

Research on Fast Identification Method of Tunnel Cable State Based on Sparse Online Hybrid Gaussian Multi-Classification Algorithm

Tian Guo^{1, a}, Yang Zhao^{1, b}, Zongwu Huang^{2, c} and Jirong Fu^{2, d 1}

¹State Grid Beijing Electrical Power Company, Beijing 100031, China;

²School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China.

ABSTRACT.

In this paper, a fast tunnel cable state identification method based on a sparse online hybrid Gaussian multi-classification algorithm is proposed for real-time monitoring of the state of tunnel cables and their changes. The method first reduces the size of the training set using the data subset approximation method, then uses the hybrid Gaussian multi-classification algorithm to train the tunnel cable state fast identification model, and finally performs the parameter update online when necessary. The proposed method takes less time in the model training phase, and the trained tunnel cable state fast identification model can maintain good recognition accuracy under complex operating conditions and can be updated quickly.

Keywords: Cable condition recognition; Mixed Gaussian process; Multi-classification; Online model update; Sparse technology.

1. INTRODUCTION

With the development of urbanization, cables are more and more widely used in power systems. Compared with primary equipment such as transformers, switches, and circuit breakers, which are more centralized in the power grid, cables have a large number, long mileage, and wide distribution, and it is not easy to realize comprehensive monitoring and exact perception of their status. Moreover, the operation and maintenance management of cables has not yet formed a system, especially for 10 kV distribution cables, which are mostly randomly placed in cable trenches, which are often full of water and have poor air circulation [1]. Such a harsh operating environment is likely to cause damage to the cables, accelerating the aging of the cable insulation and leading to failures.

Therefore, it is necessary to accurately diagnose faults on faulty cables in order to formulate appropriate maintenance and disposal plans, especially for early faults, which need to be paid attention to and handled before they develop into permanent faults, so as to prevent the deterioration of cable insulation from causing serious faults [2]. For normal cables in service, it is necessary to accurately assess the state of their insulation and grasp their current state of health in order to provide a reasonable maintenance program [3]. Accurate diagnosis of cable faults and reliable assessment of cable insulation status, and comprehensive monitoring of cable status are of great significance for the timely detection of potential cable faults [4], slowing down the occurrence of permanent faults, reducing the risk of grid operation, and ensuring the reliability of power supply in the power system.

Along with the development of artificial intelligence technology in recent years, some scholars have begun to apply data-driven ideas to the research of cable fault state diagnosis, and have also achieved certain results [5]. Southwest Jiaotong University realized the identification of early cable faults based on the probabilistic neural network [6] and support vector machine [7] respectively; Wang and Lu applied a restricted Boltzmann machine and stacked autoencoder to classify cable faults in distribution networks [8]; Professor Shin Yong-jun of Yonsei University in South Korea took fault waveforms and their reflection waveforms as the research object, and extracted the three features of fault waveforms, namely, the time delay, peak voltage, and time-frequency phase difference, as the inputs of generalized regression neural network as the input of the generalized regression neural network, so as to diagnose a variety of cable fault states [9]. From this type of research, it is not difficult to find that in the field of cable fault identification, artificial intelligence has a greater potential for development [10].

¹ ^a guotian124com@163.com, ^b 15101691629 @163.com, ^c 2017302290070@whu.edu.cn, ^d 2020302191016@whu.edu.cn

Compared with the traditional electrical threshold method and other state diagnosis methods, the data-driven cable fault state method is more adaptable and robust [11]. However, the current data-driven cable fault state diagnosis method still has shortcomings such as less research on cable weak fault diagnosis [12] and lack of generalization ability of the cable fault state diagnosis model [13]. Therefore, in this paper, the identification of weak cable faults is incorporated into the diagnostic model to achieve more comprehensive cable fault state identification; the data subset approximation method is used to reduce the size of the training set and improve the model training speed. At the same time, a hybrid Gaussian multi-classification algorithm is used to train the tunnel cable state fast identification model, and the correlation between cable states is deeply explored. Finally, an online updating mechanism of the model is proposed to improve the generalization performance of the cable fault state diagnosis model.

