


## Article

# Predicting the Life of Varistors via a Nonlinear Coefficient Based on a Small-Scale Data Model

Xiaofei Ruan, Weidong Cheng, Yihao Bai \*  and Weigang Wen

School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China; ruanxiaofei@rails.cn (X.R.); wdcheng@bjtu.edu.cn (W.C.); wgwen@bjtu.edu.cn (W.W.)

\* Correspondence: 19116001@bjtu.edu.cn

**Abstract:** With the great leaps forward in the development of the railway, the importance of safe operation and maintenance has become increasingly prominent. Varistor is the key component in surge protective devices for railway communication and signaling equipment, it is necessary to study a description model of the varistor degradation process and predict its lifetime through condition monitoring. Among the monitoring parameters, nonlinear coefficient is an important index to measure the health of varistors. Considering that the degradation process of varistors is the cumulative effect of surge history, and its nonlinear coefficient has a time-series relationship, a life prediction model for varistors based on BiLSTM is proposed. The model innovatively uses nonlinear coefficient as the characterization of the deterioration degree of varistors and uses a small-scale network to predict the development trend of nonlinear coefficient automatically and accurately. Verified by surge impact experiments, the model can accurately predict the state of nonlinear coefficient according to historical data and has the potential for engineering applications in predicting the life of varistors.

**Keywords:** varistor; nonlinear coefficient; life prediction; BiLSTM



**Citation:** Ruan, X.; Cheng, W.; Bai, Y.; Wen, W. Predicting the Life of Varistors via a Nonlinear Coefficient Based on a Small-Scale Data Model. *Appl. Sci.* **2024**, *14*, 171. <https://doi.org/10.3390/app14010171>

Academic Editor: Mohamed Benbouzid

Received: 13 November 2023

Revised: 11 December 2023

Accepted: 18 December 2023

Published: 24 December 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The metal oxide varistor (MOV) is a nonlinear resistance element with zinc oxide (ZnO) as the main component. Due to its high surge current tolerance and very large current absorption capacity, it is the key component in surge protective devices (SPD). At present, it has been widely used in systems to protect railway communication from. Its performance directly affects the safety of the whole system [1]. In their applications, due to the influence of environmental factors such as high voltage and high temperature, varistors are more likely to deteriorate and often fail to reach their calibrated working life. At present, the operation and maintenance of SPD is generally periodic maintenance, which is mainly conducted by staff on a regular basis to observe whether the appearance is damaged or whether the thermal fuse device is fused. The efficiency is relatively low, and only SPDs that have completely failed can be detected; the SPDs that are about to fail or have just failed are often unable to respond. With the advancement of intelligent railways, it is necessary to conduct effective and accurate online evaluations of lifetime surge capacity to prevent disasters caused by SPDs and improve the level of lightning protection operation and maintenance. With the development of electronic technology, monitoring methods have become increasingly mature and advanced. It is feasible to carry out online life predictions for surge protective devices under the existing technical conditions [2–6]. At present, the breakdown voltage  $U_{1mA}$  and leakage current  $I_{Leakage}$  are often used to judge the degree of deterioration in varistors. However, these two indicators can only reflect the working state of the varistor and cannot reflect the degree of deterioration, and there will be a certain amount of lag in their judgment on the failure of the varistor [7,8].

The nonlinear coefficient is an important index reflecting the working performance of the varistor. The nonlinear coefficient is a calculated parameter that reflects the ratio of

the current to the voltage after the voltage exceeds a threshold. The greater the value, the more effective the protection of the varistor on the circuit. At present, it has been shown that there is a strong correlation between the nonlinear coefficient and the degree of deterioration in a varistor. When the varistor works normally, its range is generally 20–100 [9]. Yi Guilai et al. [10] found that the nonlinear coefficient of a varistor changes regularly under the impact current. Through a varistor impact experiment, Yang Zhongjiang et al. [11] found that the nonlinear coefficient of the varistor would decrease with its degradation. In recent years, some research work has explored the factors affecting the nonlinear coefficient. He Jinliang et al. [12] found that the formation of a liquid phase during sintering has a significant influence on the nonlinear coefficient, which is related to the chemical reaction between the liquid phase and the mixed elements at the grain boundary. Nahm et al. [13] found that the nonlinear coefficient has a great relationship with the morphology of the grain boundary. At the same time, there are many researchers and scholars committed to improving the nonlinear coefficient of varistors [14–18], but there are few studies that have applied the nonlinear coefficient as the main parameter to construct the lifetime model.

In practice in engineering, there is often a need for a way to make automatic and accurate judgments on the deterioration of varistors. Nowadays, with the continuous development of the computer industry and the maturity of artificial neural network technology, this goal has become possible. At present, the trend of development in neural networks is deeper and more complex structures, which have a good fitting effect in the face of big data. However, lightning strikes occurring under natural conditions are difficult to predict, so the acquisition cycle is long and it is extremely difficult to obtain the key parameters in the deterioration process of varistor for a full life cycle. Even in a laboratory setting, collecting data for a full life cycle requires several hours of repeated surge impacts, and each experiment results in the complete failure of a varistor, which creates high experimental costs. In the face of limited sample data, too large network models will produce serious overfitting problems. So, we need to find a model of the right scale.

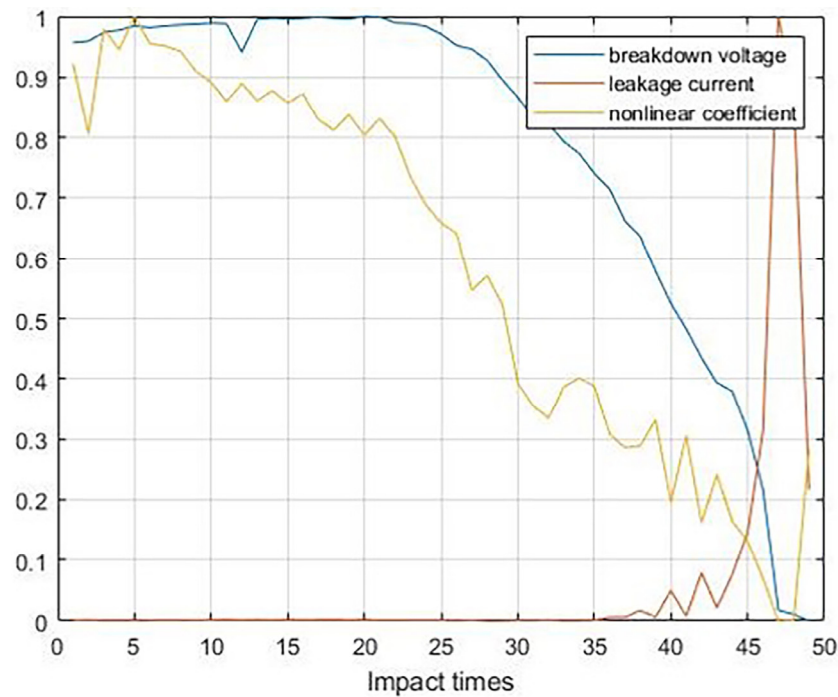
By establishing feedforward links between neurons, the recurrent neural network (RNN) can mine the temporal information between data and has good adaptability in the face of temporal data tasks. On this basis, LSTM can effectively improve the gradient explosion problem in RNN by constructing the gate structure, and it has a very good fitting effect in the face of small-scale data. BiLSTM is an improved LSTM with a higher timing prediction accuracy, which has been widely used in text recognition, semantic recognition and other aspects [19–21]. This model has good application prospects in predicting the life of a varistor.

