



Research on multi-sensor data fusion and optimization algorithm in grid flooding intelligent warning system

Mengyuan Huang^{1,*}, Guiqiao Huang¹, Guang Wang¹, Jiawen Huang¹, Hai Li¹, Guoliang Li¹, Liwen Huang¹ and Chuming Mo¹

¹ Guangdong Power Grid Yunfu Luoding Power Supply Bureau, Yunfu, Guangdong, 527300, China

SUMMARY: *Aiming at the deficiencies of the traditional covariance cross-fusion algorithm in terms of computational complexity and fusion accuracy, this paper proposes an improved multi-sensor data fusion and optimization algorithm for enhancing the accuracy and speed of the grid flooding intelligent warning system. The method combines the Drosophila optimization algorithm to optimize the fusion coefficients, the dynamic weight allocation method based on information entropy and the attention mechanism to enhance the ability of capturing the key data features to improve the robustness and adaptability of the system. The efficient fusion of data and decision layers is achieved by designing a two-stage fusion method combined with Kalman filtering, adaptive weighted averaging and ELM neural network with ARO optimization. The results show that the proposed algorithm outperforms traditional methods in terms of root mean square error, fusion time, and warning accuracy, especially in high-noise environments and large-scale sensor networks. The study provides a reliable technical support for the grid flooding intelligent warning system, which has important theoretical value and practical application prospects.*

KEYWORDS: *multi-sensor fusion; dynamic weight allocation; attention mechanism; two-level fusion; grid flood warning*

1 Introduction

In recent years, with the rapid development of social and economic development and industrialization, the demand for electricity has increased dramatically, and the grid intelligent early warning system has become the key to ensure the safe and stable operation of power equipment [1]. The traditional manual inspection and static monitoring methods have been difficult to meet the demand for real-time, accuracy and reliability of the modern power grid. Internet of Things (IoT) technology, with its efficient data acquisition, transmission and processing capabilities, has gradually become the core technology for power grid monitoring and management, and the introduction of intelligent early warning system can realize the early identification of faults and rapid response, and effectively improve the stability and security of the power grid [2, 3].

The technology for real-time collection of grid condition data relies on Wireless Sensor Networks (WSNs), which utilize distributed sensor nodes to acquire electrical and environmental data from various monitoring points [4]. Sensor nodes are generally designed with low power consumption for long-term stable operation. The transmission of data after

*huangmengyuan9326@163.com
<https://doi.org/10.65102/is2026650>

acquisition is carried out with the help of wireless communication protocols, commonly used communication protocols include LoRa, NB-IoT, and Zigbee. LoRa (Long Range) technology is suitable for data transmission between sensor nodes that are widely distributed in the grid due to its long range, low power consumption, and low bandwidth characteristics [5]. NB-IoT (Narrowband Internet of Things) technology with the help of transmitting data in low frequency bands has stronger penetration and wide network coverage, which is suitable for grid monitoring in urban and remote areas [6, 7]. Zigbee protocol is suitable for low-power, short-distance data transmission in localized areas [8]. Communication protocols for grid monitoring systems need to be selected according to the actual application scenarios in order to guarantee the reliability, real-time and energy-saving of data transmission [9, 10]. Real-time data analysis is an indispensable part of the intelligent early warning system, which relies on IoT technology to assess the operational status of the power grid based on continuously collected real-time data. In terms of data analysis, the use of time series data analysis method, dynamic modeling of various indicators of the power grid, timely detection of abnormal fluctuations, based on the trend analysis of the data and the combination of historical data and real-time data, can achieve the prediction of the state of the power grid and the early warning of faults [11, 12].

Due to the differences in physical principles, working environments, and performance of different sensors, the data collected are usually redundant, contradictory, or even conflicting, thus affecting the effective utilization and accurate analysis of the data. In order to overcome these deficiencies and improve the accuracy and efficiency of data processing, multi-sensor data fusion algorithms have emerged. Multi-sensor data fusion aims at realizing data complementarity and enhancement by integrating information from different sensors, thus forming a more comprehensive, accurate and reliable understanding of the monitoring object or environment [13, 14]. The role of data fusion is mainly reflected in [15, 16]: first, the main energy consumption of sensor nodes is in data acquisition, data computing, and data transmission, and the energy consumed by sensor nodes in a wireless sensor network for data transmission is much larger than that consumed by computation within the nodes. Second, the sensor network consists of a large number of low-cost deployment in a variety of environments sensor nodes, sensor nodes transmitted data has a high degree of unreliability, the harsh operating environment may damage the function of the node, making the node in the abnormal state of work can not send the correct data. Multi-sensor data fusion algorithms and optimization can effectively eliminate redundancy and conflict between data, and currently become one of the key technologies to achieve high-precision sensing and decision-making in the field of intelligent monitoring and early warning of power grid faults [17, 18]. To this end, the research is carried out on the basis of multi-sensor data fusion technology, introducing the fruit fly optimization algorithm to improve the traditional covariance cross-fusion algorithm, and then calculating the quality index of each sensor data based on the data fusion algorithm of dynamic weight allocation and attention mechanism. At the same time, for the problem of diverse sensor data types and inconsistent quality in the grid flooding intelligent warning system, a two-level fusion method is proposed to improve the reliability and robustness of the system.

2 Related work

2.1 Multi-sensor data fusion technology

Multi-sensor data fusion technology is a method of comprehensively processing multi-source information, which realizes information complementarity and redundancy by integrating data collected by multiple sensors, thus obtaining more comprehensive and accurate monitoring results than a single sensor. With the development of Internet of Things and artificial

intelligence technology, this technology has turned out to be the key to solving complex environmental monitoring problems, and it has a wide range of applications in the fields of electric power system safety monitoring, industrial equipment fault diagnosis, and intelligent transportation.