2. ACQUISITION OF MODEL TRAINING DATA

By analyzing the mechanism of weak and permanent faults of cables, suitable models are selected to simulate weak and permanent faults of cables. Use MATLAB or PSCAD/EMTDC to establish the cable model [14] and build the cable weak fault [15], permanent fault and overcurrent [16], and overvoltage perturbation model [17]. By simulating different cable fault conditions, different grounding resistances and forms of perturbation, the fault current, grounding current, partial discharge and other data under different conditions are obtained, and at the same time, the cable network operation temperature is obtained through the cable modem and the cable trench modeling simulation, and the monitoring quantities of the fault current, grounding current, partial discharge, and the cable network operation temperature are used as the input variable x , and the cable state is used as the output variable m , and the value of m is taken as follows:

$$\left\{ \begin{array}{l} 0 \leq m < 1 \quad (\text{normal}) \\ 1 \leq m < 2 \quad (\text{Early Failure 1}) \\ 2 \leq m < 3 \quad (\text{Early Failure 2}) \\ \dots \\ n_1 \leq m < n_2 \quad (\text{Permanent fault 1}) \\ n_2 \leq m < n_3 \quad (\text{Permanent fault 1}) \\ \dots \end{array} \right. \quad (1)$$

At the same time, cable factory experimental data, online monitoring data, and power failure detection data can be added to the database. The data under each operating condition and fault condition form the original data set. K samples are uniformly selected from the original dataset as the new dataset by the data subset approximation method.

3. FAST IDENTIFICATION OF TUNNEL CABLE STATUS BASED ON SPARSE ONLINE HYBRID GAUSSIAN MULTI-CLASSIFICATION ALGORITHM

3.1 Gaussian mixture modeling

The input variable x and the output variable m are combined into a new vector of variables $y = [x^T m^T]$, then the training set is $Y = [y_1, y_2, \dots, y_K]$. A Gaussian mixture model of the joint probability density of the inputs and outputs is built [18]:

$$p(y) = \sum_{j=1}^M p(j)p(y/j) \quad (2)$$

where M is the number of unknown Gaussian components, j represents the j th Gaussian component, $p(y)$ represents the probability density function of y , and $p(j)$ represents the probability density function of j , which is calculated as follows:

$$p(j) = \gamma_j f(j|\delta_j) \quad (3)$$

where γ_j is the probability weight of the j th Gaussian component, $\delta_j = \{\eta_j, \Sigma_j\}$, $f(j|\delta_j)$ is calculated as follows:

$$f(j|\delta_j) = \frac{1}{\sqrt{(2\pi)^e |\Sigma_j|}} \exp\{-1/2(j - \eta_j)^T \Sigma_j^{-1} (j - \eta_j)\} \quad (4)$$

where e is the dimension of y , where η_j is the mean vector of the j th Gaussian component, and Σ_j is the covariance matrix of the j th Gaussian component. $p(y/j)$ represents the conditional probability density function of y with respect to j :

$$p(y/j) = \frac{p(y)}{p(j)} \quad (5)$$

3.2 Acquisition of model parameters

It is first necessary to determine the range of M , i.e., to determine the maximum value, M_{\max} , and the minimum value, M_{\min} , of M . The absolute increasing log-likelihood criterion (AIL) is chosen to select the optimal Gaussian component [19]. The AIL criterion is defined as

$$\text{AIL}(\hat{\alpha}_{(M)}, M) = \log F(Y, \hat{\alpha}_{(M)}) - \log F(Y, \hat{\alpha}_{(M-1)}) \quad (6)$$

where $\log F(Y, \hat{\alpha}_{(M)})$ and $\log F(Y, \hat{\alpha}_{(M-1)})$ denote the maximum log-likelihood function of a Gaussian mixture model with M and $M-1$ Gaussian components, respectively, and the maximum log-likelihood function of a Gaussian mixture model with M Gaussian components is computed as follows:

$$\log F(Y, \hat{\alpha}) = \sum_{n=1}^K \log(\sum_{j=1}^M \gamma_j f(y_n | \delta_j)) \quad (7)$$

The maximum log-likelihood function for a Gaussian mixture model with $M-1$ Gaussian components is similar. where K is the number of samples and $\hat{\alpha} = \{\{\gamma_1, \eta_1, \Sigma_1\}, \{\gamma_2, \eta_2, \Sigma_2\}, \dots, \{\gamma_M, \eta_M, \Sigma_M\}\}$ is the total parameters in the complete Gaussian mixture model with M components [20].

