Membrane Technology
ISSN (online): 1873-4049

Research on the Prediction Method of MOV Deterioration State based on Principal Component Analysis and Grid Search-Optimized Support Vector Machine Regression Algorithm

Xu Zheng Wang, Zhihengzhu *, Huang Zhou

School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan,
corresponding authorEmail:102200209@hbut.edu.cn

SUMMARY: Background: Accurate diagnosis of faults in metal oxide varistors (MOV) is crucial for the safe operation of power systems, and the deterioration of MOV under continuous pulse impacts can be more severe. To effectively improve the fault diagnosis rate, this paper proposes a fault diagnosis algorithm based on Principal Component Analysis (PCA) and Grid Search-optimized Support Vector Regression (GS-SVR).Objective, The objective of this study is to propose an effective fault diagnosis algorithm that accurately predicts the fault state of MOV under single and continuous pulse impacts, while reducing the correlation between indication indicators through dimensionality reduction. Method: The proposed experiment involves conducting a comparative test on MOV with different time intervals between impacts, on the order of 10 seconds. The data collected from this experiment, with a time resolution of 10 seconds, will be subjected to dimensionality reduction using PCA to reduce the correlation between the original indicators. Finally, the GS-SVR model will be employed to analyze and predict the effects of single and continuous pulse impacts on MOV. Results: Experimental results demonstrate that the GS-SVR model achieves a mean square error of less than 0.00057 in predicting single pulse impacts and still exhibits certain effectiveness for irregular pulse impacts, such as continuous pulses. Conclusion: The proposed fault diagnosis algorithm based on PCA and GS-SVR can effectively improve the fault diagnosis rate of MOV, and accurately predict the fault state of MOV under single and continuous impulse shock. This is of great significance to the safe operation of power system.

key words: *Metal Oxide Varistor； Principal Component Analysis； Grid Searc； Support Vector Machine, Continuous Pulse Impact*

1. INTRODUCTION

Metal oxide varistors (MOV) are widely used lightning protection devices in power and electrical systems. They possess the characteristic of varying resistance with voltage sensitivity and can rapidly respond to discharge surge currents, interrupting sustained current flow [1]. However, these devices are prone to aging or deterioration issues. After each protection against lightning or overvoltage threats, their electrical performance gradually undergoes irreversible degradation until complete failure. Therefore, in practical production, extensive research has been conducted to real-time monitor the performance status of MOV devices, identify highly deteriorated devices, and mitigate operational risks to critical system locations caused by damaged protection devices [2,3].

Researchers have evaluated the degree of deterioration of MOV devices by analyzing their microphysical properties under impacts, cumulative energy data, temperature rise data, and static sampling data. In references [4,5], researchers studied the thermal characteristics and shock resistance of surge arresters under lightning current impacts. In reference [6], the aging characteristics of zinc oxide surge arresters under multiple lightning strikes were investigated. In references [7,8], the failure modes of varistors under impulse voltage were studied using multiple pulse waveforms. In references [9,10], thermal effects on surge arresters under lightning strikes were simulated using finite element methods. In reference [11], multiple pulse impact experiments on varistors were conducted, and the variation patterns of their electrical parameters were summarized, revealing a significant decline in their withstand characteristics after reaching the threshold. Reference [12] supplemented the study of the double Schottky barrier attenuation phenomenon in varistors and explained the influence of defect structures caused by manufacturing processes on the deterioration mechanism.

In recent years, many researchers have chosen data-driven algorithms combined with sampled data from MOV devices for statistical modeling and exploring the degree of device deterioration. In reference [13], researchers analyzed the harmonic characteristics of leakage current during the deterioration process of varistors, proposed an optimized leakage current component extraction algorithm based on fast Fourier transform, and compared it through simulation tests. They successfully achieved precise detection of the harmonic resistive component, thereby diagnosing the degree of device deterioration. In reference [14], researchers analyzed parameter data obtained from experiments that characterized MOV deterioration and used neural network algorithms to establish a unified model, enabling numerical prediction of specific device indicators. In reference [15], researchers proposed a monitoring model for MOV based on genetic algorithms, optimizing the parameters reflecting device deterioration in the equivalent model to monitor the degree of deterioration. In references [16-19], researchers further applied various optimization algorithms based on multi-pulse MOV experimental data, discussed and compared the

advantages and disadvantages of different methods, providing effective references for the practical deployment of device deterioration monitoring.

However, most of the aforementioned studies on continuous surges mainly focus on lightning pulses with time intervals ranging from tens to hundreds of milliseconds. Under such time intervals, the heat dissipation of MOV valve discs can be negligible, and the process can be considered adiabatic, having minimal effect on MOA electrical performance. However, for continuous impacts at the magnitude of tens of seconds, different time intervals can result in significant differences in temperature rise and noticeable changes in electrical performance. If the final temperature rise is the same under different impact intervals, the differences in electrical performance may not be significant when solely judged by the number of impacts or the amplitude.

Therefore, this experiment simulates the scenario where MOV is subjected to continuous overvoltage impacts in the laboratory to study the relationship between different interval times and MOV temperature rise, aiming to provide a basis for determining the MOV.

2 Methodology and Experimental Design

2.1 Experimental Design and Data Collection

In this study, to quantitatively evaluate the health status of Metal Oxide Varistors (MOV), it is necessary to obtain various electrical parameters of the devices during the degradation process. Therefore, an experimental setup was designed using a laboratory-based lightning impulse generator to simulate lightning surges and perform consecutive impact experiments. Data sampling and recording were conducted before and after each impact. The entire impact experiment was carried out in the electrical laboratory of Hubei University of Technology. A total of 28 new MOV devices were selected for the impact tests, labeled as A1 to A11, B1 to B7, C1 to C7, and D1 to D7. All devices had the same specifications: 7D241K, with a nominal varistor voltage V1mA of 240V and a maximum discharge current In of 1.2kA. Under the actual experimental conditions and equipment, the collected electrical parameters of the MOVs included varistor voltage before and after impact, leakage current (total current) before and after impact, and nonlinearity coefficient.

During the simulation of overvoltage impacts, continuous impacts with the same amplitude but different time intervals were applied to each device. The time interval between impacts was set to 10 seconds, and five groups of experiments were conducted, consisting of 1, 2, 3, 4, and 5 consecutive impacts. Each group of experiments involved five consecutive impacts, with data collection and recording of the device's electrical parameters before and after each impact. After each impact, the tested device was allowed to cool to room temperature before the next impact.

