Article

Pattern Recognition of Development Stage of Creepage Discharge of Oil–Paper Insulation under AC–DC Combined Voltage Based on OS-ELM

Fubao Jin [1,*]**, Shanjun Zhang** [2] **and Yuanxiang Zhou** [3]

1 School of Water Resources and Electric Power, Qinghai University, Xining 811600, Qinghai, China
2 Department of Information Science, Kanagawa University, Yokohama 222-0033, Japan; zhang@info.kanagawa-u.ac.jp
3 State Key Laboratory of Control and Simulation of Power Systems and Generation Equipment, Department of Electrical Engineering, Tsinghua University, Beijing 10084, China; zhouyx@tsinghua.edu.cn
* Correspondence: jinfubao@qhu.edu.cn; Tel.: +86-189-9718-8893

Abstract: The recognition of the creepage discharge development process of oil–paper insulation under AC–DC combined voltage is the basis for fault monitoring and diagnosis of converter transformers; however, only a few related studies are available. In this study, the AC–DC combined voltage with a ratio of 1:1 was used to develop a recognition method for the creepage discharge development process of an oil–paper insulation under a cylinder–plate electrode structure. First, the pulse current method was used to collect the discharge signals in the creepage discharge development process. Then, 24 characteristic parameters were extracted from four types of creepage discharge characteristic spectra after dimensionality reduction. Finally, based on the online sequential extreme learning machine (OS-ELM) algorithm, these characteristic parameters were used to recognize the development stage of the creepage discharge of the oil–paper insulation. The results showed that when the size of the sample training set used in the OS-ELM algorithm is close to the number of hidden layer neurons, a high recognition accuracy can be obtained, and the type of activation function has little influence on the recognition accuracy. Four stages of the creepage discharge development process were recognized using the OS-ELM algorithm; the trend was the same as that of the characteristic parameters of the entire creepage discharge development process. The recognition accuracy was 91.4%. The algorithm has a high computing speed and high accuracy and can train data in batches. Therefore, it can be widely used in the field of online monitoring and evaluation of electrical equipment status.

Keywords: AC–DC combined voltage; oil–paper insulation; creepage discharge; OS-ELM; pattern recognition

1. Introduction

The insulating structure and operation conditions of converter transformers differ from those of conventional AC transformers. The valve-side winding of a converter transformer can withstand AC–DC combined voltage (ADCV) and harmonic voltage components. The international council on large electric systems (CIGRE) statistics report that the failure rate of converter transformers is about twice that of AC transformers and that considerable cases of insulation failure are caused by creepage discharge at the oil–paper interface [1–3]. Thus, the oil–paper interface remains the vulnerable part of a converter transformer insulation and has therefore been studied widely. Studies have been conducted on the creepage discharge characteristics of oil–paper insulation under ADCV using partial discharge measurements and other methods, and beneficial results have been obtained [1–3]. The partial discharge phenomenon can reflect the insulation failure process. Therefore, it can (a) be used as an indicator of the performance degradation of a transformer

oil–paper insulation, (b) determine the severity of discharge development by extracting and analyzing partial discharge signals, and (c) be used for effective condition assessment of oil–paper insulation. Therefore, in-depth research on the development characteristics of the creepage discharge of oil–paper insulation under ADCV can effectively reveal the condition of oil–paper insulation. Furthermore, it is an exigent problem to be resolved in the operation and maintenance of ultrahigh voltage DC converter transformers; in addition, it has great theoretical and engineering significance.

Existing studies on the creepage discharge characteristics of oil–paper insulation under ADCV focus mainly on two aspects: (1) they investigate the factors affecting creepage discharge, including different voltage types and electric field components [4], discharge characteristics [5–8], aging [9,10], and other factors. It has been revealed that the temperature rise, aging, and DC electric field components accelerate the creepage discharge development and (2) they intend to mainly find the characteristic parameters and recognition methods that can effectively characterize the creepage discharge development under ADCV. The currently available techniques for the recognition of partial discharge development process mainly include fingerprint recognition of discharge gray image [11], support vector machines (SVM) [12,13], back-propagation neural networks (BPNN) [14], and fuzzy clustering analysis of radar spectra [15]. Nevertheless, only limited research has been conducted on the pattern recognition of the creepage discharge development process of oil–paper insulation under ADCV. For the recognition of the creepage discharge development process, the fundamental requisite is to divide the development stage of the creepage discharge. However, in this method, some subjective components are available only based on the changing trend of the discharge characteristic parameters with time. Moreover, the aforementioned recognition methods can perform pattern recognition only after collecting all data samples. Thus, if new data are added, all the data must be reevaluated, which is time consuming and inefficient. Therefore, to resolve this problem, this study builds a creepage discharge test platform for oil–paper insulation under ADCV. It comprehensively analyzes four two-dimensional characteristic spectra during the development of the creepage discharge. These characteristic spectra reflect the relationship between discharge interval time, discharge amplitude, and number of discharges during the development. Twenty-four characteristic parameters of the creepage discharge development process are extracted from the statistics of the discharge signal during creepage discharge. Based on the variation law of the characteristic parameters of creepage discharge, a method of recognizing the creepage discharge development process is established. Then, based on the online sequential extreme learning machine (OS-ELM) algorithm, the discharge development stage is identified in the creepage discharge development process of the oil–paper insulation.

2. Test System and Method

2.1. Sample and Electrode Model

Karamay 25# transformer oil was used in the test. Prior to the test, the gas volume fraction of the transformer oil was set to be less than 2% and the microwater volume fraction to be less than 10^{-5}, through the process of degassing, drying, and slag removal, which meets the requirements of GB/T7595. The insulating paper was cut into square specimens of dimensions 90 mm × 90 mm × 1 mm during the test and dried at high temperature (100 °C) for 72 h. Then, the insulating paper was impregnated with a pretreated transformer oil for 72 h under a vacuum of approximately 10 Pa and temperature of 60 °C.