In this paper, life prediction for varistors is realized by building a life prediction model via the change in the nonlinear coefficient during the deterioration of varistors based on BiLSTM. The structure of the paper is as follows. The first section is an introduction. In Section 2, the life prediction method for varistors is presented, including parameter selection, data enhancement and model construction. In Section 3, the correctness of the model is verified by experiments. The conclusion is in Section 4.

## 2. Materials and Methods

### 2.1. Electrical Parameters of the Varistor

In research, breakdown voltage, leakage current and nonlinear coefficient are common electrical parameters used to measure the performance of a varistor. They can be measured directly by professional equipment. Breakdown voltage refers to the voltage across the varistor when 1 mA of direct current flows through it. Leakage current refers to the current flowing through the varistor at a specified temperature and maximum direct current voltage. Figure 1 shows the changes in these three electrical parameters with various times of impacts in a measured experiment. As there is a large difference in the range, in order to compare the trends of changes in the three electrical parameters more intuitively, the data have been normalized.



**Figure 1.** The changes in electrical parameters with the time of impact (the values were normalized).

It can be seen that there is no correlation between the leakage flow curve and the impact times. The level remains low until the varistor is completely damaged and then increases sharply. Therefore, it can only reflect the working state of the varistor and not its deterioration. The breakdown voltage and nonlinear coefficient are both negatively correlated with the number of impacts. But it is clear that the change in the breakdown voltage lags behind the change in the nonlinear coefficient. Even after the 35th impact, when the varistor is close to damage, the breakdown voltage is still more than 70% of its peak value, which cannot reflect the deterioration of the varistor over time. Therefore, the nonlinear coefficient is chosen as the electrical parameter used to construct the life prediction model.

Nonlinearity is an inherent characteristic of MOV. When the terminal voltage is lower than a certain threshold, the flux of the varistor is about zero. When the terminal voltage exceeds this threshold, the through current in the varistor increases sharply with the increase in the terminal voltage, and its volt–ampere characteristic curve is shown in Figure 2.

Within a certain range greater than the terminal voltage threshold, the volt–ampere characteristic curve is approximately a straight line, and its curve can be approximately expressed as

$$lgI = AlgU - b, \tag{1}$$

where  $lgI$  is the vertical coordinate in Figure 2,  $lgU$  is the horizontal coordinate in Figure 2, and  $A$  is the slope of the straight line, i.e.,

$$A = lg\theta, \tag{2}$$

$\theta$  is the angle at which the approximate line intersects the coordinate axis, and  $b$  is the intercept of the straight line.

Let  $b = AlgC$ , where  $C$  is the material parameter of the varistor, i.e.,

$$I = (U/C)^A, \tag{3}$$

It can be seen that the volt–ampere characteristic of the varistor in this range is determined by the parameters  $A$  and  $C$ .  $A$  is the nonlinear coefficient, which is an electrical

parameter describing the nonlinear degree of MOV. The larger the  $A$  value, the steeper the volt-ampere characteristic curve in this voltage range, and the stronger the absorption ability of the varistor for large currents.

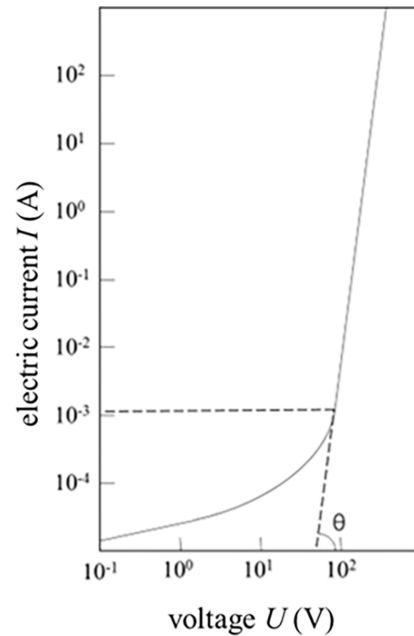


Figure 2. Volt-ampere characteristic curve for the varistor.

After experiencing multiple surge impacts, the varistor deteriorates; ions on both sides of the forward and reverse bias barriers migrate, resulting in a drop in the varistor voltage and a deterioration in the volt-ampere characteristics of the varistor, as shown in Figure 3. At the same time, the nonlinear coefficient  $A$  also decreases, and the current absorption ability of the varistor decreases. When nonlinear coefficient  $A$  decreases to a certain extent, the varistor fails and cannot play the role in overcurrent protection. This paper focuses on the trend of change in the nonlinear coefficient in the process of deterioration, and establishes a life prediction model based on experimental data.

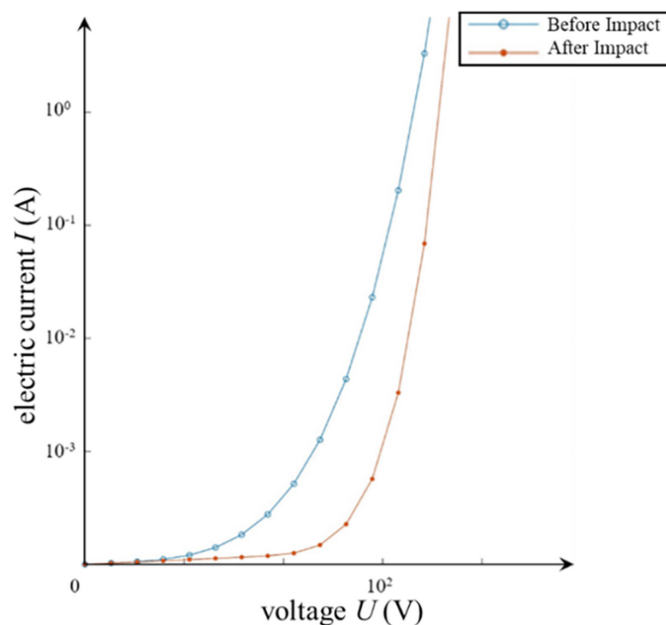


Figure 3. Change in volt-ampere characteristic curve before and after surge impact.

