33rd International Conference on Lightning Protection
25 - 30 September 2016
Estoril, Portugal

A New Method to Evaluate the Lightning Flashover Risk of Transmission Line Based On Data Mining

Jian Li[1], Yao Yao[2,3], Xueming Zhou[2,3], Danhui Hu[2,3], Ling Ruan[2,3]
1: Wuhan NARI Co., Ltd. of State Grid Electric Power Research Institute;
2: State Grid Hubei Electric Power Research Institution, Wuhan, China;
3: SGCC Key Laboratory of Field Testing Technology, Wuhan, China;
hust-lijian@qq.com

***Abstract*—Lightning flashover risk evaluation is the basis for choosing proper measures to improve the lightning performance of the transmission lines. Thus, the effectiveness of the measures adopted is directly determined by the evaluation accuracy. At present, the risks are mainly evaluated through comparing the calculated lightning flashover rate (LFOR) of every tower. However, there are so many assumptions and simplifications in the process of calculation. This way will often cause some discrepancies between the evaluated results and operating experience. A new evaluation method based on data mining will be proposed to study the characteristics of the lightning failure history. The association relationship between the lightning failure and altitude, height difference, relative height difference, ground flash density, etc. will be considered to build classification rules library to determine the tower lightning flashover risk. Comparison of the proposed evaluation method with the traditional methods shows that the results deduced through the proposed method is more agree with the operating experience.**

Keywords-data mining; operating experience; lightning flashover; risk evaluation;

I. INTRODUCTION

The operating experience of power transmission lines shows that the lightning flashover is the main reason for line tripping. Though the SGCC (State Grid Corporation of China) has taken many lightning protection measures, the LFOR remains high [1-3]. Using lightning arrester is a very effective method. But the arrester cannot be popularized in large areas due to the very expensive cost. The present researches show that the lightning performance of lines having different tower structure, or in different areas, or in different ground flash density grade, is different [3]. Thus, installing the arrests only on towers with high lightning flashover risk will greatly reduce the LFOR and also has the best economy at the same time. So how to evaluate the lightning flashover risk becomes a crucial question attracting researcher's attention.

In order to evaluate the risk, most of the researchers focused on the simulation model of lightning flashover [4-5]. In order to improve the simulation precision and to reproduce the actual lightning fault features, several simulation algorithm including electro-geometric model, leader progression model and monte-carlo method were proposed. Based on these simulation models, researchers have studied lots of risk evaluation methods using some certain risk analysis methods.

According to the lightning flashover mechanism, the intensity of lightning activity, the insulation configuration, the line structure as well as the geomorphic conditions all have significant influence on the lightning performance of transmission lines. However, the recent evaluation methods cannot consider all the factors precisely leading to the disagreements between calculated results and operating experiences.

During the development and operation of the power system in China, lots of lightning flashover faults have been accumulated. These towers on which once lightning flashover happened, must have relatively high risk. But these valuable experiences have not been considered as guidance in the recent evaluation methods. In this paper, the data mining technology was adopted to study the characteristics of all the lightning flashover towers including the association relationship between the lightning fault and the altitude, height difference, relative height difference, flash density of the place in where the towers were located. Through studying the lightning flashover probability of every multiple factors combined interval, a classification rules library was built to determine the lightning flashover risk of every tower of the transmission line.

II. QUANTITATIVE ANALYSIS OF OPERATING DATA

Although the ground flash density, the ground slope angle and the terrain conditions (mountains or plains) are already considered in the calculation of LFOR, the effects of the actual running environment are more complicated. According to the operating experiences, these towers located on the top of hill, or having large crossing will be struck by the lightning more easily. In order to find out the degree to which the environment factors affects the probability of lightning flashover, the relationship between the lightning performance and the environment factors are firstly studied separately through analyzing the actual operating data in different areas.

A. Lightning performance in different lightning grades

According to the ground flash density, *Ng,* Hubei province could be divide into six lighting grades, A~D2. The definitions are as follows.