Creepage discharge occurs easily at the spacer and barrier positions, where the longitudinal electric field is stronger than the radial electric field. Therefore, in the test, the cylinder-plate electrode model specified by CIGRE Method II and IEC60243 was used, and an oil-immersed insulating paper was placed between the high-voltage electrode and the ground electrode. The electrode structure and specific parameters are illustrated in Figure 1 [1,16–19].
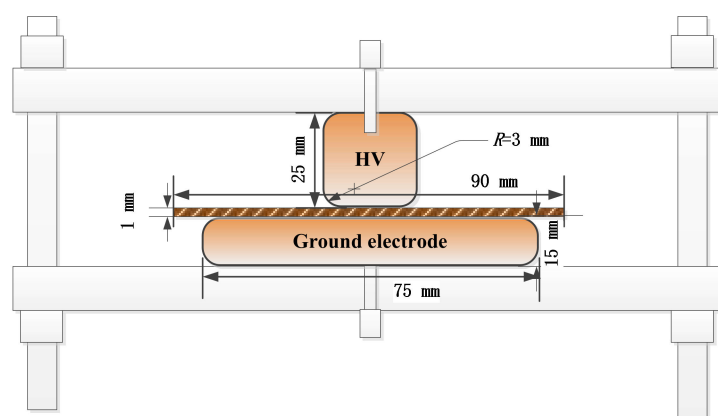
Figure 1. Cylinder-plate electrodes configuration.

2.2. Experimental Platform

The structure of the test system is shown in Figure 2. The AC and DC power used for the test were produced by nonpartial discharge (PD) equipment. The test circuit design was the same as that in Refs. [8,9]. The PD measuring instrument used the pulse current method for measurement; further, the sampling frequency of the system was 5 MHz, and the PD bandwidth was 0.04–1 MHz. An ADCV was applied to the test sample via a protective resistor. When the entire test system was loaded to 70 kV (peak value) without the test sample, no partial discharge signal was present, and the sensitivity of the test system was 0.1 pC.
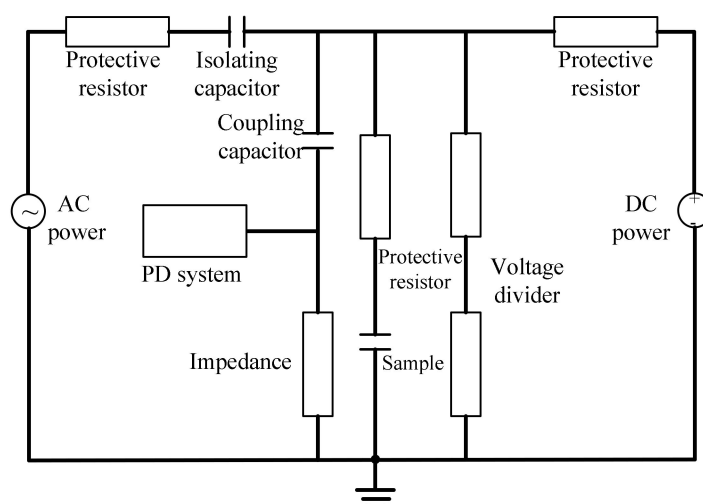

Figure 2. Schematic of the experimental platform.

2.3. Experimental Method

The ratio of the ADCV during the test was selected according to the general criteria and test requirements proposed by IEEE and CIGRE for oil-immersed converter transformer tests. IEEE has specified that the ratio of the ADCV must be 1:1, 1:3, 1:5, or 1:7, where AC takes a valid value and DC takes the average [20]. To obtain a relatively stable and sustainable development of the creepage discharge signal, the ratio of ADCV was selected as 1:1 in this work to study the creepage discharge development process of the oil–paper insulation.

Two methods are currently available for applying voltage, namely, constant voltage method and the step-up method, which have the same effect on the insulation breakdown process. Herein, the constant voltage method was used because it could simulate the actual operation condition of converter transformers better. The voltage ratio was maintained

unchanged at 1:1 during the application of voltage, and the boost step was set to 1 kV for both AC and DC voltages. The constant voltage was applied after boosting the voltage to 1.1 times the AC discharge inception voltage, where both the AC and DC voltages were 37 kV.

To accurately record the creepage discharge development process of the oil–paper insulation under ADCV, a set of data was recorded at an interval of 1 min from the beginning of application of the constant voltage, and each recording was performed continuously for 1 min until breakdown occurred. This test procedure was repeated 12 times, and a total of 1210 sets of data were obtained, mainly using which the characteristic parameters of creepage discharge would be subsequently extracted.

3. Pattern Recognition Based on OS-ELM

The conventional extreme learning machine (ELM) transmits all data to the system and then trains and calculates them. However, if new sample data are added, it is required to retrain all sample data, resulting in prolonged learning time and reduced learning rate. To resolve this problem, Liang et al. proposed the aforementioned OS-ELM algorithm, which is an improved version of the ELM approach. In the OS-ELM algorithm, the used data are discarded and are not reused. Further, this algorithm can add new data to the model in individual succession or in batches for training to avoid repeated training of the used sample data; therefore, this algorithm possesses a higher learning rate and higher accuracy [21–25].