The research on multi-sensor data fusion in power grid systems started earlier and is therefore relatively perfect. In recent years, the research mainly focuses on the problem of grid fault monitoring. Jiao et al [19] (2018) established a transmission line fault localization model based on multi-sensor data fusion in response to the impact of power line faults on the stability of the power system, and launched performance tests in different fault types, fault locations, and fault resistances, and achieved excellent experimental results. De Farias et al [20] (2019) proposed a multi-sensor data fusion technique that can divide monitoring intervals into multiple non-overlapping and overlapping intervals and effectively identify potential correlations in abstract sensors, and the results of simulation experiments and node tests proved the value of the application of this technique in terms of saving energy and extending the operational life. CAO et al [21] (2022) in power system inspection UAVs incorporates a multi-sensor fusion data analysis model, which has an early warning performance ratio and efficiency ratio of more than 97%, and the failure rate of fault detection and processing time are reduced by 10.2% and 7.8%, respectively. Jiang [22] (2022) utilizes data fusion of multiple sensors in combination with the output plausibility ranking for location localization of interruption fault scenarios in the distribution system, and the computational efficiency of the proposed method is higher than that of the existing methods, and the judgment accuracy is much higher than the state-of-the-art technology. Kang et al [23] (2024) used the constructed multivariate data fusion fault analysis model based on the fault dataset of the distribution network for the warning training, and through the experimental analysis, it can be seen that the grid equipment model based on the multisensor data fusion algorithm has a better fit and effect on the warning of faults, and the prediction accuracy reaches over 99%, which is highly practical. , which has a high degree of practicality.

The basic principle of multi-sensor data fusion is to detect, combine, correlate, estimate and combine multi-source data through multi-level and multi-faceted data processing process to achieve accurate state estimation and real-time assessment. According to different fusion levels, it can be divided into three levels: data layer fusion, feature layer fusion and decision layer fusion. Data layer fusion directly processes the original sensor data, feature layer fusion fuses the extracted features, and decision layer fusion fuses each sensor to make decisions independently. From the mathematical point of view, multi-sensor data fusion can be expressed as:

$$Y = F(X_1, X_2, \dots, X_n) \quad (1)$$

where X_n denotes the first n sensor output, F denotes the fusion function, and Y is the fusion result. The covariance cross-fertilization algorithm is a state fusion method based on Bayesian estimation, which is suitable for processing multi-sensor data with correlation, and the basic form of this algorithm is:

$$P = (\sum_{i=1}^n P_i^{-1})^{-1} \quad (2)$$

where X_i and P_i are the state estimation and covariance matrix of the i th sensor, respectively, and \hat{X}_i and \hat{P}_i are the state estimation and covariance matrix after fusion, respectively.

Dynamic weight assignment is a method to dynamically adjust the weights according to the quality of sensor data, and its mathematical expression is:

$$w_i = \frac{\exp(-\alpha \cdot \sigma_i^2)}{\sum_{j=1}^n \exp(-\alpha \cdot \sigma_j^2)} \quad (3)$$

where w_i is the weight of the i th sensor, σ_i^2 is the measurement variance of this sensor, and α is the adjustment parameter. In the study of multi-sensor data fusion in IoT environment, the fusion optimization algorithm model is proposed by combining dynamic weight allocation, attention mechanism and conflict elimination strategy. The attention mechanism can be expressed as:

$$A_i = \text{soft max}(W_q X_i \cdot (W_k X_i)^T) \cdot (W_u X_i) \quad (4)$$

where W_q , W_k , W_u are the weight matrices of the query, key and value, respectively, and X_i are the input features.

Neural networks are widely used in multi-sensor data fusion due to their powerful nonlinear fitting ability and adaptive learning ability. Commonly used neural network models include BP neural network, RBF neural network and ELM neural network.

The attention mechanism can be expressed as:

2.2 Application of Optimization Algorithms in Data Fusion

Optimization algorithms play a key role in the field of data fusion by finding the best combination of parameters to enhance the fusion effect, effectively solving the problems of difficult parameter selection and insufficient fusion accuracy in traditional data fusion methods. With the increasing complexity of multi-sensor systems, optimization algorithms are increasingly widely used.

Intelligent algorithms have made rapid development in recent years, and various intelligent algorithms such as back propagation neural networks and convolutional neural networks have made significant progress in the direction of pattern recognition, data prediction, etc., and some researchers have taken an alternative approach to apply intelligent algorithms to multi-sensor data fusion technology [24]. Zhen et al [25] (2018) calculated the multi-sensor data fusion algorithm in the power grid system by asynchronous clustering algorithm sensor most group number to reduce the data flow of the grid, based on the principle of multi-sensor information fusion and detection to achieve multi-source information fusion in the grid system, the method demonstrates high information fusion accuracy and stability, and can reduce the loss of the grid. Sun et al [26] (2019) realized the abnormal and noisy data through the improved genetic algorithm degree multi-sensor data fusion algorithm and you character optimization of effective elimination, dramatically improving the accuracy of data fusion and reducing the execution time of the algorithm. Li et al [27] (2025) combined multi-sensor data fusion and deep learning methods to break the barriers of uneven point distribution and environmental interference, and the recognition accuracy of the grid pylons in different environments reaches 90.8%, which provides a strong support for the intelligent detection and management of power grid facilities.

Nowadays, intelligent algorithms have been widely used in a variety of data fusion scenarios, however, they are still difficult to use in practice.

