J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 7861--7884

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

# A Multidimensional Data Mining and Analysis Approach for Grid Operation and Maintenance Integrating Computing and Information Theory

Songyao Feng[1], Zhengyan Huang[1,✉], Junhao Song[1], Xuexia Quan[1]

[1] *Information Center of Guangxi Power Grid Co., Ltd., Nanning, Guangxi, 530012, China*

ABSTRACT

Smart grid technology is developing rapidly around the world and is gradually applied to the operation and maintenance management of power systems, and its main advantage lies in its integration capability, which can effectively realize the high efficiency, security and reliability of power system operation and maintenance. This paper explores the integration of grid operation and maintenance by integrating computing and information theory using multidimensional data mining and analysis methods. The operation data of smart grid is first preprocessed, including resampling and PCA dimensionality reduction of multidimensional data signals. Then, a CNN-based power operation state prediction model and an R-CNN-based grid fault diagnosis model are constructed to ensure the stable operation and timely maintenance of the smart grid, and the predicted and actual values of the smart grid operation state of the CNN model are basically consistent with each other, with the MAE, MSE, and RMSE of 0.00104, 0.00014, and 0.012, respectively, and the prediction results are good. The effect is good. Compared with CNN and SVM, the performance of R-GNN model is better, and after PCA dimensionality reduction, the fault identification rate of R-GNN model is as high as 98.91%. And the delay of the R-GNN method for fault diagnosis is only 0.04s, while it can realize the comprehensive and accurate localization of the fault area. This paper provides methodological reference for the utilization of multidimensional data mining and analysis technology to realize the operation and maintenance integration of smart grid.

*Keywords:* operation and maintenance integration, CNN, R-GNN, operation state prediction, fault diagnosis

## 1. Introduction

With the rapid development of the electric power industry and the continuous expansion of the scale of the power grid, the importance of the electric power industry has become more and more

✉ Corresponding author.
*E-mail address:* hzy900529@163.com (Z. Huang).
Received 25 March 2024; Revised 05 June 2024; Accepted 15 December 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-433

prominent with the continuous development of the social economy and the improvement of people's living standards. The power grid, as the core facility of the power industry, carries the important task of power transmission, distribution and supply. Guaranteeing the normal operation and stable power supply of the power grid in the electric power industry is of great significance for the development of the national economy and the convenience of people's life. Therefore, grid operation and maintenance has become an important link to guarantee the safety, reliability and stability of power supply [1]. Grid O&M is not simply daily inspection and equipment maintenance, it includes grid emergency repair, grid overhaul, power equipment maintenance, grid new construction, safety production, technical transformation, line inspection, outage management, grid optimization, equipment replacement, maintenance and reconstruction, grid dismantling, meter reading and inspection, equipment maintenance, and defects handling [2-3]. However, in the field of power system operations and maintenance, there is an increasing number of challenges that involve not only the complexity of the system, but also the constant quest for high efficiency and reliability. With the expansion of power networks and the growth of energy demand, the grid structure has become more complex, including the access of various new generation methods and the application of smart grid technologies [4-5]. These changes have brought new challenges to the stable operation of the power system, and the first problem is the stability and reliability of the power system.

With the access of more renewable energy sources, the load fluctuation of the power grid increases, which puts forward higher requirements for power scheduling and load balance management [6]. In addition, the power system needs to cope with various emergencies, such as equipment failures and natural disasters, which pose a threat to the stability and reliability of the system [7]. The operation and maintenance cost of power systems is also an important issue, and effective cost control involves not only the saving of operation and maintenance costs, but also the optimization of capital investment [8]. In the process of dealing with these challenges, the necessity of data analysis becomes more and more prominent. With the development of information technology, a large amount of data is generated in the power system, including power generation, load changes, equipment status and environmental information. By analyzing these data, the operational status of the system can be more accurately understood, potential problems can be identified, and more effective decisions can be made [9]. Data analysis is especially important in operation and maintenance, by analyzing the parameters of the equipment operation data and historical fault records and other pairs of dimensions, it is possible to find potential failures of the equipment, so as to carry out maintenance and replacement in advance, and to reduce the losses caused by unplanned shutdowns and failures. It can not only improve the operation efficiency and reliability of the power system, but also effectively control the operation and maintenance costs, and provide support for the stability and sustainable development of the power system [10-11].

This paper explores the integrated realization of grid operation and maintenance by integrating computation and information theory, i.e., the monitoring and prediction of power operation status and the effective identification of faults are realized by mining and analyzing the multidimensional data of smart grids. First, the multidimensional grid data are preprocessed and downsized using resampling method and principal component analysis (PCA) to eliminate redundant data and retain the required effective data. Second, the CNN-based power operation state monitoring and prediction model is constructed to ensure timely detection of system anomalies. Again, a fault diagnosis model based on R-GNN is constructed to realize effective maintenance of the smart grid through timely diagnosis of faults. Finally, the data processing methods used and the constructed models are evaluated experimentally.

## 2. Overview

In multidimensional data analysis for power grids, OLAP (Online Analytical Processing) is a technique used for multidimensional data analysis to analyze the relationships between fact and dimension tables.Kovacic et al [12] mentioned that OLAP technology can be shared across different domains to be able to access best practice solutions, but a single OLAP query task does not have this capability. Moreover, this technology has a large space requirement for data storage due to the ability to perform relational parsing on both intra- and inter-organizational data, which leads to a slowdown in parsing speed.Pan et al [13] conducted a multi-dimensional data analysis of distribution network loads in terms of spatial-temporal, weather, and other natural or social environmental factors to mine the correlation between each factor and the loads, and also constructed a load-factor model to seek for a change rule for future loads. Qiao et al [14] introduced a tensor network model to analyze the multidimensional data of the equipment in the power system under the parameters of equipment type, voltage level, and equipment state, in order to obtain the reliability and energy consumption state of the operation of the power equipment.Liu et al [15] created a multidimensional data analysis model of the economic operation of the distribution network at different levels and perspectives, and analyzed the historical economic operation data and its different dimensional factors to explore its key factors. The historical economic operation data and its different dimensional factors were analyzed and their key economic operation influencing factors were mined.Tian and Ma [16] carried out real-time mining and analysis of multidimensional data in the power system operation through information visualization information interaction technology and data mining technology, and the relationship between these data can be presented visually.

## 3. Smart grid-based integration of power operation and maintenance

### *3.1. Integration of power operation and maintenance*

Power operation and maintenance integration is to break the traditional mode of power operation and control and power operation and maintenance of the division of labor mode, but the combination of operation and maintenance, the combination of power control and monitoring, all-day intelligent monitoring and control, to increase the mutual collaboration between the two, without increasing the number of employees at the same time to do each employee in the receipt of instructions, can be the fastest speed to accurately and efficiently complete the instructions, in the absence of an increase in the employee's labor intensity at the same time to achieve the purpose of saving energy and improving efficiency, so that the distribution center power management highly intelligent, centralized and efficient. While not increasing the labor intensity of employees, to achieve the purpose of saving energy and improving efficiency, so that the distribution grid center power management is highly intelligent, centralized and efficient. As a new generation of grid technology, smart grid has greatly promoted the progress of power operation and maintenance integration by virtue of its powerful information processing capability and adaptive adjustment characteristics.

### *3.2. Power operation and maintenance content of smart grid*

The smart grid power operation and maintenance management contains many aspects. First of all, data collection and monitoring are required, and various sensor facilities are deployed to monitor the operating data of the power grid, such as current, voltage and power, in order to carry out real-time monitoring and evaluation of the power grid condition. Afterwards, based on the collected data information, it is necessary to optimize and control the smart grid, such as adjusting the corresponding intelligent algorithms, simulating and evaluating the operating state of the power system, and evaluating the operating condition of the power grid, so as to take effective management measures to improve the operating level and efficiency of the power grid. In addition, the smart grid operation and maintenance management work also includes fault diagnosis and analysis, which can make use of data analysis and machine learning technology to carry out scientific and effective diagnosis and evaluation of power grid faults, discover potential problems, and take effective repair programs to ensure that the power grid system can operate normally. In addition, in the operation and maintenance management process, it is necessary to focus on remote supervision and control of electric power grid facilities, timely detection of equipment anomalies, efficient maintenance, such as the combination of Internet of Things technology, the introduction of remote control technology, the evaluation of equipment conditions, to achieve real-time supervision and real-time control. In addition, in the smart grid operation and maintenance management process, the relevant units should also establish a perfect security management system to assess and supervise the safety of the power grid, and at the same time, quickly respond to emergencies, to ensure that the power grid system can operate stably and reliably.

### *3.3. Key technologies for power operation and maintenance integration*

The realization of power operation and maintenance integration cannot be separated from many key technical supports, among which, data acquisition and processing technology and prediction and diagnosis technology are the two most central parts. Through these two technologies, the efficient operation and timely maintenance of smart grid can be realized, and the operational security and service quality of smart grid can be improved.

3.3.1. Data acquisition and processing techniques. Data acquisition and processing technology is the foundation of power operation and maintenance integration. The task of data acquisition and processing is to extract valuable information from numerous smart grid operation data, which needs to be realized through various sensors and monitoring equipment. Among them, power parameters such as voltage, current and frequency, as well as information on the operating status of power equipment, are important data to be collected. Data collection is only the first step, the collected data also need to be effectively processed to play its value, which requires the use of big data analysis technology. The main steps of data processing are as follows:

1) Data cleaning and pre-processing. The original data collected may have various problems, such as missing values, outliers, etc., which will affect the subsequent data analysis results. Through data cleaning and preprocessing, the raw data can be transformed into structured data, laying the foundation for subsequent analysis.