During the impacts, consecutive pulses with the same amplitude but different time intervals were applied to each MOV. The time intervals for Groups A, B, C, and D were set to 60 seconds, 10 seconds, 20 seconds, and 30 seconds, respectively. Group A measured the parameters after each individual impact, while the other groups underwent 5 consecutive impacts followed by a 60-second cooling period before collecting and recording the electrical parameters of the MOVs and continuing with the consecutive impacts.

The sampled parameters in this experiment included the baseline electrical parameters of the MOVs: P1 leakage current magnitude I, P2 varistor voltage U at room temperature after impacts, P3 nonlinearity coefficient α, P4 parallel resistance RP, and P5 parallel capacitance CP. Additionally, the total number of impacts and the temperature rise during the impact duration were recorded. The varistor voltage, leakage current, and nonlinearity coefficient were measured using a varistor tester, while the parallel resistance and capacitance were measured using a bridge tester.

2.2 Data Processing for Single-Impact Group

After conducting multiple single-pulse impact experiments on MOVs labeled as A1 to A11, a total of 485 sets of experimental data were obtained. After organizing and summarizing the data, the data from A1 to A8 were selected as the training set for subsequent analysis and modeling, while the data from A9 to A11 were used as the test set to evaluate the performance of the constructed model. Each dataset in the training set included five electrical parameter features: P1 to P5. The parameters measured before and after each impact were averaged to obtain the trend of each electrical parameter for devices A1 to A11 over the number of impacts, as shown in Fig. (**1**).

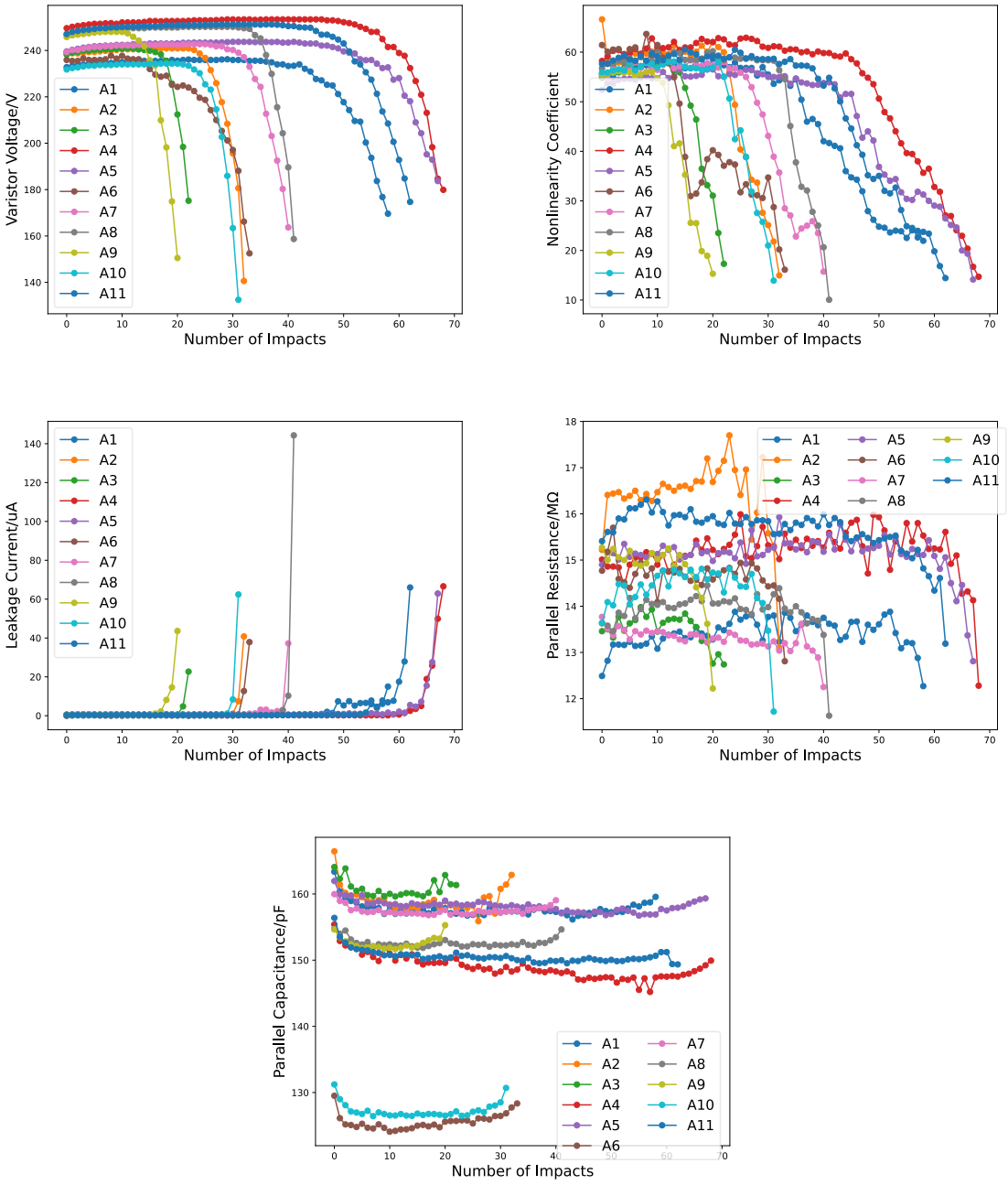

Fig. (1). Variation Trend of Electrical Parameters for A1-A11 with Number of Impacts

For each pulse, the current magnitude is the same, but the energy absorbed by different varistors is evidently different. In the A group experiments, each MOV device was impacted every 60 seconds, and the device cooled down to room temperature before the next impact. However, the span of the number of impacts ranged from 20 to 68. It is unreasonable to evaluate the degradation of varistors based solely on individual parameters like the number of impacts. From the figure, we can observe that the trends of electrical parameters for each device with the number of impacts are quite similar, and the data distribution is relatively dense. As the number of impact tests increases, the room temperature varistor voltage, nonlinearity coefficient, and parallel resistance of the MOV devices show a decreasing trend. The variation of parallel capacitance initially fluctuates and then slowly increases. The leakage current is concentrated in the range of 0-1μA, and after a certain impact, there is a noticeable positive gradient increase in the leakage current. However, there are also some anomalies, such as an increase in the nonlinearity coefficient after the 17th impact for device A6. Therefore, considering the comprehensive changes in all measured parameters and analyzing the overall trend is a reasonable approach to diagnose device degradation.

