Investigation of Arcing Time Prediction for Secondary Arc Based on Statistical Theory and Multivariable Regression Algorithms

Baihe Su,1 Hongshun Liu,1 Ziyue Zhang,1 Luyao Liu,1 and Zhiyuan Zhang2

1School of Electrical Engineering, Shandong University, Jinan 250061, China
2State Grid Jinan Power Supply Company, Jinan 250012, China

Correspondence should be addressed to Zhiyuan Zhang; zhiyuanz@mail.sdu.edu.cn

Received 6 April 2022; Accepted 23 July 2022; Published 30 August 2022

1. Introduction

With the increasing deployment of UHV lines, worldwide research on extinguishing time of the secondary arc is an important issue that needs to be addressed. Studies have shown that more than 90% of faults in UHV transmission lines are single-phase ground faults [1]. When a single-phase ground fault occurs in a power system, the circuit breaker (CB) and single-phase automatic reclosing (SPAR) operate together to cut off the faulty phase. However, during the operation of SPAR, the electromagnetic coupling between lines generates a secondary arc. The secondary arc continues to burn, resulting in the failure of the SPAR [2, 3]. SPAR is an important equipment for relay protection, and its operation depends on accurately predicting the extinction time of secondary arc [4, 5]. Therefore, a clear and detail understanding of this is crucial to correcting the operation of SPAR.

A considerable amount of research has been conducted on the parameters, mechanism of generation, and kinematic law of the secondary arc in recent years. An algorithm has been proposed for detecting the edges in images of the secondary arc to solve the problem of discontinuous and false edges in conventional edge detection operators. The radius, length, and area of the secondary arc were analyzed by processing the edge coordinates thus obtained [6]. The variations in voltage harmonics generated during single-phase AC fault clearance were used to estimate the secondary arc extinction [7]. The secondary arc model has been developed in ATP, which brings together the strengths of the main existing models and implements a pseudorandom variation of the secondary arc length, to provide accurate results for the determination of extinction times of secondary arc in studies of monopolar reclosures [8]. The stochastic model of secondary arc with different initial positions was applied into the arc chain model to calculate the arcing time with dispersion [9]. The extinction characteristics of secondary arc near an outdoor insulator under different secondary arc currents are studied experimentally [10]. EMTP (Electro-Magnetic Transient Program) was used...
to establish the secondary arc current calculation model of a 500 kV Power Station’s transmission line [11]. No study to date has sought to predict the extinguishing time of the secondary arc, resulting in the theoretical guidance for SPAR is lacking. This study thus proposes a method to predict the arcing time based on statistical theory and multiple algorithms.

In this study, key parameters are selected via statistical theories and further concrete expression on arcing time is determined from multivariable regression algorithms. In comparison with the current method of determining the arcing time of secondary arc, we propose a novel method for predicting the arcing time. This method can provide a theoretical basis for the fast and accurate operation of SPAR.

The remainder of this study is organized as follows: Simulations of the secondary arc and the means of acquisition are presented in Section 2. In Section 3, analysis of variance model is used to identify the key parameters that significantly influence the arcing time and verify it through residual analysis. The DBSCAN algorithm is used to remove outliers in the data in Section 4. Two regression models are used in Section 5 to fit the arcing time.

2. Preparation of Secondary Arc

2.1. Experimental Setup. Low-voltage simulations of the secondary arc were carried out at the China Electric Power Research Institute. The circuit that was used is shown in Figure 1. The extinguishing time of the secondary arc is determined by the recovery voltage gradient (kV/m) of the arc channel instead of the amplitude of the recovery voltage. Therefore, a short gap arc can be used to simulate secondary arcs on UHV/EHV overhead lines with the same recovery voltage gradient. It is feasible to ignite the secondary arc with a small current (1000 A) instead of a large current at the same voltage gradient. It is feasible to ignite the secondary arc with a small current (1000 A) instead of a large current at the same recovery voltage gradient with a short arc gap of 0.68 m. Thus, the low-voltage simulation experiment can equivalently simulate the secondary arc on the UHV transmission line [2, 12, 13].