2) Data Mining. Use data mining technology to extract useful information and knowledge from structured data. For example, it can be used to mine the operating laws of equipment to understand under what circumstances problems may occur in the equipment, and it can also be used to mine the failure patterns to predict the type and time of failure that may occur in the equipment, and this information and knowledge has an important guiding role in improving the efficiency and quality of power operation and maintenance.

3) Data Visualization. The application of data visualization technology can show the analysis results in an intuitive form, and the operation and maintenance personnel can understand the analysis results more easily and make more accurate decisions.

In general, data acquisition and processing technology is an important tool for realizing the integration of smart grid power operation and maintenance, and its application effect will directly affect the effect of power operation and maintenance.

3.3.2. Forecasting and diagnostic techniques. Prediction and diagnosis technology is another important electric power operation and maintenance technology, which mainly utilizes data analysis and machine learning algorithms to predict and diagnose the operating state of the smart grid. Prediction technology is mainly used to build a prediction model of the operating state of the smart grid by analyzing historical data and predicting the possible state of the smart grid in the future period. This technology can help operation and maintenance personnel make preparations in advance to prevent possible problems. Diagnostic technology, on the other hand, is to detect and diagnose abnormalities in the smart grid system in time by monitoring the operating status of the smart grid in real time. Once abnormal system operation is found, the diagnostic system will immediately issue an alarm, and at the same time, provide possible causes and solutions to help operation and maintenance personnel to deal with the problem in time to avoid the occurrence of failures.

## 4. Multidimensional data mining and analysis model for smart grid operation and maintenance

Based on the relevant contents and key technologies of smart grid operation and maintenance integration, this chapter mines and analyzes the multidimensional data of smart grid operation and maintenance. Based on the collection and processing of multidimensional data, the operational state prediction model and fault diagnosis model of smart grid are constructed, so as to realize the integration of smart grid operation and maintenance.

### *4.1. Grid data processing based on resampling and principal component analysis*

This paper utilizes data cleaning method to prioritize the original basic data of smart grid, and then uses down-sampling technology and principal component analysis technology to improve the sparsity and dimensionality reduction of the cleaned data, so as to realize the data reconstruction and improve the sparsity and sample class balance of the data, and lay the foundation of the subsequent evaluation of the grid operation status and fault diagnosis.

4.1.1. Smart grid raw data preprocessing methods:

1) Data Cleaning. For the smart grid original basic data with repetition, missing, abnormal and other defects to make the cleaning process, the data cleaning methods used in this paper are repeated value processing, vacant value processing and abnormal value processing.

(1) Repeated value processing. In this paper, the duplicate values in the smart grid data are directly deleted, and the duplicate data are deleted through the global search and comparison of the data.

(2) Vacant value processing. Due to the smart grid operation process there will be faults, maintenance and other reasons that lead to missing grid data, the missing value of the commonly used methods include direct deletion, Lagrange interpolation, mean value method, this paper according to the actual operation of the smart grid to take the vacant value processing method for the mean value method. Matlab software is used to process the basic data of the smart grid, firstly, global search is taken to find the missing points, according to the missing values above and below the position of the data to find the mean value so as to replace the missing values, and finally, the samples are saved.

(3) Abnormal value processing. Aiming at the abnormal values of the data in the smart grid, this paper adopts the mean value method to replace the value of the abnormal position in the grid data according to the actual operation of the smart grid. Set the upper and lower lines of smart grid data fluctuation in the software Python, and observe the outliers according to the set upper and lower lines, delete the abnormal values of the outliers and take the mean value method to replace them, so as to realize the abnormal value processing of data.

2) Data normalization. The state prediction in the actual operation of the smart grid needs to include a variety of quantities, and different physical quantities will have different impacts on the prediction results, so it is necessary to normalize the data. At present, the commonly used normalization methods include standardization and interval deflation. The standardization method is to make the data obey

the standard normal distribution, and the interval scaling method is to specify the data to a specific range.

(1) Standardized method. The data processed using the standardized method has a mean of 0 and a standard deviation of 1 in accordance with the standard normal distribution. The formula for standardization is:

$$x^* = \frac{x - \bar{X}}{S} \tag{1}$$

where, $x$ is the original data, $\bar{X}$ is the data mean, $S$ is the standard deviation and $x^*$ is the standardized data.

(2) Interval scaling method. Scaling law is the use of characteristics of the boundary value of the processing scaling process, so that after processing the data in the range between the provisions of the commonly used interval scaling method for the maximum and minimum interval scaling method, after processing the data for the $[0,1]$ between the data scaling formula for:

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{2}$$

where $x$ is the original feature data. $x_{\max}$ is the maximum value of the feature. $x_{\min}$ is the minimum value of the feature. $x^*$ is the processed feature data.

Therefore, in order to improve the optimization accuracy of the neural network, reduce the learning time of the model, and further enhance the effectiveness of the training and optimization results, the original distribution network $N$ attribute variables of the $X_1, X_2, \ldots, X_p$ normalization process, and the processed values fall within the range of $[0,1]$:

$$\tilde{X} = \frac{X(\cdot,i) - Min\left(X(\cdot,i)\right)}{Max\left(X(\cdot,i)\right) - Min\left(X(\cdot,i)\right)} \tag{3}$$

where $Max\left(X(\cdot,i)\right)$ and $Min\left(X(\cdot,i)\right)$ are the maximum and minimum values of the $i$ rd attribute in the sample, respectively.

4.1.2. Multi-dimensional signal resampling and dimensionality reduction for smart grids:

1) Multidimensional signal resampling for smart grids. Resampling method is an essential tool in modern statistics, which repeatedly takes samples from the training set and refits the model of interest for each sample to obtain valid information about the fitted model. Resampling is mainly categorized into up-sampling and down-sampling, up-sampling mainly increases the sampling frequency on the frequency of the original samples to increase the required data, while down-sampling reduces the sampling frequency on the basis of the original sampling frequency.

For a sequence of sample values spaced several samples sampled once, so that the new sequence is the original sequence of the downsampling, which is mainly used in signal processing, super-resolution, neural networks and other fields, through the use of downsampling for data processing has the following advantages: downsampling to reduce the number of parameters to be learned by the network, to prevent overfitting, to increase the perceptual field, and to achieve invariance, translation invariance, rotation invariance, scale invariance. Considering the huge amount of smart grid data, this paper adopts downsampling method to process the raw smart grid data.

2) Smart grid multidimensional signal principal component analysis. Principal Component Analysis (PCA) is a multivariate statistical method, which can effectively select the most important elements

and structures from a large amount of data information by downscaling the data through linear transformation, i.e., selecting the most important variables from multiple variables and removing the non-essential factors to downscale the original basic data, so as to reveal the simple structure behind a large amount of data. The principal components in principal component analysis are arranged in order of variance, so when analyzing the problem, it is possible to discard part of the principal components and only take several principal components with larger variance to represent the original variables, thus reducing the workload of calculation. In the comprehensive evaluation function, the weight of each principal component is its contribution rate, which reflects the proportion of the amount of information of the original data contained in that principal component to the proportion of all the information, and determines the objective and reasonable weights, which overcomes the defects of determining the weights that are considered to be determined in some evaluation methods.

Assuming that the standardization of the original data collects $P$ -dimensional random variables $x=\left(X_1, X_1, \ldots, X_{ip}\right)^T$, $n$ samples $x_j=\left(x_{11}, \ldots, x_{i2}, x_{ij}\right)^T$, $i=1,2,\cdots,n$ and constructs a sample matrix to standardize the sample matrix.

The main steps of principal component analysis are as follows:

(1) Standardized matrix $Z$ :

$$Z_{ij}=\frac{x_{ij}-\bar{x}_j}{s_j} \tag{4}$$

$$\bar{x}_j=\frac{\sum_{i=1}^{n} x_{ij}}{n} \tag{5}$$

$$s_j^2=\frac{\sum_{i-1}^{n}\left(x_{ij}-\bar{x}_j\right)}{n-1} \tag{6}$$

where: $\bar{x}_j$ is the mean, $s_j^2$ is the sample variance, $i=1,2,\ldots,n$, $j=1,2,\ldots,p$.

(2) Find the correlation coefficient matrix $R$ for the standardized matrix $Z$ :

Step1: Find the sample variance matrix thereby creating the sample variance matrix:

$$b_{ij}=Cov\left(X_i, X_j\right)=\sum_{k-1}^{n}\left(X_i^k-\bar{X}_i^k\right)\left(X_j^k-\bar{X}_j^k\right) \tag{7}$$

$$B=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \tag{8}$$

Step2: Based on Eq. (8), the correlation coefficient matrix is established:

$$\rho_{ij}=\frac{b_{ij}}{\sqrt{D\left(X_i\right)}\sqrt{D\left(X_j\right)}} \tag{9}$$

$$R=\begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{2n} & \cdots & \rho_{nn} \end{bmatrix} \tag{10}$$

Step3: Solve the characteristic equation $\left|R-\lambda I_P\right|=0$ of the correlation matrix $R$ to obtain the $p$ characteristic roots, determine the principal components, and determine the $m$ values according to $\frac{\sum_{j=1}^{m}\lambda_j}{\sum_{j=1}^{p}\lambda_j} >= 0.85$ so that more than 85% of the information is utilized. For each $\lambda_j$, $j=1,2,\ldots,m$, solve the system of equations $R_b=\lambda_j b$ to obtain the unit eigenvector $b_j^o$.

Step4: Convert the standardized indicator variables into principal components:

$$U_{ij}=z_i^T b_j^o, j=1,2,\ldots m \tag{11}$$

$U_1$ is called the first principal component, $U_2$ is called the second principal component, ..., $U_p$ is called the $p$ th principal component.