According to the ground flash density, *Ng,* Hubei province could be divide into six lighting grades, A~D2. The definitions are as follows

Grade A: N_g<0.78 flashes/km^2/year

Grade B1: 0.78 <=N_g<2 flashes/km^2/year

Grade B2: 2 <=N_g<2.78 flashes/km^2/year

Grade C1: 2.78 <=N_g<5 flashes/km^2/year

Grade C2: 5<=N_g<7.98 flashes/km^2/year

Grade D1: 7.98<=N_g<11 flashes/km^2/year

Grade D2: N_g>=11 flashes/km^2/year

The lightning performance of the 500kV AC transmission lines in Hubei province in different lightning grades is shown in figure 1. In this paper, the LFOR is defined as the number of lightning flashovers per unit tower per unit time.

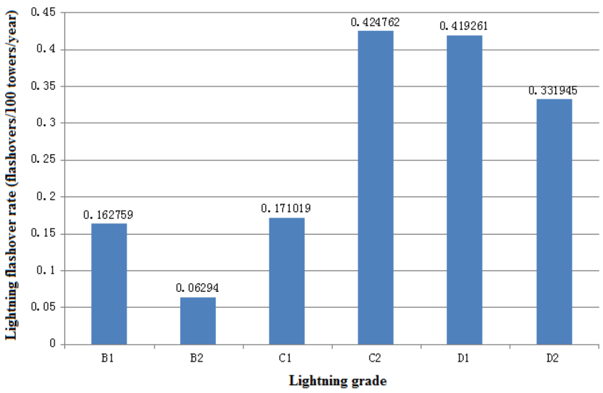

Figure 1. LFOR in different lightning grades

On the whole, the LFOR increases with the ground flash density as shown in figure 1. The transmissions lines located in C2~D2 areas have a significantly higher flashover rate than that located in B1~C1 areas. However, there are also some exceptions that the LFOR in B1 areas is higher than that in B2. A further analysis shows that there are 9 lightning flashovers in B1areas, and all the flashover tower are on the top of hill with the ground slope angle greater than 52 degree. These exceptions show that the lightning performance is affected not only by the ground flash density, but also by some other factors.

B. Lightning performance at different altitudes

In order to study the effect of altitude on the lightning performance, the altitudes (H) of towers in 500kV AC transmission lines in Hubei province are divided into 5 classes using natural break method. And the break points are 163.1m, 406.7m, 695.2m and 1043m. The lightning performance of the towers altitude classes is shown in figure 2.

It is shown that the LFOR does not present a consistent relationship with the altitude. Towers in both the high altitude and the low altitude areas have rather low LFOR. The highest LFOR towers are at the altitude between 406.7 m and 695.2 m. At the altitude of class 3, there are 1785 towers and 28 towers of them have been struck to flashover in the recent 5 years. The LFOR is 0.341 flashes/100 towers/year. The lowest is the area with the altitude in class 1.

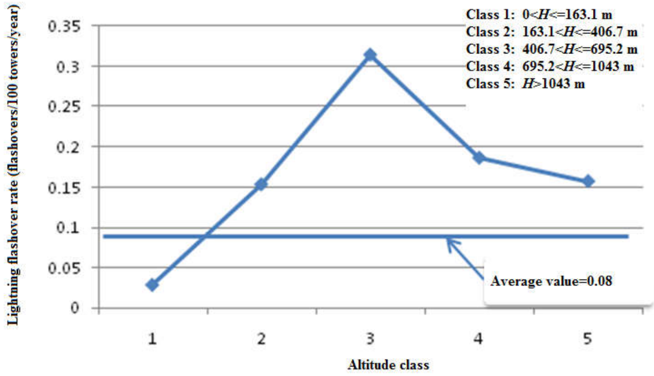

Figure 2. LFORs in different altitude areas

C. Lightning performance on the hills with different ground slope angles

Statistics show that most of the lightning flashes of 500 kV transmission line are caused by shielding failure. The ground slope angle, *S*, has significant influence on the shielding performance of the ground since the average shielding angle should be obtained by subtracting the hill angle relative to horizontal from the shielding angle [6]. Similarly, the slope angles can also be divided into 5 classes using natural break method. The lightning performances of the towers in different slope angel areas are shown in figure 3.