It was composed of an AC power source (peak voltage of 11.6 kV), equivalent capacitance \( C_q \), reactor \( L_s \), an insulator string, and circuit breakers \( CB_1, CB_2 \) and \( CB_3 \). \( L_s \) was used to generate 1 kA of short-circuit arc current and the two ends of the insulator string were connected by a nichrome wire to generate the short-circuit arc to form an arc channel. A clear image was obtained through a high-speed camera for analysis. The equivalent capacitance can be expressed as

\[
C_q = \left( 2C_p + \frac{C_g}{2} \right) I_0. 
\]

\( C_p \) and \( C_g \) respectively, represent the phase-to-phase capacitance and the phase-to-ground capacitance per unit length, respectively, and \( I_0 \) represents the length of the overhead line. Secondary currents of 15 A, 30 A, and 45 A were generated by changing the value of \( C_q \). The secondary current here was the effective value of the current calculated through the circuit and can be expressed by \( I_E \). The values of \( L_s \) and \( C_q \) and the corresponding values of \( I_E \) are shown in Table 1.

A voltage divider (1000:1 voltage ratio) and a current transformer CT were used to record the voltage and current of the secondary arc, respectively, and gradients of the recovery voltage between the electrodes and the secondary arc current in the loop were selected based on the results [12, 14, 15]. An air fan was placed around the insulator and artificial wind was set by changing the position and adjusting gear of the fan. Once the arc had been extinguished, a recovery voltage was generated at both ends of the insulator string. The waveforms of secondary current and recovery voltage were recorded by using an oscilloscope.

2.2. Parameter Acquisition. The combustion process of secondary arc was very complicated and is shown in Figure 2. Under a short-circuit large current, the nichrome wire was quickly broken at \( t = 0.008 \) s. The air around the insulator was strongly ionized and a bright arc plasma channel was formed under a short-circuit large current. The arc channel was diffused over time. When the arc plasma channel diffuses to the greatest extent of the area at \( t = 0.08 \) s, the arc channel is narrowed and the brightness is weakened. Secondary arc was generated in the plasma channel. With the extinction and reignition of arc, the morphology of arc changed greatly. In the last reignition at \( t = 0.194 \) s, the arc was very bright and its shape were complex, which makes it easy to extract the morphological characteristics of secondary arc [13, 16]. When the arcing time was short, the arc channel was not completed diffused and thus can be captured by camera as the arc area. The arc area also reflects the intensity of discharge, which is significant to determining the arcing time. As a
morphological characteristic of arc, arc area can also contribute to the determination of the arcing time.

Raw images of the secondary arc were captured by a high-speed camera in Figure 3(a). To acquire the arc morphology characteristic, corresponding binary images were processed by using the Color Thresholder in Figure 3(b). Pixel values of the nichrome wire and secondary arc can be read by binary images in Image Region Analyzer in MATLAB. The ratio can be obtained from the pixel value of the nichrome wire and the actual length. The actual arc length can be calculated by the ratio of pixel to actual length. Therefore, from binary images of secondary arc, the length, area, and diameter of the arc were obtained for further analysis.

The recovery voltage and secondary current are shown in Figure 4 when the current is 30 A and capacitive. From Figure 4(a), when secondary arc extinguished, the maximum recovery voltage appeared at the first peak of the waveform (point A in Figure 4). The maximum secondary current appeared before arc extinguished and a strong arc discharge was produced at the same time (point B in Figure 4).

3. Key Parameter Selection

The combustion of secondary arc is complicated. It is not determined which parameter influences the arcing time significantly. To determine the correlation between the arcing time and the parameters, analysis of covariance (ANCOVA) method in statistics is used to find out the key parameters [17, 18]. The key parameters are selected out for subsequent fitting.
In the secondary arc experiment platform, length of the electrode, wind speed and direction, secondary current value, and properties are artificially set to different values. These parameters are fixed factors in the ANCOVA method. After the experiment, parameters such as maximum recovery voltage, transient secondary current when the maximum recovery voltage was generated, and the maximum secondary current were obtained. Some parameters cannot be artificially designed and controlled in the experiment platform such as the arc area, arc length, and arc diameter. These parameters can be viewed as covariates in the statistical analysis.