Step5: Evaluation of the $m$ principal components

The $m$ components are weighted and summed up, i.e., the final evaluation value, and the weights are the variance contribution ratio of each principal component.

### *4.2. CNN-based power operation state prediction for smart grids*

In this section, a smart grid power operation condition monitoring and prediction method based on convolutional neural network (CNN) [17] is proposed, aiming to improve the real-time responsiveness of condition monitoring and the accuracy of prediction.

4.2.1. Smart Grid Power Data Characteristics. The core characteristics of smart grid power data are reflected in its high-dimensionality, spatio-temporal, multivariate and nonlinear nature.

1) High-dimensionality. The high dimensionality of power data originates from the multi-parameter information collected by a large number of distributed sensors. The data stream generated by each sensor is a multidimensional time series, where each dimension corresponds to a monitoring parameter.

2) Temporal and spatial. Power data not only changes over time, reflecting the dynamics of the grid state, but also has significant spatial correlation. For example, load changes within the same geographic area affect each other, while temporal continuity is reflected in the cyclical and seasonal changes in load.

3) Multi-variability. The state of grid operation is affected by a variety of factors, such as seasonal variations, weather conditions, and user behavior, which leads to the variability of power data. Fluctuations in power demand and changes in equipment performance are reflected in the data.

4) Non-linearity. The operating mechanisms and load changes of power grids are often nonlinear, and traditional linear models are difficult to accurately capture and predict the complex behavior of power systems.

The advantage of CNN-based processing of this type of data lies in its ability to automatically extract spatio-temporal features and reveal the intrinsic relevance of the data through deep structure.

4.2.2. Convolutional Neural Network Modeling. In constructing a convolutional neural network model for smart grid power operation state monitoring and prediction, the key lies in accurately capturing and analyzing high-dimensional, spatio-temporally correlated power data features. The structure of the model design proposed in this paper is shown in Fig. 1. The core of this CNN model design lies in adapting the complexity of smart grid data to realize real-time, high-accuracy monitoring and prospective prediction of power system operation status.

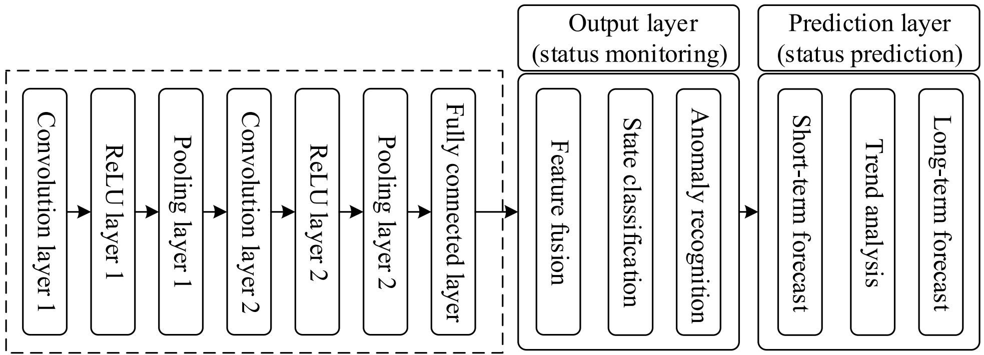



**Fig. 1.** Convolutional neural network model structure

4.2.3. Model flow. The specific process of the CNN-based smart grid power operation state monitoring and prediction model designed in this paper is as follows:

Step1: Data Acquisition and Preprocessing

First, raw power data are collected from multiple sensors distributed throughout the grid. These data can be represented as a multidimensional time series collection, denoted as:

$$D_{raw} = \{D_1, D_2, \cdots, D_n\} \tag{12}$$

where $D_i$ denotes the data sequence collected by the $i$ nd sensor.

Multi-dimensional data preprocessing and dimensionality reduction are carried out by applying the methods described in the previous section to ensure the quality and consistency of the data and to lay the foundation for the subsequent feature extraction and pattern recognition.

Step2: Input layer

In the smart grid monitoring scenario, the data processed in the input layer are preprocessed power system data that have been standardized, normalized, and converted into a format suitable for model processing.

Assume that the preprocessed data set is $X = \{X_1, X_2, \cdots, X_n\}$ , where each $X_i$ is a multidimensional time series data. Let $X_{input}$ represent the data in the input layer, which is mathematically expressed as:

$$X_{input} = reshape(X) \tag{13}$$

where $reshape$ is an operation that adapts each $X_i$ to the desired dimension and format of the network.

The main role of the input layer is to ensure that the data enters the CNN in the proper form.

Step3: Convolutional Layer 1 (Conv1)

The main role of convolutional layer 1 is to extract primary features from the input data. The operation of the convolutional layer can be represented as:

$$Z_1 = W_1 * X_{input} + b_1 \tag{14}$$

where $*$ denotes the convolution operation, $W_1$ is the weight matrix of the convolution kernel, and $b_1$ is the bias vector of the convolution layer. The convolution operation extracts features by sliding the convolution kernel over the input data and computing the dot product.

Step4: Activation Layer 1 (ReLU1)

The activation layer applies the ReLU function to the output $A_1 = \max(0, Z_1)$ of the convolutional layer with the expression:

$$A_1 = \max(0, Z_1) \tag{15}$$

This operation increases the nonlinearity of the decision function.

Step5: Pooling Layer 1 (Pooling1)

The pooling layer enables the network to focus more on the key features in the power data and ignore the less important changes, thus capturing the key indicators of the grid state more effectively.

Let $A_1$ be the output of activation layer 1, then the operation of the pooling layer can be expressed as:

$$P_1 = pool(A_1) \tag{16}$$

where pool is a pooling function which can be either maximum pooling or average pooling.

Step6: Convolutional Layer 2 (Conv2)

This layer continues to extract higher level abstract features from the features extracted from the previous layer.

$P_1$ is the output of pooling layer 1, then the operation of convolution layer 2 can be expressed as:

$$Z_2 = W_2 * P_1 + b_2 \tag{17}$$

where $W_2$ and $b_2$ are the weights and bias of convolutional layer 2, respectively.

Step7: Activation layer 2 (ReLU2)

The activation layer 2 applies the ReLU function to the output $Z_2$ of the convolutional layer 2 with the expression:

$$A_2 = \max(0, Z_2) \tag{18}$$

The ReLU activation function enhances the network's ability to learn complex functions while keeping the computation simple by setting all negative values to zero.

Step8: Pooling Layer 2 (Pooling2)

Pooling layer 2 further reduces the spatial dimensions of the features extracted by convolutional layer 2, which helps to reduce computational complexity and enhances the generalization of the model.

Let $A_2$ be the output of activation layer 2, then the operation of pooling layer 2 can be expressed as:

$$P_2 = pool(A_2) \tag{19}$$

Through convolution, activation and pooling operations, the model can gradually extract and simplify key features in the power data.

Step9: Fully Connected Layer (FC)

The full connectivity layer is responsible for synthesizing all the features extracted from the raw data to form a comprehensive understanding of the current state of the power grid.

Let $P_2$ be the output of Pooling Layer 2, then the operation of the Full Connectivity Layer can be expressed as:

$$F = \mathrm{Re}LU\left(W_{fc} \cdot flatten(P_2) + b_{fc}\right) \tag{20}$$

where $W_{fc}$ and $b_{fc}$ represent the weights and biases of the fully connected layer, respectively. $flatten$ is an operation that spreads the output $P_2$ of the pooling layer into a one-dimensional vector.

Step10: Output layer (state monitoring)

The main task of the output layer is to determine whether the grid is normal, there is an anomaly, or a specific type of anomaly.

The operation of the output layer can be expressed as:

$$Y_{status} = soft\max\left(W_{out} \cdot F + b_{out}\right) \tag{21}$$

where $W_{out}$ and $b_{out}$ are the weights and biases of the output layer, respectively. softmax function is used to convert the output of the fully connected layer into a probability distribution representing the probability of various different grid states.

Step11: Prediction layer (state prediction)

This layer focuses on predicting the future state of the grid based on current and historical data, including short- and long-term power loads, potential fault risks, and other key performance indicators. The prediction layer can be represented by the following equation:

$$Y_{predict} = \sigma\left(W_{pred} \cdot F + b_{pred}\right) \tag{22}$$

where $W_{pred}$ and $b_{pred}$ are the weights and biases of the prediction layer.

### *4.3. R-GNN based fault diagnosis scheme for smart grid*

In this paper, we propose a spatio-temporal recurrent graph neural network (R-GNN) [18] obtained by combining recurrent neural network (RNN) [19] and graph neural network (GNN) [20] for smart grid fault diagnosis using voltage measurement data as input.

4.3.1. Graph Neural Networks. A smart grid is defined as an undirected graph $g = (v, \varepsilon, A)$, where $v$ is the set of vertices, $|v| = N$ and each vertex in the graph represents a node (bus) in the smart grid. $X = \{X_1, X_2, \cdots, X_N\}$ is the tuple of node features. $\varepsilon$ is the set of edges, $|\varepsilon| = M$ and each edge represents a branch line connecting 2 buses. $E = \{E_1, E_2, \cdots, E_M\}$ is the tuple of edge features. $A \in \mathbb{R}^{N \times N}$ is the adjacency matrix of the distribution network. $X_{i=1,2,\cdots,N}$ is the node features of the image learning input data, and $E_{i=1,2,\cdots,M}$ is the edge features.

The learning objective is to generalize the mapping model between input and output of node or edge attributes. The output of graph learning can be a classification or regression task at the node or graph level. The input and output model of graph neural network (GNN) can be represented as:

$$\hat{y} = F(g, v, \varepsilon, \omega) \tag{23}$$

where: $\omega$ is the training weights, $\varepsilon$ is the set of edges, $v$ is the set of nodes, $g$ is the graph neural network, and $\hat{y}$ is the inferred output. The training weights are iteratively updated by backpropagation on the minimized loss function $L(\hat{y}, y)$, where $y$ is the output label.