Figure 4 shows the variation in the nonlinear coefficients of three different varistors in the same model but in different batches in the surge impact experiment. It can be seen that the nonlinear coefficient has a clear correlation with the number of impacts. But the correlation between different varistors is significantly different. Therefore, it is difficult to obtain a reliable model with generalization ability through nonlinear fitting and other traditional statistical methods.

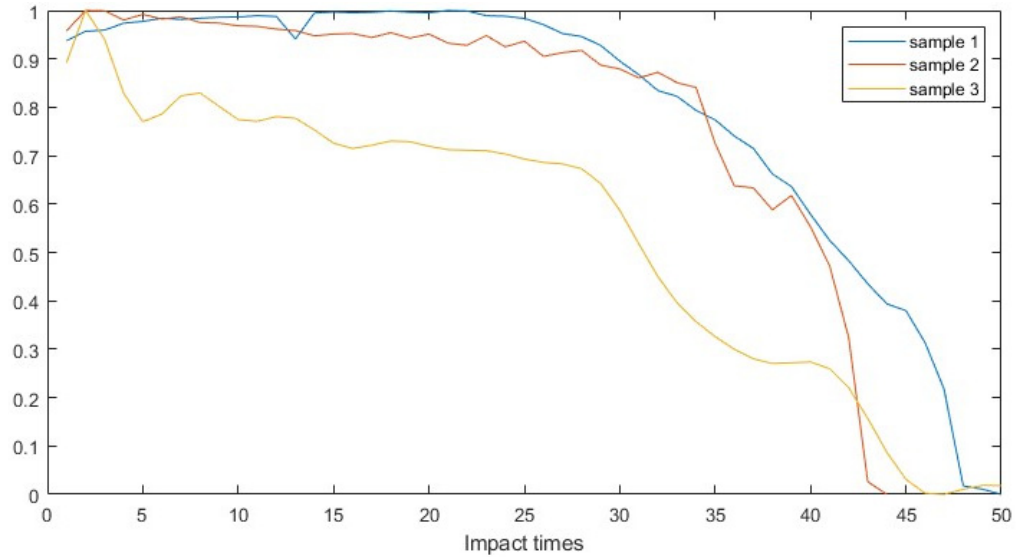


Figure 4. The variation in the nonlinear coefficients of three samples (the values were normalized).

2.2. BiLSTM Network

Long short-term memory network (LSTM) is a special recurrent neural network (RNN) which was proposed by Hochreiter and Schmidhuber in 1997 [22]. It solves the problem of long-term dependence on information in time series learning through network design and is widely used in engineering practices.

The structure of the network in the single-layer LSTM is shown in Figure 5.  $x_t$  is the input at the  $t$  training, and  $y_t$  is the output at the  $t$  training. The time-series information is transmitted between each training through two transmission states: cell state  $C$  and hidden state  $h$ . For example, during the  $t$  training, the states  $C_{t-1}$  and  $h_{t-1}$  obtained from the  $t-1$  training will participate in the training, and the new states  $C_t$  and  $h_t$  will be calculated to participate in the  $t + 1$  training.

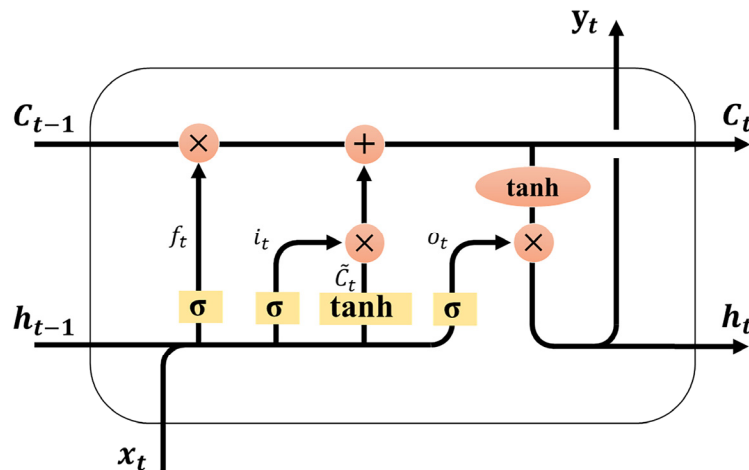


Figure 5. Schematic diagram of LSTM layer structure.

The LSTM network contains three gate structures and four intermediate state parameters  $f_t, \tilde{C}_t$  and  $o_t$ .

- Forgotten gate

This structure is responsible for controlling the influence of historical information on the current training. Its state parameter  $f_t$  represents the degree of forgetting the cell state information at the previous time. When it is equal to 1, it means that the cell state information at the previous time is completely retained, and when it is equal to 0, it means that the cell state information at the previous time is completely forgotten.

$$f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f), \tag{4}$$

where  $\sigma$  is sigmoid activation function.

- Input gate

This structure is responsible for controlling the new information in the cell state.

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i), \tag{5}$$

$$\tilde{C}_t = \tanh(W_C \times [h_{t-1}, x_t] + b_c), \tag{6}$$

$$C_t = f_t \times C_{t-1} + i_t * \tilde{C}_t, \tag{7}$$

- Output gate

This structure is responsible for controlling the output information in the cell state.

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o), \tag{8}$$

$$y_t = o_t \times \tanh(C_t), \tag{9}$$

The weight matrix  $W$  and the bias matrix  $b$  are obtained through the supervised learning of the training data and updating with the loss function as a metric.

Bi-directional long short-term memory (BiLSTM) is a complex LSTM model proposed by Mike Schuster and Kuldip K. Paliwal [23]. A single BiLSTM layer is composed of two LSTM layers, one forward processing sequence and one reverse processing sequence, as shown in Figure 6. Together, these two LSTM outputs form the BiLSTM output. Compared with the LSTM structure, the BiLSTM structure increases the transmission of reverse information and enhances the calculation accuracy by paying attention to the relationship between the before and after moments and the present moment at the same time.