Principal Component Analysis (PCA) is a statistical method for dimensionality reduction, which reveals the internal structure of multiple variables by examining their correlations. It helps to extract a few principal components that represent the comprehensive variation patterns of multiple parameters in the original data. Thus, for the collected MOV electrical parameter data, PCA can be used to represent the multi-parameter variation patterns using one-dimensional data. For each numbered MOV device, PCA is applied to the data obtained from its impact experiments, and the first principal component is extracted, representing the original 11-dimensional data with a new 1-dimensional representation. The average proportion of the first principal component in the total variance for different numbered MOV devices reaches 0.73. This indicates that the one-dimensional data obtained after dimensionality reduction effectively capture the hidden patterns and information in the original high-dimensional data.

The first principal component data corresponding to devices A1-A11 are analyzed and normalized. It can be observed that the values of the first principal component for each device show an increasing trend with the number of impacts, and the specific trends are generally similar. Taking the normalized first principal component of A11 as an example in Fig. (**2**), PCA achieves the projection of data onto a new feature space, where each principal component is a linear combination of the original features. This makes the principal components easier to interpret and helps us understand the key factors and patterns in the data. We can consider this trend as representing the real-time physical degradation of MOV devices after experiencing impacts, consistent with the understanding that the performance of the devices deteriorates with an increasing number of impacts. The value of the first principal component corresponding to a certain number of impacts reflects the device's degradation state at that moment, where a larger value indicates a more severe damage to the device.

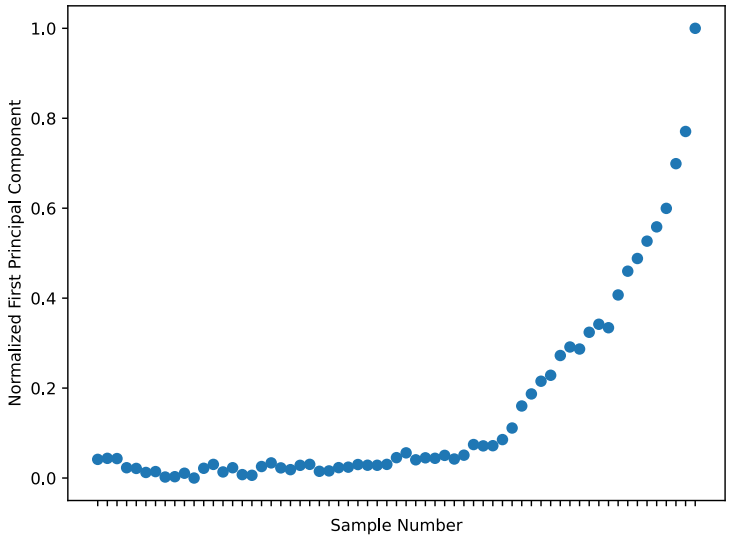

Fig. (2): Trend of Normalized First Principal Component for A11 with Number of Impacts

2.3 Data Processing for Continuous Impact Group

After the entire impact procedure, all devices were thoroughly damaged. Based on the judgment of multiple experimenters and real-time examination of the electrical parameters and appearance of the devices after impacts, devices that exhibited significant degradation or visible damage to the casing were identified as having lost their protective capability. This state was defined as a highly degraded condition and used as a label for the input of the raw data. Each device in this state corresponded to a numerical value representing the degree of degradation, calculated using PCA. In addition to the basic electrical parameters, the electrical performance of the MOVs is closely related to temperature.

Li Pengfei concluded from experiments conducted in the microsecond and millisecond time scales that the internal modules of MOVs accumulate heat almost instantaneously, causing a rapid increase in temperature. From the perspective of energy convection exchange, this process can be regarded as an adiabatic temperature rise process [20].

The temperature rise is related to the energy absorbed by the varistor. The energy absorbed by the varistor's valve disk (E, measured in joules) is given by:

$$E = \int_0^{t_0} uidt \tag{1}$$

In the equation, u represents residual voltage in volts (V), i represents the impulse current in amperes (A), and t0 represents the duration of the current in seconds (s).

The temperature rise of the valve disk (ΔT, measured in Kelvin) is related to the injected energy E by the following relationship:

$$\Delta T = \frac{E}{mc_p} \tag{2}$$

Formula (2) can be alternatively expressed as:

$$\Delta Q = \frac{\Delta W}{\Delta V \rho c_p} \tag{3}$$

In Formula (3), ΔV represents the volume of the MOV unit, ρ represents the volume mass of the MOV (averaging at 5600 kg/m^3), and cp represents the specific heat capacity of the MOV (made of zinc oxide material) at 20°C, approximately 500 J/(kg·°C).

In this experiment, the temperatures before and after each of the five consecutive impacts were recorded. The experiment aimed to discuss the degradation state and lifespan distribution of MOV in its normal condition. Under the same magnitude of impact current and pulse duration, a higher temperature indicates that the pulse injected more energy into the MOV. During a single impact process, the accumulation of heat is not very significant due to factors such as heat dissipation speed. The temperature rise is regulated by various factors, resulting in a relatively small temperature increase. However, during the process of continuous impacts, energy injection is continuous, and the MOV does not have enough time for effective heat conduction and dissipation. Previous impacts cause the internal temperature of the MOV to continue to rise. Before the next impact arrives, the MOV does not have enough time to cool down. This continuous accumulation of heat leads to a continuous increase in the temperature of the MOV, which in turn triggers more severe degradation. Therefore, the temperature rise during the final pulse process of the continuous impacts, which can be considered as an adiabatic process, is selected as the parameter to extract, as it exhibits a more pronounced temperature rise effect.

Fig. (**3**) takes the temperature rise data of A1-D1 as examples. The data from groups B, C, and D show clear regularity and better fit the degradation characteristics of the MOV.