ANCOVA is a statistical algorithm which combines the characteristics of variance analysis and regression analysis. The main ingredient of ANCOVA consists of two parts:

\[ y = (X_1\beta + X_2\gamma) + \epsilon. \]  

The independent variables corresponding to \( X_1\beta \) are factors, and \( X_1 \) is a 0-1 matrix. \( X_1\beta \) represents the part of the variance analysis model in ANCOVA. The independent variables corresponding to \( X_2\gamma \) are covariates, and \( X_2\gamma \) stands for the part of the regression analysis model. \( \epsilon \)
represents for random noise. $X_1\beta_1, X_2\gamma$, and $\varepsilon$ can be calculated by statistical software.

Covariance analysis is used to obtain the modified average arcing time through linear regression. By converting the covariates to equal values, the influence of confounding covariates is controlled. Then, variance analysis is used to compare the differences between the modified means.

In the regression analysis, collinearity in the covariates can lead to the results inconsistent with objective facts. This process can be achieved during regression analysis. Collinearity of the covariates can be determined by the variance inflation factor (VIF) as follows:

$$\text{VIF} = \frac{1}{1 - R_i^2}.$$  

(3)

$R_i$ is the negative correlation coefficient in regression analysis between one covariate and other covariates. The larger the VIF is, the greater is the likelihood of collinearity between the independent variables. If the VIF is between 0.1 and 10, collinearity is not serious between the parameters. The results of diagnosis are shown in Table 2. The VIFs of all parameters are bigger than 0.1 and lower than 10, which means the covariates have passed the collinearity diagnosis.

The significance of each parameter in ANCOVA model is shown in Table 3, where * indicates that $\alpha$ is less than 0.05 and ** indicates that $\alpha$ is less than 0.01. Table 3 shows that the $F$ value of current, the properties of current, arc length, maximum recovery voltage, transient secondary current, and wind direction are, respectively, 0.015, 0.007, 0.0016, 0.022, 0.007, and 0.047. They thus have a significant effect on arcing time. The significance of the transient secondary current and properties of current is below 0.01, which means that these variables have an extremely significant effect on arcing time than the others. The significance of the maximum secondary current, and area, and diameter of the arc is above 0.05, which means that they do not have a significant effect on arcing time.

The ANCOVA needs to be verified by residual analysis. In (2), the expected value of $\varepsilon$ is assumed to be zero, and it is a random variable that obeys a normal distribution. The essence of residual analysis is to test whether the assumption regarding $\varepsilon$ is true. There should be no explanatory or predictable information in $\varepsilon$. The residual is the difference between the observed value $y_i$ and the prediction obtained from the regression analysis $\hat{y}_i$. The residual is expressed as $e_i$. The standardized residual is the value obtained from dividing the residual by its standard deviation and it can be expressed as $z_e$. The residual of the $i$th observed value is

$$e_i = (\hat{y}_i - y_i),$$

$$z_{ei} = \frac{e_i}{s_e}$$

(4)

where $s_e$ is the estimated residual of the standard deviation. If the assumption about $\varepsilon$ is true, the distribution of $z_e$ obeys the standard normal distribution. Theoretically 95.44% of the values of the standardized residual are in the interval $[-2, 2]$ and 99.74% are in the interval $[-3, 3]$. The results of residual analysis are shown in Figure 5. The residuals of the parameters exhibit no significant increase or decrease, and the residuals are distributed on both sides of the zero randomly and unpredictably. Therefore, the data for the ANCOVA model satisfies the normal test and the homogeneity of variance and is correct.

