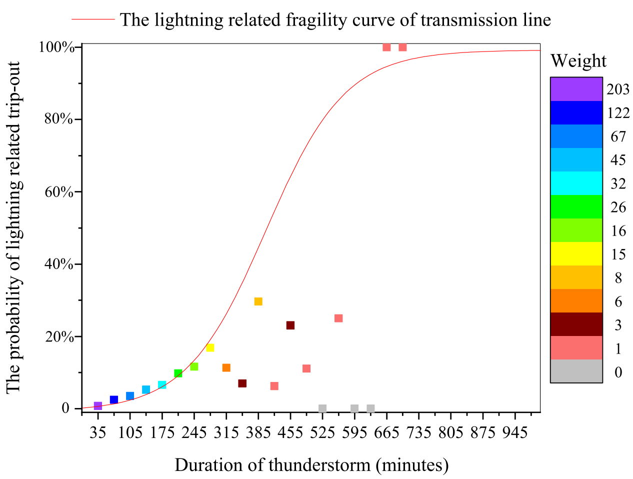


Fig. 7. The lightning related fragility curve of transmission line.

seen that the longest duration of thunderstorm is over 10 hours, but more than half of the trip-outs occurred within 1 hour.

Furthermore, the proportion of trip-out corresponding to different thunderstorm durations is shown by scatter points in Fig. 7. Considering the distribution of the scattered points, this paper uses the Dose-response curve fitting, which is one of the sigmoidal function in nonlinear curve fitting [32]. The formula of Dose-response function is defined as follow:

$$\hat{y} = \mathrm{A1} + \frac{\mathrm{A2} - A1}{1 + 10^{(\mathrm{LOGx_0} - \mathrm{x}) \times \mathrm{p}}} \tag{18}$$

where $\mathrm{A1}$ is the bottom asymptote. $\mathrm{A2}$ is the top asymptote. $\mathrm{LOGx_0}$ is the center of the Dose-response curve. p is the hill slope of the Dose-response curve.

Since the range of proportion of trip-outs is between 0 to 1, the bottom asymptote $\mathrm{A1}$ is set to 0 and the top asymptote $\mathrm{A2}$ is set to 1. The formula of Dose-response function can be rewritten as follow:

$$\hat{y} = \frac{1}{1 + 10^{(\mathrm{LOGx_0} - \mathrm{x}) \times \mathrm{p}}} \tag{19}$$

The estimate proportion $\hat{P}_t$ of trip-outs corresponding to a thunderstorm duration t can be expressed as follows:

$$\hat{P}_t = \frac{1}{1 + 10^{(\mathrm{LOGx_0} - \mathrm{t}) \times \mathrm{p}}} \tag{20}$$

Considering that the cumulative number of trip-outs corresponding to diverse durations is different, the actual impact on the transmission line is also different. Therefore, this paper uses weight coefficient w to reflect the importance. The weighted residual sum of squares S can be expressed as follows:

$$\mathrm{S} = \sum_{\mathrm{t}=1}^{\mathrm{n}} \mathrm{w}_t (\mathrm{P}_t - \hat{P}_t)^2 \tag{21}$$

where the weight coefficient w_t is the cumulative number of trip-out for different thunderstorm durations from 2014 to 2016.

The target of the function fitting is to minimize the weighted residual sum of squares S.

$$\mathrm{S} = \sum_{\mathrm{t}=1}^{\mathrm{n}} \mathrm{w}_t \left(\mathrm{P}_t - \frac{1}{1 + 10^{(\mathrm{LOGx_0} - \mathrm{t}) \times \mathrm{p}}}\right)^2 = \min \tag{22}$$

Then, the fragility function is fitted based on the partial derivative equations for $\mathrm{LOGx_0}$ and p. Finally, the fragile function expression considering weight is fitted as follow:

$$\mathrm{P(t)} = \frac{1}{1 + 10^{(0.005 \times (400 - \mathrm{t}))}} \tag{23}$$

In Fig. 7, there are some data points is far off the fitted curve between 315-630 on the x-axis. The weight coefficient of these data points is less than 8. It means that thunderstorms lasting from 315 to 630 minutes have less than 8 trip-outs. Most lightning related trip-outs actually occur in the thunderstorms that last less than 315 minutes. The thunderstorm with a large number of trip-outs is more important in function fitting. Therefore, for most lightning related trip-outs, the weighted fitting function can more effectively reflect the probability of trip-outs.