As a single-hidden layer feedforward neural network, the OS-ELM algorithm is composed of three layers: input layer, single hidden layer, and the output layer. As shown in Figure 3, the connection weight between the input layer and hidden layer, and the threshold value of the hidden layer neurons are randomly generated, which need not be repeatedly iterated and adjusted in the training process. Compared with the pattern recognition effects of the conventional ELM, SVM, and BPNN methods, the OS-ELM algorithm has the advantages of higher speed, learning rate, generalization performance, and accuracy [26,27]. Therefore, the OS-ELM algorithm was selected in this study for pattern recognition of the creepage discharge development process of the oil–paper insulation under ADCV.
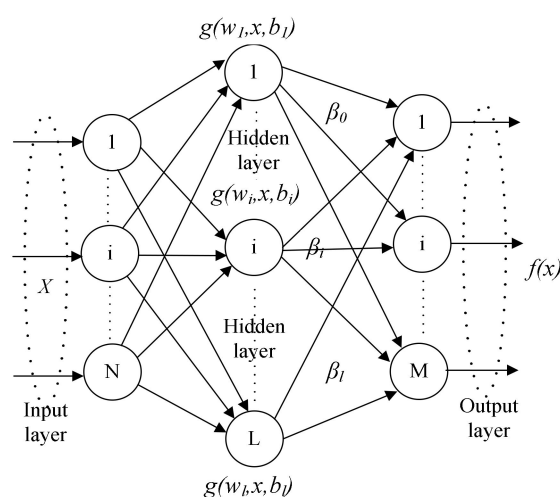

Figure 3. Structure of online sequential extreme learning machine (OS-ELM).

The key part of the OS-ELM algorithm is the learning process of the output weight of the single hidden layer network, and it comprises two stages: (1) the initial stage, wherein the output weight β of the single hidden layer is obtained by training a small number of sample data, and (2) online learning stage, wherein the updated sample data are used to update the output weight β obtained in stage (1).

In stage (1), N sample data (X_i) are present, where $X_i = [x_1, x_2, \ldots, x_n]^T \in R^m$, m denotes the dimension of the sample dataset, $T_i = [t_1, t_2, \ldots, t_n]^T \in R^n$, and n denotes the dimension of the output vector. First, the number of hidden layer neurons L is determined; then, the connection weight w of the input layer and the threshold b of the hidden layer neurons are randomly set; subsequently, the appropriate activation function $g(w, x, b)$ is selected; finally, the initial output matrix H_0 of the hidden layer is calculated. According to the ELM theory, the minimum β_0 that can satisfy $||H_0 \beta - T_0||$ was calculated, where

$$H_0 = \begin{bmatrix} g(w_1, x_1, b_1) & \cdots & g(w_L, x_1, b_L) \\ \vdots & \ddots & \vdots \\ g(w_1, x_N, b_1) & \cdots & g(w_L, x_N, b_L) \end{bmatrix}_{N \times L} \tag{1}$$

$$T_0 = \begin{bmatrix} t_1 & \cdots & t_L \end{bmatrix}_{M \times L} \tag{2}$$

According to the generalized inverse matrix algorithm, the initial output weight β_0 was calculated, following which the initial learning stage ends, as represented below:

$$\beta_0 = P_0^{-1} H_0^T T_0, P_0 = H_0^T H_0 \tag{3}$$

Subsequently, stage (2) begins. Assuming that N_1 new sample data enter the model, it is desirable to obtain β_1 such that Equation (4) can be the minimum.

$$\left|\left| \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \beta - \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} \right|\right| \tag{4}$$

$$H_0 = \begin{bmatrix} g(w_1, x_{N+1}, b_1) & \cdots & g(w_L, x_{N+1}, b_L) \\ \vdots & \ddots & \vdots \\ g(w_1, x_{N+N_1}, b_1) & \cdots & g(w_L, x_{N+N_1}, b_L) \end{bmatrix}_{N_1 \times L} \tag{5}$$

According to the generalized inverse matrix algorithm, the output weight β_1 is calculated as follows:

$$\beta_1 = P_1^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} T_0 \\ T_1 \end{bmatrix}, P_1 = \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} \tag{6}$$

The output weight β_{k+1} can be obtained using the Woodbury formula and the recursive relationship as follows:

$$\beta_{k+1} = \beta_k + P_{k+1}^{-1} H_{k+1}^T (T_{k+1} - H_{k+1}\beta_k) P_{k+1}^{-1} = P_k^{-1} - P_k^{-1} H_{k+1}^T \left(I + H_{k+1} P_k^{-1} H_{k+1}^T \right)^{-1} \times H_{k+1} P_k^{-1} \tag{7}$$

Equation (7) provides the basic principle of the OS-ELM algorithm, and weight β of the single-hidden layer neural network can be adjusted through two stages. In the OS-ELM training process, if the data obtained in stage (1) are the only data, i.e., no updated sample data are present, OS-ELM is the same as the standard ELM algorithm. Therefore, the OS-ELM algorithm can be understood as a special ELM algorithm in which all training data arrive simultaneously.

4. Experimental Results and Discussion

4.1. Creepage Discharge Development Process of Oil–Paper Insulation under ADCV

Because the partial discharge of the oil–paper insulation under ADCV is more random than that under AC voltage, the signal representation method for the former significantly differs from that for the latter. In this study, the interval time Δt of two adjacent discharges is used instead of the phase ϕ in the partial discharge under AC voltage. Then, the statistical characteristic spectra are formed by combining the interval time Δt, partial discharge magnitude Q, and number of discharges n. Time-resolved pulse sequence analysis (TPRSA) was used to study the creepage discharge development process of oil–paper insulation under ADCV [8,9].