| Variable                  | F-value | Significance level |
|---------------------------|---------|--------------------|
| Electrode length          | 2.088   | 0.129              |
| Wind speed                | 0.013   | 0.910              |
| Wind direction            | 3.410   | 0.047*             |
| Current value             | 4.333   | 0.015*             |
| Current property          | 5.174   | 0.007**            |
| Arc length                | 5.952   | 0.016*             |
| Maximum recovery voltage  | 5.380   | 0.022*             |
| Transient secondary current| 7.693  | 0.007**            |
| Maximum secondary current | 0.289   | 0.592              |
| Area                      | 0.006   | 0.940              |
| Diameter                  | 1.126   | 0.291              |

Table 3: Results of ANCOVA.

Figure 5: Result of residual analysis.
4. Outlier Removal

It is concluded in Section 2 that the key parameters are arc length, maximum recovery voltage, transient secondary current, and transient power. However, accidental data in the experiment affect the subsequent fitting of the results. Outliers thus need to be removed to improve the accuracy of the fitting model. There are two models to remove outliers. The first is to remove them through statistical theory, such as Grubbs’ criterion and Pauta criterion. However, statistical analysis cannot eliminate individual data that do not conform to the law of overall changes. The second method to remove the outliers involves using an appropriate machine learning method. DBSCAN algorithm is used to remove outliers, which is a density-based clustering method.

Density-based clustering algorithm (DBSCAN) is an unsupervised machine learning clustering algorithm. DBSCAN does not need to specify the number of clusters, and it works very well in clusters of any shape and size [19, 20]. Therefore, DBSCAN algorithm can be used to remove the outliers in the data.

Epsilon and the minPts are firstly defined for the DBSCAN algorithm. Epsilon is the maximum radius of the cluster. If the mutual distance between data points is less than or equal to the specified value of epsilon, these points will be grouped into one cluster. A larger value of epsilon produces a cluster with more data. In general, a smaller value of epsilon should be chosen because a cluster with a dense data distribution is needed. minPts is the least number of points which a cluster contains. A lower minPts leads to more clusters and thus more noise or outliers. A higher minPts makes a more robust cluster. But if the cluster is too large, outliers can be included in it. This is not conducive to removing them.

To determine a suitable value of epsilon, the average distance calculated at every point and its nearest neighbor is set to be the benchmark for it. MinPts should be greater than or equal to the number of dimensions of the dataset, but this should be adjusted according to the experimental data at hand. DBSCAN can be accessed by the results of the visualization of the data. The better the interpretability of the results is, the better is the effect of DBSCAN. The advantage of using DBSCAN algorithm for outlier removal is that it can perform exploratory classification. Moreover, DBSCAN can cluster nonlinear relationships.

Epsilon was set to be four, and minPts to three. In [13, 16], arc length is positively correlated with arcing time, and recovery voltage is negatively correlated with arcing time. Clusters which do not fit the correlation need to be removed as outlier. The results of clustering the recovery voltage and arcing time are shown in Figure 6. The red x in the figure are the calculated outliers, and the blue dots represent values that were considered normal. These outliers can be explained as accidental values that occurred for various reasons in the experiment.

Figure 6(a) shows the results of the recovery voltage with arcing time after DBSCAN. Points in the upper-right corner were sparsely distributed and were considered outliers. Once they had been removed, arcing time at most points was shorter than 0.3 seconds. A total of 82.2% of the recovery voltage was in the range of 25 kV to 32 kV. Figure 6(a) also shows a negative correlation between recovery voltage and arcing time. The higher the maximum recovery voltage is, the more likely the secondary arc extinguishes.

The results of the transient secondary current with arcing time after the use of DBSCAN are shown in Figure 6(b). Due to the accuracy of the oscilloscope, the transient secondary current recorded discrete current values of 0.0667, 0.1667, 0.6667, 1.3333, and 3.3333. Only three points had a value of 3.3333, one of which was considered an outlier owing to its distance to the other points. Most points had a value of 1.333 A and 1.667 A.