Fruit fly optimization algorithm (FOA) is a kind of intelligent algorithm simulating the foraging behavior of fruit flies, which has the characteristics of few parameters, simple implementation and fast convergence, and its basic mathematical model is:

$$X_{axis} = X_0 + \text{Random} \cdot \text{Value} \quad (5)$$

$$Y_{axis} = Y_0 + \text{Random} \cdot \text{Value} \quad (6)$$

$$Dist_i = \sqrt{X_i^2 + Y_i^2} \quad (7)$$

$$S_i = \frac{1}{Dist_i} \quad (8)$$

$$Smell_i = \text{Function}(S_i) \quad (9)$$

where X_{axis} and Y_{axis} represent the population position, $Dist_i$ is the distance of *Drosophila* individuals to the origin, S_i is the value of the odor concentration determination, and $Smell_i$ is the value of the fitness function. .

Particle Swarm Optimization (PSO) algorithm is a stochastic search algorithm based on group intelligence, which simulates the foraging behavior of bird flocks and uses the information shared by the particles in the group to find the optimal solution, and its iterative update formula is:

$$v_{id}^{t+1} = w \cdot v_{id}^t + c_1 \cdot r_1 (p_{id} - x_{id}^t) + c_2 \cdot r_2 (p_{gd} - x_{id}^t) \quad (10)$$

where v_{id}^{t+1} and x_{id}^t denote the d -dimensional velocity and position of the i th particle at the t th iteration, respectively, w is the inertia weight, c_1 and c_2 are the learning factor, r_1 and r_2 are random numbers between $[0,1]$, p_{id} is the individual optimal position, and p_{gd} is the global optimal position.

2.3 Application of Neural Networks in Data Fusion

Neural network has become an important technology in the field of multi-sensor data fusion because of its strong nonlinear mapping ability, good adaptive learning and high fault tolerance. It can effectively deal with the uncertainty, incompleteness and conflict problems in sensor data and realize high-precision fusion. The development of deep learning technology makes neural networks more and more widely used in data fusion, especially in the processing of high-dimensional complex data, BP neural network is a kind of classical feed-forward neural network, trained with the error back propagation algorithm, with strong nonlinear fitting ability. However, the traditional BPNN encounters dimensional catastrophe when dealing with high-dimensional data, and it is easy to fall into the local optimum, to address these problems, a BP neural network model based on rough set theory (RS-BPNN) is proposed. This model uses rough set attribute approximation algorithm to feature mine the system data and reduce the system data

dimensionality without affecting the fusion effect, and then uses BP neural network to fuse the minimal approximation after the approximation.

The RS-BPNN model is mathematically expressed as:

$$y_i = f\left(\sum_{i=1}^m \varpi_{ij} \cdot x'_i + \theta_j\right) \quad (11)$$

where x'_i is the input feature after rough set approximation, ϖ_{ij} is the connection weight, θ_j is the threshold, and $f(\cdot)$ is the activation function. The rough set approximation process can be expressed as:

$$\gamma_C(D) = \frac{|POS_C(D)|}{|U|} \quad (12)$$

where $\gamma_C(D)$ is the dependence of the attribute set C on the decision attribute D , $|POS_C(D)|$ is the positive domain of C with respect to D , and U is the argument domain.

The case validation of water quality class evaluation shows that the RS-BPNN model improves 43.2% and 35.8% in fusion accuracy and speed, respectively, than the traditional BPNN. Radial basis function neural network is a kind of feed-forward neural network with strong local approximation ability, and its output is:

$$y_k = \sum_{j=1}^h \omega_{jk} \cdot \phi\left(\|x - c_j\| / \sigma_j\right) \quad (13)$$

Among them, $\phi(\cdot)$ is the radial basis function, usually chosen as Gaussian function, c_j is the center vector of the node of the j th hidden layer, σ_j is the width parameter, and ω_{jk} is the weight of the connection from the hidden layer to the output layer.

The traditional RBFNN parameters are difficult to choose to affect the network performance. Zhiwei Xue proposes a particle swarm optimization based RBFNN (PSO-RBFNN) model, which uses PSO algorithm to find the optimal network parameters including center vector, width parameter and weights. The PSO-RBFNN optimization objective function is:

$$J = \frac{1}{N \sum_{i=1}^N \sum_{k=1}^m (y_{ik} - \hat{y}_{ik})^2} \quad (14)$$

where y_{ik} is the actual output, \hat{y}_{ik} is the network output, N is the number of samples, and m is the number of output dimensions. The air quality index case validates the effectiveness of the method, and the fusion accuracy is improved by 39.4% over the traditional RBFNN.

3 Modeling

3.1 Improvement of the covariance cross-fertilization algorithm

Covariance cross-fertilization algorithms are widely used in the field of multi-sensor data fusion, especially when dealing with correlated sensor data, the advantages are obvious. However, the traditional algorithm is too complicated to compute when facing large-scale sensor networks, and the fusion accuracy is not satisfactory enough. Therefore, this study improves the covariance cross-fusion algorithm by introducing the fruit fly optimization algorithm in to

optimize the fusion coefficients, which improves the accuracy and computational speed of data fusion. With the increase in the number of sensors in practical applications, the matrix inverse operation leads to a sharp increase in computational complexity, and it is also easy to introduce numerical errors.

In this paper, the covariance cross-fusion algorithm is improved, and the state update equation is expressed as in the Kalman filter framework:

$$k|k = \hat{X}_{k|k-1} + K_k (Z_k - H_k \hat{X}_{k|k-1}) \quad (15)$$

The covariance matrix update equation is:

$$P_{k|k} = (I - K_k H_k) P_{k|k-1} \quad (16)$$

where K_k is the Kalman gain, H_k is the observation matrix, and $P_{k|k-1}$ is the prediction covariance matrix.

This study transforms the multi-sensor data fusion problem into a convex combinatorial optimization problem:

$$\sum_{i=1}^n \omega_i = 1, \omega_i \geq 0 \quad (17)$$

where ω_i is the fusion coefficient of the i th sensor. The key problem is to determine the optimal fusion coefficients so that the fusion result achieves the minimum mean square error.