In this study, four sets of data in the initial stage, two middle stages, and the adjacent end stage of one test were selected for statistical analysis, and the Q-Δt_{pre}-n three-dimensional characteristic spectra are formed—here, Δt_{pre} denotes the time interval between the current discharge and the previous discharge. Figure 4 shows the variation law of the Q-Δt_{pre}-n characteristic spectra of the abovementioned four sets of data during the creepage discharge development process.

(**a**) First stage

(**b**) Second stage

(**c**) Third stage

(**d**) Fourth stage

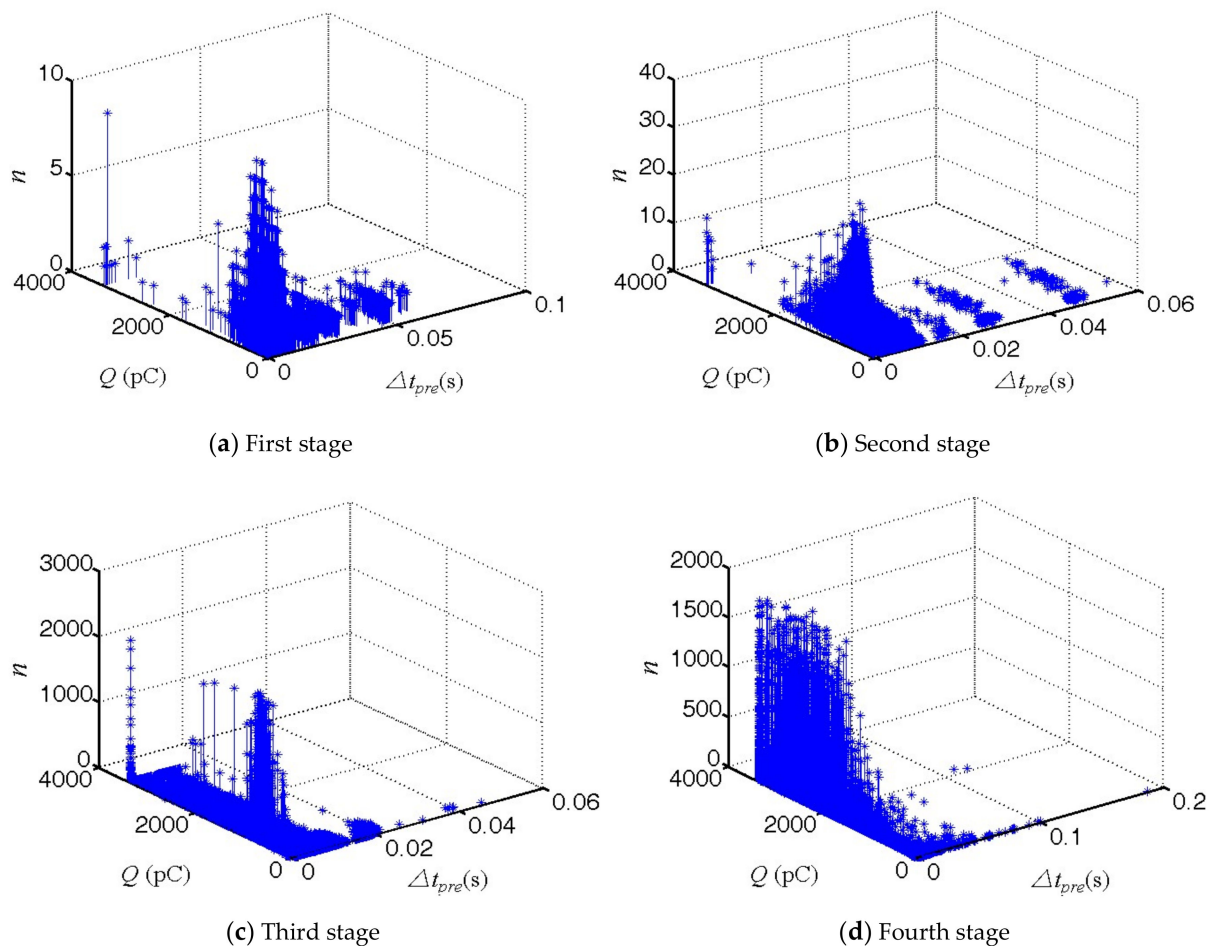**Figure 4.** Q-Δt_{pre}-n characteristic spectra of four sets of data during the creepage discharge development process.

Figure 4 clearly indicates the existence of significant differences between the characteristic spectra at each stage. In the first stage (Figure 4a), n is small, Δt_{pre} is large, and the distribution range is 0.01–0.06 s. Further, Q is mainly concentrated at approximately 400 pC, and the distribution of Q is highly scattered. In the second stage (Figure 4b), n increases gradually, and Q is mainly distributed in the region less than 1300 pC; however, Q greater than 2000 pC appears gradually. Further, Δt_{pre} is gradually shortened and is distributed within the range of 0.04 s. In the third stage (Figure 4c), n increases significantly compared with that in the second stage. Further, Q is mainly distributed around 1700 pC but gradually enters the range of 2000–3000 pC, and Δt_{pre} is mainly concentrated in the range of 0.02 s. In the fourth stage (Figure 4d), n continues to increase gradually; further, Q is mainly concentrated in the region greater than 2000 PC, and Δt_{pre} is mainly concentrated in the range of 0.01 s.

Although the TPRSA spectra can reflect the time domain distribution of the creepage discharge development process better, directly using these spectra for pattern recognition is difficult because they contain large amounts of information. Therefore, to alleviate the difficulties in the recognition of the creepage discharge development process, the

original data must be transformed, the dimension must be reduced, and the characteristic parameters reflecting the creepage discharge development process must then be extracted.