Note that the duration of thunderstorms can usually be forecasted from the meteorological department. Therefore, based on the fragility model and the forecasted thunderstorm duration, a power company can predict the possibility of lightning related trip-out.

B. Transmission Line Real-Time Risk Assessment Under Thunderstorm

The accuracy of the fragility model is affected by the forecast of thunderstorm duration from the meteorological department. To improve the probability accuracy of lightning related trip-out, the real-time risk assessment is proposed based on real-time lightning location instead of the weather forecast.

Based on LITA algorithm that calculates the lightning location system data, the thunderstorm data in Southwest China from 2014 to 2016 can be obtained. More specifically, the thunderstorms in Southwest China from 2014 to 2015 are used as historical thunderstorm data, while the thunderstorms in Southwest China in 2016 as test data for real-time risk assessment. Moreover, as shown in Fig. 6, more than 70% of lightning related trip-outs occurred within 100 minutes. In addition, there is too

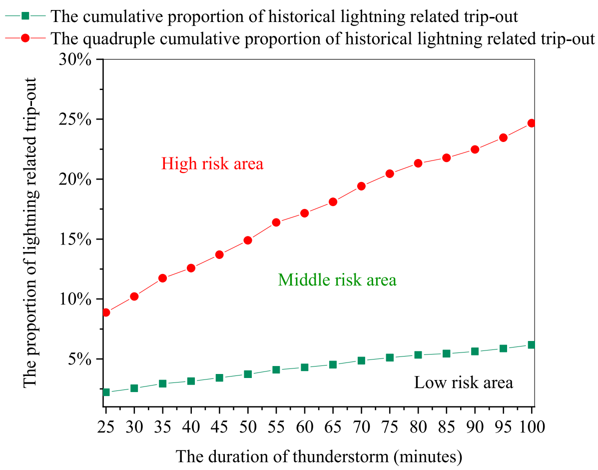


Fig. 8. The cumulative trip-out proportion of different thunderstorms whose duration are over t from 2014 to 2015 and the risk area classification of real-time thunderstorms in 2016.

little information to assess risk at the beginning of a thunderstorm. Therefore, the time point for real-time risk assessment of the test is between 25 and 100 minutes.

As mentioned in section IV, the risk value of a real-time thunderstorm is set identical to the trip-out proportion of the same type of historical thunderstorm it belongs to. Therefore, the trip-out proportion of historical thunderstorm is a key index to risk assessment. The cumulative trip-out proportion of different thunderstorms whose duration are over t from 2014 to 2015 is shown in Fig. 8. t is the abscissa value corresponding to the scattered point of proportion. For example, the first green dot from left to right in Fig. 8 represents that 2.2% of thunderstorms whose duration are over 25 minutes caused transmission line trip-out.

When calculating the real-time risk value of each thunderstorm in 2016, the cumulative trip-out proportion from 2014 to 2015 is used as a key index to risk assessment. For a real-time thunderstorm with current duration t, it belongs to

- the low risk group if its risk value is lower than the cumulative trip-out proportion of the same duration from 2014 to 2015.
- the middle risk group if its risk value is between the cumulative and quadruple cumulative trip-out proportion of the same duration from 2014 to 2015.
- the high risk group if its risk value is higher than quadruple cumulative trip-out proportion of the same duration from 2014 to 2015.

The risk value assessed for a real-time thunderstorm is just a number. The cumulative trip-out proportion of historical thunderstorms in 2014-2015 is used as the baseline for the classification of risk, which is actually just a reference. Whether it is really high-risk or low-risk depends on whether the corresponding high-risk group really has a higher proportion of trip-out than the low-risk group. The corresponding calculation time of the AP algorithm and the test set accuracy of the two-layer feed-forward network in classification is shown in Table I. Moreover, in the two-layer feed-forward network, the number of hidden neurons is 100. All calculations are performed on a personal computer using MATLAB 2018a with Intel Core i7-9700 CPU (3.00GHz) and 16 GB of memory. The calculation time of the AP algorithm in the off-line part is much longer than that in the real-time part. The latter meets the requirements of real-time risk assessment for calculation time. In addition, the historical thunderstorms data are randomly divided into three set, i.e., the training set, the validation set and the test set. The proportions of the three sets are 70%, 15%, and 15% of the whole data, respectively. The confusion matrix shows the true accuracy of the two-layer feed-forward network. For example, the confusion matrix of the two-layer feed-forward network at time point 25 minutes is shown in Fig. 9. The test set accuracy is based on the test confusion matrix. As shown in Table I, the test set accuracy of the two-layer feed-forward network in classification is over 99% at every time point. Therefore, the neural network satisfies the requirements of real-time risk assessment for the accuracy of classification.