To solve this problem, we introduce the fruit fly optimization algorithm (FOA) to optimize the fusion coefficients. The algorithm simulates the foraging behavior of *Drosophila*, with few parameters, simple implementation and fast convergence.

The process of the improved covariance cross-fertilization algorithm includes initializing the *Drosophila* population position and parameters, randomly generating individual *Drosophila* positions, calculating the distance of individuals to the origin and the judgment value of odor concentration, substituting into the fitness function to calculate the fitness values of each individual, finding the optimal fitness value and the corresponding *Drosophila* individual, retaining the optimal *Drosophila* position, and the other *Drosophila* fly towards the optimal position, repeating the above steps until reaching the maximum number of iterations, and finally outputting the optimal fusion coefficients. The above steps are repeated until the maximum number of iterations is reached, and finally the optimal fusion coefficient is output.

3.2 Dynamic weight allocation and attention mechanisms

In multi-sensor data fusion systems, the measurement accuracy, reliability and environmental adaptability of different sensors differ significantly, and in order to give full play to the advantages of each sensor and suppress the defects, the dynamic weight allocation and attention mechanism has become a key technology to optimize the effect of data fusion.

The traditional static weight allocation method usually relies on a priori knowledge to preset fixed weights, which is difficult to adapt to the dynamic changes in sensor performance and environmental conditions. The dynamic weight allocation mechanism, on the other hand, can adaptively adjust the weight coefficients according to the real-time state of the sensor and data quality, and its basic mathematical model is:

$$\omega_i = f(q_i, e_i, r_i, t) \quad (18)$$

where ω_i represents the first i sensor weight, q_i denotes the data quality factor, e_i denotes the estimation error, r_i denotes the sensor reliability, and t is the time variable.

In this study, the dynamic weight assignment method based on measurement variance is used, and the mathematical expression is:

$$\omega_i^{\text{var}} = \frac{1/\sigma_i^2}{\sum_{j=1}^n 1/\sigma_j^2} \quad (19)$$

where σ_i^2 represents the measurement variance of the i th sensor, which gives higher weights to low variance (high accuracy) sensors but performs poorly in the presence of outliers or systematic biases.

To solve the above problem, we introduce a dynamic weight assignment method based on information entropy:

$$\omega_i^{\text{ent}} = \frac{1-H_i}{\sum_{j=1}^n (1-H_j)} \quad (20)$$

where H_i is the normalized information entropy of the i th sensor data, calculated as:

$$H_i = -\sum_{k=1}^m p_{ik} \log p_{ik} \quad (21)$$

where p_{ik} is the probability of the i th sensor's k state.

The attention mechanism originates from the human visual system, which can automatically focus on important information, and highlights the contribution of key sensor data by calculating the attention weights in multi-sensor data fusion. Combining the attention mechanism, we propose an improved dynamic weight distribution formula. When α is larger, the weight distribution is more concentrated on the sensors with high attention scores, and when α is smaller, the weight distribution is more uniform. The data fusion algorithm based on dynamic weight allocation and attention mechanism. The process includes, calculating the quality indicators (variance, information entropy, etc.) of each sensor's data, using the attention mechanism to calculate the attention score of each sensor, calculating the dynamic weights based on the attention score, and finally using the dynamic weights for weighted fusion.

3.3 Two-level fusion methods

Aiming at the problem of diverse sensor data types and inconsistent quality in the grid flooding intelligent warning system, this study proposes a two-level fusion method to improve the system reliability and robustness. The method utilizes Kalman filtering and adaptive weighted average algorithm for one-level fusion processing when dealing with the same type of sensor data. The mathematical expression of Kalman filtering for noise reduction processing of single sensor data is:

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k (Z_k - H_k \hat{X}_{k|k-1}) \quad (22)$$

$$P_{k|k} = (I - K_k H_k) P_{k|k-1} \quad (23)$$

where $\hat{X}_{k|k}$ represents the state estimate, $\hat{X}_{k|k-1}$ is the predicted value, K_k denotes the Kalman gain, Z_k is the observed value, H_k represents the observation matrix, and $P_{k|k}$ and $P_{k|k-1}$ denote the estimated and predicted covariance, respectively.

For the same type of multi-sensor data, the adaptive weighted average algorithm is used for fusion:

$$\hat{X}_1 = \sum_{i=1}^n w_i \hat{X}_i \quad (24)$$

where \hat{X}_1 is the first level fusion result, \hat{X}_i denotes the filtered data of the i th sensor, and w_i stands for the adaptive weights, which are calculated as:

$$w_i = \frac{1/\sigma_i^2}{\sum_{j=1}^n 1/\sigma_j^2} \cdot \alpha_i \quad (25)$$

where σ_i^2 is the measurement variance of the i th sensor, and α_i denotes the reliability coefficient dynamically adjusted according to the historical performance of the sensor. The second-level fusion then uses the ARO-optimized ELM neural network for decision fusion, and the basic ELM model is:

$$f(x) = \sum_{i=1}^L \beta_i g(w_i \cdot x + b_i) = H\beta \quad (26)$$

where $g(\cdot)$ is the activation function, w_i and b_i represent the input weights and bias, respectively, β_i denotes the output weights, and H stands for the hidden layer output matrix.