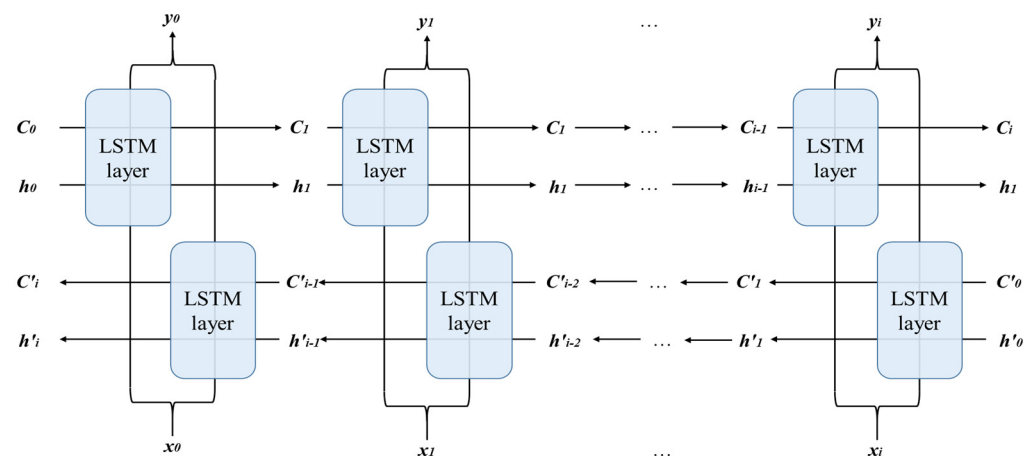
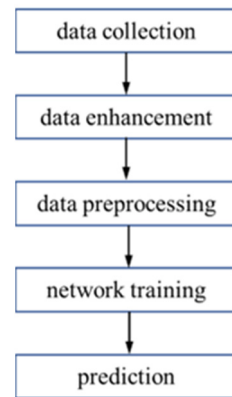


Figure 6. Structure of BiLSTM.

### 2.3. Life Prediction Model

A flow chart for the proposed life prediction model is shown in Figure 7. The model consists of five steps: data collection, data enhancement, data preprocessing, network training and prediction.



**Figure 7.** Flow chart for the proposed life prediction.

- Data collection

A certain current is used to impact the varistor at a certain time interval to simulate its actual working conditions. A varistor undergoing several impacts from full health until complete damage is recorded as one experiment. In the experiment, the nonlinear coefficient is calculated directly by the experimental equipment. The nonlinear coefficient of the varistor in one experiment was recorded as one set of data A.

- Data enhancement

The nonlinear coefficient data obtained from the surge impact test provide the original data for establishing a life prediction model for a varistor. However, due to the long test cycle and high cost, there is the problem that the amount of data available is small, which may affect the prediction effect of the model. In order to further improve the accuracy and generalization ability of the network, this paper uses data enhancement to improve the sample quality. It should be considered that the experimental data are discrete and have time sequence regularity. Common methods, such as flipping, clipping and interpolation, would break the regularity. Finally, add-noise is selected for data enhancement.

Gaussian white noise with SNR 50 is added to the original data to enhance the data in a ratio of 1:1. Some of the generated data is shown in Figure 8.

- Data preprocessing

In the degradation process of different varistors in the impact test, the variation curves of the nonlinear coefficients are bound to be different. In order to make the established life model have better generalization and migration, the nonlinear coefficients are normalized by 0–1 to be the life parameters  $\tilde{A}$ , i.e.

$$\tilde{A} = (A - A_{\min}) / (A_{\max} - A_{\min}), \quad (10)$$

where  $A_{\max}$  is the maximum value in this set of data, namely, the nonlinear coefficient in full health state;  $A_{\min}$  is the minimum value in this set of data, that is, the nonlinear coefficient after complete damage.

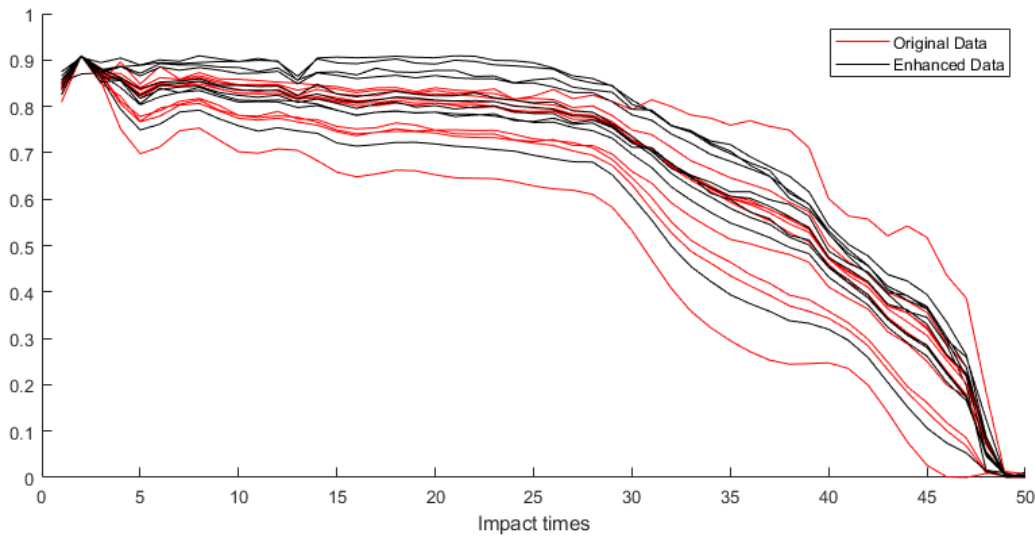


Figure 8. The data are enhanced by adding white Gaussian noise.

The data set  $\tilde{A}$  of sample length  $N$  is divided and reconstructed, and its continuous  $n$  data  $\tilde{A}_i, \tilde{A}_{i+1}, \dots, \tilde{A}_{i+n-1}$  are taken as a sample,  $\tilde{A}_{i+n}$ , to serve as the supervision label of the sample as shown in Figure 9, i.e.

$$x_i = [A_i, A_{i+1}, \dots, A_{i+n-1}]^T, \tag{11}$$

$$X = [x_1, x_2, \dots, x_{N-n}]^T, \tag{12}$$

$$Y = [y_1, y_2, \dots, y_{N-n}]^T = [\tilde{A}_{n+1}, \tilde{A}_{n+2}, \dots, \tilde{A}_N]^T, \tag{13}$$

where  $X$  is the sample set and  $Y$  is the sample label set.  $X$  and  $Y$  constitute a data set, and each experiment generates a data set.