TABLE I
THE CALCULATION TIME OF THE AP ALGORITHM AND THE TEST SET ACCURACY OF THE TWO-LAYER FEED-FORWARD NETWORK AT DIFFERENT TIME POINTS

Time point for real-time risk assessment	Calculation time of the AP algorithm in the off-line part	The average calculation time of the AP algorithm in the real-time part	The test set accuracy of two-layer feed-forward network in classification
25	768.29 s	11.27 s	99.1%
30	655.12 s	7.53 s	99.7%
35	354.67 s	5.51 s	99.9%
40	260.43 s	3.35 s	99.9%
45	122.13 s	2.75 s	99.7%
50	104.28 s	2.01 s	99.7%
55	100.93 s	1.86 s	100%
60	108.52 s	1.69 s	99.8%
65	100.10 s	1.35 s	99.4%
70	79.47 s	1.12 s	99.3%
75	66.59 s	0.92 s	99.7%
80	69.90 s	0.82 s	100%
85	65.11 s	0.68 s	99.7%
90	63.31 s	0.56 s	100%
95	56.44 s	0.44 s	99.9%
100	54.94 s	0.43 s	99.7%

To demonstrate the performance of the proposed risk assessment model, this paper introduces two other kinds of risk assessment models. In the comparison of three models, the real-time risk assessment model proposed in this paper is named the RT model. The two other kinds of risk assessment models are the One-time RT model and the Analytic Hierarchy Process (AHP) model, respectively [33], [34].

The One-time RT model is the RT model without updating at every time point. The RT model dynamically adjusts risk value based on the development of real-time thunderstorm. However,

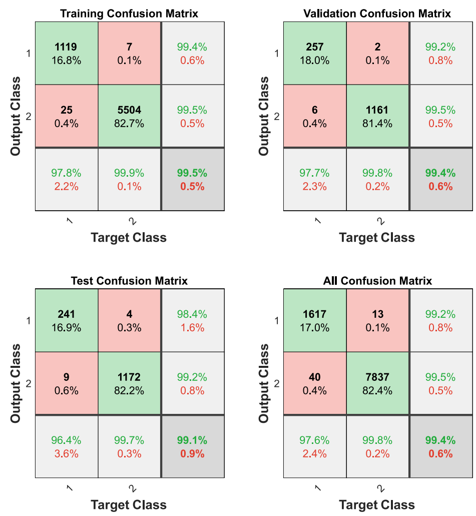


Fig. 9. The confusion matrix of the two-layer feed-forward network at time point 25 minutes.

the One-time RT model only assessed the risk value at the initial time point. The initial time point is 25 minutes in this paper. Therefore, in the case study of One-time RT model, the risk value of the thunderstorms are all based on the risk assessment at time point 25 minutes. Moreover, the classification standard of real-time thunderstorm risk is consistent with the RT model.

The AHP model is based on the number of the lightning strikes, the sum of the peak current, and the number of the subsequent strokes. The mean values of the above three parameters for all historical thunderstorms were calculated as the reference values. For a real-time thunderstorm, the risk value of the three parameters can be obtained by comparing the three parameters with the reference value of all historical thunderstorms. The parameter of real-time thunderstorm is larger than the reference value, the higher the risk value of the corresponding parameter. The risk values of the three parameters are weighted and summed as the real-time thunderstorm risk value. The weight values of the above three parameters in the risk assessment model are based on Analytic Hierarchy Process. Moreover, the classification standard of real-time thunderstorm risk is consistent with the RT model. The general assessment idea of the AHP model can be summarized as follows: the real-time thunderstorm with a larger number of the lightning strikes, the sum of the peak current, and the number of the subsequent strokes, the higher the risk of the corresponding real-time thunderstorm.

As shown in Table II, the proportion causing transmission line trip-out after real-time risk assessment in the high risk group is higher than the low risk group. The likelihood of trip-outs in high risk group vs. in low risk group represents the efficiency of picking out the thunderstorms which cause trip-out events. For the different time points of risk assessment, the average likelihood of the RT model, the One-time RT model, and the AHP model is 10, 3.8, 2.89, respectively. It shows that the RT model is the most efficient in picking out the thunderstorm that causes the actual trip-out.