Let M be equal to M_{\max} and use the K-means algorithm to initialize the parameters of the Gaussian mixture model with M Gaussian components, denoted as $\hat{\alpha}_M^{(1)}$.

3.2.1 EM algorithm [21] for obtaining estimated model parameters $\hat{\alpha}_M$

E step, calculate the posterior probability of the i th training sample for the j th Gaussian component at the k th iteration:

$$\varphi^{(k)}(y_i, \alpha^{(k)}) = \frac{\gamma_j^{(k)} f(y_i | \eta_j^{(k)}, \Sigma_j^{(k)})}{\sum_{l=1}^M \gamma_l^{(k)} f(y_i | \eta_l^{(k)}, \Sigma_l^{(k)})} \quad (8)$$

M step, the probability weight $\gamma_j^{(k+1)}$, mean $\eta_j^{(k+1)}$ and covariance $\Sigma_j^{(k+1)}$ of the j th Gaussian component at the $(k+1)$ st iteration are computed:

$$\gamma_j^{(k+1)} = \frac{\sum_{i=1}^K \varphi^{(k)}(y_i, \alpha^{(k)})}{K} \quad (9)$$

$$\eta_j^{(k+1)} = \frac{\sum_{i=1}^K \varphi^{(k)}(y_i, \alpha^{(k)}) y_i}{\sum_{i=1}^K \varphi^{(k)}(y_i, \alpha^{(k)})} \quad (10)$$

$$\Sigma_j^{(k+1)} = \frac{\sum_{i=1}^K \varphi^{(k)}(y_i, \alpha^{(k)}) (y_i - \eta_j^{(k+1)})(y_i - \eta_j^{(k+1)})^T}{\sum_{i=1}^K \varphi^{(k)}(y_i, \alpha^{(k)})} \quad (11)$$

Let $\alpha^{(k+1)} = \{\{\gamma_1^{k+1}, \eta_1^{k+1}, \Sigma_1^{k+1}\}, \{\gamma_2^{k+1}, \eta_2^{k+1}, \Sigma_2^{k+1}\}, \dots, \{\gamma_M^{k+1}, \eta_M^{k+1}, \Sigma_M^{k+1}\}\}$, denote the $(k+1)$ th total model parameter. The total parameters obtained from each iteration are combined to obtain the sequence $\{\alpha\}$, where each element is α . Maximize the log-likelihood function to find the model parameters $\hat{\alpha}_M$:

$$\hat{\alpha}_M = \arg \max_{\alpha} (\log F(Y, \alpha)) \quad (12)$$

3.2.2 Calculate the corresponding absolute increasing log-likelihood criterion $\text{AIL}(\hat{\alpha}_{(M)}, M)$

Remove the least probable Gaussian component and merge it to the closest Gaussian component. Select the least likely component by finding the component with the smallest mixing probability:

$$r = \arg \min_j \{\gamma_j, j = 1, 2, \dots, M\} \quad (13)$$

Then the component s closest to the r th component is chosen.

$$s = \arg \min_i \{B_s[f(y|\eta_r, \Sigma_r), f(y|\eta_i, \Sigma_i)], i = 1, 2, \dots, M, i \neq r\} \quad (14)$$

where the B_s function is the symmetric KL scatter between the r th Gaussian component and the other components, and the symmetric KL scatter is a common measure of the difference between probability density functions:

$$B_s[f(y|\eta_r, \Sigma_r), f(y|\eta_i, \Sigma_i)] = \frac{1}{2} \text{tr}[(\Sigma_r - \Sigma_i)(\Sigma_i^{-1} - \Sigma_r^{-1})] + \frac{1}{2}(\eta_r - \eta_i)^T(\Sigma_r^{-1} + \Sigma_i^{-1})(\eta_r - \eta_i), i = 1, 2, \dots, M, i \neq r \quad (15)$$

The r th and s th Gaussian components are combined into a single component whose mixed probability, mean, and covariance are updated as

$$\gamma' = \gamma_r + \gamma_s \quad (16)$$

$$\eta' = \frac{\gamma_r \eta_r + \gamma_s \eta_s}{\gamma_r + \gamma_s} \quad (17)$$

$$\Sigma' = \frac{\gamma_r(\Sigma_r + \eta_r \eta_r^T) + \gamma_s(\Sigma_s + \eta_s \eta_s^T)}{\gamma_r + \gamma_s} - \eta' \eta'^T \quad (18)$$

Obtain a Gaussian mixture model with $M-1$ Gaussian components with initial parameters $\hat{\alpha}_{(M-1)}^{(1)}$.