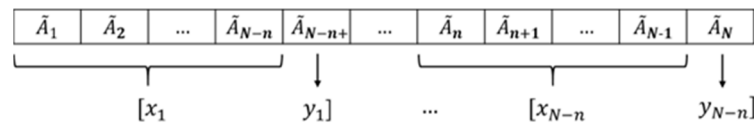


Figure 9. Divide and reconstruct  $\tilde{A}$ .

- Network training and prediction

The network includes several BiLSTM layers, a full connection layer and a regression layer. The overall structure is shown in Figure 10. The normalized and divided data set is taken as the input of the network. The output of the network is the value predicted by the lifetime model.

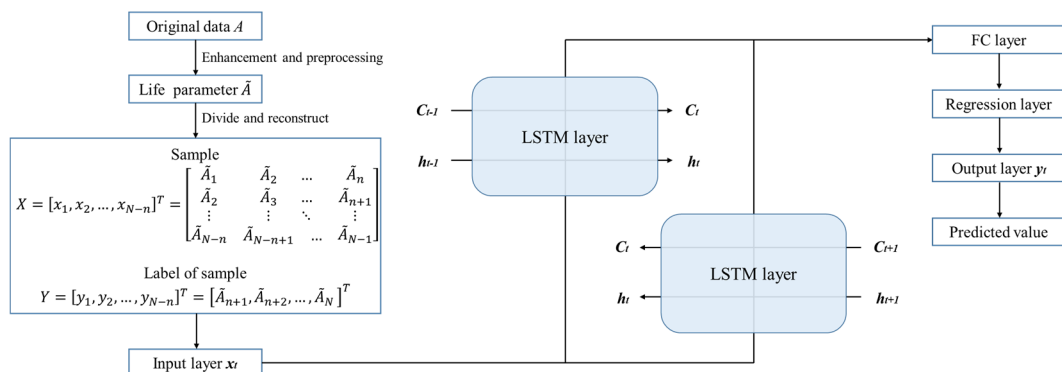


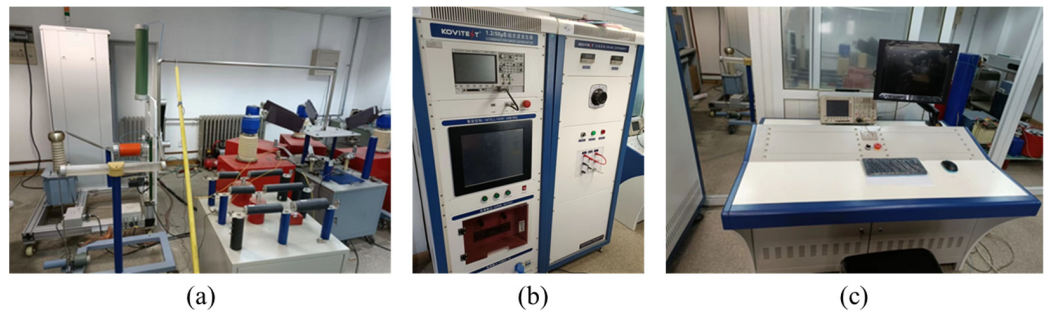
Figure 10. Structure of life prediction model.



### 3. Experiment and Analysis

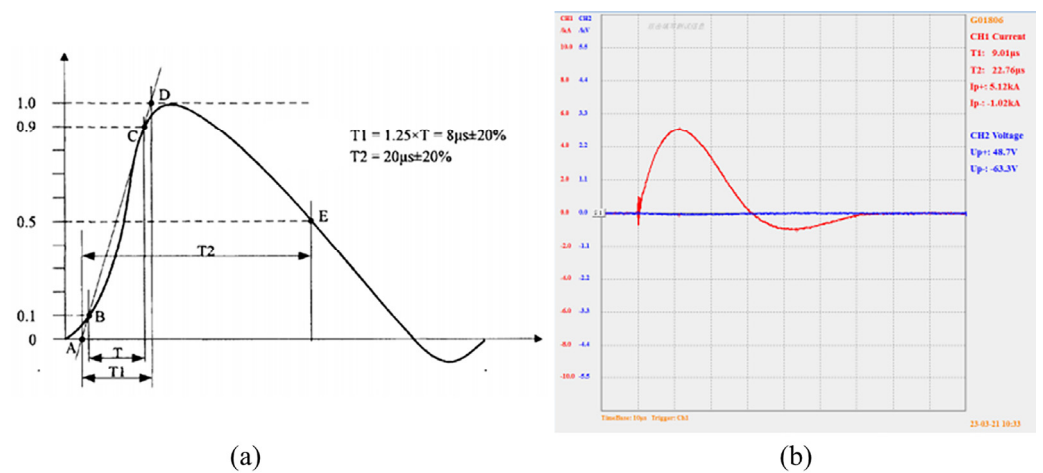
#### 3.1. Experimental Setup and Data Processing

The experimental sample used is a certain brand of LKD 22S181K R03 model ZnO varistor, with a breakdown varistor voltage of  $U_{1mA} = 180\text{ V}$ , a nominal discharge current of  $I_n = 12\text{ KA}$  and a maximum discharge current of  $I_{Max} = 25\text{ KA}$ . We selected these types of varistors from the same batch for the surge impact test. The experimental equipment is shown in Figure 11, including a test bench, a surge generator and a data acquisition terminal.



**Figure 11.** The experimental equipment. (a) Test bench, (b) surge generator, (c) data acquisition terminal.

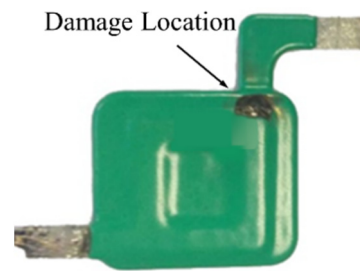
The test sample is impacted with a current of 8/20  $\mu\text{s}$  waveform, which is stipulated by national standards and railway industry standards, The theoretical impact waveform and the actual impact waveform of the experimental device are shown in Figure 12.



**Figure 12.** Diagram of 8/20  $\mu\text{s}$  impact waveform. (a) Theoretical impact waveform; (b) actual impact waveform.

The interval between each surge impact is 5 min, and its state parameters are recorded until it is completely damaged. Complete damage often takes more than 40 impacts. After complete damage, the appearance clearly changes, one corner of the epoxy resin layer bursts and granular black powder appears, as shown in Figure 13.