The average likelihood of the RT model represents a 246% improvement compared to that of the AHP model. It demonstrates

TABLE II
THE RESULTS OF THREE REAL-TIME RISK ASSESSMENT MODELS AT DIFFERENT TIME POINTS

Time point for real-time risk assessment	Risk model	The proportion causing transmission line trip-out after real-time risk assessment			Likelihood of trip-out in high risk group vs. in low risk group
		High risk group	Middle risk group	Low risk group	
25	RT model	7.84%	4.36%	2.34%	**3.35**
	One-time RT model	7.84%	4.36%	2.34%	3.35
	AHP model	5.84%	4.93%	2.33%	2.51
30	RT model	12.73%	5.61%	1.74%	**7.32**
	One-time RT model	8.64%	5.15%	2.84%	3.05
	AHP model	7.83%	5.33%	2.88%	2.72
35	RT model	20.00%	6.52%	1.59%	**12.58**
	One-time RT model	9.86%	5.62%	2.71%	3.64
	AHP model	6.86%	6.04%	3.22%	2.13
40	RT model	25.00%	6.39%	2.11%	**11.88**
	One-time RT model	11.11%	6.05%	3.25%	3.42
	AHP model	10.11%	6.76%	3.26%	3.10
45	RT model	37.50%	6.68%	2.87%	**13.06**
	One-time RT model	10.71%	6.80%	3.20%	3.35
	AHP model	10.39%	7.28%	3.33%	3.12
50	RT model	33.33%	7.34%	2.80%	**11.89**
	One-time RT model	12.24%	7.36%	3.11%	3.94
	AHP model	9.23%	7.74%	3.58%	2.58
55	RT model	30.00%	8.05%	3.20%	**9.38**
	One-time RT model	14.63%	7.59%	3.65%	4.01
	AHP model	10.91%	7.64%	4.32%	2.53
60	RT model	18.18%	9.23%	2.54%	**7.17**
	One-time RT model	18.18%	9.23%	2.54%	7.17
	AHP model	9.30%	9.09%	4.38%	2.12
65	RT model	36.36%	9.43%	4.21%	**8.64**
	One-time RT model	17.65%	9.01%	4.37%	4.04
	AHP model	12.12%	9.07%	5.23%	2.32
70	RT model	50.00%	10.22%	5.05%	**9.89**
	One-time RT model	18.75%	9.30%	4.82%	3.89
	AHP model	17.24%	9.17%	5.32%	3.24
75	RT model	20.00%	11.25%	2.40%	**8.32**
	One-time RT model	16.67%	9.47%	5.19%	3.21
	AHP model	18.52%	9.20%	5.20%	3.56
80	RT model	44.44%	10.68%	5.15%	**8.63**
	One-time RT model	18.52%	9.94%	5.73%	3.23
	AHP model	16.13%	9.66%	6.09%	2.65
85	RT model	50.00%	13.60%	4.15%	**12.06**
	One-time RT model	20.00%	11.04%	5.92%	3.38
	AHP model	16.00%	11.32%	5.91%	2.71
90	RT model	20.00%	13.77%	3.07%	**6.52**
	One-time RT model	22.73%	11.68%	4.73%	4.81
	AHP model	15.38%	12.55%	4.81%	3.20
95	RT model	25.00%	14.07%	2.42%	**10.33**
	One-time RT model	22.73%	12.30%	5.30%	4.29
	AHP model	18.18%	12.33%	5.37%	3.39
100	RT model	50.00%	14.29%	3.05%	**16.38**
	One-time RT model	20.00%	12.11%	5.79%	3.46
	AHP model	19.23%	11.50%	5.80%	3.32
Average	RT model	30.00%	9.50%	3.00%	**10.00**
	One-time RT model	15.60%	8.60%	4.10%	3.80
	AHP model	12.70%	8.70%	4.40%	2.89

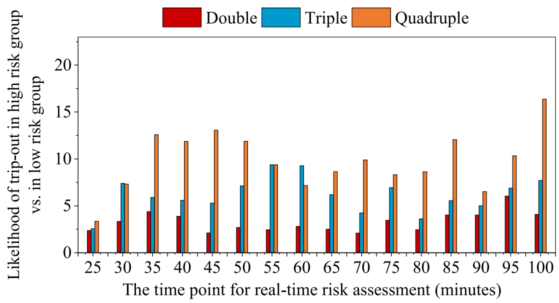


Fig. 10. The comparison of different high risk thresholds.

the advantage of the RT model on real-time risk assessment. Compared with the One-time RT model, which only assesses the risk value at the initial time point, the RT model has significantly improved the average likelihood by 163%. Therefore, it is necessary to dynamically adjust the risk value as the real-time thunderstorm develops.