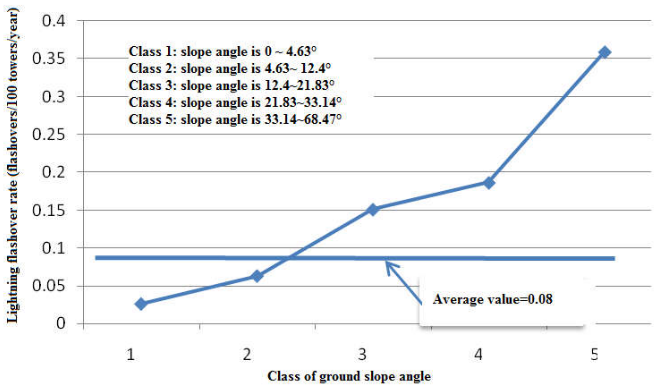

Figure 3. LFOR in different groud slope angle areas

It is shown in figure 3 that the LFOR increases with the ground slope angle. Because an increasing of ground slope angle will weaken the shielding effect of the ground since the exposed distance of the phase conductor outside of the slope is increased [6].

D. Lightning performance in an areas with different orographic fluctuation

Trees and structures beside the line right-of-way are beneficial to the lightning performance, because they increase the height of the effective earth plane, and may sometimes reduce the exposure distance of the phase conductor. Thus, the ground bumps near towers are also beneficial and lines located in the bottom of hill are hardly stroke. In order to the consider this effect deducing by the rise and fall of the earth plane, local height difference, ΔH, and the relative height difference, ΔH_r are defined in following equations, as shown in figure 4. It is noted that a local area with a radius of 200m around the tower is considered in next analysis.

$$\Delta H = H_{max} - H_{min} \tag{1}$$

$$\Delta H_r = (H - H_{min})/(H_{max} - H_{min}) \tag{2}$$

Where,

H_{max} is the altitude of the highest point in the local area.

H_{min} is the altitude of the lowest point in the local area.

H is the altitude of the tower

Thus, the relative height difference gives the relative position in local space and it can be divided into 4 classes as shown in figure 4.

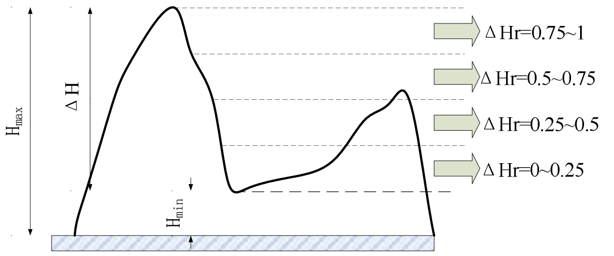

Figure 4. diagram of the definition of relative height difference

All the 500kV towers in Hubei province are divided into 5 classes based on their height difference using natural break method. And the LFOR of towers in areas with different height difference and different relative height difference are given in figure 5 and figure 6.

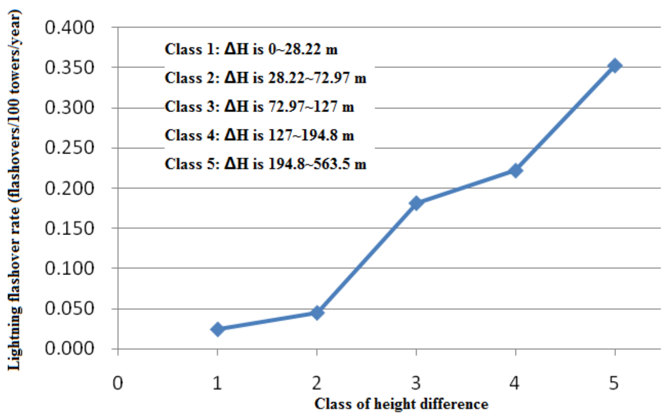

Figure 5. LFOR of towers with different height difference

It can be found that towers with greater height difference have higher flashover rates. And the LFOR increases with the increasing of the relative height difference. That is mean the more the tower near the highest point of local area, the more easily the lightning flashover occurs. These results can be easy to understand. The reason is the same with that lines located along hilltops are especially vulnerable since the effective height above ground is increased [6].