Let $M = M-1$ and repeat EM algorithm flow if $M \geq M_{\min}$, otherwise proceed to the next step.

Obtain the optimal M_{opt} by maximizing the AIL.

$$M_{\text{opt}} = \arg \max_M \{ \text{AIL}(\hat{\alpha}_{(M)}, M), M = M_{\min}, M_{\min} + 1, \dots, M_{\max} \} \quad (19)$$

The final parameter estimates of the model $\hat{\alpha}_{K_{\text{opt}}}$ are obtained.

3.3 Computational model output

Assume that the new input data is x_{new} and the true value of the corresponding output is m_{new} . divide the mean vector and covariance matrix of each Gaussian component into input and output parts as follows:

$$\eta_{\text{new},j} = \begin{bmatrix} \eta_{\text{new},j}^I \\ \eta_{\text{new},j}^O \end{bmatrix}, \Sigma_{\text{new},j} = \begin{bmatrix} \Sigma_{\text{new},j}^{II} & \Sigma_{\text{new},j}^{IO} \\ \Sigma_{\text{new},j}^{OI} & \Sigma_{\text{new},j}^{OO} \end{bmatrix} \quad (20)$$

Estimate the posterior probability of each Gaussian component of the new input data:

$$p(j|x_{\text{new}}) = \frac{p(x_{\text{new}}|\delta_j)p(j)}{p(x_{\text{new}}|\delta_j)} \quad (21)$$

Compute the conditional distribution of the output of each Gaussian component of the new input data with respect to the input:

$$p(m_{\text{new},j}|x_{\text{new},j}) \sim f(m_{\text{new},j}|\hat{m}_{\text{new},j}, \hat{\Sigma}_{\text{new},j}) \quad (22)$$

where $\hat{m}_{\text{new},j}$ and $\hat{\Sigma}_{\text{new},j}$ are the mean and covariance parameters of the j th conditional Gaussian distribution, calculated as follows:

$$m_{\text{new},j} = \eta_{\text{new},j}^O + \Sigma_{\text{new},j}^{OI}(\Sigma_{\text{new},j}^{II})^{-1}(m_{\text{new}}^I - \eta_{\text{new},j}^I) \quad (23)$$

$$\hat{\Sigma}_{\text{new},j} = \Sigma_{\text{new},j}^{OO} - \Sigma_{\text{new},j}^{OI}(\Sigma_{\text{new},j}^{II})^{-1}\Sigma_{\text{new},j}^{IO} \quad (24)$$

The conditional distributions of the test set output with respect to the inputs are combined into a single Gaussian distribution. Finally, each component is weighted and combined to obtain the final prediction of the output:

$$\hat{m}_{\text{new}} = \sum_{j=1}^{M_{\text{opt}}} p(m_{\text{new},j}|x_{\text{new},j})\hat{m}_{\text{new},j} \quad (25)$$

Set the threshold function to

$$\hat{y} = \begin{cases} 0, & -\infty < \hat{m}_{new} < 1 \\ 1, & 1 \leq \hat{m}_{new} < 2 \\ 2, & 2 \leq \hat{m}_{new} < 3 \\ \dots & \dots \\ n_1, & n_1 \leq m < n_2 \\ n_2, & n_2 \leq m < n_3 \\ \dots & \dots \end{cases} \quad (26)$$

Tunnel cable equipment fault state diagnosis, online monitoring equipment, and inspection robots collected cable equipment running current, grounding current, partial discharge, cable network temperature, and other monitoring quantities as the input to the cable fault state diagnosis model, to get the model output value \hat{y} . When $\hat{y}=0$, the cable is in the normal state, when $\hat{y}=1$, the cable is in a weak fault 1 state, when $\hat{y}=2$, the cable is in a weak fault 2 state, when $\hat{y}=n_1$, the cable is in a permanent fault 1 state, when $\hat{y}=n_2$, the cable is in a permanent fault 2 state, and so on.