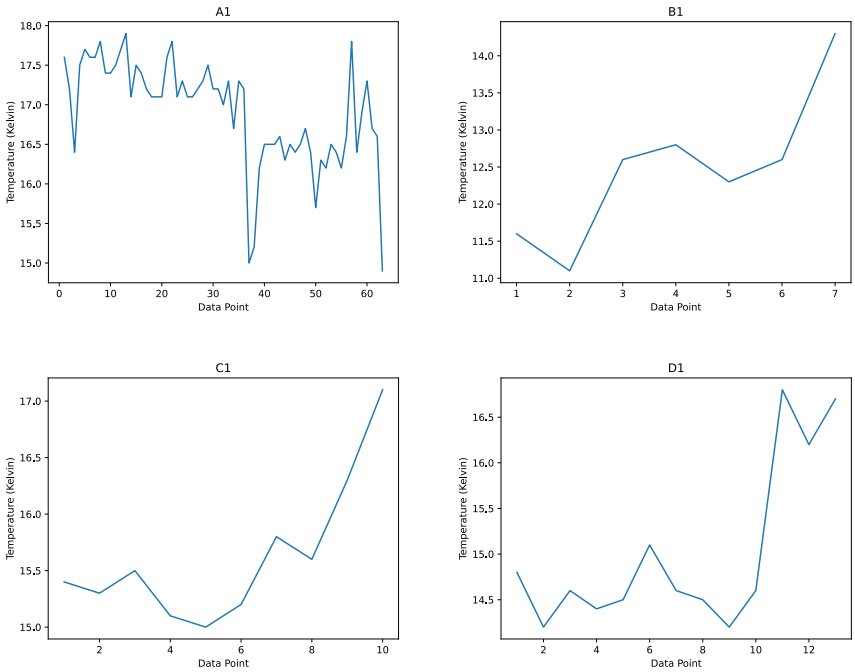

Fig. (3) The trend of temperature rise during the fifth pulse of impacts (A1-D1) with the number of impacts.

Based on the results mentioned above, the temperature rise data for groups B, C, and D were added and standardized. Seven sets of data within each group were selected, and the degradation level of each data sample was calculated. This degradation level was represented by an array ranging from 0 to 1, indicating the degree of degradation on a linear scale (normalized PCA), where 0 represents no degradation and 1 represents maximum degradation. These degradation level values were stored in a designated list. Additionally, the group labels (1 to 7) to which each data sample belongs were stored in another list. All the data were then sorted based on the degradation level list, and the merged dataset was rearranged using the sorted indices. This was done to better understand and visualize the data in subsequent analyses. Finally, the source information of each sample was added as a new column to the sorted dataset, labelling the group affiliation of the new data points.

Fig. (**4**) shows the scatter plot of the degradation distribution for each group after undergoing the aforementioned processing. The degradation distribution trends for groups B, C, and D are similar to that of group A.

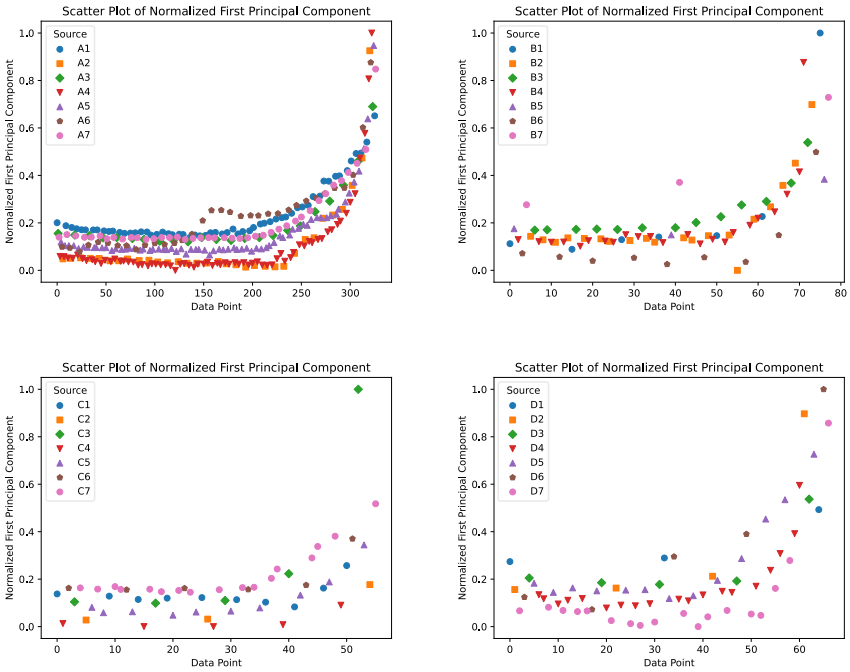

Fig. (4) The scatter plot of the deteriorated distribution for groups A to D

3 Model Validation and Analysis

3.1 Algorithm Selection

3.1.1 Support Vector Machine (SVM)

Support Vector Regression (SVR) is a branch of Support Vector Machine (SVM) that was proposed. Its basic idea is to map the input space to a high-dimensional space using a kernel function and then find the optimal classification plane for the original sample set. This plane minimizes the errors between all samples and the optimal classification plane. SVR has the following characteristics:

Efficient Nonlinear Classifier: SVR constructs a hyperplane in a high-dimensional space, making it suitable for solving nonlinear classification problems. By using a kernel function, SVR can map the low-dimensional input space, such as the 6 to 7 groups representing the degradation of MOV, to a high-dimensional feature space, enabling it to handle nonlinear relationships effectively. Robustness: SVM selects the decision boundary by maximizing the margin, making it robust to noise and outliers. In the case of MOV, there may be outliers or noise due to manufacturing processes and other factors. However, SVM can still find a good classification boundary. Small Sample Size: SVR is an algorithm based on structural risk minimization, with a focus on support vectors, which are the samples closest to the decision boundary. Since the degradation parameters of the MOV involve only a small number of samples for model construction, SVR is highly effective for such small sample problems[21,23].

Let the linear regression function established in the high-dimensional feature space be represented by Formula (4):

$$f(x) = w^{\mathrm{T}}\varphi(x) + b = \sum_{i=1}^{m} \alpha_i y_i \varphi(x_i)^{\mathrm{T}}\varphi(x) + b \tag{4}$$

In the equation, $\boldsymbol{x_i}$ represents the input vector of the i-th sample, $\boldsymbol{y_i}$ represents the output value of the sample, w = (w_1, w_2, ..., w_n)$^{\mathrm{T}}$ represents the weight coefficient vector, $\boldsymbol{\varphi(x_i)}$ represents the nonlinear mapping function, $\boldsymbol{f(x)}$ represents the estimated value of y, and b represents the constant term.