Figure 6(c) shows the result of arc length with arcing time after DBSCAN was applied. After the outliers are removed, a significant positive correlation between arc length and arcing time is shown in the scatterplot. This means that the longer the arc length was before the secondary arc extinguished, the longer the arcing time was. During secondary arc combustion, the movement of secondary arc was irregular in the vertical and horizontal directions owing to thermal buoyancy and the electromagnetic force. When the secondary arc was about to extinguish, the arc was quickly drawn out. Therefore, the longer the duration of combustion of the secondary arc was, the longer was the arc length before arc extinguishment.

In summary, the DBSCAN is used and conclusions of Section 3 to analyze the relationship between the arcing time and the scatterplots of the parameters with the arcing time. However, the conclusions drawn by using only the scatterplot are intuitive, but not quantitative. Only the correlation and the scatterplots cannot yield the precise functional relationship between the arcing time and the relevant parameters.

5. Establishing and Analyzing Regression Model of Secondary Arc

The arcing time is significantly influenced by recovery voltage, secondary current and arc length. The conclusion can help predict the arcing time. There are two ways to predict the arcing time. One is to create a fit with the given data. The advantage of such fitting is that it is simple and easy to implement, but the predicted result may differ greatly from the true response. To improve the accuracy of fitting, more complicated formulas such as derivatives are often needed that greatly increase the difficulty of fitting and the formulation of the formulas. The other is to use machine learning or deep learning algorithms to make predictions. These algorithms are accurate but lack a simple and clear expression. The two methods are compared to obtain accurate prediction of the arcing time.

5.1. Fitting the Multivariate Model of Secondary Arc

Before fitting, the key factors in the data should be processed first. There are three factors, namely, current value, current property, and wind direction. Due to the insufficient sample size, the result of grouping by factors and fitting is not ideal.

One-hot encoding is a way to convert factors into numerical values. In regression, one-hot encoding can
introduce factor into regression equation. One-hot encoding sets one attribute of a factor to 1 and the other attributes to 0.

Maximum recovery voltage and corresponding transient current were generated after the arc extinguished. It is not very helpful to predict the arcing time by using maximum recovery voltage and transient current. The arc length increases during the arc combustion. It is more reasonable to predict the arcing time by arc length.

Linear regression is the most used regression method, which is far faster than other regression models. The factors after one-hot encoding and arc length are set as independent variables to conduct multiple linear regression as follows:

\[ t = b_1 + B_2 I_e + B_3 I_p + B_4 W_d + b_5 L_a. \]  

(5)

\( I_e, I_p, \) and \( W_d \) are all column vectors of \( n \times 1 \), representing the numerical matrix of corresponding current value, current property, and wind direction after one-hot encoding. \( B_2, B_3, \) and \( B_4 \) are the row vectors of \( n \times 1 \).

Results of multiple linear regression is shown in Figure 7. Twenty percent of the points lie below the red line. Most of the predicted values are greater than the actual values. When the actual arcing time is short, the predicted value is much higher than the actual. In the regression model Equation (7), there is a positive correlation between arc length and arcing time. In the early stage of secondary arc, the arc length is short and elongates with arc combustion. The arc reaches the maximum length at the last reignition. If the predicted value provided by the regression model is less than the
actual value at the early stage of the secondary arc, it is highly likely to lead to failure of SPAR. The linear regression can provide a safe margin for SPAR closing at the initial stage of secondary arc.

5.2. Gaussian Progress Regression. Gaussian process regression (GPR) is a machine learning-based regression algorithm. It performs well on complex problems involving many dimensions and non-linearity [21, 22]. GPR is based on the kernel function. The Gaussian process firstly establishes the prior probability, and then completes the transformation from prior probability to the posterior probability under the Bayesian framework. The Gaussian process can calculate hyperparameters during regression; hence, there is no need to find the kernel function.