The random generation of input weights and biases by traditional ELM leads to unstable network performance, so the ARO algorithm is introduced to optimize the parameters, and the ARO algorithm simulates the rabbits' behavior of foraging, alerting and escaping, and the mathematical model of foraging phase is:

$$X_{i,d}^{t+1} = X_{i,d}^t + r_1 \cdot (X_{best,d}^t - X_{i,d}^t) + r_2 \cdot (X_{rand,d}^t - X_{i,d}^t) \quad (27)$$

The alert phase is:

$$X_{i,d}^{t+1} = X_{i,d}^t + r_3 \cdot (X_{best,d}^t - X_{i,d}^t) \cdot F_i \quad (28)$$

Escape stage for:

$$X_{i,d}^{t+1} = X_{i,d}^t + r_4 \cdot L \cdot (X_{rand,d}^t - X_{i,d}^t) \quad (29)$$

where $X_{i,d}^t$ represents the d -dimensional position of the i rabbit at the t th iteration, $X_{best,d}^t$ is the current optimal position, $X_{rand,d}^t$ represents the random position, and r_1 to r_4 is the number of $[0,1]$ intervals. random numbers, F_i denotes the vigilance factor, and L is the escape step.

The ARO-ELM optimization objective function is:

$$J = \frac{1}{N} \sum_{i=1}^N \|f(x_i) - y_i\|^2 + \lambda \|\beta\|^2 \quad (30)$$

where N is the number of samples, y_i denotes the target output, and λ is the regularization parameter.

4 Experiments and analysis of results

4.1 Experimental design

In order to verify the effectiveness of the multi-sensor data fusion and optimization algorithm proposed in this paper in the grid flooding intelligent early warning system, a complete set of experimental scheme is designed, including the design of the experimental environment, data acquisition, data preprocessing and algorithm implementation.

The experimental environment consists of two parts: hardware and software. The hardware environment includes a computing platform with Intel Core i7-12700K CPU, 32GB RAM and NVIDIA RTX 3080Ti GPU, a distributed sensor network consisting of 20 water level sensors, 15 temperature sensors, 15 humidity sensors and 10 pressure sensors, and an NI-DC with a sampling frequency of 100Hz on a NI-DAQ data acquisition card.

For the software environment, MATLAB R2023a was used for algorithm implementation and simulation, Python 3.9 was responsible for data preprocessing and visualization, and LabVIEW 2022 was used for sensor data acquisition and control. The data acquisition work is divided into two parts: the laboratory simulation data and the actual grid environment data, the former through the 1:10 scale grid equipment miniature model simulation of different degrees of flooding scenarios, the latter from a provincial power company to obtain the three substations, the sensor arrangement follows the principle of “multi-level, multi-angle, full coverage,” around the key equipment, set up high-density sensor arrays, the non-sensor array, and the non-sensor array is set up. The sensor arrangement follows the principle of “multi-level, multi-angle, full coverage”, with high-density sensor arrays around key equipment and low-density arrangements in non-critical areas. Specific sensor parameters are shown in Table 1, the data acquisition period is 10 minutes, each acquisition lasts for 30 seconds, the sampling frequency is 100Hz, in order to simulate the interference and abnormal conditions in the real environment, some experiments artificially introduce noise and sensor failure, the data set contains normal state data (70%), slight flooding data (20%) and severe flooding data (10%), the total collection time is 30 days, the cumulative data volume is about The total collection time is 30 days and the accumulated data volume is about 250GB.

Table 1: Sensor Parameters Used in the Experiment

Sensor type	Model	Measurement range	Precision	Quantity
Water level sensor	WL-100	0-100 cm	$\pm 0.5\text{cm}$	20
Temperature sensor	PT100	$-50^\circ\text{C} - 150^\circ\text{C}$	$\pm 0.1^\circ\text{C}$	15
Humidity sensor	HS-220	0-100%RH	$\pm 2\%\text{RH}$	15
Pressure sensor	PS-500	0-500kPa	$\pm 0.5\%$	10

The original sensor data have problems such as noise, outliers and missing values, which need to be preprocessed to improve data quality. The data preprocessing flow is shown in Figure

1, and the specific preprocessing steps include noise removal, detection and processing of outliers, filling of missing values, data normalization, time alignment and feature extraction. Noise removal adopts wavelet transform denoising method, choosing dB5 wavelet, the number of decomposition layers is 4, and high-frequency coefficients are processed by soft thresholding method. Outliers detection and processing combines the quartile distance and local anomaly factor methods, and the thresholds are set to $Q_1 - 1.5 \times IQR$ and $Q_3 + 1.5 \times IQR$, and detected outliers are corrected by local linear interpolation method.

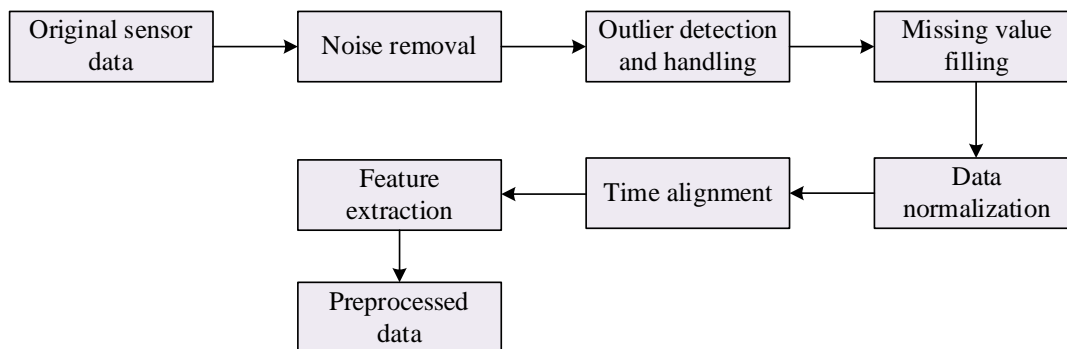


Figure 1: Data Preprocessing flow

The filling of missing values was handled by the multiple interpolation method, and the filling accuracy reached more than 95%. Data normalization was performed using the Z-score normalization method with the formula $x' = (x - \mu) / \sigma$, μ is the mean and σ is the standard deviation. Time alignment uses linear interpolation to align the data to a uniform time axis, and the sampling interval is 10 ms. feature extraction extracts time-domain features, frequency-domain features, and time-frequency features from the preprocessed data to form feature vectors for subsequent fusion algorithms.