The construction of the optimal classification plane can be transformed into a quadratic programming problem, as shown in Formula (5):

$$max\alpha, \alpha^* \left[\begin{array}{c} -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \\ k(x_i, x_j) \\ \sum_{i=1}^{n} (\alpha_i + \alpha_i^*)\varepsilon + \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) y_i \end{array} \right] \quad (5)$$

The constraint conditions can be represented by Formula (6):

$$\begin{cases} \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) = 0 \\ 0 \leqslant \alpha_i \leqslant C \\ 0 \leqslant \alpha_i^* \leqslant C \end{cases} \quad (6)$$

In the equation, $\boldsymbol{\alpha_i}$ represents the Lagrange multiplier associated with each sample, ε represents the linear insensitive loss error, and C represents the penalty factor. A larger value of C indicates a stronger penalty for samples with errors greater than ε, while a smaller value indicates a weaker penalty.

Finally, SVR can be represented by Formula (7):

$$f(x) = \sum_{i=1}^{m} \left(\hat{\alpha}_i - \alpha_i \right) k(x_i^{\mathrm{T}} x) + b \quad (7)$$

In the equation, $\boldsymbol{k(x_i^{\mathrm{T}} x) = \varphi(x_i)^{\mathrm{T}} \varphi(x_j)}$ represents the kernel function.

3.1.2 GS-SVR Algorithm

The grid search algorithm was originally used to solve constrained nonlinear optimization problems. The basic steps of the algorithm are as follows: first, set a larger step size and a range for the parameters C and g by adjusting the grid division multiple times. This helps roughly determine the approximate range of the optimal parameter combination, thereby narrowing down the search range and reducing the computational complexity after grid division, thus improving the computational speed of the model. Then, set a reasonable range to constrain the parameters C and g, and perform fine grid division with a smaller step size. Combine the values of each parameter and generate a grid of all possible combination results. Next, use each combination for SVR training and employ cross-validation with the minimum root mean square error on the training set as the fitness function to search for optimal parameters [24]. Finally, when the minimum root mean square error is achieved, the optimal result graph under the selected search is obtained, and the optimal combination parameter values are calculated. The advantages of this algorithm are as follows: 1) It can simultaneously search multiple parameters and obtain the "optimal" combination. 2) It can avoid the problem of multiple solutions caused by parameter coupling. 3) It is computationally efficient as it can be executed in parallel [25].

First, the required input variables are determined using dimensional analysis. On this basis, different input combinations are selected and normalized. The SVM model is optimized using the grid search and cross-validation methods for different training sample ratios and input combinations. The GS-SVR model with the best input combination is established, and the distribution patterns of the model's hyperparameters C and g are analyzed. The entire workflow of the GS-SVR model is shown in Fig. (5).

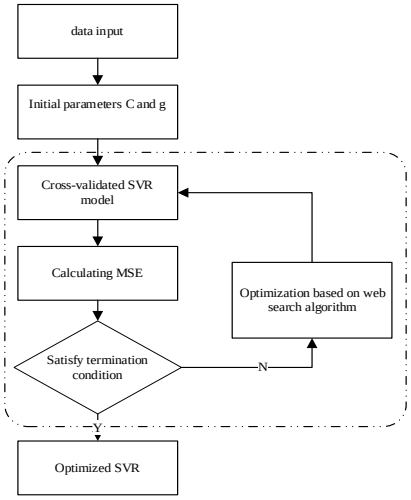

Fig. (5): GS-SVR Model Process

3.2 Degradation Model

3.2.1 Model Construction Process

By applying the GS-SVR algorithm, a numerical regression modeling can be established for the degradation state of MOV devices using the A, B, C, and D groups' MOV electrical parameter data and their corresponding first principal component data obtained from impact experiments. Once the model is constructed, it can take a new set of electrical parameters as input and output a numerical value reflecting the degree of device degradation, ranging from 0 to 1. When this value exceeds a threshold (estimated based on the distribution of the first principal component values corresponding to completely damaged MOVs in the respective groups), it is considered that the device is completely damaged after the impact. The overall construction process of this study is illustrated in Fig. (6).

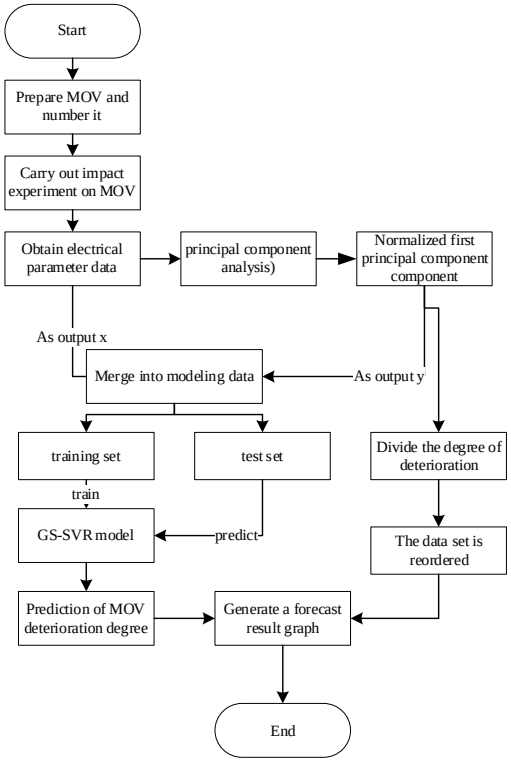

Fig. (6): Overall Process for MOV Degradation State Diagnosis

As the degradation state of a device changes, its corresponding single electrical parameter exhibits different distribution patterns in the data space, indicating its inherent regularity. As shown in the flowchart, when the number of measurements of variability (MOV) reaches a certain threshold, more data on the sampled electrical parameters can be obtained, leading to a more comprehensive understanding of the distribution patterns of each electrical parameter. Therefore, from a statistical perspective, the more data available, the stronger the robustness of the model constructed based on the data. This results in more accurate predictions of the degradation state values for new data, aligning closely with the actual degradation level of the device itself. This process achieves the evaluation and warning of the actual state and degradation level of MOV through monitoring and analysis of the sampled data.