This paper attempts to compare the effects of different high risk thresholds. As shown in Fig. 10, the high risk thresholds are set to double, triple, and quadruple cumulative trip-out proportion. By comparing the likelihood of trip-outs in high risk group vs. in low risk group of different thresholds in Fig. 10, the threshold of the quadruple cumulative trip-out proportion picks out the thunderstorms which cause trip-out events more effectively. Therefore, the high risk threshold is the quadruple cumulative trip-out proportion in this paper.

Moreover, the benefit of real-time risk assessment is to help the electric power company make dispatch decisions. For the transmission lines near high risk thunderstorms, the power company can reduce the operating power to improve the lightning protection performance. It isn't necessary to make dispatch decisions for every thunderstorm. Because most of the thunderstorms does not cause the trip-out event. Focusing on the high risk thunderstorms, which are assessed from the RT model, can improve the efficiency of dispatch decisions.

VI. Conclusion

This paper proposed the fragility model and real-time risk assessment of thunderstorms. The conclusions can be summarized as follows.

1) Based on the statistical analysis of actual trip-out records, this paper proposed a fragility model to describe the relationship between the duration of thunderstorms and the probability of lightning related trip-out. Meanwhile, the corresponding fragility function expression is proposed, which can help a power company to predict the possibility of lightning related trip-out as the duration of thunderstorms can usually be forecasted from the meteorological department. Meanwhile, the proposed fragility function expression can be applied to research on the resilience of power systems under the background of the thunderstorm.
2) The proposed real-time risk model combines the location, peak current, and subsequent stroke of real-time thunderstorm so it can dynamically adjust risk value based on the development of real-time thunderstorm. The case study shows that, in comparison with the AHP model and the One-time RT model, the RT model can improve the average likelihood of trip-out in high risk group vs. in low risk group by 246% and 163%, respectively. It demonstrates the advantage of the RT model on real-time risk assessment.
3) The case study based on the real data in Southwest China shows that the average risk of trip-out in high risk group is 10 times that in low risk group. It clearly demonstrates that the proposed real-time risk assessment model can efficiently pick out the trip-out risks of different real-time thunderstorms.

References

[1] M. Panteli and P. Mancarella, "Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies," *Electric Power Syst. Res.*, vol. 127, pp. 259–270, Oct. 2015.

[2] Y. Liu and C. Singh, "Reliability evaluation of composite power systems using Markov cut-set method," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 777–785, May 2010.

[3] G. Li *et al.*, "Risk analysis for distribution systems in the north-east U.S. under wind storms," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 889–898, Mar. 2014.

[4] Y. Wang, C. Chen, J. Wang, and R. Baldick, "Research on resilience of power systems under natural disasters—A review," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1604–1613, Mar. 2015.

[5] M. A. Mohamed, T. Chen, W. Su, and T. Jin, "Proactive resilience of power systems against natural disasters: A literature review," *IEEE Access*, vol. 7, pp. 163778–163795, Nov. 2019.

[6] S. Ma, N. Carrington, A. Arif, and Z. Wang, "Resilience assessment of self-healing distribution systems under extreme weather events," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Atlanta, GA, USA, 2019, pp. 4–8.

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

[8] M. Kirsty and K. Bell, "Wind related faults on the GB transmission network," in *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst.*, Durham, U.K., 2014, pp. 1–6.

[9] S. Espinoza, A. Poulos, H. Rudnick, J. Carlos, M. Panteli, and P. Mancarella, "Risk and resilience assessment with component criticality ranking of electric power systems subject to earthquakes," *IEEE Syst. J.*, vol. 14, no. 2, pp. 2837–2848, Jun. 2020.

[10] L. Gong, Y. Fu, M. Shahidehpour, and Z. Li, "A parallel solution for the resilient operation of power systems in geomagnetic storms," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3483–3495, Jul. 2020.