For a given set of data \( D = \{(x_i, y_i)\}_{i=1}^n \), the input matrix is \( x_i \in \mathbb{R}^d \) and the output matrix is \( y_i \in \mathbb{R} \). In the given data, \( f(x^{(1)}), f(x^{(2)}), \ldots, f(x^{(n)}) \) can form a set of random variables. This represents the column vector composed of significant variables and has a joint Gaussian distribution. Statistical features in GPR are the mean function \( m(x) \) and covariance function \( k(x, x') \). It is considered that the noise in the target \( y \) is built for a general model for the Gaussian regression problem:

\[
 f(x) \sim \text{GP}(m(x), k(x, x')), \quad y = f(x) + \varepsilon. \tag{6}
\]

In each experiment, the influence of the surroundings can be considered random. Therefore, such influences can be regarded as Gaussian noise, and are marked as \( \varepsilon \sim \mathcal{N}(0, \sigma^2) \). When \( f(x) \) was of Gaussian distribution, \( y \) was also of Gaussian distribution. The covariance function \( C(X, X) \) can be expressed as follows:

\[
 C(X, X) = K(X, X) + \sigma^2_n I. \tag{7}
\]

According to the Bayes principle, Gaussian prior distribution \( Y \) was established for the set of data \( D = \{(x_i, y_i)\}_{i=1}^n \) and transformed into posterior distribution for \( n \) given data set \( D_n = \{(x_i, y_i)\}_{i=n+1}^n \). The given output Gaussian prior distribution \( Y \) and the joint Gaussian prior distribution of \( Y \) with the predicted value \( y \) are as follows:

\[
 Y \sim \mathcal{N}(0, K(X, X) + \sigma^2_n I), \quad \begin{bmatrix} Y \\ y \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(X, X) + \sigma^2 1 & K(x_n, X) \\ K(x_n, X) & K(x_n, x_n) \end{bmatrix} \right) \tag{8}
\]

\[
 Y \sim \mathcal{N}\left(0, \begin{bmatrix} K(x_n, x_n) & K \end{bmatrix} \begin{bmatrix} K(x_n, x_n) + \sigma^2_n I & K(x_n, X) \\ K(x_n, X) & K \end{bmatrix} \right). \]

\( K(X, X) \) is matrix of covariance between the training sets and \( I_n \) is an n-dimensional identity matrix. \( K(x^*, X) = K(X, x^*) \) is the matrix of covariance between the test set \( x^* \) and the training set \( X \), and \( K(x^*, x^*) \) is the matrix of covariance between the training sets \( x \). The posterior distribution of the predicted value \( y \) is as follows:

\[
 \begin{bmatrix} y |Y \end{bmatrix} \sim \mathcal{N}(\bar{y}, \sigma^2), \quad \begin{bmatrix} y \\ \bar{y} \end{bmatrix} = K(x_n, x_n) + K(x_n, X) K(X, X) + \sigma^2_n I \tag{9}
\]

\[
 \sigma^2 = K(x_n, x_n) - K(x_n, X) K(X, X) K(x_n, X) K(x_n, x_n). \]

An important problem to solve when applying the GPR model is to test whether the predicted value reflects the true correlation between the arcing time and the variables influencing it. Usually, the coefficient of determination \( R^2 \) can be used to test the quality of fitting the model. \( R^2 \) is determined by the SSR and SST, where the former is the sum of squares of the difference between the predicted value and the mean of the input data, and the latter is the sum of squares of the difference between the input data and the mean.

\[
 \text{SSR} = \sum_{i=1}^n (\bar{y} - y_i)^2, \quad \text{SST} = \sum_{i=1}^n (y_i - \bar{y})^2, \quad R^2 = \frac{\text{SSR}}{\text{SST}} \tag{10}
\]

**Figure 7**: Results of linear regression.

\textit{R-square} is defined as the ratio of SSR to SST. The value of \textit{R-square} is between zero and one. It is a statistical indicator of the reliability in the regression model.