After each impact test, the value of the nonlinear coefficient is read and recorded from the experimental terminal. The nonlinear coefficients of each varistor from initial to complete failure were recorded as a data set. The data set is processed as described in Section 2.3 and becomes the input to the lifetime prediction model.



**Figure 13.** Obvious damage of varistor.

### 3.2. Prediction Performance of Life Model under Different Hyper-Parameters

In the life prediction model proposed in this paper, in addition to the network training parameters, there are still a large number of hyper-parameters, among which the training sample length  $n$  and the number of BiLSTM network layers have the greatest impact on the model. In order to verify the influence of different hyper-parameters on the life model, the root mean square error (RMSE) was originally used to evaluate the prediction accuracy of the model, namely:

$$\text{RMES} = \sqrt{(\sum_1^{N-n} (y_i - \hat{y})^2) / (N - n)}, \quad (14)$$

where  $\hat{y}$  is the value predicted by the model.

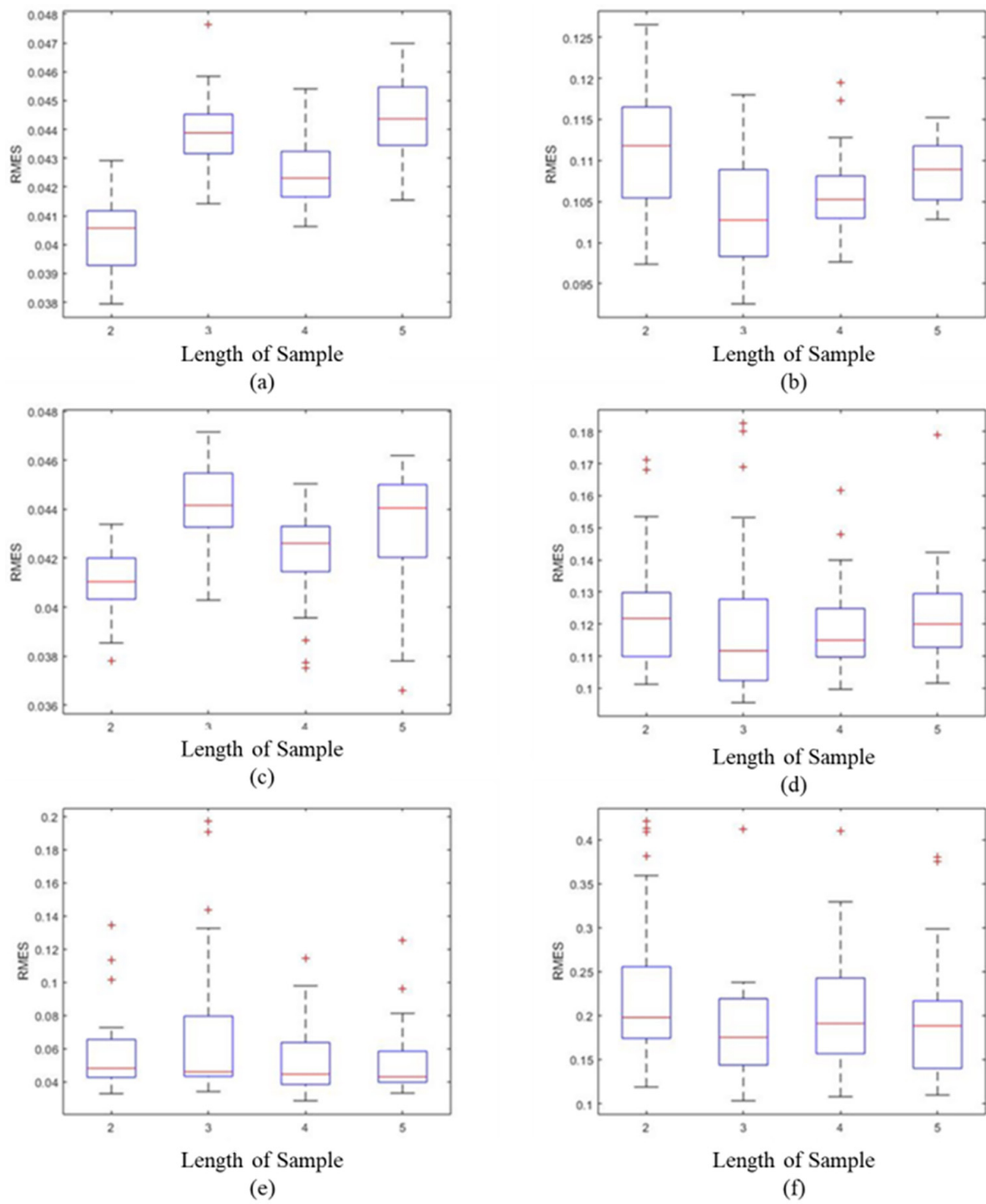
The data set is divided into the training set and the test set according to the ratio of 80% to 20%. The training set data are used to train the life model under different preset hyperparameters by cyclic iteration method. Test set data include test set 1 and test set 2. The data from test set 1 and the training set are collected from the same batch of varistors of the same model, while the data from test set 2 and the training set are from different batches.

Figure 14 shows statistics for the prediction errors in the life curves of test set 1 and test set 2 generated by the models under different hyper-parameters after 100 independent trainings, in which the prediction accuracy of test set 2 reflects the generalization performance of the model.

From the statistics for 100 independent test errors under each hyper-parametric condition, it can be seen that the model has good stability and reliability. For test set 1, when the network contains a 1-layer BiLSTM layer and a 2-layer BiLSTM layer, the performance is close, but the performance decreases when there are three layers. For test set 2, layer one performance is the best. In the case of the same number of layers, when the training sample length is two, it performs better on test set 1, but it performs worse on test set 2. When the sample length is three and four, the model has a relatively balanced performance on the two data sets, and when the sample length is four, the performance is more stable. If the sample length continues to increase, the accuracy of the model will decline. Clearly, the multi-layer BiLSTM layer in this model cannot further increase the accuracy of prediction and will definitely occupy more computing resources. An appropriate increase in the sample length will help to enhance the stability of the model, but if it is too long, it will lead to an excessively high degree of model fitting and a lack of generalization and mobility.

### 3.3. Comparison between Predicted Curve and Observed Curve

Finally, the number of BiLSTM layers in the network is one, and the sample length is four. Other hyper-parameters are selected according to the empirical formula. If we input the limited length experimental data into the built varistor life model, the model converges quickly, as shown in Figure 15. The predicted life curve is obtained and compared with the observed life curve, as shown in Figure 16. The results show that the prediction trend from the model is smooth, and the predicted value at the key node where the life parameter decreases is consistent with the observed value. The local error mainly occurs at the node after the varistor is damaged, and the overall error meets the expectation. The output of this model can predict the lifetime of varistors.