The main difference between the traditional neural network structure and the graph neural network structure is that the graph neural network structure consists of a graph structure through an adjacency matrix $A$ of an undirected graph $g = (v, \varepsilon, A)$. In this case, the set of edges $\varepsilon$ and the adjacency matrix $A$ remain unchanged.

4.3.2. Troubleshooting program. This paper focuses on the voltage measurement based fault diagnosis scheme for R-GNN. Here, only bus voltage measurements are considered as input node features. Each bus has three phase voltages $V_{a,K}$, $V_{b,K}$ and $V_{c,K}$ in the time series, where $K$ is the time index. Therefore, the node feature in node $i$ is of the form:

$$X_i = \begin{bmatrix} V_{a,1} & V_{a,2} & \cdots & V_{a,K} \\ V_{b,1} & V_{b,2} & \cdots & V_{b,K} \\ V_{c,1} & V_{c,2} & \cdots & V_{c,K} \end{bmatrix}^T \tag{24}$$

The node conductance matrix of the smart grid may be a weighted adjacency matrix. The graph data $g=\left(v:\{X_1,X_2,\cdots,X_n\},A\right)$ comprises voltage measurements of nodes in the smart grid system and an adjacency matrix $A$ representing the graph.

The inferred outputs $\hat{y}$ of the GNN model $F(\cdot)$ are fault categories and fault localization, where the graph classification is used to discriminate the fault categories and the node classification is used to localize the faults. Fault classification is categorized into 2 labels: fault type and fault stage. Fault types are categorized into 6 types: no fault (NF), single-phase ground (LG), two-phase (LL), two-phase ground (LLG), three-phase (3L) and three-phase ground (3LG).

Thus, $y_{type} \in B^{1\times 6}$, where $y_{phase}[i]$ denotes the $i$ rd element, $y_{type}:y_{type}[i]=1$ denotes that the $i$ th fault category has occurred, and all other $y_{phase}[i]=0$. The fault phase is determined by $y_{type} \in B^{1\times 3}$, where $y_{phase}[i]=1$ denotes that the fault occurs in case of asymmetric fault types, i.e., the fault occurs in phases $A$, $B$, and $C$, or $AB$, $BC$, and $CA$ for $LG$, $LL$, and $LLG$, respectively. The fault location is denoted by $y_i=1$ and if the fault occurs in the $i$ th bus, then $i=1,2,3,\cdots,N$, otherwise $y_i=0$. Fault localization detection in node level classification, the recursive graph neural network based fault diagnosis scheme is shown in Fig. 2.

Node features $\{X_1,X_2,\cdots,X_N\}$ and hidden features $\{H_1,H_2,\cdots,H_N\}$ are extracted through the R-GNN layer. From these different hidden features, on the one hand, the dense layer in node classification can be used to capture the fault localization output $y_i$. On the other hand, the unification of the graph features is achieved by pooling operation, and then the fault type and fault phase are determined by the dense layer. The fault type is detected first and when asymmetric faults are detected, fault phase identification is performed.

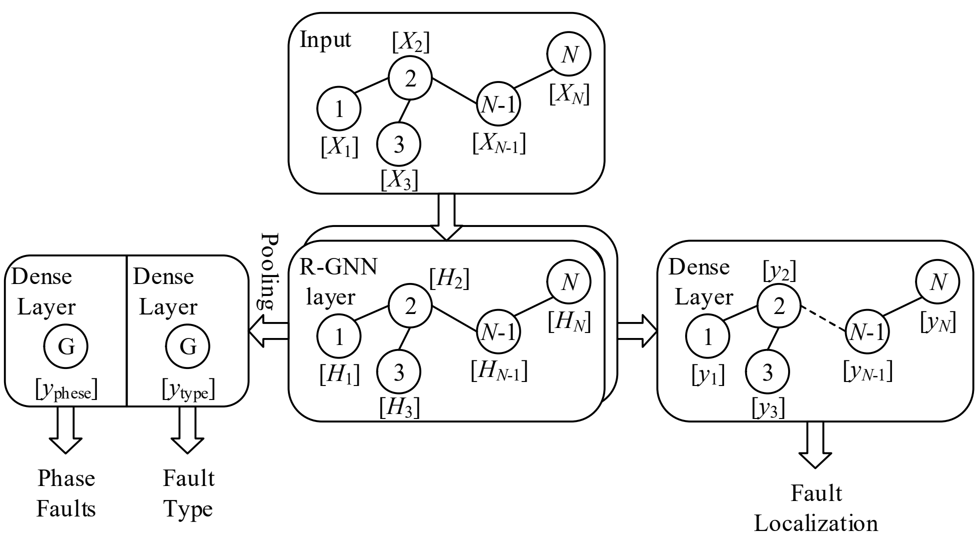



**Fig. 2.** Fault diagnosis scheme

The overall flow of fault diagnosis based on spatio-temporal recurrence graph neural network is shown in Fig. 3. The flowchart summarizes the offline training and online diagnosis of the fault diagnosis neural network model.

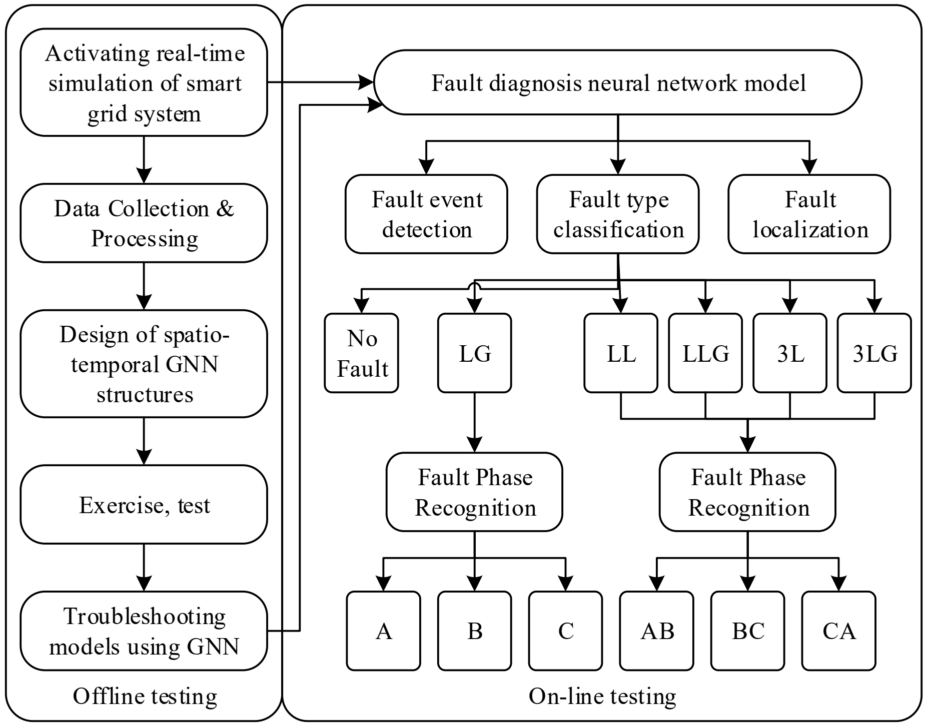



**Fig. 3.** Overall fault diagnosis process

4.3.3. Spatiotemporal Recurrence Map Convolutional Network Layer. In this paper, we propose temporal and spatial recursive GNN layer for fault diagnosis modeling. The temporal R-GNN layer captures the temporal and spatial correlation of the input data. The fault diagnosis model is represented by a classification function $F(\cdot)$ mapping the input time series $\{X_i^1, X_i^2, \cdots, X_i^K\}$ with the following fault labels on $g$:

$$\hat{y} = F\left(v : \{X_1^1, X_1^2, \cdots, X_1^K\}, A\right) \tag{25}$$

where: $v$ is the node set, $X_1 = \{X_1^1, X_1^2, \cdots, X_1^K\}$ is the node features, $i$ is the node index, $K$ is the time index and $A$ is the adjacency matrix. The structure of $g$ is reflected by the adjacency matrix $A \in \mathbb{R}^{N\times N}$.

The fault diagnosis timing R-GNN structure is shown in Fig. 4, which consists of an RNN unit and a GNN layer for feature extraction. First, the RNN unit is utilized to extract temporal features from the time series voltage of each node. Second, the GNN layer is utilized to identify the spatial correlation between smart grid bus voltages. Finally, a global pooling operation concentrates the hidden features of all nodes and trains the dense layer to classify fault types and fault phases.

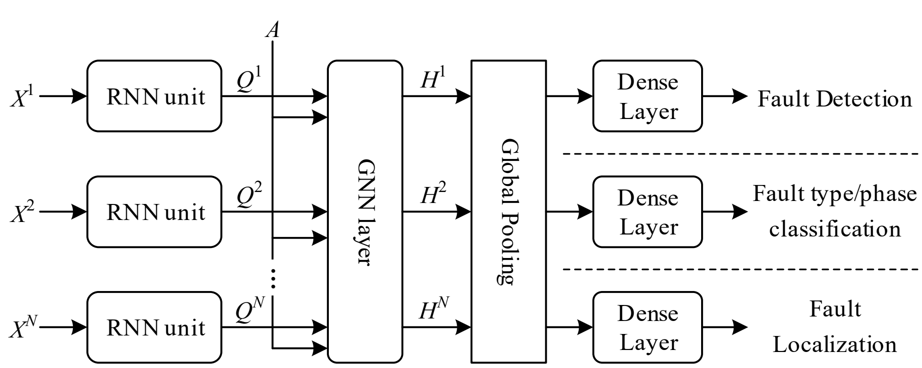


**Fig. 4.** Timing R-GNN structure for fault diagnosis

1) Graph neural network layer. The node features at each time index are processed by the GNN layer, which can be represented as

$$H_{(l+1)}^{i}=\sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H_{(l)}^{i}W_{(l)}\right) \tag{26}$$

where: $\tilde{A}=A+I_N$ is the adjacency matrix with self-connections. $I_N$ is the unit matrix. $\tilde{D}$ is the diagonal matrix from $\tilde{A}$ to $\tilde{D}_{ii}=\sum_i \tilde{A}_{ij}$ and $\tilde{D}_{ii}=0$. $H_{(l)}^{i}$ is the output of layer $l$. $H_{(0)}^{i}=V^{i}$ is the initial value of nodes. $W_{(l)}$ is the weight matrix of layer $l$. $\sigma(\cdot)$ is the nonlinear activation function.