[11] K. Alvehag and L. Soder, "A reliability model for distribution systems incorporating seasonal variations in severe weather," *IEEE Trans. Power Del.*, vol. 26, no. 2, pp. 910–919, Apr. 2011.

[12] V. Rakov and F. Rachidi, "Overview of recent progress in lightning research and lightning protection," *IEEE Trans. Electromagn. Compat.*, vol. 51, no. 3, pp. 428–442, Aug. 2009.

[13] J. He, X. Wang, Z. Yu, and R. Zeng, "Statistical analysis on lightning performance of transmission lines in several regions of China," *IEEE Trans. Power Del.*, vol. 30, no. 3, pp. 1543–1551, Jun. 2015.

[14] R. Zeng *et al.*, "Survey of recent progress on lightning and lightning protection research," *High Voltage*, vol. 1, no. 1, pp. 2–10, Apr. 2016.

[15] E. Volpov and E. Katz, "Characterization of local environmental data and lightning-caused outages in the IECo transmission-line network," *IEEE Trans. Power Del.*, vol. 31, no. 2, pp. 640–647, Apr. 2016.

[16] V. Rakov and M. Uman, "Review and evaluation of lightning return stroke models including some aspects of their application," *IEEE Trans. Electromagn. Compat.*, vol. 40, no. 4, pp. 403–426, Nov. 1998.

[17] H. Wen, Z. Chun, S. Yun, and Z. Jun, "Evaluation of whole span lightning shielding flashover risk based on 3-D laser scanning technology," *IEEE Trans. Power Del.*, vol. 29, no. 2, pp. 794–801, Apr. 2014.

[18] P. Mikropoulos and T. Tsovilis, "Estimation of lightning incidence to overhead transmission lines," *IEEE Trans. Power Del.*, vol. 25, no. 3, pp. 1855–1865, Jul. 2010.

[19] P. Mikropoulos and T. Tsovilis, “Statistical method for the evaluation of the lightning performance of overhead distribution lines,” *IEEE Trans. Dielectic Electric Insul.*, vol. 20, no. 1, pp. 202–211, Feb. 2013.

[20] S. Visacro and F. Silveira, “The impact of the frequency dependence of soil parameters on the lightning performance of transmission lines,” *IEEE Trans. Electromagn. Compat.*, vol. 57, no. 3, pp. 434–441, Jun. 2015.

[21] C. Xie, J. Bai, W. Zhu, G. Luo, and H. Wang, “Lightning risk assessment of transmission lines based on D-S theory of evidence and entropy-weighted grey correlation analysis,” in *Proc. IEEE Conf. Energy Internet Energy Syst. Integration*, Beijing, China, Nov. 2017, pp. 1–6.

[22] J. Bao, X. Wang, Y. Zheng, F. Zhang, X. Huang, and P. Sun, “Lightning performance evaluation of transmission line based on data-driven lightning identification, tracking, and analysis,” *IEEE Trans. Electromagn. Compat.*, vol. 63, no. 1, pp. 160–171, Feb. 2021.

[23] B. Frey and D. Dueck, “Clustering by passing messages between data points,” *Science*, vol. 315, no. 5814, pp. 972–976, Feb. 2007.

[24] M. Ester *et al.*, “A density-based algorithm for discovering clusters in large spatial databases with noise,” in *Proc. 2nd Int. Conf. Knowl. Discov. Data Mining*, Portland, PD, USA, 1996, pp. 226–231.

[25] M. Moller, “A scaled conjugate gradient algorithm for fast supervised learning,” *Neural Netw.*, vol. 4, no. 4, pp. 525–533, 1993.

[26] S. Henry, “Analysis of thunderstorm lifetime as a function of size and intensity,” in *Proc. 26th Conf. Radar Meteorol.*, Oklahoma, OK, USA, 1993, pp. 138–140.

[27] R. Brown, C. Kaufman, and D. MacGorman, “Cloud-to-ground lightning associated with the evolution of a multicell storm,” *Geophys. Res.*, vol. 107, no. 19, pp. 1–13. Oct. 2002.

[28] J. Gill. “Weighted Least Squares,” in *Generalized Linear Models: A Unified Approach*, Los Angeles, CA, USA: SAGE, 2001, pp. 42–44.

[29] J. Chen *et al.*, “Present status and development trend of lighting detection and protection technology of power grid in China,” *High Voltage Eng.*, vol. 42, no. 11, pp. 3361–3375. Nov. 2016.