**Figure 14.** Model prediction error under different hyper-parameters. (a) 1 BiLSTM layer + test set 1. (b) 1 BiLSTM layer + test set 2. (c) 2 BiLSTM layer + test set 1. (d) 2 BiLSTM layer + test set 2. (e) 3 BiLSTM layer + test set 1. (f) 3 BiLSTM layer + test set 2.

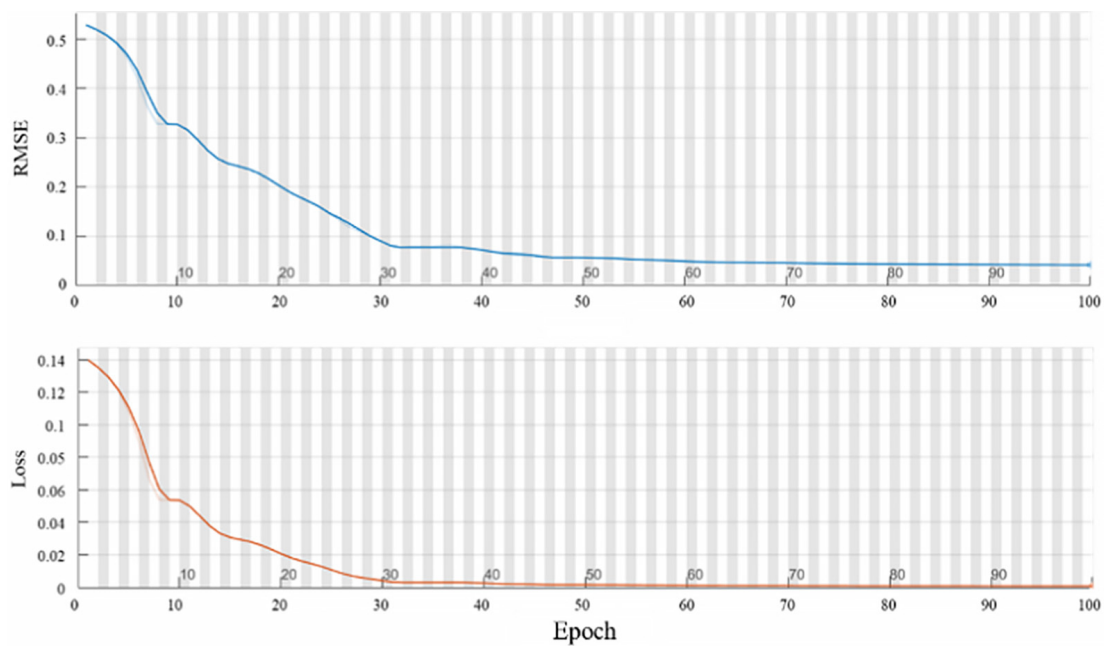


Figure 15. Model training.

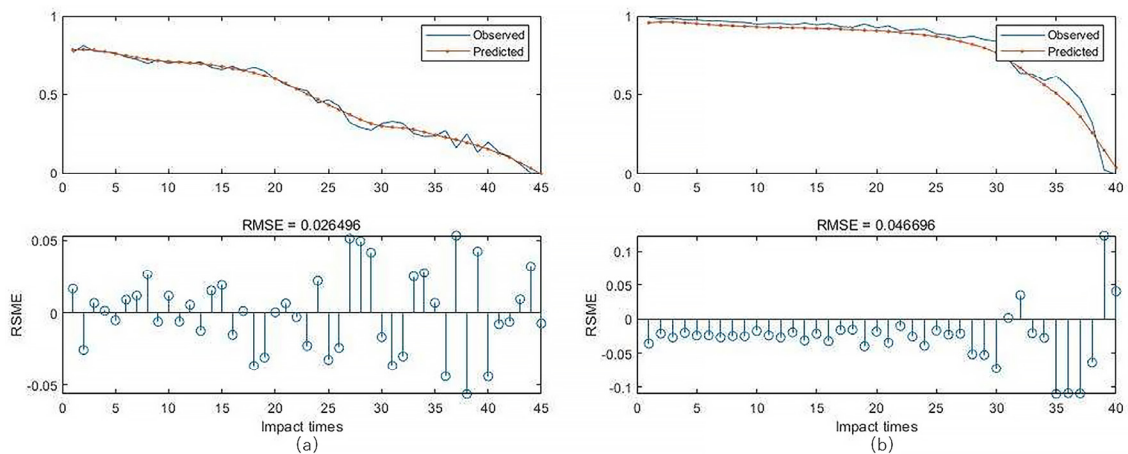


Figure 16. Compare the predicted life curve with the observed life curve. (a) On test set 1. (b) On test set 2.

### 3.4. Performance Comparison of Different Structures

In addition to the LSTM structure and BiLSTM structure, this paper also considers a lightweight LSTM structure called GRU for the characteristics of small-scale data.

The Gate Recurrent Unit (GRU) is a simplified LSTM model. Compared with LSTM, GRU uses an update gate instead of a forgetting gate and an input gate, thus reducing the number of parameters by about one third. In addition, GRU omits the intermediate parameter  $\tilde{C}$  in LSTM, which ignores the long-term memory and focuses more on the association between the previous moment and this moment. And the GRU structure cancels the second-order nonlinear function during output. Therefore, GRU tends to have a higher computational efficiency and a stronger generalization in the face of data with small samples and simple structures.

In this paper, three different structures are tested on the original data and enhanced data, and the test results are shown in Table 1. In order to more clearly compare the computational efficiency of different algorithms, the training was set up with 100 epochs. In practice, five to ten epochs can already achieve very good results which require less than two seconds.

**Table 1.** Test results of three structures.

Structure	Training Time of 1000 Epochs (s)	RMSE on Test Set 1	RMSE on Test Set 2
LSTM + original data	18	0.040	0.101
GRU + original data	17	0.038	0.074
BiLSTM + original data	19	0.026	0.045
LSTM + enhanced data	18	0.036	0.086
GRU + enhanced data	17	0.035	0.071
BiLSTM + enhanced data	20	0.023	0.039

The experimental results show that the training time required by the three structures is close to each other. However, compared with other structures, the BiLSTM structure has a higher prediction accuracy on both test set 1 and test set 2.