The characteristic spectra, such as *n*-*Q*, *n*-Δ*t*, Δ*t*$_{pre}$-*Q*, and *Q*$_{mean}$-Δ*t*$_{pre}$, were obtained by dimensionality reduction. The Δ*t*$_{pre}$-*Q* two-dimensional (2D) characteristic spectra after dimensionality reduction corresponding to each stage of Figure 4 are shown in Figure 5. From the characteristic spectra of each stage shown in Figure 5, the variation law between *Q* and Δ*t*$_{pre}$ in each stage of the creepage discharge development process can be obtained more clearly. Δ*t*$_{pre}$ is relatively dispersed from the initial stage and the development stage, the critical breakdown is relatively concentrated, and Δ*t*$_{pre}$ is gradually reduced. Further, *Q* transforms from being a small value initially to a large value at the critical breakdown—its distribution transforms from being a dispersed distribution initially to a concentrated distribution at the critical breakdown. Moreover, it can be used to analyze the *n*-Δ*t* and *Q*$_{mean}$-Δ*t*$_{pre}$ characteristic spectra in a reduced dimension, which is not discussed in detail in this paper. A comprehensive analysis of the variation law between the discharge characteristic parameters in the abovementioned characteristic spectra can effectively reflect the creepage discharge development process. Accordingly, the characteristic parameters were extracted from the abovementioned four characteristic spectra and were used to analyze the creepage discharge development process.
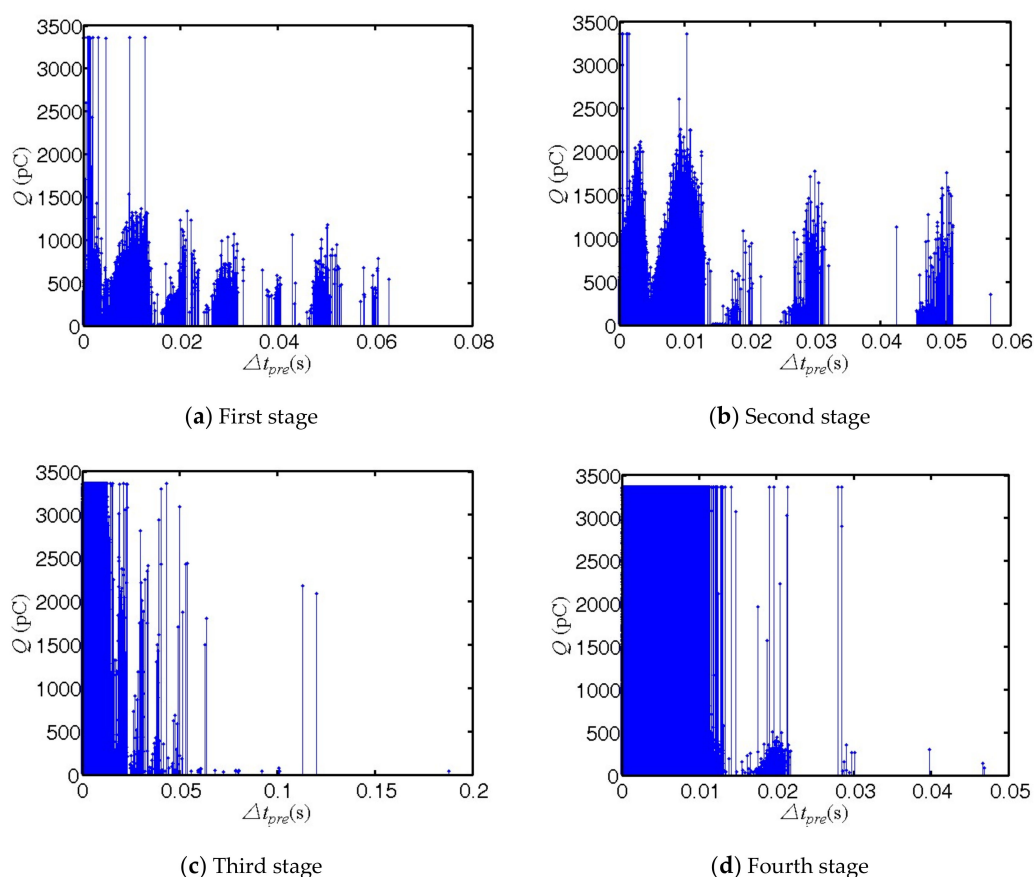

(**a**) First stage

(**b**) Second stage

(**c**) Third stage

(**d**) Fourth stage

Figure 5. Δ*t*$_{pre}$-*Q* two-dimensional characteristic spectra after dimensionality reduction corresponding to each stage of Figure 4.

4.2. Pattern Recognition of Creepage Discharge Development Process Based on OS-ELM

For the *n*-*Q*, *n*-Δ*t*, Δ*t*$_{pre}$-*Q*, and *Q*$_{mean}$-Δ*t*$_{pre}$ characteristic spectra, their skewness (S_k), kurtosis (K_u), asymmetry (A_{sy}), cross-correlation coefficient (C_c), and the Weibull distribution parameters *α* and *β* were extracted. Twenty-four characteristic parameters were extracted from each set of data from the aforementioned four 2D characteristic spectra (Table 1).

Table 1. Statistical characteristic parameters of creepage discharge.