The output of each time indicator layer of the GNN is the input of a recurrent neural network (RNN), where the RNN unit can be a gated recurrent unit (GRU) or a long short-term memory network (LSTM).The GRU structure is simpler and more computationally efficient than the LSTM structure. While LSTM can remember longer sequences and performs better in remote time tasks.

2) Long Short-Term Memory Network. An LSTM cell computes the hidden state $h_t$ and the cell state $c_t$ for each time step based on the input $x_t$, the previous hidden state $h_{t-1}$ and the previous cell state $c_{t-1}$. In each LSTM, there are intermediate states for the oblivion gate $f_t$, the cell candidate gate $g_t$, the input gate $i_t$ and the output gate $o_t$. The relationship between these state variables is represented as follows:

$$f_t=\sigma_g\left(W_f x_t+R_f h_{t-1}+b_f\right) \tag{27}$$

$$g_t=\sigma_c\left(W_g x_t+R_g h_{t-1}+b_g\right) \tag{28}$$

$$i_t=\sigma_g\left(W_i x_t+R_i h_{t-1}+b_i\right) \tag{29}$$

$$o_t=\sigma_g\left(W_o x_t+R_o h_{t-1}+b_o\right) \tag{30}$$

where: $W_f$, $W_g$, $W_i$, $W_o$ is the weight matrix. $R_f$, $R_g$, $R_i$, $R_o$ are the training weights. $b_f$, $b_g$, $b_i$, $b_o$ are the bias. $\sigma_g(\cdot)$ is the sigmoid function. $\sigma_c(\cdot)$ is the tanh function. The cell state and hidden state are calculated as follows:

$$c_t = f_1 \circ c_{t-1} + g_t \circ i_t \tag{31}$$

$$h_t=f_t\circ\sigma_c\left(c_{t-1}\right) \tag{32}$$

where: $\circ$ is element-by-element multiplication.

3) Gating cycle unit. The GRU unit contains only reset gate $r_t$, update gate $z_t$ and is calculated as follows:

$$r_t=\sigma_r\left(W_r x_t+R_r h_{t-1}+b_r\right) \tag{33}$$

$$z_t=\sigma_z\left(W_z x_t+R_z h_{t-1}+b_2\right) \tag{34}$$

The hidden states of the last layer of RNN units are collected. Then, the output is computed using the dense layer. Each fault diagnosis task has 3 dense blocks. Therefore, there are 3 different models for each task. The dense layer formula is as follows:

$$\hat{y}=\sigma_d\left(W_d\left[h_1,h_2,\cdots,h_L\right]+b_d\right) \tag{35}$$

where: $W_d$, $b_d$ are the trainable weights and bias. $\left[h_1,h_2,\cdots,h_L\right]$ is the flat matrix of all hidden states collected from the RNN units. $\sigma_d\left(\cdot\right)$ is the activation function.

4) Multi-task fault diagnosis migration and fine-tuning. Fault diagnosis includes three tasks of fault detection, fault type/phase classification and fault localization identification, and each deep neural network model is trained independently for each task, which is a simple but inefficient method. Therefore, this paper proposes a transfer learning technique for multi-task fault diagnosis as shown in Fig. 5.

First, the R-GNN model for fault classification is pre-trained. Then, the R-GNN layer of trained node feature extraction in the model is transferred to the fault type/phase identification model in the new R-GNN. The 2 new R-GNN models consist of trained R-GNN layers and new additional dense layers. During the new model training process, the trained R-GNN layer weights are frozen so that only the weights of the dense layers are trained. Then, the transmitted R-GNN layers are unfrozen and fine-tuned to train all layers using a small learning rate of 0.001. Using this approach reduces the training effort and greatly improves the network efficiency.

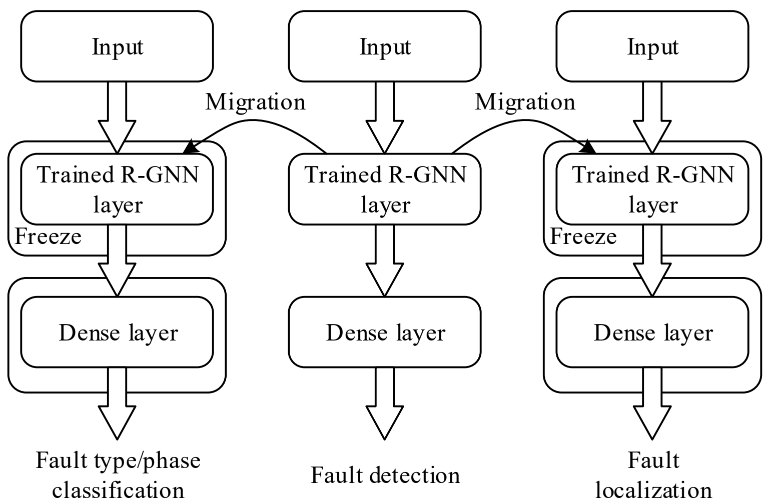
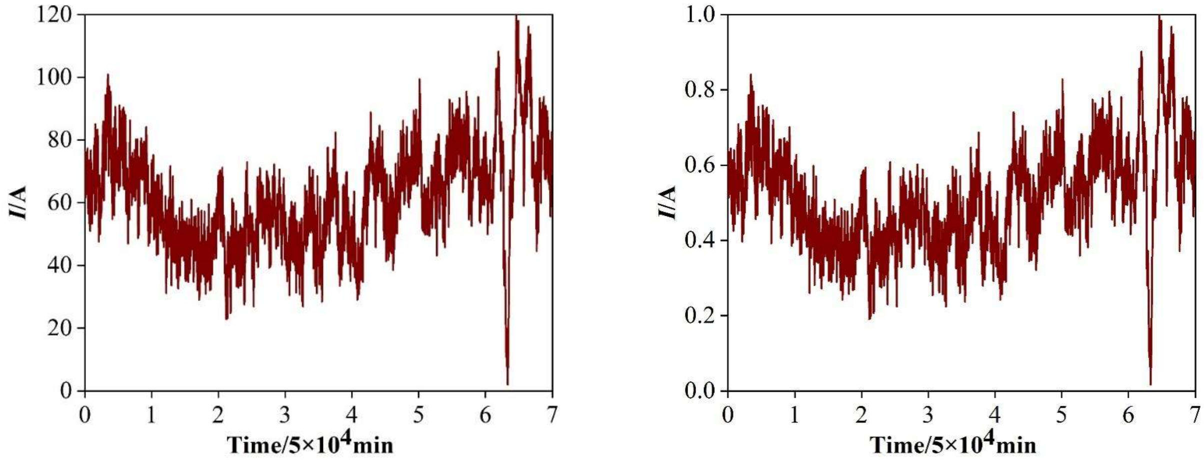



**Fig. 5.** Multitask fault diagnosis layer transmission system

## 5. Experimental results and analysis

### *5.1. Smart Grid Multidimensional Data Processing*

5.1.1. Data pre-processing. This paper takes a 10kV smart microgrid in a region as an example, preprocesses the multidimensional raw data of the grid's operation and maintenance, cleanses the multidimensional data and normalizes it. Taking the raw current data in the power operation of the grid as an example, the comparison between the raw data and the preprocessed data is shown in Figure 6. Where Fig. (a) and Fig. (b) represent the original data and normalized data, respectively. The normalized current is within 0~1, which enhances the validity of the training and optimization results while facilitating viewing, and lays the foundation for the subsequent research on the prediction of smart grid operation state and fault diagnosis.




(a) Raw data (b) Normalized data

**Fig. 6.** Comparison of current data before and after normalization

5.1.2. Data resampling and dimensionality reduction:

1) Data resampling. In order to verify the effectiveness of the downsampling method, the normalized smart grid current data is processed based on Matlab by applying the downsampling method, and the downsampled processed current data is shown in Fig. 7. Comparing with Fig. 6(b), it can be seen that the normalized current data is time unit of $5\times10^4$min, while the time unit of the data for which downsampling is performed is days, which enhances the sparsity of the curve.