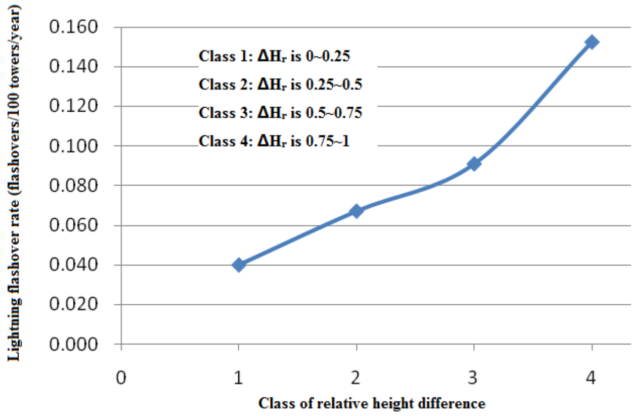

Figure 6. LFOR of towers with different relative height difference

III. ANALYSIS USING DATA MINING

From the above analysis, it can be found that the altitude, ground flash density, ground slope angle, height difference and the relative height difference of the place where a tower located in all affect lightning performance. And it is impossible to find an equation to calculate the flashover rate of a specific tower based on its environment conditions. At the same time, the factors, such as the altitude and ground flash density, affect each other. Thus, using some simple statistic method can hardly give out the LFOR.

The purpose of data mining is to find the hidden rule from a large amount of mutual coupling data. Thus, the data mining technology provides a powerful tool to analyze the operating experiences in lightning protection. In order to find the characteristic of the towers that have great probability to flashover, the data mining method is investigated to analyze the lightning operation data.

A. Seletion of the data mining technology

Data mining technology mainly includes two kinds of analysis algorithm: classification and regression analysis [7]. Regression analysis is to give an estimation value of a continuous variable, while classification is to predict the property of a discrete variable. In this paper, in order to discover the hidden rule from the history of the lighting flashover occurred in 500kV transmission lines, the purpose of data mining is to predict the lightning flashover risk. Thus, the classification algorithm is used to distinguish the high risk and low risk towers. Through analysis and comparison, a widely used classification method, decision-tree algorithm [7], is chosen to classify the risk grade. Using the actual lightning experiences of all the 500kV towers in Hubei province as samples to train the decision-tree, a risk decision-tree classification model can be created.

B. Building the decision- tree

The operating data of the 500 kV AC transmission lines in recent five years is selected in the training process. When building the decision-tree to evaluate the flashover risk, the towers are taken as analysis objects. All the factors including altitude, ground slope angle, height difference, relative height difference, ground flash density are chosen as the input properties, and the object value is the flashover risk to be classified.

Because the generation of decision-train is a kind of supervised learning process, the samples need to be pre-classified before the training. The towers on which lightning flashover have occurred can be classified as high risk while the other towers cannot all be considered to be low risk. In order to classify, the object value (flashover risk) is set to “1” (high risk) if lightning flashover has happened, and the object value is set to “0” if lightning flashover has not happened.

The generation process of traditional decision-tree is to improve the recognition accuracy of both “high risk” and “low risk”, thus the weights of the recognition accuracy of “1” and “0” are the same. However, because most of the towers have not been struck to flashover, if using regular division growth method, the decision-tree will not grow since there are 22022 towers having the risk value of “0” and only 86 towers having the risk value of “1”. In regard to the total 22108 towers, even if all the samples are classified as "0", the accuracy rate has reached 99.611%, which already meet the general decision tree classification accuracy. Obviously, to classify all the towers as "0" cannot meet the requirement of lightning risk assessment. Thus, during the growth, the decision-tree need to be trimmed artificially aiming at searching for the optimal classification method to maximize the difference of lightning flashover rate in different classes.

Through leaning the operating data of 500 kV ac transmission lines in Hubei province and artificially trimming, the preliminary part of the decision-tree to classify the lightning flashover risk is generated as shown in figure 7.