Characteristic Spectra	Characteristic Parameters					
	S_k	K_u	A_{sy}	C_c	α	β
n-Q	√	√	√	√	√	√
n-Δt	√	√	√	√	√	√
Q-Δt_{pre}	√	√	√	√	√	√
Q_{mean}-Δt_{pre}	√	√	√	√	√	√

Figure 6 shows the variation law of each characteristic parameter in the Δt_{pre}-Q characteristic spectra during the entire creepage discharge development process. According to the variation law of the characteristic parameters, such as S_k, K_u, and β, shown in the figure, the creepage discharge development process can be divided into four stages.
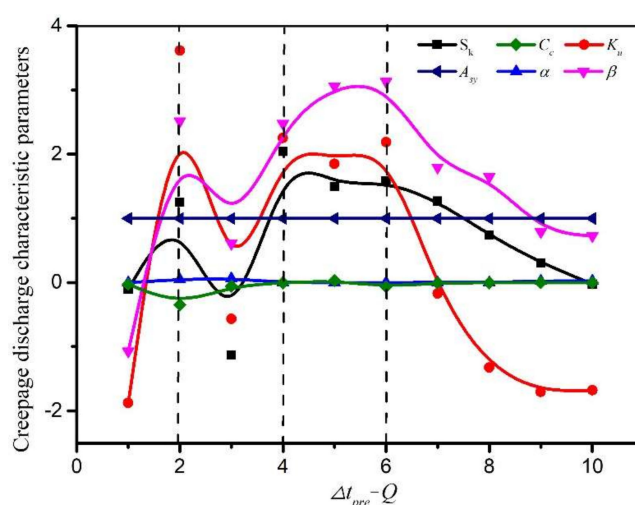
Figure 6. Variation law of characteristic parameters in the Δt_{pre}-Q characteristic spectra during the entire creepage discharge development process.

In the first stage, the characteristic parameters rise rapidly. The changes in the characteristic parameters indicate that Q increases gradually at this stage, Δt_{pre} shortens gradually, and the creepage discharge enters the initial stage. In the second stage, the parameters initially exhibit a decreasing trend and then rise slowly. This indicates that Q initially rises and then stabilizes, Δt_{pre} is gradually shortened, and the creepage discharge enters the development stage. The third stage maintains the current state, and creepage discharge continues to develop. In this stage, Q increases and tends to stabilize, Δt_{pre} is stable, and the creepage discharge enters a stable state. The fourth stage exhibits a trend of gradual decline, and the creepage discharge enters the critical breakdown stage. The variation law of the characteristic parameters of the Δt_{pre}-Q discharge characteristic spectra can aid in better analyzing the relationship between Q and Δt_{pre} during the creepage discharge development process and also in better understanding the physical development process of creepage discharge.

In this study, the OS-ELM algorithm code was written using MATLAB. The characteristic parameters extracted during the entire creepage discharge process were used as sample sets for the recognition of the creepage discharge development process.

First, the influence of the number of hidden layer neurons, type of activation function, and size of the sample data on the training accuracy of the initial stage were analyzed. The corresponding results can be used to rationally select the parameters of the OS-ELM algorithm. Figure 7 shows the relationships among the training accuracy of the initial stage, activation function type, and number of hidden layer neurons. The number of training samples is 750. The horizontal axis is the number of hidden layer neurons, and the vertical

axis is the training accuracy under different activation functions. The figure shows that the difference between the accuracies of the same neuron number is small under different activation functions. This indicates that the activation function has less influence on the accuracy under creepage discharge sample. The training accuracy improves gradually as the number of hidden layer neurons increases and tends to stabilize when the number of hidden layer neurons becomes close to the number of training samples. Therefore, the relationship between the sample size, number of hidden layer neurons, and activation function type should be collectively considered to select the algorithm parameters for realizing a better pattern recognition.
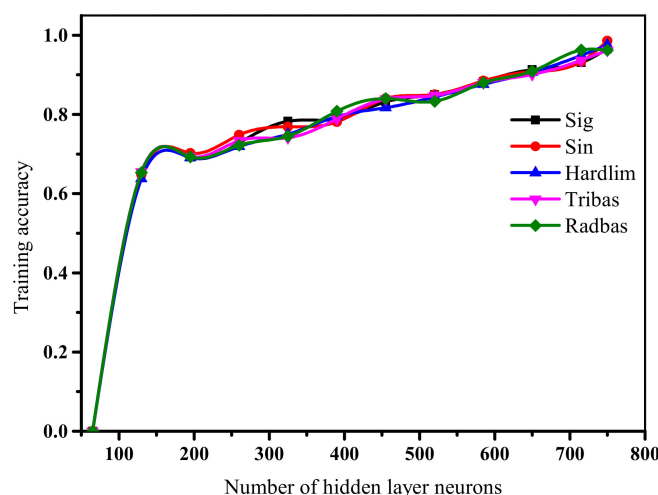

Figure 7. Relationship between activation function, number of hidden layer neurons, and training accuracy.

In this study, the sigmoid function was selected as the activation function, and the number of hidden layer neurons was 650. There were 12 groups of experimental data, of which seven were randomly selected as sample training sets to train the OS-ELM, ELM, SVM, and BPNN algorithms. The remaining five groups were used as sample test sets for recognizing the creepage discharge development process, and a confusion matrix was created based on the recognition results. The confusion matrix is shown in Figure 8. The figure displays that the entire creepage discharge development process can be recognized by four stages, namely, C1, C2, C3, and C4, by the OS-ELM algorithm. In the confusion matrix (Figure 8), the horizontal and vertical directions represent four states: creepage discharge initiation (C1), creepage discharge development (C2), creepage discharge acceleration (C3), and creepage discharge critical flashover (C4).