4 Conclusion

This experiment mainly tests the performance of the algorithm in the analysis and processing of actual MOV degradation states, as well as its advantages compared to other similar algorithms. For this purpose, random forest regression (RF), traditional support vector regression (SVR), gradient boosting decision tree (GBDT), and grid search support vector regression (GS-SVR) are specifically chosen as control groups. The evaluation metrics include mean squared error (MSE), mean absolute error (MAE), and coefficient of determination (R^2) for each model. The formulas for calculating these three metrics are as follows:

$$MSE(y,\hat{y}) = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 \tag{8}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2/m}{\sum_{i=1}^{m}(y_i - \overline{y}_i)^2/m} \tag{10}$$

Mean Squared Error (MSE) is a common metric used to measure the difference between the predicted values and the true values in regression models. It computes the squared differences between the predicted and true values and takes the average over all samples. Mean Absolute Error (MAE) is another common metric used to measure the difference between the predicted values and the true values in regression models. It computes the absolute differences between the predicted and true values and takes the average over all samples. Both of these metrics are used to assess the prediction accuracy of regression models, where smaller values indicate smaller differences between the predicted and true values and better predictive capabilities of the model. The coefficient of determination (R^2) ranges from 0 to 1 and represents the model's explanatory power for the target variable. A higher R^2 value indicates a better ability of the model to explain the target variable, indicating a better fit. When the R^2 value is equal to 1, it means the model perfectly captures the variations in the target variable, indicating that the model can explain all the variability in the target variable.

For the total training set of data in Group A, the calculated MSE is 0.00056, MAE is 0.01909, and R^2 is 0.98802. The prediction results for other algorithms are shown in Table 1.

Table 1 Testing Metrics Results for Four Algorithms in Group A

Algorithm	MSE	MAE	R^2
GS-SVR	0.00056788	0.01908692	0.98802007
SVR	0.00107279	0.02630365	0.97736878
RF	0.00092899	0.01866634	0.98040241
GB	0.00071055	0.02039731	0.98501037

For a visual comparison, Fig(8-10) shows the numerical values of the testing metrics results for the four algorithms. The data has been scaled to facilitate their display on the same scale.

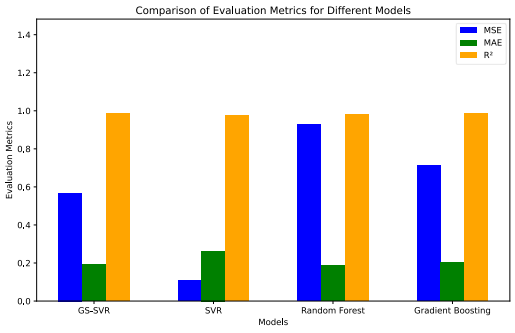

Fig. (7): Performance Evaluation Metrics Comparison of Four Algorithms in Group A

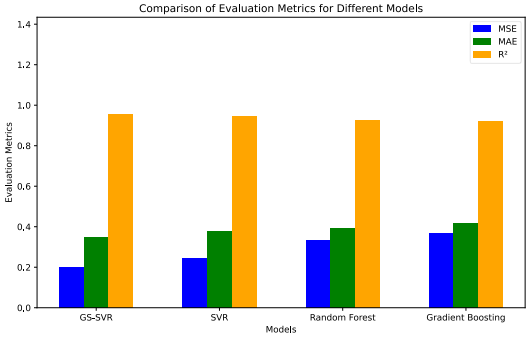

Fig. (8): Performance Evaluation Metrics Comparison of Four Algorithms in Group B

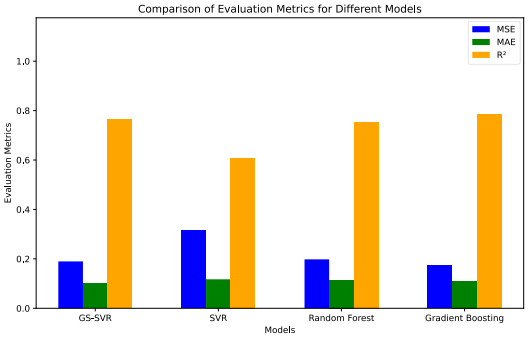

Fig. (9): Performance Evaluation Metrics Comparison of Four Algorithms in Group C

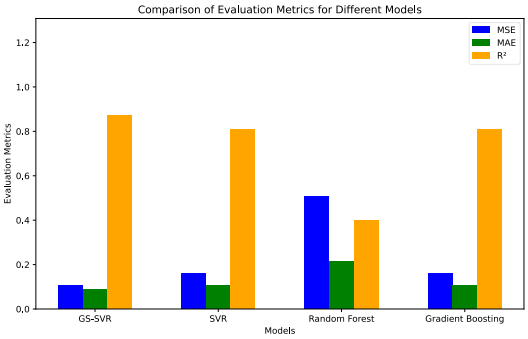

Fig. (10): Performance Evaluation Metrics Comparison of Four Algorithms in Group D

In datasets A, B, C, and D, the GS-SVR model performed the best, with the lowest mean squared error (MSE) and mean absolute error (MAE), as well as the highest coefficient of determination (R^2). The SVM model followed closely, with slightly higher metrics than the random forest model and gradient boosting decision tree model. Although the GS-SVR model still performed the best in datasets B, C, and D, compared to dataset A, its MSE and MAE increased slightly, while the coefficient of determination decreased slightly. The SVM model also showed relatively good performance on these datasets, but there is still some gap compared to the GS-SVR model. The random forest model and gradient boosting decision tree model performed less satisfactorily.

To provide a more visual explanation of the prediction results, Figure 11 is presented, where each point's x-coordinate represents the actual degradation state value in Group A, and the y-coordinate represents the predicted degradation state value given by the GS-SVR model. It can be observed that all the scatter points are well-distributed around the red reference line, indicating a small error between the predicted and actual values. The model can effectively diagnose and predict the degradation state of MOV devices.

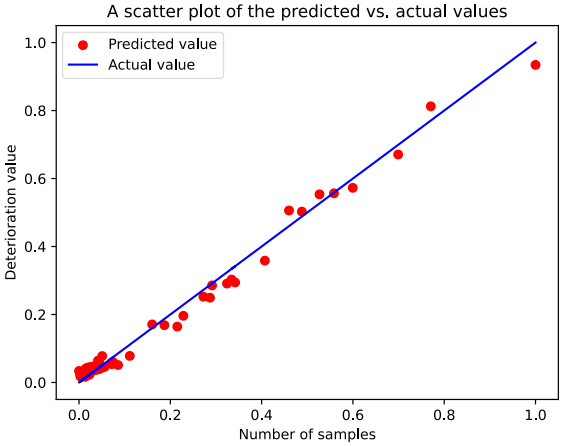

Fig (11): Comparison of Predicted and Actual Degradation Values by the GS-SVR Model

Additionally, the scatter plots showing the distribution of predicted results versus the true values for groups A, B, C, and D are presented in Fig (**12-15**).