The results of GPR model are shown in Figure 8. The data in each group is plotted on the horizontal axis and is in the range [1, 97]. The vertical axis represents the arcing time, which is arranged from small to large. The discrete dots represent the true response, and the red line is the arcing time as predicted by the GPR model. When \( n \) is less than 50, the predicted value of GPR is larger than the true value. When \( n \) is greater than 50, the predicted value is smaller than the true value. The slope of GPR in the figure is smaller than the true response.
R-square of GPR model is 0.273 and R-square of linear model is 0.377. When \( n \) is greater than 50, the difference between the linear regression prediction curve and the real value is small, and linear regression have a good prediction in this range. Both regression models have certain significance. The imitative effect of linear regression is better than that of nonlinear GPR regression model. Overall, the predicted value of linear regression is greater than that of GPR. The GPR prediction curve is relatively gentle. The predicted value of linear regression is larger than the real value, and the curve fluctuation of linear regression prediction is relatively large.

In this section, linear regression and nonlinear GPR methods are respectively used to fit the arcing time. The linear relationship between arc length and arc time is stronger than the nonlinear. The results show that the linear model has a good imitative effect and can provide the margin of SPAR at the initial stage of secondary arc combustion.

6. Conclusion

This study used the data obtained from a low-voltage simulation of the secondary arc. The data were analyzed to obtain the key factors influencing the arcing time. After removing outliers by DBSCAN algorithm, a linear multi-variate regression and GPR model was performed, and the fitting results were compared. The following conclusions can be drawn from the analysis:

1. The calculations of ANCOVA showed that the most significant parameters affecting the arcing time are direction of wind, recovery voltage, secondary current, and arc length.

2. Scatter plots were drawn after outlier removal, which shows the recovery voltage was negatively correlated with arcing time. By contrast, the arc length and arcing time showed a positive correlation. These correlations may help in the development of suppression technologies for secondary arcs.

3. After comparing the R-square between two regressions, the results indicate that there was a linear relationship between the arcing time and the key parameters when the arcing time was long. Linear regression can provide a safe closing margin at the initial stage of arc and can avoid the further fault caused by premature closing.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

No potential conflicts of interest were reported by the authors.

Acknowledgments

This work has been supported by the Natural Science Foundation of Shandong Province (No. ZR2020ME196) and the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (NCEPU, LAPS22001).

References

[1] Working Group on Protective Relaying Performance Criteria of the Power System Relaying Committee, J. W. Ingelson, and P. F. Arnold, “Protective relaying performance reporting,” IEEE Transactions on Power Delivery, vol. 7, no. 4, pp. 1892–1899, 1992.

[2] I. M. Dudurych, T. J. Gallagher, and E. Rosolowski, “Arc effect on single-phase reclosing time of a UHV power transmission line,” IEEE Transactions on Power Delivery, vol. 19, no. 2, pp. 854–860, 2004.

[3] S.-P. Ahn, Chul-Hwan Kim, R. K. Aggarwal, and A. T. Johns, “An alternative approach to adaptive single pole auto-reclosing in high voltage transmission systems based on variable dead time control,” IEEE Transactions on Power Delivery, vol. 16, no. 4, pp. 676–686, 2001.

[4] M. C. Tavares, J. Talaisys, and A. Camara, “Voltage harmonic content of long artificially generated electrical arc in out-door experiment at 500 kV towers,” IEEE Transactions on Dielectrics and Electrical Insulation, vol. 21, no. 3, pp. 1005–1014, 2014.

[5] T. Tsuboi, J. Takami, S. Okabe, K. Aoki, and Y. Yamagata, “Study on a field data of secondary arc extinction time for large-sized transmission lines,” IEEE Transactions on Dielectrics and Electrical Insulation, vol. 20, no. 6, pp. 2277–2286, December 2013.

[6] S. Sun, F. Liang, F. Wang, H. Cong, Q. Li, and J. Yan, “Investigation on the geometrical characteristics of secondary arc by image edge detection,” IEEE Transactions on Plasma Science, vol. 46, no. 6, pp. 2016–2025, June 2018.