In this matrix, each row represents the actual state data, and each column represents the identified state data. The elements on the main diagonal represent the correct number of the creepage discharge development stage identified based on the test data, and the nonmain diagonal elements represent the misidentified number. In total, 500 sets of data were used as test samples. In the first row, the total number of states is 120, and the four states are 113, 5, 2, and 0. These data indicate that the total number of real states should be 120. After the diagnosis of the algorithm, the number of diagnosed C1 is 113, and the number of misdiagnosed C2, C3, and C4 is 5, 2, and 0, respectively. The last column represents the recall rate, which reflects the ratio of the correct number of samples to the actual number of samples. The last row of elements represents the precision rate, which reflects the ratio of the predicted number of correct samples to the number of predicted samples. The location at which the recall rate and accuracy rate intersect represents the recognition accuracy, which reflects the ratio of the correct number to the total number of samples. This shows that both the recall rate and the precision of the creepage discharge development process are significantly high (>88%) based on the OS-ELM algorithm, and

the sample recognition accuracy is 91.4%. The recognition accuracy of ELM is lower than that of OS-ELM, and the recognition accuracy of SVM is 85.6%, which is lower than ELM.

Figure 8. Pattern recognition result of four algorithms.

In contrast, the recognition accuracy of BPNN is 77%, which is much lower than that of OS-ELM. The recognition result of OS-ELM is the same as the result of dividing the creepage discharge development process according to the variation law of the characteristic parameters of the creepage discharge development process depicted in Figure 6. Based on this method, the OS-ELM algorithm can theoretically support the pattern recognition of the creepage discharge development process of oil–paper insulation.

In the ELM algorithm, the weight matrix and threshold from the input layer to the hidden layer are generated randomly; this avoids the iterative process of repeated training and threshold modification in SVM and BPNN. By contrast, the weight matrix from the hidden layer to the output layer uses the least squares method to find the optimal solution to minimize the loss function; thus, it does not require an iterative process. Therefore, the OS-ELM algorithm can shorten the training adjustment time and has a very fast training speed. In the training time comparison part, 200 out of 1210 sets of experimental data were randomly selected as sample training sets to train four algorithms. OS-ELM and ELM algorithms selected the same number of samples and hidden layers; SVM algorithm cross-validation obtains the best parameter penalty factor and RBF kernel function variance. The training time of the four algorithms is shown in Table 2. Both OS-ELM and ELM have short training times, while BPNN requires a longer time. SVM adopts a cross-validation method; therefore, its training time is the longest. It can be observed that, combined with the experimental results, the OS-ELM algorithm is better than the other three methods in terms of recognition rate and training speed, and the OS-ELM algorithm has a significantly short training time.

Table 2. Comparison of the training time of the four algorithms.

Algorithm	Training Time/s
OS-ELM	0.0532
ELM	0.0532
SVM	574.1774
BPNN	264.7327

The recognition accuracy of OS-ELM has improved, compared with that of the other methods, such as SVM, BPNN, and ELM. The OS-ELM algorithm can separately train the data arriving in batches, avoiding repeated training of all data, thereby greatly reducing the time required for network training adjustment and effectively improving the training speed. Owing to the abovementioned advantages, the OS-ELM algorithm is applicable in the field of electrical equipment condition monitoring and evaluation engineering.

5. Conclusions

A creepage discharge test platform for oil–paper insulation under ADCV was developed in this study for the pattern recognition of the creepage discharge development process. According to the variation law of the characteristic parameters of the creepage discharge development process of the oil–paper insulation, the process was divided into four stages. In the first stage, the characteristic parameters increased rapidly; in the second stage, they decreased first and then gradually increased; in the third stage, they stabilized, and then gradually decreased in the fourth stage.

For the creepage discharge data sample type, the hidden layer activation function type had only a slight effect on the training accuracy. In the OS-ELM algorithm parameter selection process, when the sample set size was close to the number of hidden layer neurons, the training rate improved and a higher training accuracy rate was obtained.

The sample recognition accuracy was 91.4% in this study. Furthermore, the recall rate and the precision of the creepage discharge development process were both very high (>88%). The OS-ELM algorithm has an advantage that it can train samples in batches and, therefore, has good engineering application prospects in the field of electrical equipment condition monitoring and evaluation.

References

1. Jin, F.B.; Zhang, S.J.; Zhou, Y.X.; Ma, H.; Tapan, K.S. Pattern Recognition of Development stage of Creepage Discharge of Oil–Paper Insulation Under AC–DC Combined Voltage Based on OS-ELM. In Proceedings of the 2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE 2020), Beijing, China, 6–10 September 2020.
2. China South Power Grid. *Typical Fault Analysis of HVDC Transmission System Equipment*; China Electric Power Press: Beijing, China, 2009; pp. 15–173. (In Chinese)
3. Zhou, Y.X.; Sha, Y.C.; Li, J.Z.; Wang, J.Y. Partial Discharge Characteristics in Oil-paper Insulation Under Combined AC-DC Voltage. *IEEE Trans. Dielectr. Electr. Insul.* **2014**, *21*, 1529–1539.
4. Li, J.H.; Zhang, L.; Han, X.T.; Yao, X.; Li, Y.M. PD Detection and Analysis of Oil-Pressboard Insulation Under Pulsed DC Voltage. *IEEE Trans. Dielectr. Electr. Insul.* **2017**, *24*, 324–330. [CrossRef]

5. Wei, Z.; Li, C.R. Discharge of Oil-pressboard Insulation in AC-DC Composite Field: Phenomenon and Characteristics. *IEEE Trans. Dielectr. Electr. Insul.* **2016**, *23*, 237–245.

6. Wu, H.; Li, C.R.; Qi, B.; Zhao, X.L.; Lv, J.Z.; Zhao, L.J. The Electric Field Distribution in Oil-Paper Insulation Under Combined AC-DC Voltage. In Proceedings of the IEEE International Conference on Condition Monitoring and Diagnosis (CMD 2012), Bali, Indonesia, 23–27 September 2012; pp. 1097–1101.