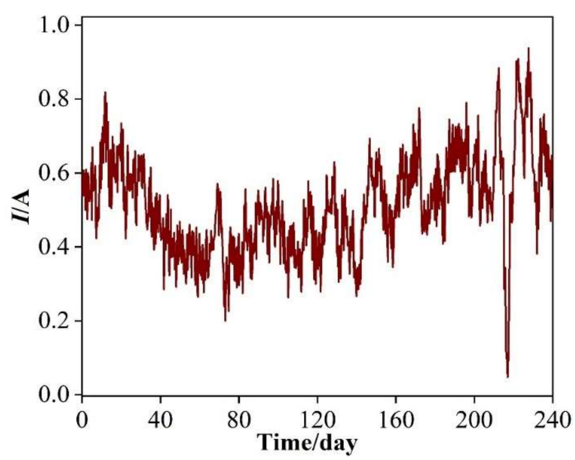
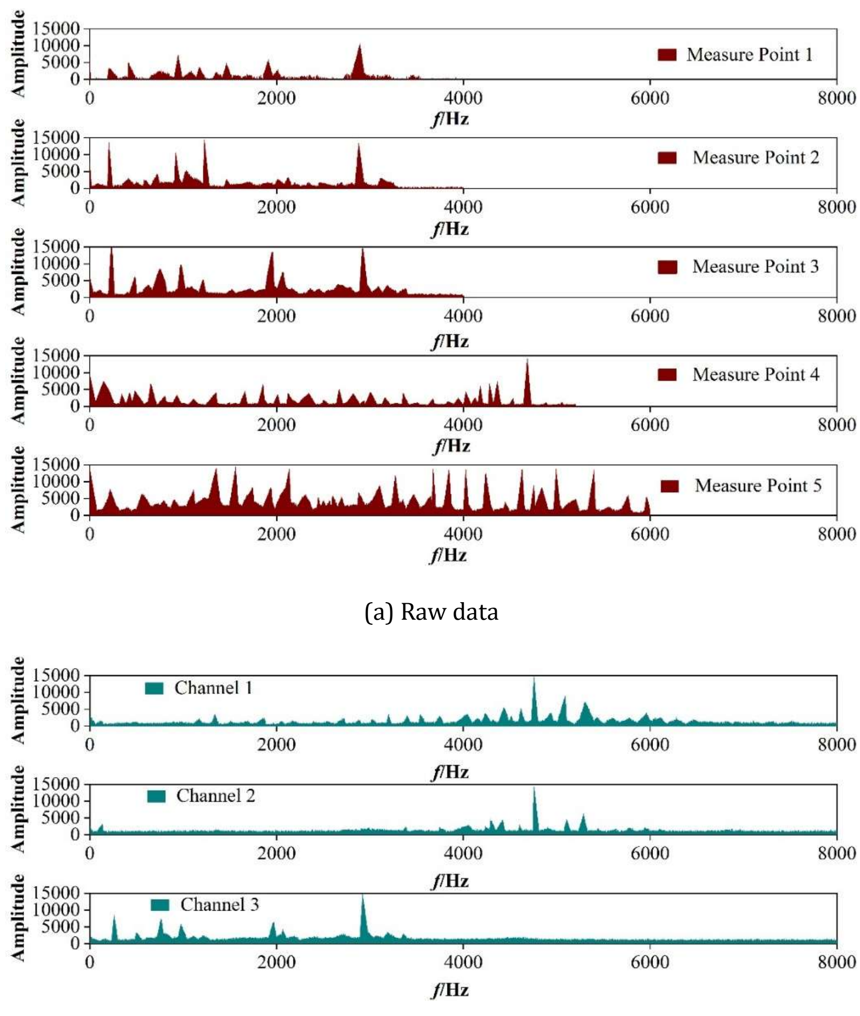



**Fig. 7.** Current data after down-sampling processing

In order to eliminate possible accidental errors during the experiment, 10 experiments were conducted, and by comparing the original data with the reconstructed data, the average error of the 10 experiments was 1.862%, which indicates that the smart grid operation data can be effectively reconstructed by applying the downsampling method in this paper. The downsampling method on the one hand reduces the redundant information in the current data, increases the information density about the grid operation state features in the current data, and facilitates the subsequent feature extraction and network training. On the other hand, it increases the data volume, which has the effect of data enhancement.

2) PCA-based data dimensionality reduction. In addition, this paper adopts Principal Component Analysis (PCA) to process the preprocessed grid operation signal data, reduce the number of data channels, and extract appropriate feature data.The principle of PCA is to utilize the method of dimensionality reduction, maximize the preservation of inter-sample variance by means of orthogonal linear combinations, and transform multiple indexes into a few comprehensive indexes.

The frequency domain plots before and after PCA dimensionality reduction of the multidimensional data of a smart grid line operation with five measurement points are shown in Fig. 8. Where (a) and (b) represent the original data and the data after dimensionality reduction, respectively.PCA reduces the original five data channels into three data channels, which automatically extracts and retains the main frequency domain features of the original five data channels while greatly reducing the amount of data participating in training.




(a) Raw data

(b) Data after PCA dimensionality reduction

**Fig. 8.** Frequency domain diagram of signal before and after dimensionality reduction

### *5.2. Smart Grid Power Operation State Prediction Experiment*

The multidimensional operation data of the smart grid after data processing is imported into the CNN-based power operation state monitoring and prediction model, and the real-time detection results of the state affiliation function of the whole smart microgrid on a certain day are obtained by simulation as shown in Figure 9. The normal operation state affiliation function of the smart microgrid in the region is the maximum value during the simulation time, so it can be known that the state belonging to the smart microgrid in the region is the normal operation state.

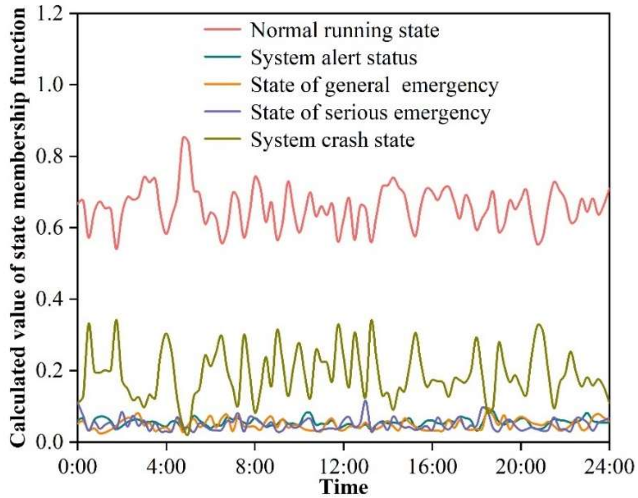
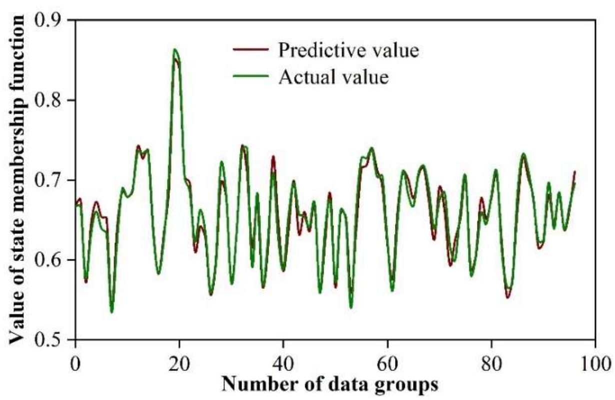
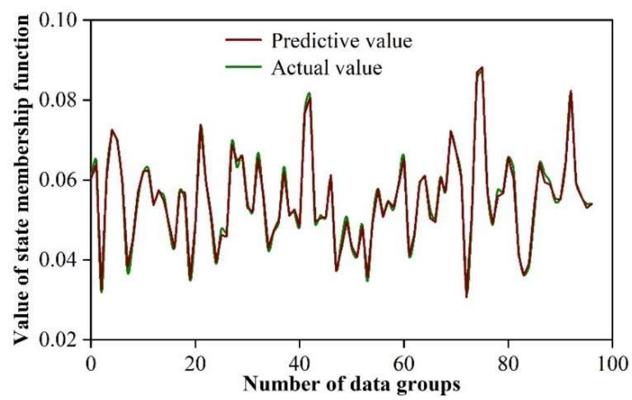



**Fig. 9.** Comprehensive status membership function of smart microgrid

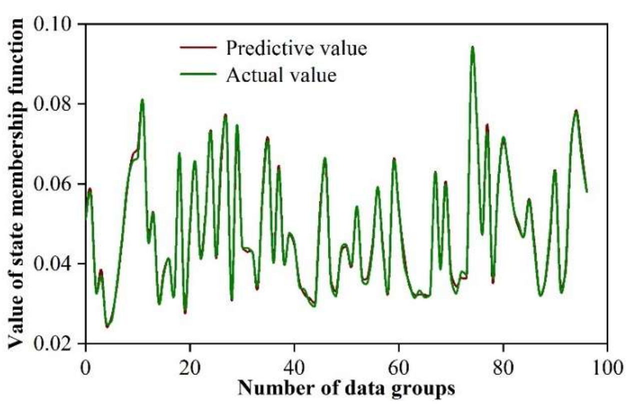
The comparison between the model prediction results and the actual values of the grid operation states is shown in Fig. 10, where (a) to (e) are the five operation states of normal operation, system alert, general emergency, severe emergency, and system collapse, respectively. It can be seen that there is a basic match between the CNN model predicted values and the actual values for the five smart grid operation states, with a small difference.