The model is updated online when new training points are added, and the process can be regarded as a process of adding new training points one by one. When the number of samples in the training set is increased to $K+1$, let the added samples be $y_{K+1}=(x+1, m_{K+1})$, then the mean and covariance parameter of the j th conditional Gaussian distribution are updated as:

$$\eta_j^{K+1} = (\Sigma_j^K)^{-1}\eta_j^K + \Sigma_j^K(\Sigma_j^K + \sigma_n^2)^{-1}y_{K+1} \quad (27)$$

$$\Sigma_j^{K+1} = \Sigma_j^K + \Sigma_j^K(\Sigma_j^K + \sigma_n^2)^{-1} \quad (28)$$

where σ_n^2 is Gaussian white noise with zero variance. The overall flow of the algorithm is shown in Fig. 1.

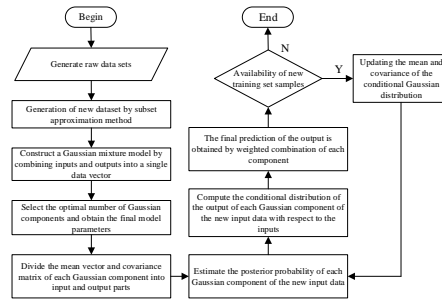


Figure. 1 General flow of the methodology

4. SIMULATION ANALYSIS

4.1 Evaluation indicators

The common metrics of the recognition model are accuracy, precision, recall, F1 comprehensive evaluation metrics, etc., among which accuracy is the most commonly used evaluation metric in fault recognition, which is calculated as shown below:

$$Ac = \frac{N_T}{N_T + N_F} \quad (29)$$

where N_T is the number of samples correctly predicted by the recognition model and N_F is the number of samples incorrectly predicted by the recognition model.

The accuracy can only describe the recognition results of the recognition model for all test samples, but cannot reflect the recognition performance of the model in specific categories. In order to comprehensively assess the performance of the recognition model, the F1 comprehensive evaluation index is introduced for comprehensive evaluation.

The accuracy is calculated as follows:

$$Pr_k = \frac{N_{Tk}}{N_{Tk} + N_{Fk}} \quad (30)$$

where N_{Tk} is the number of samples correctly predicted by the discriminative model in the k th category sample, and N_{Fk} is the number of samples incorrectly predicted by the discriminative model in the k th category sample.

The F1 composite evaluation index is calculated as follows:

$$F_1 = n \frac{\prod_{k=1}^n Pr_k}{\sum_{k=1}^n Pr_k} \quad (31)$$

The larger the value of the F1 comprehensive evaluation index, the stronger the comprehensive performance of the identification model.

4.2 Experimental setup

There are 8 cable operation states set up, which are coded as shown in Table 1. The data obtained through the cable and fault modeling simulation consists of the original dataset, which has 12000 samples, of which 2000 samples obtained by changing the states other than the eight cable operation states in the table are used as test set 2, 8000 samples are selected from the remaining 1000 samples as the training set, and the remaining 2000 samples are used as test set 1.

Table 1. Cable operation status code

Operational state	Encodings
Normal state	1
Capacitor switching	2
Constant impedance grounding	3
Early semi-circular wave failure	4
Multi-Cycle Early Failure	5
Load Shedding	6
Motor switching	7
Excitation surge	8

SVM, DT, RF, and KNN diagnostic models are used as a comparison of the proposed method. The simulation is carried out on MATLAB and the processor of the simulation device is 5800X with 32G RAM.

4.3 Simulation results and analysis

Five different methods of tunnel cable equipment fault state diagnosis models were trained using the training set, and the diagnostic results using test set 1 as the test set are shown in Fig. 2 and Table 2.