7. Zhou, Y.X.; Huang, M.; Chen, W.J.; Lu, L.C.; Jin, F.B.; Huang, J.W. Space Charge Behavior Evolution with Thermal Aging of Oil-paper Insulation. *IEEE Trans. Dielectr. Electr. Insul.* **2015**, *22*, 1381–1388. [CrossRef]

8. Zhou, Y.X.; Jin, F.B.; Sha, Y.C.; Huang, M.; Huang, J.W.; Liu, Z.H.; Lu, L.C. Effects of DC Prestressing on Partial Discharge in Oil-impregnated Pressboard Insulation. *IEEE Trans. Dielectr. Electr. Insul.* **2016**, *23*, 4668. [CrossRef]

9. Zhou, Y.X.; Jin, F.B.; Huang, M.; Le, T.H.; Huang, J.W.; Liu, Z.H.; Lu, L.C. Effects of Thermal Aging on Creepage Discharge in Oil-Impregnated Pressboard Under Combined AC-DC Voltage. *IEEE Trans. Dielectr. Electr. Insul.* **2015**, *22*, 2737–2746. [CrossRef]

10. Li, J.; He, Z.M.; Grzybowski, S. Electrical Aging Lifetime Model of Oil-Impregnated Paper Inder Pulsating DC Voltage Influenced by Temperature Insulation. *IEEE Trans. Dielectr. Electr. Insul.* **2013**, *20*, 1992–1997. [CrossRef]

11. Li, S.M.; Si, W.; Li, Q.Q. Partition and Recognition of Partial Discharge Development Stages in Oil-Pressboard Insulation with Needle-Plate Electrodes Under Combined AC-DC Voltage Stress. *IEEE Trans. Dielectr. Electr. Insul.* **2017**, *24*, 1781–1793. [CrossRef]

12. Wang, K.; Liao, R.J.; Yang, L.; Li, J. Optimal Features Selected by NSGA-II for Partial Discharge Pulses Separation Based on Time-Frequency Representation and Matrix Decomposition. *IEEE Trans. Dielectr. Electr. Insul.* **2013**, *20*, 825–838. [CrossRef]

13. Ma, Z.Q.; Yang, X.; Lin, C.Y.; Zhou, D.; Chen, J.G. Partial Discharge Developing Stages Identification Based on Cluster-Hierarchical Decision SVM in Oil-Paper Insulation. In Proceedings of the 2016 IEEE International Conference on High Voltage Engineering and Application (ICHVE 2016), Chengdu, China, 19–22 September 2016.

14. Wang, K.; Li, J.Z.; Zhang, S.Q.; Gao, F.; Grzybowski, S. A New Image-Oriented Feature Extraction Method for Partial Discharges. *IEEE Trans. Dielectr. Electr. Insul.* **2015**, *25*, 1015–1024. [CrossRef]

15. Li, S.M.; Li, Q.Q.; Si, W.; Yao, J.Y. An Image-Oriented Recognition Method for PD Development Stage in Oil-Pressboard Insulation with Needle-Plate Model Under AC-DC Mixed Voltage. In Proceedings of the International Conference on Electrical Materials and Power Equipment (ICEMPE 2017), Xi'an, China, 14–17 May 2017; pp. 275–282.

16. International Electrotechnical Commission. *Specification for Pressboard and Press Paper for Electrical Purposes*; IEC 60641; IEC: Geneva, Switzerland, 2008.

17. International Electrotechnical Commission. *Electrical Strength of Insulating Materials Test Methods*; IEC 60243; IEC: Geneva, Switzerland, 1998.

18. International Electrotechnical Commission. *High Voltage Test Technique-Partial Discharge Measurement*; IEC 60270; IEC: Geneva, Switzerland, 2000.

19. Kennedy, W.M. Recommended Dielectric Tests and Test Procedures for Converter Transformers and Smoothing. *IEEE Trans. Power. Deliv.* **1986**, *3*, 161–166. [CrossRef]

20. IEEE Std C57. *129TM IEEE Standard for General Requirements and Test Code for Oil-Immersed HVDC Converter Transformers*; IEEE: Piscataway, NJ, USA, 2008.

21. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. In Proceedings of the International Joint Conference on Neural Networks (IJCNN 2004), Budapest, Hungary, 25–29 July 2004; pp. 985–990.

22. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme Learning Marching: Theory and Application. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]

23. Zong, W.E.; Huang, G.B.; Chen, Y.Q. Weighted Extreme Learning Machine for Imbalance Learning. *Neurocomputing* **2013**, *101*, 229–242. [CrossRef]

24. Liang, N.Y.; Huang, G.B.; Saratchandran, P.; Sundararajan, N. A Fast and Accurate Online Sequential Learning Algorithm for Feedforward Networks. *IEEE Trans. Neural Netw.* **2006**, *17*, 1411–1423. [CrossRef] [PubMed]

25. Huang, G.B.; Chen, L.; Siew, C.K. Universal Approximation Using Incremental Constructive Feedforward Networks with Random Hidden Nodes. *IEEE Trans. Neural Netw.* **2006**, *17*, 879–892. [CrossRef] [PubMed]

26. Zhang, Q.Q.; Song, H.; Yong, J.; Sheng, G.H. Online Sequential Extreme Learning Machine for Partial Discharge Pattern Recognition of Transformer. In Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference, Denver, CO, USA, 16–19 April 2018.

27. He, J.; Liu, C.H.; Tang, Q.H.; Zhang, C.H.; Cao, M.; Xu, X. Partial Discharge Pattern Recognition Algorithm Based on Sparse Self-coding and Extreme Learning Machine. In Proceedings of the 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2 2018), Beijing, China, 20–22 October 2018.