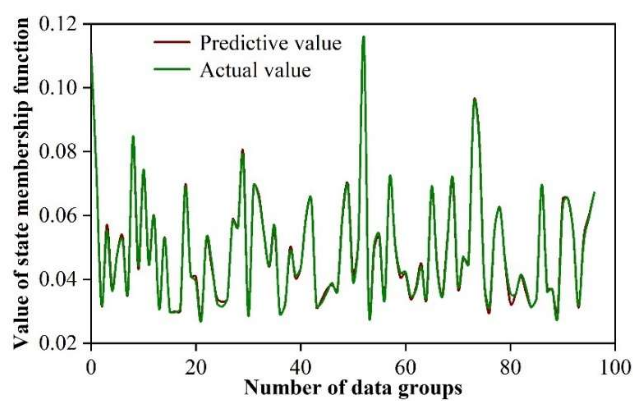

(a) Normal running state



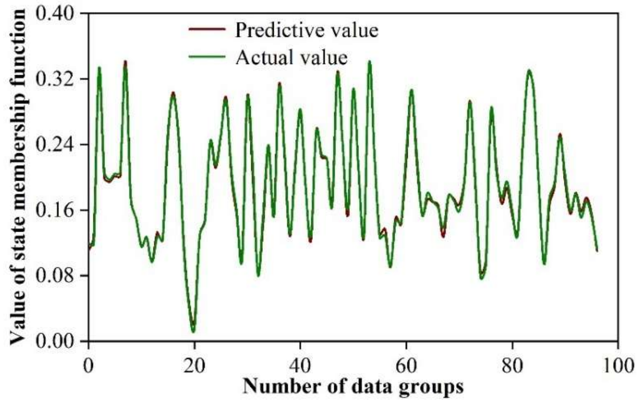

(b) System alert status




(c) State of general emergency




(d) State of serious emergency




(e) System crash state

**Fig. 10.** Comparison between CNN prediction data and original data

The normal operation state affiliation function prediction and the actual value are selected to do the error analysis, and the results of the error analysis are shown in Fig. 11. From the error value in Fig. 11, the mean absolute error (MAE), mean square error (MSE) and root mean square error (RMSE) are 0.00104, 0.00014 and 0.012, respectively, which are within the reasonable range, and the prediction effect is good and meets the actual demand of smart grid operation state prediction.

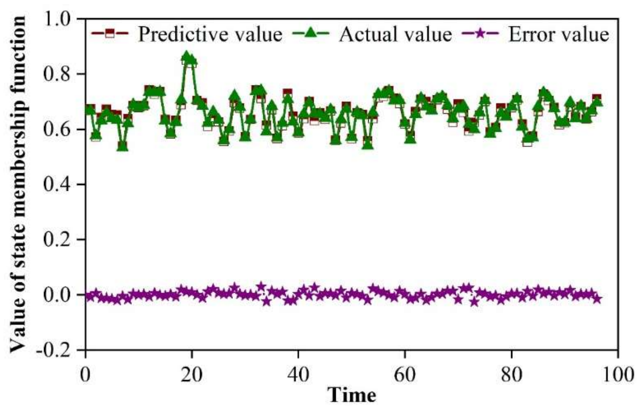



**Fig. 11.** Error analysis of state membership function based on CNN

### *5.3. Smart Grid Fault Diagnosis Experiments*

In order to validate the overall effectiveness of the R-GNN-based smart grid fault diagnosis method, it needs to be tested. In this section, the 10kV smart microgrid described in the previous section is chosen as an arithmetic example to verify the performance of the fault diagnosis model.

5.3.1. Accuracy of fault diagnosis. The PCA algorithm is utilized to reduce the dimension of the original 1000-dimensional sampling points by retaining 99% of the variance of the original dataset, which improves the efficiency of the model without affecting the performance of the model, and according to the experiments, it is proved that the accuracy of the model test is high and the operation rate is greatly improved when the mapped dimension is 50 dimensions. At the same time, three-phase current and voltage data without any processing are selected to compare CNN and support vector machine (SVM) for fault diagnosis to verify the advantages of the R-GNN method proposed in this paper.

The number of 1500, 3000 and 4500 training samples are selected to train the model, and 500 sets of test samples are used to test the accuracy of the model fault diagnosis. The fault diagnosis test results are shown in Fig. 12, which shows that as the number of training samples increases, the accuracy of the fault diagnosis model increases. In addition, under the same number of training samples, the performance of R-GNN model is higher than that of CNN and SVM, but it is slightly less than that of PCA+ R-GNN, and the fault identification rate of R-GNN model after PCA dimensionality reduction is as high as 98.91%. And after PCA dimensionality reduction the model running rate is greatly improved, the average training time is about 50s, which is more conducive to timely fault detection.

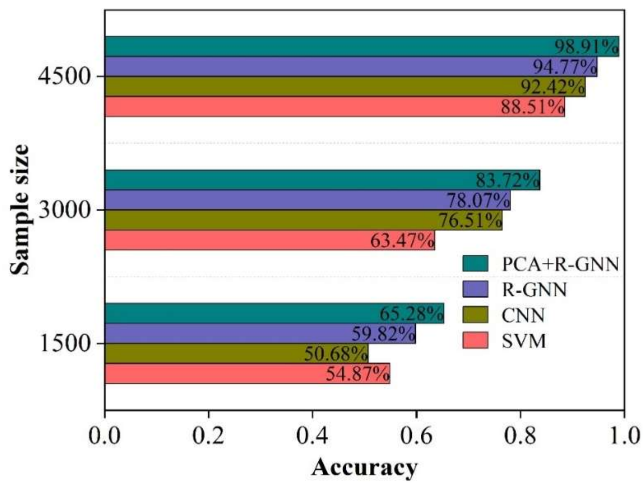



**Fig. 12.** Accuracy of fault type diagnosis

5.3.2. Time delay for fault diagnosis. The proposed R-GNN method and SVM, LSTM, CNN and GNN methods are used respectively to diagnose the smart grid faults of the above test objects, and record the time at which the different methods diagnose the faults of the grid (the data are pre-processed, re-sampled, and PCA downscaled). By obtaining the difference between the time when the different methods diagnosed the fault and the time when the fault actually occurred, it is judged whether there is a time delay in fault diagnosis by different methods. The difference between the time to diagnose the fault and the actual occurrence time of the fault by different methods is shown in Fig. 13.

When the proposed R-GNN method is used to diagnose complex grid chain faults, the difference between the diagnosis-to-fault time and the actual occurrence time of the fault is only 0.04s, and there is almost no time delay, which indicates that the fault diagnosis efficiency of the proposed R-GNN method is high. This is because the proposed method obtains grid fault diagnosis results by establishing a smart grid fault diagnosis model and combining the fault features within the model with the R-GNN algorithm. This diagnostic result obtained based on classification is more credible and efficient. Using CNN, GNN, LSTM and SVM to diagnose smart grid faults, the diagnosis-to-fault time and the actual occurrence time of the faults are all greater than 1s, with a large time delay, and the fault diagnosis efficiency is low. After the above comparison, it can be seen that the proposed R-GNN method's diagnosis efficiency of smart grid faults is significantly better than the traditional methods.

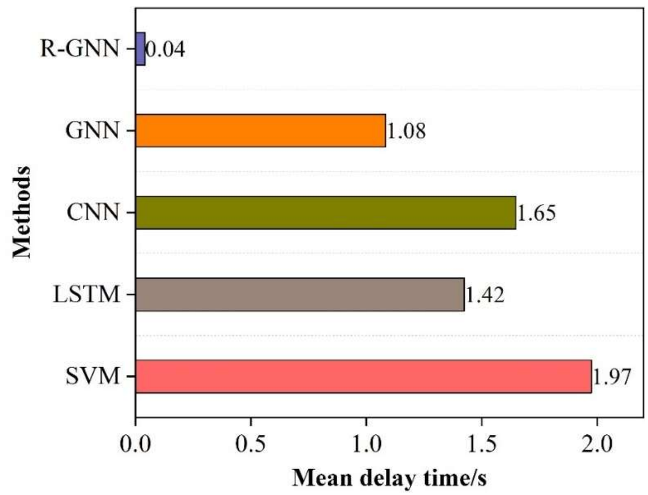



**Fig. 13.** Time delay of fault diagnosis by different methods

5.3.3. Ability to localize fault areas. The local grid nodes and links are traced using a computer and plotted on a 2D plan view. The 2D plan view of the local grid nodes and links of the test subject is shown in Fig. 14.