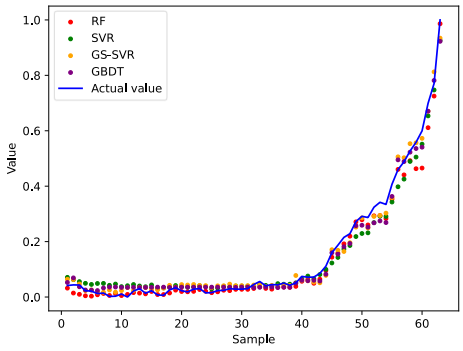

Fig. (12): Scatter Plot of Predicted Results versus True Values for Group A in Various Algorithms

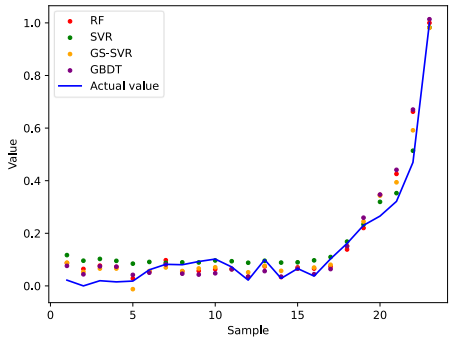

Fig. (13): Scatter Plot of Predicted Results versus True Values for Group B in Various Algorithms

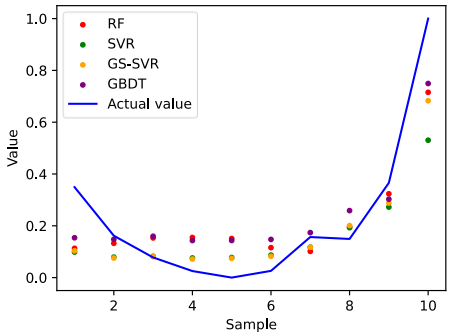

Fig. (14): Scatter Plot of Predicted Results versus True Values for Group C in Various Algorithms

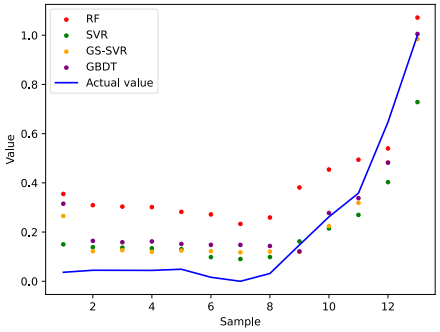

Fig. (15): Scatter Plot of Predicted Results versus True Values for Group D in Various Algorithms

5 Discussion

This study designed experiments to obtain electrical parameter data for MOV devices in different experimental groups, subjected to single-pulse and continuous-pulse impacts. The data was then used to characterize the degradation state of the MOV devices using principal component analysis (PCA). For the continuous-pulse group, the temperature information from the last impact was included as supplementary data. Finally, a numerical prediction model for the degradation state of the MOV devices was built using grid search and support vector regression (SVR) algorithm. The feasibility of the model's approach for device performance monitoring applications was verified based on the predicted results obtained from the training dataset. The model demonstrated the ability to accurately predict the degradation state values of MOV devices after experiencing impacts. In terms of extending the analysis to continuous-pulse impacts, the experiment tested MOV devices with continuous pulses and recorded their parameter data. The GS-SVR model was used to train a predictive model specifically for single-pulse data. The experimental results showed that this method still achieved certain effectiveness in predicting continuous-pulse data. Since the degradation assessment values are scaled between 0 and 1, it provides an intuitive basis for evaluating the actual degradation level of the devices. The overall approach provides a new feasible solution for online monitoring of varistors and their engineering applications in electrical systems.

However, there are areas for improvement and limitations in the entire research process. Due to limitations in experimental time and conditions, the collected experimental data was limited, resulting in a less comprehensive and representative distribution of the data samples in the data space. Consequently, the robustness of the constructed model is somewhat lacking. Based on this research process, in future practical applications, it is important to collect more diverse data for modeling to improve the accuracy of diagnosing the degradation state of MOV devices and enable efficient response to more complex pulse impact scenarios in real-world applications. Additionally, although the GS-SVR model performed relatively well on different datasets, with low mean squared error, mean absolute error, and high coefficient of determination, this is only a preliminary analysis of the data, and more factors and evaluation metrics need to be considered to comprehensively assess the performance of the model. Furthermore, many other excellent and suitable algorithms have not been fully explored. Finally, further exploration of the physical characteristics of the devices should be strengthened to gain a more scientific understanding and discover sensitive features that affect device degradation. By incorporating these features as new sampling parameters and following the methodology outlined in the paper, the one-dimensional numerical representation of the device degradation level obtained through PCA will also be more reliable and credible.

LIST OF ABBREVIATIONS

MOV	=	Metal Oxide Varistors
PCA	=	Principal Component Analysis
GS	=	Total Harmonic Distortion Grid Search
SVR	=	Support Vector Regression
MSE	=	Mean Squared Error
MAE	=	Mean Absolute Error

REFERENCES

[1] Jie, G., Shuaishuai, Y., Mengyuan, W., Yang, Z., & Hanxin, X. (2019). Research Progress on Pulse Discharge Characteristics of Zinc Oxide Varistor Ceramics. Functional Materials, 50(02), 2001-2005.

[2] Eda, K. "Zinc oxide varistors." IEEE Electrical Insulation Magazine, vol. 5, no. 6, pp. 28-30, 1989.

[3] Gupta, T. K. (1990). Application of Zinc Oxide Varistors. Journal of the American Ceramic Society, vol. 73, no. 7, pp. 1817-1840.

[4] Y. Liao, R. Li, C. Gao, S. Li, H. Wang, J. He. "Study on residual voltage characteristics of whole arrester under multiple lightning impulses." Proceedings of the 2014 Annual Conference of the Chinese Society for Electrical Engineering, Hefei, China Society for Electrical Engineering, 2014, pp. 1-6.

[5] Lin Chubiao, Liao Yongli, Li Ruihai, Li Shengtao, Wang Hui, Gao Chao, He Jinqiang. "Study on aging characteristics of 110kV whole zinc oxide arrester under lightning impulse." Southern Power Grid Technology, 2015.

[6] L. Zhou, R. Wei, C. Liu, W. Chen, L. Huang, and Y. Ma, "Impact Aging Characteristics of Zinc Oxide Surge Arrester under Multiple Lightning Strikes," High Voltage Engineering, vol. 48, no. 9, pp. 3507-3516, 2022.