[30] S. Chen, Y. Du, L. Fan, H. He, and D. Zhong, “A lightning location system in China: Its performances and applications,” *IEEE Trans. Electromagn. Compat.*, vol. 44, no. 4, pp. 555–560, Nov. 2002.

[31] *Lightning Density Based on Lightning Location Systems (LLS) - General Principles*, IEC Standard 62858, 2015.

[32] OriginLab Corporation. Origin Help. 2012. [Online]. Available: https://www.originlab.com/doc/Origin-Help/DoseResp-FitFunc

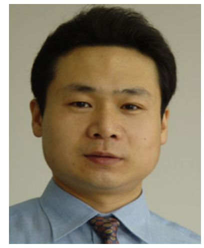
[33] X. Mou, Z. Guo, G. Lin, Q. He, Z. Wang, and J. Cao, “The optimization of regional lightning risk assessment based on AHP and fuzzy comprehensive evaluation,” in *Proc. 7th Asia-Pacific Int. Conf. Lightning*, Chengdu, China, 2011, pp. 331–334.

[34] R. Duan *et al.*, “500kV transmission line lightning transformation evaluation based on improved analytic hierarchy process,” in *Proc. IEEE PES T&D Conf. Expo.*, Chicago, IL, USA, 2014, pp. 1–5.

[35] M. Dixon *et al.*, “TITAN: Thunderstorm identification, tracking, analysis, and nowcasting-a radar-based methodology,” *J. Atmospheric Ocean. Technol.*, vol. 10, no. 6, pp. 785–797, Dec. 1993.

Jie Bao received the B.S. degree in electrical engineering from Hohai University, Nanjing, China, in 2016, and is currently working toward the Ph.D. degree with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China. He is also a Visiting Researcher with the Robert W. Galvin Center for Electricity Innovation, Illinois Institute of Technology, Chicago, IL, USA. His research interests include data-driven natural disaster risk evaluation and resilience-oriented fragility modeling.

Xin Wang received the B.S. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1993, and the master’s and Ph.D. degrees in control theory and engineering from Northeastern University, Shenyang, China, in 1998 and 2002, respectively. From 2002 to 2004, he was a Postdoctoral Research Fellow with Shanghai Jiao Tong University and currently an Associate Professor. His current research interests include smart grid, microgrid, energy internet, distributed generation, transmission, storage and forecasting technology of renewable energy resource, and power equipment online monitoring technology.



Yihui Zheng received the B.S., M.S., and Ph.D. degrees in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1993, 2002, and 2010, respectively. He is currently a Professor of electrical research with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University. His interests include electrical power quality and intelligent control technology in power system.

Feng Zhang (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Xi’an Jiao Tong University, Xi’an, China, in 1991 and 1996, respectively. From 2000 to 2002, he was a Postdoctoral Researcher with the Kyushu University, Kyushu, Japan. He is currently a Professor with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China. His current research interests include the electrical theory and new technology, intelligent diagnosis of electrical equipment, and model of metro transportation. He is a Member of the Higher Education Steering Committee of the National Ministry of Education and a Member of Chinese Society of Electrical Engineering of Theoretical Electrical Engineering Professional Committee.

Xuyong Huang received the B.S. degree in electrical engineering from the Wuhan University of Hydraulic and Electrical Engineering, Wuhan, China, in 1996, and the master’s and Ph.D. degrees in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2002 and 2008, respectively. He is currently a Senior Engineer with the Electric Power Research Institute, Yunnan Power Grid Company Limited, Kunming, China. His current research interests include power system monitoring, big data application, and early warning and prediction of power grid natural disasters.

Peng Sun received the Ph.D. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2003. He is currently a Professorate Senior Engineer with the Dispatching Control Center, Yunnan Power Grid Company Limited, Kunming, China. His current research interests include the transmission and transformation equipment status assessment, and operation and maintenance strategy.

Zuyi Li (Senior Member, IEEE) received the B.S. degree from Shanghai Jiao Tong University, Shanghai, China, in 1995, the M.S. degree from Tsinghua University, Beijing, China, in 1998, and the Ph.D. degree in electrical engineering from the Illinois Institute of Technology (IIT), Chicago, IL, USA, in 2002. He is currently a Professor with the Electrical and Computer Engineering Department, IIT. His current research interests include economic and secure operation of electric power systems, renewable energy integration, energy internet, cybersecurity in smart grid, microgrid operation, electricity internet of things, and smart consumption grid.