Based on the preprocessed data, the multi-sensor data fusion and optimization algorithm proposed in this paper is implemented, and the algorithm implementation framework is shown in Fig. 2, and the main algorithm implementation includes improved covariance cross-fertilization algorithm, dynamic weight allocation and attention mechanism, neural network model, and two-stage fusion method. Improved covariance cross-fertilization algorithm is based on the traditional covariance cross-fertilization algorithm, and Drosophila optimization algorithm is introduced to optimize the fusion coefficients, with a population size of 50, and a maximum number of iterations of 100. Dynamic weight allocation and attention mechanism implements dynamic weight allocation method based on information entropy, and the dimensions of weight matrix of query, key value, and value in the attention mechanism is 64×64 , and the number of attention heads is 8. Neural network model implementation Two neural network models RS-BPNN and PSO-RBFNN, rough set attribute approximation in RS-BPNN using heuristic algorithm based on positive domain, BP neural network structure is input layer-hidden layer (20 nodes)-output layer, and the number of nodes of hidden layer of RBF neural network in PSO-RBFNN is 30. The two-stage fusion method in the first-stage fusion using Kalman filter and adaptive weighted averaging, and the ARO-ELM neural network structure in the second level fusion is input layer-hidden layer (50 nodes)-output layer (3 nodes).

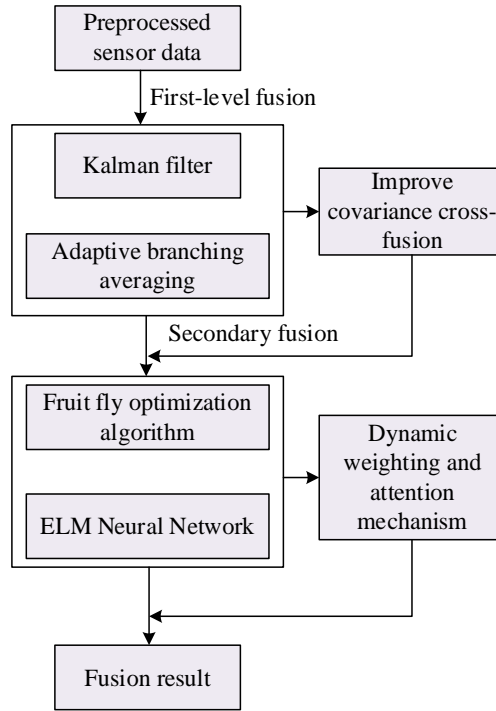


Figure 2: Algorithm implementation framework

The experiments use 5-fold cross-validation to evaluate the performance of the algorithms, and the performance metrics include root-mean-square error, fusion time, warning accuracy, false alarm rate, and response time, etc. Meanwhile, the algorithms are compared with the traditional covariance cross-fusion algorithms, the convex combinatorial fusion algorithms, and the static weighting method to validate the superiority of the algorithms proposed in this paper. Figure 3 demonstrates the percentage performance improvement of the three main algorithms proposed in this paper compared with the traditional methods, and the improved covariance cross-fertilization algorithm has the most significant improvement in fusion time and warning accuracy, reaching 43% and 52%, respectively. The dynamic weighting and attention mechanism excels in terms of false alarm reduction, reaching 43.8%. The two-level fusion method, on the other hand, shows a steady improvement in all indicators.

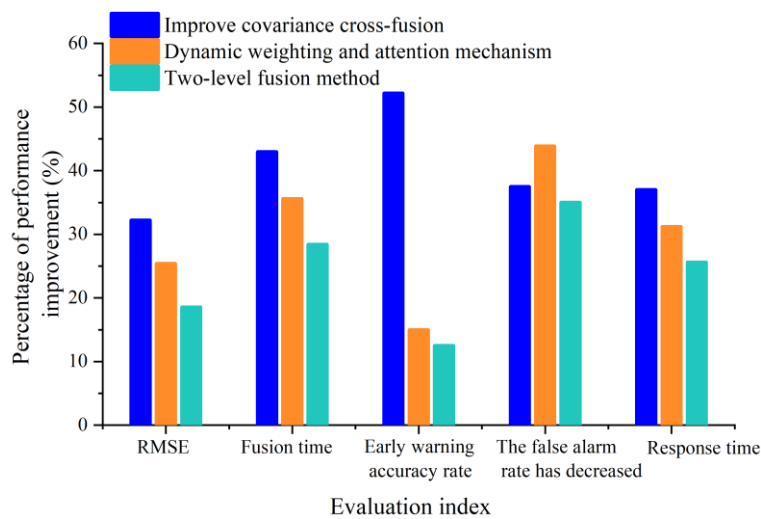


Figure 3: Comparison of algorithm performance

4.2 Experimental results and analysis

In this section, the effectiveness of the proposed multi-sensor data fusion and optimization algorithms is verified through a series of simulation experiments, focusing on the fusion accuracy and speed of each algorithm and comparing the performance of different algorithms. The comparison results of the root mean square error (m) of different fusion algorithms under different conditions are shown in Table 2, and the experimental data show that the improved covariance cross-fertilization algorithm exhibits the lowest root mean square error under various conditions, especially in the high-noise environment, and the advantage is more obvious, compared with the traditional covariance cross-fertilization algorithm, the improved algorithm reduces the RMSE by about 34.5% in the low noise environment, and reduces the RMSE by about 36.8% in the high-noise environment. Compared with the traditional covariance cross-fusion algorithm, the improved algorithm reduces the RMSE by about 34.5% in low noise environment and about 36.8% in high noise environment.