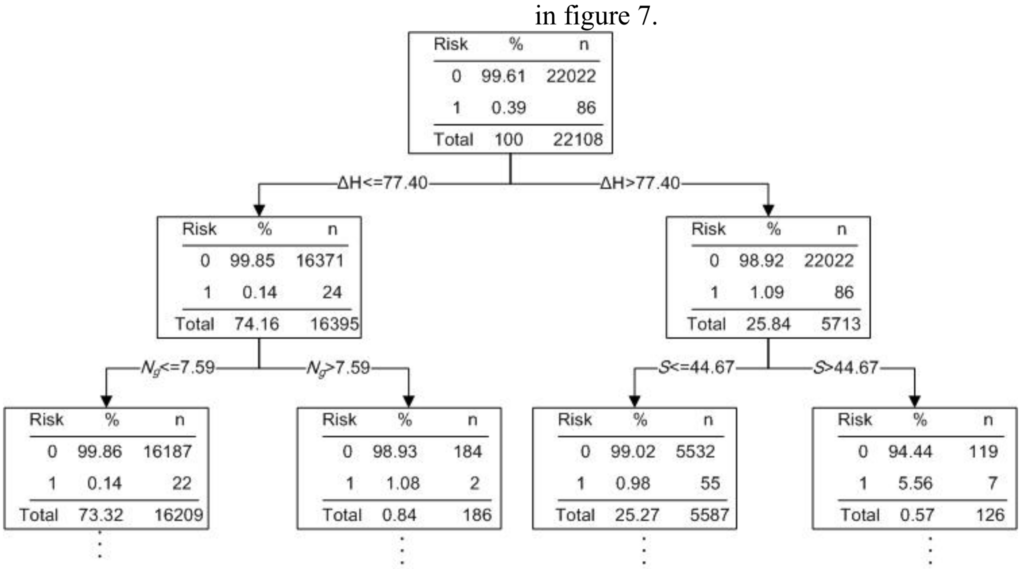

Figure 7. Part of the decision-tree model to evaluate the lightning flashover risk of transmission line tower

IV. RISK CLASSIFICATION RULES LIBRARY

A. Building of the risk classification rules library

According to the decision-tree generated above, the classification rules can be easily deduced through listing all the classification principles from the top-most root node to the bottom-most node of the tree. Then, all the towers can be classified into 11 sub-classes as shown in table I.The classification rules as well as the LFOR value are also presented in the table I.

TABLE I. THE LIGHTNING FLASHOVER RISK CLASSIFICATION RULES EXTRACTED FROM THE DECISION-TREE

Sub-class	Classification rules	LFOR
1	if ΔH <= 77.40 and N_g <= 7.59 and S <= 15.92 and H <= 336.032	0.021
2	if ΔH <= 77.40 and N_g <= 7.59 and S <= 15.92 and H > 336.03	0.143
3	if ΔH <= 77.40 and N_g <= 7.59 and S > 15.92	0.114
4	if ΔH <= 77.40 and N_g > 7.59 and H <= 242.70	0.111
5	if ΔH <= 77.40 and N_g > 7.59 and H > 242.70	3.333
6	if ΔH> 77.40 and S <= 44.67 and $ΔH_r$<= 0.73 and N_g <= 4.47	0.096
7	if ΔH > 77.40 and S <= 44.67 and $ΔH_r$ <= 0.73 and N_g > 4.47	0.221
8	if ΔH > 77.40 and S <= 44.67 and $ΔH_r$ > 0.73 and N_g <= 6.56	0.343
9	if ΔH > 77.40 and S <= 44.67 and $ΔH_r$ > 0.73 and N_g > 6.56	1.724
10	if ΔH > 77.40 and S > 44.67 and $ΔH_r$ <= 0.44	0
11	if ΔH > 77.40 and S > 44.67 and $ΔH_r$ > 0.44	1.458

B. *The determination of risk level of every subclass*

According to the provision deduced by SGCC, the LFOR control standard for 500kV transmission lines is 0.14 flashovers/100km/year (converted to 40 thunderstorm days/year), considering the average span of 0.426km, therefore, the equivalent control LFOR for towers is 0.05 flashovers/100 towers/year (converted to the actual number of thunderstorm days/year in Hubei province). The risk level of every sub-class is defined in table II.

TABLE II. THE DEFINITION OF LIGHTNING FLASHOVER RISK LEVEL OF DIFFERENT SUBCLASSES

LFOR interval	**(0,0.05]**	**(0.05,0.1]**	**(0.1,0.25]**	**(0.25,1)**
Sub-class	1, 10	6	2, 3, 4, 7	5,8,9,11
Risk level	A	B	C	D

C. *Verification of the evaluation methods*

In regard to a specific 500 kV transmission line, table 3 gives a comparison of the risk evaluation results using both the decision-tree model and traditional method in which the flashover rates are calculated using the method recommended by IEEE [6]. In table 3, R1 refers to evaluated results using traditional method, and R2 refers to that deduced by the decision-tree model.