[7] Zhang, Chunlong, Hongyan Xing, Pengfei Li, Chunying Li, Dongbo Lv, and Shaojie Yang. "An Experimental Study of the Failure Mode of ZnO Varistors Under Multiple Lightning Strokes." Electronics 8, no. 2 (2019): 172.

[8] B. Wang, J. Lu, P. Gao, Z. Fu, and Z. Jiang, "Study on the high impulse current withstand properties and failure mechanism of ZnO varistors with different Bi2O3 content," Journal of Materials Science: Materials in Electronics, vol. 33, no. 33, pp. 25446-25462, 2022.

[9] Feng Q ,Wei C ,Xutong W , et al.Transient Response Characteristics of Metal Oxide Arrester under High-Altitude

Electromagnetic Pulse[J].Energies,2022,15(9):3303-3303.

[10] Attila S ,Mihai F F ,Mihaela F , et al.Service Limits for Metal Oxide Varistors Having Cylindrical Symmetry as Function of the Ambient Temperature [J].Symmetry, 2022,14(7):1351-1351.

[11] Qibin Z ,Xin H ,Ting C , et al.Research on electrothermal characteristics of metal oxide varistor based on multi-physical fields[J].IET Generation, Transmission & Distribution,2022,16(18):3636-3644.

[12] Wang X ,Liu Y ,Lv Y , et al.Investigation of Impulse Aging of Energy-Absorption Elements for Hybrid DC Circuit Breakers[J].Applied Sciences,2023,13(17):J. Li, Y. Jia, "Research on coal burst forecast based on improved GS-SVM," World Science and Technology Research and Development, vol. 2016, no. 4, pp. 758-762.

[13] Shu H ,Shao Z ,Dai Y .Research on the improved hybrid DC circuit breaker with voltage-limiting and current-limiting capability[J].International Journal of Electrical Power and Energy Systems,2024,158109943-.

[14] Wang K ,Jia C .A high-reliability voltage balancing circuit using MOV for series-connected HV-IGBT[J].Journal of Physics: Conference Series,2025,2936(1):012008-012008.

[15] Z. Yin, P. Xie, B. Wang, et al., "Finite element analysis of zinc oxide surge arrester under impulse current," Insulators and Surge Arresters, vol. 2023, no. 01, pp. 216-222, 2023.

[16] W. Lu, X. Xiao, X. Wei, et al., "Aging characteristics of valve discs in zinc oxide surge arresters under long duration impulse current," Insulators and Surge Arresters, vol. 2023, no. 04, pp. 90-95, 2023.

[17] C. Zhang, C. Li, D. Lv, H. Zhu, H. Xing, "An Experimental Study on the Effect of Multiple Lightning Waveform Parameters on the Aging Characteristics of ZnO Varistors," Electronics, vol. 9, p. 930, 2020.

[18] J. He, C. Cheng, J. Hu. "Electrical degradation of double-Schottky barrier in ZnO varistors." AIP Advances, vol. 6, no. 3, pp. 030701, 2016.

[19] H. Huang, B. Yan, C. Zhao, "Optimized FFT analysis for leakage current detection of zinc oxide arrester," Journal of Electronic Measurement and Instrumentation, vol. 2019, no. 4, pp. 87-94.

[20] W. Wang, T. Guo, H. Yang, "Study on metal oxide varistor fault diagnosis based on principal component analysis and BP neural network," Insulators and Surge Arresters, vol. 2019, no. 06, pp. 20-25.

[21] Z. Yang, H. Cao, P. Li, "Online monitoring of metal oxide arrester using genetic algorithm," High Voltage Engineering, vol. 2015, no. 9, pp. 3104-3109.

[22] W. Wang, Z. Yang, Z. Yu, Z. Yu, Y. Sun, X. Yao, H. Li., "Research on online monitoring index of aging deterioration of MOA under multi-pulse," Insulators and Surge Arresters, vol. 2019, no. 01, pp. 31-37, 2019.

[23] C. Niu, Z. Yu, Z. Yang, "Evaluation of MOV power loss based on ANFIS," Insulators and Surge Arresters, vol. 2020, no. 01, pp. 1-6, 2020.

[24] Z. Yu, J. Liang, W. Li, "Discussion on the problems of MOA online monitoring," Insulators and Surge Arresters, vol. 2020, no. 06, pp. 144-150, 2020.

[25] F. Mao, Q. Gui, J. Jiang, et al., "Test and simulation of ZnO varistor under multiple pulses," Insulators and Surge Arresters, vol. 2018, no. 02, pp. 73-76, 82.

[26] P. Li, C. Zhang, D. Lv, "Destruction modes of metal oxides under multiple pulse lightning impulse," High Voltage Engineering, vol. 2017, no. 11, pp. 3792-3799. DOI: 10.13336/j.1003-6520.hve.20171031040.

[27] Y. Liu, L. Wang, K. Gu, "A support vector regression (SVR)-based method for dynamic load identification using heterogeneous responses under interval uncertainties," Applied Soft Computing, vol. 110, pp. 107599, 2021. ISSN 1568-4946.

[28] J. Niu, "Dam deformation monitoring model based on singular spectrum analysis and SVM optimized by PSO," Advances in Science and Technology of Water Resources, vol. 2020, no. 6, pp. 60-65.

[29] Y. Zhang, Q. Yang, "Research on thermal error prediction of CNC machine tools based on GS-SVM," Mechanical Engineer, vol. 2019, no. 11, pp. 36-38.

[30] J. Wang, L. Zhang, G. Chen, et al., "A parameter optimization method for an SVM based on improved grid search algorithm," Applied Science and Technology, vol. 2012, no. 3, pp. 28-31.

[31] J. Li, Y. Jia, "Research on coal burst forecast based on improved GS-SVM," World Science and Technology Research and Development, vol. 2016, no. 4, pp. 758-762.

Xu Zhengwang ceived the B.S. and M.S. degrees in Hubei University of Technology. Since July 1998, he has been working in the Department of Electronic Information at Hubei University of Technology, His current and past research work includes CNC cutting, non-destructive testing, AC variable frequency, soft start, wireless power transmission, and solar-powered LED lights.

Zhiheng Zhu received his Bachelor of Science degree in Electronic Information Engineering from Hubei University of Technology in 2021. Since 2022, he has been pursuing a Master's degree in Power Electronics and Power Transmission at Hubei University of Technology.

Zhou Huang obtained both his Bachelor's and Master's degrees from Hubei University of Technology. He is currently engaged in teaching activities at Hubei University of Technology.