Table 2: Comparison of Root mean square Error (m) of Different Fusion Algorithms under different conditions

Algorithm	Three sensors		Five sensors	
	Low noise	Low noise	Low noise	Low noise
Traditional covariance cross-fusion	4.35	5.78	3.92	5.25
Convex combination fusion	3.42	4.52	3.18	4.23
Improve covariance cross-fusion	2.85	3.65	2.53	3.28

The comparison results of the computation time (s) of different fusion algorithms are shown in Table 3, which show that the improved covariance cross-fertilization algorithm not only improves the fusion accuracy, but also significantly reduces the computation time, especially when the number of sensors is increased, and the growth of its computation time is significantly smaller than that of the traditional algorithm, which suggests that the algorithm has a better scalability when dealing with large-scale sensor networks.

Table 3: Comparison of Computing Time (s) of Different Fusion Algorithms

Algorithm	Three sensors		Five sensors	
	100 iterations	200 iterations	100 iterations	200 iterations
Traditional covariance cross-fusion	0.76	1.45	1.24	2.36
Convex combination fusion	0.58	1.12	0.85	1.68
Improve covariance cross-fusion	0.41	0.82	0.47	0.95

The results of the comparison of the relevance and weight factors of different sensor data in the grid flood warning system are shown in Figure 4. The analysis shows that the dynamic weight allocation method combined with the attention mechanism can better reflect the relevance of the sensor data, and the weight allocation is more reasonable, higher weight values are given to the sensors with higher relevance, while the weights are appropriately reduced for the sensors with lower relevance, and this allocation effectively improves the accuracy of the fusion results.

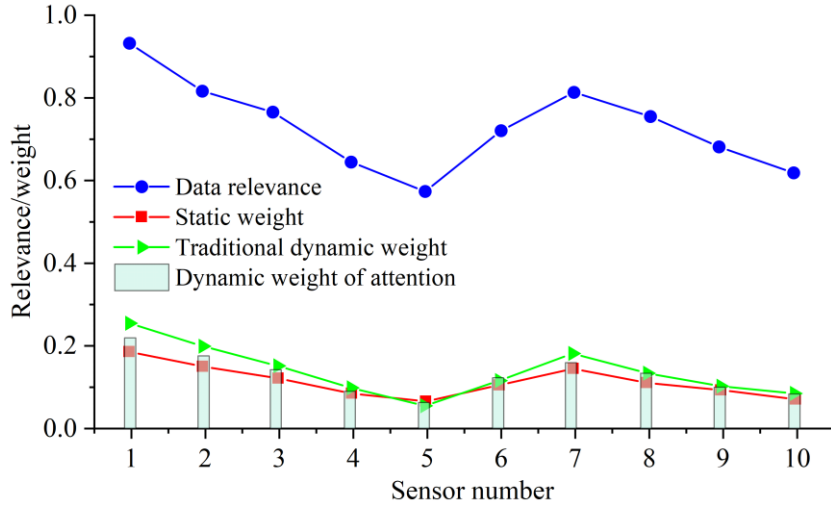


Figure 4: Comparison of data correlation and weight factors of different sensors

5 Conclusion

In this paper, an improved multi-sensor data fusion and optimization algorithm is proposed for the accuracy and speed of multi-sensor data fusion in the grid flooding intelligent warning system. By introducing the improved covariance cross-fertilization algorithm and combining with the fruit fly optimization algorithm to optimize the fusion coefficients, the fusion performance is effectively improved, and experiments show that the algorithm reduces the RMSE by about 36.8% and 34.5% and the computation time by 42.8% in high and low noise environments, respectively. Flooding intelligent warning system in the electric power, water industry already has practical application cases, the application of the State Grid provides a reference for the research and promotion of the field of power supply. At the same time, this early warning system is an effective solution for the development of regional power grid security, which has certain reference value.

About the Authors

Mengyuan Huang, born in 1993 in Yuncheng City, Shanxi Province, Master's degree holder from Taiyuan University of Technology, currently employed at Guangdong Power Grid Yunfu Luoding Power Supply Bureau, the main research direction is the application of artificial intelligence technology in the new power system.

Guiqiao Huang, born in Yunfu City, Guangdong Province in October 1982, graduated from Zhaoqing University in Guangdong Province with a bachelor's degree, currently employed at Guangdong Power Grid Yunfu Luoding Power Supply Bureau, the main research direction is the application of digital twin technology in the new power system.

Guang Wang, born in 1980 in Suizhou City, Hubei Province, Master's degree holder from South China University of Technology, currently employed at Guangdong Power Grid Yunfu Luoding Power Supply Bureau, the main research direction is the application of big data and big model technology in power grid management.

Jiawen Huang, born in 1986 in Yunfu City, Guangdong Province, Undergraduate degree from Guangdong University of Technology, currently employed at Guangdong Power Grid Yunfu Luoding Power Supply Bureau, the main research direction is the operation and maintenance optimization of distribution network based on digital twin.

Hai Li, born in 1988 in Luoding City, Guangdong Province, Master's degree holder from Guangdong University of Technology, currently employed at Guangdong Power Grid Yunfu Luoding Power Supply Bureau, the research direction is related to the application of electric power information communication and other related specialties.

Guoliang Li, born in July 1992 in Xinxing, Yunfu, Guangdong Province, member of communist party of China, he holds a bachelor's degree, currently employed at Guangdong Power Grid Yunfu Luoding Power Supply Bureau, the research direction is the application of artificial intelligence technology for the high-quality operation of distribution network.

Liwen Huang, born in 1974 in Xingning City, Guangdong Province, Bachelor's degree from South China Normal University, currently employed at Guangdong Power Grid Yunfu Luoding Power Supply Bureau, the main research direction includes the application of information system software and hardware maintenance and network security technology in new power system.