Actual state	1	363	4		1				
	2	3	232						
	3			245		1			
	4				240				
	5				3	221	2		
	6				1		274	2	
	7							252	
	8						1	5	154
		1	2	3	4	5	6	7	8

Projected State

Figure. 2 Confusion Matrix for Cable Condition Diagnostic Results

Table 2. Failure evaluation indicator values for each method

Diagnostic model	Training set accuracy	Test Set Accuracy	F1 Composite indicator
SVM	91.69%	90.35%	89.83%
DT	90.25%	89.50%	88.21%
RF	96.48%	95.75%	94.54%
KNN	98.43%	97.50%	97.06%
Proposed	99.26%	99.05%	98.47%

In order to evaluate the generalization ability of the tunnel cable equipment fault state diagnosis models trained by each method, test set 2 is used as the test set. The diagnostic results are shown in Fig. 3 and Table 3.

	1	2	3	4	5	6	7	8
Actual state	1	359	6	2	1			
	2	4	227		3		1	
	3		3	240		2	1	
	4				240			
	5				3	221	2	
	6			1	1		272	2
	7				2		1	249
	8				1		3	6
		1	2	3	1	6	150	
		Projected State						

Figure. 3 Confusion Matrix for Cable Condition Diagnostic Results

Table 3. Failure evaluation indicator values for each method

Diagnostic model	Training set accuracy	Test Set Accuracy	F1 Composite indicator
SVM	91.69%	86.65%	86.20%
DT	90.25%	85.70%	83.63%
RF	96.48%	90.25%	88.78%
KNN	98.43%	91.25%	89.37%
Proposed	99.26%	97.70%	97.14%

The simulation results of each method show that the proposed sparse online hybrid Gaussian multi-classification algorithm-based tunnel cable state fast identification method has higher identification accuracy than the SVM, DT, RF, and KNN-based tunnel cable state identification methods, especially in the new operating conditions, the cable fault identification algorithms based on MLRM, SVM, and ANN produce larger measurement errors, but the proposed recognition algorithm has a smaller increase in the cable fault state recognition error, which indicates that it is more capable of adapting to new operating conditions, i.e., the generalization ability of this soft sensor is stronger.

The hybrid Gaussian process is the union of multiple Gaussian components, and the number of Gaussian components is the number of different operating conditions of the system, so different from the general global identification, the hybrid Gaussian will automatically identify the number of operating conditions, and judge the current operating conditions during the actual measurement, so as to select the corresponding Gaussian components. Due to the huge difference between different operating conditions, global regression methods often cannot maintain high prediction accuracy over the whole operating domain, and the hybrid Gaussian process divides the whole operating domain and maintains good prediction performance in each operating condition.

Since it is generally not possible for the training set to contain all cable states, the online updating capability of the model becomes more important. In this paper, we introduce an online updating mechanism for the model, which can introduce new training data and update the model faster, without the need to start the whole model training process from scratch. Therefore, the proposed method can still maintain a good recognition accuracy under new operating conditions.

The model training time of each method is shown in Table 4:

Table 4. Model training time for each method (unit: s)

SVM	DT	RF	KNN	Proposed
5.786	4.761	4.327	6.543	1.263

From the above table, it can be seen that the model training time of the proposed method is the shortest. Since the training time of the hybrid Gaussian multi-classification algorithm has a linear relationship with the third power of the size of the training samples, the introduction of the sparse technique in the proposed method in this paper can greatly reduce the training time of the model in the case of a small increase in the error, which provides the possibility of online identification of the cable fault status.

5. SUMMARY

In this paper, a fast identification method of tunnel cable status based on a sparse online hybrid Gaussian multi-classification algorithm is proposed. The main novelty of the method is the use of a sparse online hybrid Gaussian multi-classification algorithm for modeling the fast tunnel cable condition identification model. The results show that for the tunnel cable state diagnosis problem with complex operating conditions and high dimensionality of variables, the method has high accuracy and efficiency, and can quickly obtain the current state of the tunnel cable.

6. ACKNOWLEDGEMENT

This paper was completed with the financial support from the fund of State Grid Beijing Electric Power Company's scientific and technological project "Research and Application of Key Technology for Intelligent Inspection of Large Urban Cable Tunnels Based on Panoramic Lidar Technology" (520246230002).