#### 4. Conclusions

In this paper, a life prediction method based on BiLSTM is proposed to study the cumulative effect of the degradation process caused by surge impact on varistors. Compared with the existing life prediction methods, this method uses nonlinear coefficients to construct life parameters, which can better characterize the degradation process of varistors under surge impact, and uses the BiLSTM network to build a life model which can effectively use historical data to judge the life of varistors. Experiments show that this method has a high prediction accuracy, simple implementation structure, short network training time and uses less computational resources. At the same time, the model has good generalization and mobility, and has application prospects in predicting the life of varistors. At the same time, this method may be applicable to other small-scale time series forecasting scenarios.

Subsequent studies will further explore changes in the lifetime of varistors under varying surges and build more accurate life models with the help of more complex networks.

**Author Contributions:** Methodology, X.R.; Data curation, X.R.; Writing—original draft, X.R.; Writing—review & editing, Y.B. and W.W.; Supervision, W.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study is supported by the Beijing Municipal Natural Science Foundation (Grant No. 3192025).

**Data Availability Statement:** Restrictions apply to the availability of these data. Data was obtained from China Academy of Railway Sciences Communication Signal Research Institute and are available from the authors with the permission of China Academy of Railway Sciences Communication Signal Research Institute.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Su, Y.; Sun, L.; Yang, Z.; Fan, K. Research progress of low voltage varistors. *Electron. Compon. Mater.* **2010**, *29*, 74–78.
2. Wang, C.; Lei, W.; Lin, H.; Li, C. Design and implementation of real-time monitoring system for surge protection devices in smart grid. In Proceedings of the 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Xi'an, China, 25–28 October 2016.
3. Zhong, D.; Gao, Q.; Zhang, J. Development of on-line monitoring device for UHV AC metal oxide arrester. *Electr. Porcelain Arrester* **2018**, *04*, 120–123+127.
4. Ruan, X.; Jin, S.; Wen, W.; Cheng, W. The study on life model of MOV based on various parameters and surge history. *Soft Comput.* **2022**, *26*, 7595–7600. [[CrossRef](#)]
5. Liu, T. *Design and Research of Surge Protector Online Monitoring System*; Qingdao University of Technology: Qingdao, China, 2015.
6. Jin, S. *Research on Residual Life Model of Varistors in Railway Surge Protective Devices*; China Academy of Railway Sciences: Xi'an, China, 2019.
7. Yang, Z.; Chen, L. Analysis and application of electric capacity change during deterioration of zinc oxide varistors. *High Volt. Technol.* **2010**, *36*, 2167–2172.

8. Yan, X.; Wen, Y.; Yi, X. Study on the resistive leakage current characteristic of MOV surge arresters. *IEEE PES Transm. Distrib. Conf. Exhib.* **2002**, *2*, 683–687.
9. Zhang, L. *Preparation of ZnO Nanomaterials and Study on their Microwave Dielectric Properties*; Northwestern Polytechnical University: Xi'an, China, 2007.
10. Yin, G.L.; Li, J.Y.; Yao, G.; Cheng, P.F.; Li, S.T. Study on the electron trap process of impact aging ZnO pressure sensitive ceramics. *Acta Phys. Sin.* **2010**, *59*, 6345–6350.
11. Yang, Z.; Zhang, C.; Chai, J.; Li, X.; Ru, H. Study on nonlinear coefficient change of zinc oxide varistors during aging. *Electron. Compon. Mater.* **2011**, *30*, 27–30. [[CrossRef](#)]
12. He, J.; Liu, J.; Hu, J.; Long, W. Progress of ZnO varistors for lightning arresters in power systems. *High Volt. Technol.* **2011**, *37*, 10.
13. Nahm, C.W. Effect of Sintering Temperature on Varistor Properties and Aging Characteristics of ZnO-V<sub>2</sub>O<sub>5</sub>-MnO<sub>2</sub> Ceramics. *Ceram. Int.* **2009**, *35*, 2679–2685. [[CrossRef](#)]
14. Tabasco-Novelo, C.; Cervantes-López, J.L.; González-Panzo, I.J.; Rodríguez-Gattorno, G.; Quintana, P. High non-linear electrical properties of Li<sub>3x</sub>Co<sub>7-4x</sub>Sb<sub>2+x</sub>O<sub>12</sub> a new ceramic varistor. *J. Alloys Compd.* **2021**, *878*, 160356. [[CrossRef](#)]
15. García, L.F.; Azough, F.; Parsons, G.; Quadling, A.; Wang, B.; Freer, R. Optimising the performance of SiC-based varistors through composition and microstructure control. *J. Eur. Ceram. Soc.* **2022**, *42*, 600–607. [[CrossRef](#)]
16. Begum, S.; Karim, A.N.M.; Fawzia, S.; Hashmi, M.S.J. Industrial Fabrication of ZnO Varistor: Leveraging the Powder Processing Parameters. In *Encyclopedia of Materials: Electronics*; Haseeb, A.S.M.A., Ed.; Elsevier: Amsterdam, The Netherlands, 2021; Volume 3.
17. Tian, T.; Zheng, L.; Podlogar, M.; Man, Z.; Ruan, X.; Shi, X.; Bernik, S.; Li, G. Influence of Ca-doping on the nonlinear properties of novel ZnO-Cr<sub>2</sub>O<sub>3</sub>-based varistor ceramics. *J. Eur. Ceram. Soc.* **2022**, *42*, 2268–2273. [[CrossRef](#)]
18. Li, Z.; Ren, X.; Wang, X.; You, W.; Shi, L. Effectively enhanced comprehensive electrical performance of ZnO varistors by a fast combinatorial refinement method. *Mater. Sci. Semicond. Process.* **2021**, *133*, 105945. [[CrossRef](#)]
19. Chen, T.; Xu, R.F.; He, Y.L.; Wang, X. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Syst. Appl.* **2017**, *72*, 221–230. [[CrossRef](#)]
20. Xu, G.; Meng, Y.; Qiu, X.; Yu, Z.; Wu, X. Sentiment Analysis of Comment Texts Based on BiLSTM. *IEEE Access* **2019**, *7*, 51522–51532. [[CrossRef](#)]
21. Luo, L.; Yang, Z.; Yang, P.; Zhang, Y.; Wang, L.; Lin, H.; Wang, J. An attention-based BiLSTM-CRF approach to document-level chemical named entity recognition. *Bioinformatics* **2018**, *34*, 1381–1388. [[CrossRef](#)] [[PubMed](#)]
22. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
23. Schuster, M.; Paliwal, K.K. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **1997**, *45*, 2673–2681. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.