TABLE III. AN EXAMPLE OF LIGHTNING FLASHOVER RISK EVALUATION BASED ON DATA MINING TECHNOLOGY

Tower No.	N_g	**H(m)**	**S(°)**	**ΔH(m)**	ΔH_r	**flashover or not**	**R1**	**R2**
#42	C2	250.5	13.3	142.0	0.71	0	B	C
#43	C2	420.0	8.2	145.2	0.60	0	B	C
#44	C2	551.9	38.1	259.7	0.58	0	B	C
#45	C2	546.2	48.2	222.7	0.40	0	D	A
#46	C2	775.3	11.6	66.2	0.53	1	B	C
#47	C2	792.4	19.7	99.1	0.53	0	A	C
#48	C2	801.0	16.1	92.5	0.56	0	B	C
#49	C2	778.4	18.4	92.6	0.42	0	A	C
#50	C2	783.4	25.4	119.9	0.70	0	A	C
#51	C2	686.8	9.8	214.0	0.85	0	A	C

In table III, the lightning flashover risk of #46 tower is evaluated as "C", high risk, by the decision-tree model proposed in this paper, while the traditional method gives a result of "B". Since #46 has occurred to lightning flashover, the result deduced by the decision-tree model shows more agreements with the operating experience.

In addition, an obvious difference can be found on #45 tower. The decision-tree model gives an evaluation value of "A", while the traditional method gives "D". A further analysis of the tower information shows that the tower height of #45 is very high which will lead to the increasing of the exposure distance of the phase conductor, so the lightning flashover rate is calculated to be great by traditional method. However, because the height difference around #45 is 222.7 and the relative height difference is 0.40, the tower is located in the mountain valley. So, the high terrain around the tower will attract most of the lightning. That is why the risk of the tower is evaluated as "A". Thus, #45 has not occurred to flashover yet even if its height is higher than #46.

V. CONCLUSIONS

1. A new lightning risk evaluation method is introduced to learn from the operating experience based on data mining.

2. The environment factors, including altitude, height difference, relative height difference and ground flash density in the place where tower located all have great influence on the lightning performance of a transmission line.

3. Data mining technology is good at discovering the characteristic hidden behind the big operating data.

4. The decision-tree risk evaluation model can well account for the effect of the surrounding environment on the lightning performance of the tower. The evaluation results are somewhat more in line with the operating experience.

VI. REFERENCE

[1] Chun Zhao, Jiangjun Ruan, Xiaolan Li, et al. Technology and Economy Evaluation of Comprehensive Transmission Line Lightning Protection Measures. High Voltage Engineering, 2011, 37(02):290-297

[2] Hengxin He, Junjia He, Guan-jun Qian, et al. Lightning Shielding Analysis Model of UHV AC Overhead Transmission Line. High Voltage Engineering, 2010, 36(01): 196-204

[3] Yaoheng Xie, Wei Xie, Shanqiang Gu, et al. Lightning Performance Analysis of Shielding Failure for 220 kV Transmission Lines with Double Circuit on the Same Tower Using Lightning Parameters Statistics. High Voltage Engineering, 2009, 35(11): 2657-2662.

[4] Shanqiang Gu, Jiahong Chen, Wei-jiang Chen, et al. Evaluation Method of the Time-space-difference of Lightning Protection Performance of Transmission Lines. High Voltage Engineering, 2009, 35(02):294-298.

[5] Junjia He, Xiaolan Li, Xiaogen Yin. Calculation of Shielding Failure Flashover Rate for Transmission Line Based on Revised EGM. High Voltage Engineering, 2006,32(03):42-44.

[6] IEEE,Std1243-1997. IEEE Guide for Improving the Lightning Performance of Transmission Lines.

[7] Pang-Ning Tan, Michael Steinbach, Vipin Kumar. Introduction to Data Mining. Pearson, 2005.