7. REFERENCES

- [1] Zhou, Chengke, Huajie Yi, and **ang Dong. "Review of recent research towards power cable life cycle management." *High voltage* 2.3 (2017): 179-187.
- [2] Chen, **aolong, et al. "BN-RA: a hybrid model for risk analysis of overload-induced early cable fires." *Applied Sciences* 11.19 (2021): 8922.
- [3] Song, Yuxuan, et al. "Online multi-parameter sensing and condition assessment technology for power cables: A review." *Electric Power Systems Research* 210 (2022): 108140.
- [4] Benesl, Lukas, et al. "Cable monitoring using broadband power line communication." *Sensors* 22.8 (2022): 3019.
- [5] Huo, Yinjia, et al. "Advanced smart grid monitoring: Intelligent cable diagnostics using neural networks." *2020 IEEE International Symposium on Power Line Communications and its Applications (ISPLC)*. IEEE, 2020.
- [6] Chi, Peng, et al. "A CNN recognition method for early stage of 10 kV single core cable based on sheath current." *Electric Power Systems Research* 184 (2020): 106292.
- [7] Zhu, Wenwei, et al. "Anchor Fault Identification Method for High-Voltage DC Submarine Cable Based on VMD-Volterra-SVM." *Energies* 16.7 (2023): 3053.
- [8] Wang, Ying, et al. "Cable incipient fault identification using restricted Boltzmann machine and stacked autoencoder." *IET Generation, Transmission & Distribution* 14.7 (2020): 1242-1250.
- [9] Kwon G Y, Lee C K, Shin Y J. Diagnosis of shielded cable faults via regression-based reflectometry[J]. *IEEE Transactions on Industrial Electronics*, 2018, 66(3): 2122-2131.
- [10] Kumar, Haresh, et al. "A Review on the Classification of Partial Discharges in Medium-Voltage Cables: Detection, Feature Extraction, Artificial Intelligence-Based Classification, and Optimization Techniques." *Energies* 17.5 (2024): 1142.
- [11] Tang, Zehua, et al. "Multi-source data-cooperated neutral low-resistance grounding cable grid faulty segment identification." *IEEE Transactions on Power Systems* 37.2 (2021): 1413-1424.
- [12] Laib, Abderrzak, et al. "Enhanced artificial intelligence technique for soft fault localization and identification in complex aircraft microgrids." *Engineering Applications of Artificial Intelligence* 127 (2024): 107289.
- [13] Liu, Qing, et al. "Health Assessment of HV Cable Based on Data Correlation Analysis and Fuzzy Rules Joint Algorithm." *2022 6th International Conference on Power and Energy Engineering (ICPEE)*. IEEE, 2022.
- [14] Smith, Dominic M., Lee S. Cunningham, and Lujia Chen. "Efficient finite element modelling of helical strand cables utilising periodicity." *International Journal of Mechanical Sciences* 263 (2024): 108792.
- [15] Lyu, Kehong, et al. "Mechanism analysis and dynamic model construction of intermittent fault of weak-current aviation cable under vibration stress." *Advances in Mechanical Engineering* 14.3 (2022): 16878140221074820.
- [16] Wang, Xudong, et al. "Overcurrent tests and numerical simulations on a 66-kV-class RE123 high-temperature superconducting model cable." *IEEE transactions on applied superconductivity* 22.3 (2011): 5800904-5800904.
- [17] Ren, Hongtao, and Ying Zhang. "Analysis on switching overvoltage and suppression method of cable joint in 500 kV cable line." *Energy Reports* 7 (2021): 567-575.

- [18] Yuan, Ziyun, et al. "Knowledge-informed Variational Bayesian Gaussian mixture regression model for predicting mixed oil length." *Energy* 285 (2023): 129248.
- [19] Saves, Paul, et al. "A mixed-categorical correlation kernel for Gaussian process." *Neurocomputing* 550 (2023): 126472.
- [20] Ghahroodi, Z. Rezaei, R. Aliakbari Saba, and T. Baghfalaki. "Gaussian copula-based regression models for the analysis of mixed outcomes: an application on household's utilization of health services data." *Journal of Statistical Theory and Applications* 18.3 (2019): 182-197.
- [21] Wu, Di, and **wen Ma. "A two-layer mixture model of Gaussian process functional regressions and its MCMC EM algorithm." *IEEE Transactions on Neural Networks and Learning Systems* 29.10 (2018): 4894-4904.