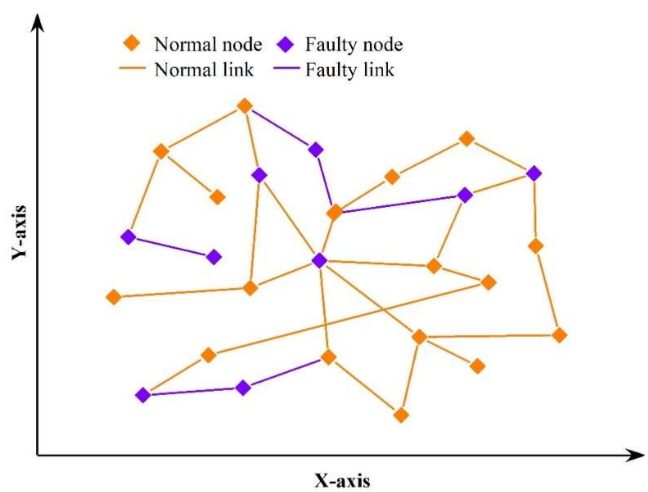
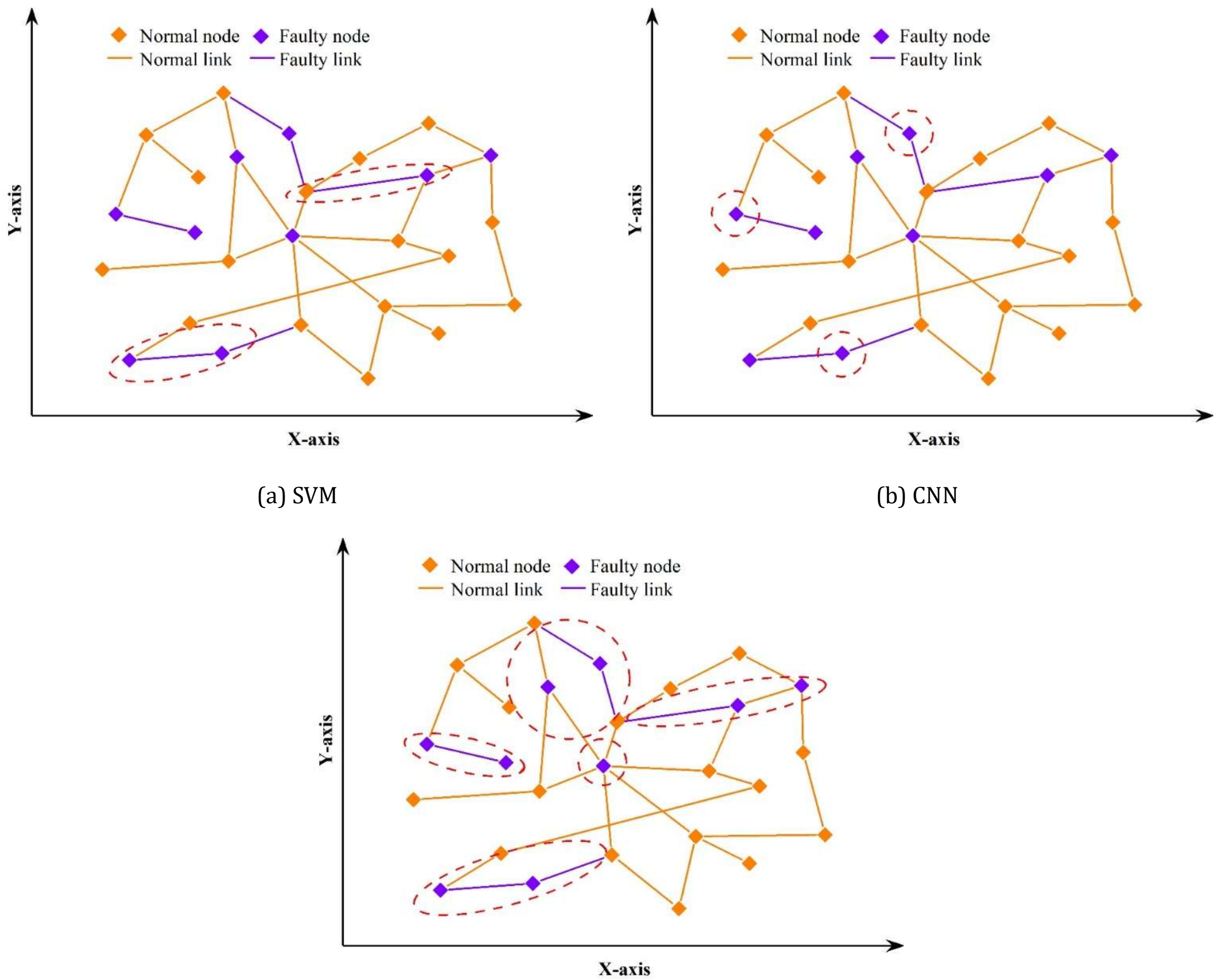



**Fig. 14.** Two-dimensional planar diagram of local power network nodes and links

The proposed R-GNN method, CNN method and SVM method are used to diagnose the grid faults of the experimental subjects and localize the fault areas, respectively. The fault region localization results of different methods are shown in Fig. 15. Where, (a)~(c) denote the three modeling methods of SVM, CNN and R-CNN, respectively.

The proposed R-GNN method is used to localize the fault region of the smart grid with high localization accuracy and high fault region coverage, which indicates the strong diagnostic performance of the R-GNN method for smart grid faults. The CNN method and SVM method are used to locate the smart grid fault region, and both of them have low localization accuracy and low fault region coverage, indicating that these two methods have poor diagnostic performance for smart grid faults. After the above comparisons, it can be seen that the proposed R-GNN method has better diagnostic performance for smart grid faults compared to traditional methods.




(a) SVM

(b) CNN

(c) R-GNN

**Fig. 15.** Fault area location results of different methods

## 6. Conclusion

In this paper, the processing of grid operation data is realized through data cleaning, normalization, resampling and PCA dimensionality reduction, and the operation state prediction model and fault diagnosis model are constructed to mine and analyze the multidimensional data to realize the integration of operation and maintenance of smart grid.

The 10kV smart microgrid in a region is selected as an example, and the normalized multidimensional data of the grid is within 0~1, which enhances the effectiveness of the training and optimization results while facilitating viewing. The resampling method, on the other hand, reduces the redundant information in the original data and increases the information density of the data about the characteristics of the grid operation state, which has the effect of data enhancement. PCA, on the other hand, reduces the dimensionality of the original data channel, which reduces the amount of training data while retaining the main frequency domain features of the original data channel.

The CNN-based power operation state monitoring and prediction model realizes the accurate prediction of five operation states: normal operation, system alert, general emergency, severe emergency and system collapse, and the difference between the model prediction value and the actual value is very small, corresponding to the average absolute error, mean square error and root mean square error of 0.00104, 0.00014 and 0.012, respectively, which are all within the reasonable range. The data after dimensionality reduction using PCA are

Using the PCA dimensionality reduction data, the fault recognition rate of the R-GNN-based smart grid fault diagnosis method is as high as 98.91%, which is better than that of CNN, SVM, and the R-GNN model using unprocessed data, illustrating the practicability of the R-GNN model for fault diagnosis and the effectiveness of the PCA data dimensionality reduction method. And the R-GNN method has only 0.04s delay for fault diagnosis and can realize comprehensive and accurate localization of the fault area, which is much better than the traditional method.

## References

[1] Rocchetta, R., Bellani, L., Compare, M., Zio, E., & Patelli, E. (2019). A reinforcement learning framework for optimal operation and maintenance of power grids. Applied energy, 241, 291-301.

[2] Qian, F., Liu, Y., Yang, Y., Gao, W., & Wu, Y. (2018). Equipment operation and maintenance management of shanghai power distribution network after power system reform. Energy Procedia, 152, 1182-1187.

[3] Keisang, K., Bader, T., & Samikannu, R. (2021). Review of operation and maintenance methodologies for solar photovoltaic microgrids. Frontiers in Energy Research, 9, 730230.

[4] Impram, S., Nese, S. V., & Oral, B. (2020). Challenges of renewable energy penetration on power system flexibility: A survey. Energy strategy reviews, 31, 100539.

[5] Bertsch, J., Brown, T., Hagspiel, S., & Just, L. (2017). The relevance of grid expansion under zonal markets. The Energy Journal, 38(5), 129-152.

[6] Brunner, C., Deac, G., Braun, S., & Zöphel, C. (2020). The future need for flexibility and the impact of fluctuating renewable power generation. Renewable Energy, 149, 1314-1324.

[7] Benidris, M., Mitra, J., & Singh, C. (2017). Integrated evaluation of reliability and stability of power systems. IEEE Transactions on Power Systems, 32(5), 4131-4139.

[8] Tavakoli, M., & Nafar, M. (2021). Reduce maintenance costs by improving human reliability in power grids. Electrical Engineering, 103(4), 1911-1921.

[9] Madadi, S., Nazari-Heris, M., Mohammadi-Ivatloo, B., & Tohidi, S. (2018). Application of big data analysis to operation of smart power systems. Big Data in Engineering Applications, 347-362.

[10] Gao, Y., Sun, K., Qiu, W., Li, Y., Li, M., & Sun, Y. (2024). Data-Driven Multidimensional Analysis of Data Reliability's Impact on Power Supply Reliability. IEEE Transactions on Industrial Informatics.

[11] Ma, Z., Ren, Y., Xiang, X., & Turk, Z. (2020). Data-driven decision-making for equipment maintenance. Automation in Construction, 112, 103103.

[12] Kovacic, I., Schuetz, C. G., Neumayr, B., & Schrefl, M. (2022). OLAP Patterns: A pattern-based approach to multidimensional data analysis. Data & knowledge engineering, 138, 101948.

[13] Pan, X., Wang, J., Li, Y., & Xie, W. (2018, December). Multidimensional data analysis of load influencing factors in smart distribution network. In IOP Conference Series: Earth and Environmental Science (Vol. 199, No. 5, p. 052015). IOP Publishing.

[14] Qiao, J., Xu, M., Zhou, A., Peng, L., Pan, S., & Yang, P. (2024, December). Research on multidimensional data analysis method of electric power based on tensor network model. In International Conference on Electronics, Electrical and Information Engineering (ICEEIE 2024) (Vol. 13445, pp. 177-183). SPIE.

[15] Liu, Z., Sheng, W., & Du, S. (2018, October). Multidimensional Data Model and Analysis Method of Economic Operation in Distribution Network. In Proceedings of the 10th Asia-Pacific Power and Energy Engineering Conference (APPEEC 2018), Kota Kinabalu, Malaysia (pp. 7-10).

[16] Tian, Y., & Ma, W. (2020, August). Real-time Multidimensional Data Mining and Analysis Technology Based on Big Data. In 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA) (pp. 747-752). IEEE.

[17] Botao Ning, Liang Zeng, Kaidi Fan & Feiyu Chen. (2025). Low-speed impact localization of wind turbine blades with a single sensor utilizing multiscale feature fusion convolutional neural networks.Ultrasonics107598-107598.

[18] Yuan Jiang, Chen Shanxiong, Mo Bofeng, Ma Yuqi, Zheng Wenjun & Zhang Chongsheng. (2024). R-GNN: recurrent graph neural networks for font classification of oracle bone inscriptions. Heritage Science(1).

[19] Ziqing Chen, Jianguo Chen, Zhicheng Zhu, Jian Chen, Taolin Lv, Dongdong Qiao & Yuejiu Zheng. (2025). Cloud-based estimation of lithium-ion battery life for electric vehicles using equivalent circuit model and recurrent neural network. Journal of Energy Storage(PA),115718-115718.

[20] Minh Ly Duc, Petr Bilik & Radek Martinek. (2025). White blood cell classification using multi-hop attention graph neural networks. Expert Systems With Applications126725-126725.