[7] M. Saad and C. H. Kim, “Unscented-Kalman-filter-based single-phase adaptive reclosing of shunt-compensated extra-
high-voltage transmission lines,” *Alexandria Engineering Journal*, vol. 61, 2022.

[8] R. M. A. Velásquez and J. V. M. Lara, “Secondary arc and critical time of fault clearance in overhead lines,” *IEEE Latin America Transactions*, vol. 16, no. 3, pp. 859–868, 2018.

[9] H. Cong, Q. Li, J. Xing, and W. H. Siew, “Modeling study of the secondary arc with stochastic initial positions caused by the primary arc,” *IEEE Transactions on Plasma Science*, vol. 43, no. 6, pp. 2046–2053, 2015.

[10] H. Cong, D. Han, S. Wang, L. Qiao, and Q. Li, “Experimental study on the extinction characteristics of secondary arc under different secondary arc currents near outdoor insulator,” *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 29, no. 1, pp. 103–110, 2022.

[11] R. Hongtao, “Analysis on secondary arc current of 500kV power station transmission line,” in *Proceedings of the 2020 7th International Forum on Electrical Engineering and Automation (IFEEA)*, pp. 428–432, Hefei, China, September 2020.

[12] Y. M. He, W. Li, K. Yang, B. Cui, and Y. Yuan, “The experimental research of secondary arc in 1000 kV UHV systems in China,” in *Proceedings of the 3rd Int. Conf. Elect. Power Equip.—Switching Technol.*, pp. 161–165, Busan, Republic of South Korea, October 2015.

[13] H. Liu, R. Li, D. He, J. Wei, and Q. Li, “Experimental study of multiple-reignition features of secondary arcs on EHV/UHV transmission lines,” *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3247–3255, 2019.

[14] J. Y. Xing, Q. M. Li, H. X. Cong, J. S. Li, Q. Chen, and Q. Y. Li, “Physical mechanism and simulation method of the arc root jumping and arc length variation of the secondary arcs with long-distance transmission lines,” *Transactions of China Electrotechnical Society*, vol. 31, no. 12, pp. 90–98, 2016.

[15] B. Khodabakhchian, “EHV single-pole switching: it is not only a matter of secondary arc extinction,” in *Proceedings of the presented at the Int. Conf. Power Syst. Transients*, vol. 18-20, Vancouver, BC, USA, July 2013.

[16] X. Tian, L. Hongshun, X. Zhitong, J. Liang, C. Haoxi, and L. Qingmin, “The Movement Characteristics and the Plasma State of Secondary Arc,” in *Proceedings of the 2018 China International Conference on Electricity Distribution (CICED)*, pp. 1615–1619, Tianjin, China, September 2018.

[17] H. J. Keselman, C. J. Huberty, L. M. Lix et al., “Statistical practices of Educational researchers: an analysis of their ANOVA, MANOVA, and ANCOVA analyses,” *Review of Educational Research*, vol. 68, no. 3, pp. 350–386, 1998.

[18] J. Jamieson, “Analysis of covariance (ANCOVA) with difference scores,” *International Journal of Psychophysiology*, vol. 52, no. 3, pp. 277–283, 2004.

[19] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, “DBSCAN revisited, revisited: why and how you should (still) use DBSCAN,” *ACM Transactions on Database Systems*, vol. 42, no. 3, pp. 1–21, 2017.

[20] K. Khan, S. U. Rehman, and K. Aziz, “DBSCAN: past, present, and future,” in *Proceedings of the fifth international conference on the applications of digital information and web technologies (ICADIW 2014)*, pp. 232–238, IEEE, Chennai, India, February 2014.

[21] E. Schulz, M. Speekenbrink, and A. Krause, “A tutorial on Gaussian process regression: modelling, exploring, and exploiting functions,” *Journal of Mathematical Psychology*, vol. 85, pp. 1–16, 2018.

[22] J. Q. Shi and T. Choi, *Gaussian Process Regression Analysis for Functional Data*, CRC Press, Florida, FA, USA, 2011.