Chuming Mo, born in 1982 in Yunfu City, Guangdong Province, Undergraduate student at Huazhong University of Science and Technology currently employed at Guangdong Power Grid Yunfu Luoding Power Supply Bureau, the main research direction of big data analysis is the application of artificial intelligence technology in the new power system.

References

- [1] Zhang, F. (2024, May). Intelligent Monitoring and Early Warning System for Electric Power Safety using Artificial Intelligence Approach. In *2024 Second International Conference on Data Science and Information System (ICDSIS)* (pp. 1-5). IEEE.
- [2] Bedi, G., Venayagamoorthy, G. K., Singh, R., Brooks, R. R., & Wang, K. C. (2018). Review of Internet of Things (IoT) in electric power and energy systems. *IEEE Internet of things Journal*, 5(2), 847-870.
- [3] Cao, L., Wang, Z., & Yue, Y. (2022). Analysis and prospect of the application of wireless sensor networks in ubiquitous power internet of things. *Computational Intelligence and Neuroscience*, 2022(1), 9004942.
- [4] Abdulwahid, A. H. (2021, February). Power grid surveillance and control based on wireless sensor network technologies: Review and future directions. In *Journal of Physics: Conference Series* (Vol. 1773, No. 1, p. 012004). IOP Publishing.
- [5] Siddique, A., Prabhu, B., Chaskar, A., & Pathak, R. (2019). A review on intelligent agriculture service platform with lora based wireless sensor network. *Life*, 100, 7000.
- [6] Migabo, E. M., Djouani, K. D., & Kurien, A. M. (2020). The narrowband Internet of Things (NB-IoT) resources management performance state of art, challenges, and opportunities. *IEEE Access*, 8, 97658-97675.
- [7] Dangana, M., Ansari, S., Abbasi, Q. H., Hussain, S., & Imran, M. A. (2021). Suitability of NB-IoT for indoor industrial environment: A survey and insights. *Sensors*, 21(16), 5284.
- [8] Aju, O. G. (2015). A survey of zigbee wireless sensor network technology: Topology, applications and challenges. *International Journal of Computer Applications*, 130(9), 47-55.

- [9] Tightiz, L., & Yang, H. (2020). A comprehensive review on IoT protocols' features in smart grid communication. *Energies*, 13(11), 2762.
- [10] Mufana, M. W., & Ibrahim, A. (2022). Monitoring with communication technologies of the smart grid. *IDOSR Journal of Applied Sciences*, 7(1), 102-112.
- [11] Rajagukguk, R. A., Ramadhan, R. A., & Lee, H. J. (2020). A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. *Energies*, 13(24), 6623.
- [12] Strielkowski, W., Vlasov, A., Selivanov, K., Muraviev, K., & Shakhnov, V. (2023). Prospects and challenges of the machine learning and data-driven methods for the predictive analysis of power systems: A review. *Energies*, 16(10), 4025.
- [13] Dong, J., Zhuang, D., Huang, Y., & Fu, J. (2009). Advances in multi-sensor data fusion: Algorithms and applications. *Sensors*, 9(10), 7771-7784.
- [14] Kong, L., Peng, X., Chen, Y., Wang, P., & Xu, M. (2020). Multi-sensor measurement and data fusion technology for manufacturing process monitoring: a literature review. *International journal of extreme manufacturing*, 2(2), 022001.
- [15] Azimirad, E., Haddadnia, J., & Izadipour, A. L. I. (2015). A COMPREHENSIVE REVIEW OF THE MULTI-SENSOR DATA FUSION ARCHITECTURES. *Journal of Theoretical & Applied Information Technology*, 71(1).
- [16] Liu, Q., Jin, D., Shen, J., Fu, Z., & Linge, N. (2016, January). A WSN-based prediction model of microclimate in a greenhouse using extreme learning approaches. In *2016 18th International Conference on Advanced Communication Technology (ICACT)* (pp. 730-735). IEEE.
- [17] Pan, Y., Hong, R., Chen, J., Singh, J., & Jia, X. (2019). Performance degradation assessment of a wind turbine gearbox based on multi-sensor data fusion. *Mechanism and machine theory*, 137, 509-526.
- [18] Kordestani, M., & Saif, M. (2017, April). Data fusion for fault diagnosis in smart grid power systems. In *2017 IEEE 30th Canadian conference on electrical and computer engineering (CCECE)* (pp. 1-6). IEEE.
- [19] Jiao, Z., & Wu, R. (2018). A new method to improve fault location accuracy in transmission line based on fuzzy multi-sensor data fusion. *IEEE Transactions on Smart Grid*, 10(4), 4211-4220.
- [20] De Farias, C. M., Pirmez, L., Fortino, G., & Guerrieri, A. (2019). A multi-sensor data fusion technique using data correlations among multiple applications. *Future generation computer systems*, 92, 109-118.
- [21] CAO, Shipeng, et al. Multi-Sensor fusion and data analysis for operating conditions of low power transmission lines. *Measurement*, 2022, 190: 110586.

- [22] Jiang, Y. (2022). Outage management of active distribution systems with data fusion from multiple sensors given sensor failures. *IEEE Transactions on Power Delivery*, 38(3), 1891-1903.
- [23] Kang, Z., Zhang, Y., & Du, Y. (2024). Application of multi-sensor information fusion technology in fault early warning of smart grid equipment. *Energy Informatics*, 7(1), 123.
- [24] Blasch, E., Pham, T., Chong, C. Y., Koch, W., Leung, H., Braines, D., & Abdelzaher, T. (2021). Machine learning/artificial intelligence for sensor data fusion—opportunities and challenges. *IEEE aerospace and electronic systems magazine*, 36(7), 80-93.
- [25] Zhen, M., & Muzaffar, H. K. T. (2018). Intelligent fusion algorithm for multi-sensor information in integrated power grid operation system. *Journal of Intelligent & Fuzzy Systems*, 35(4), 4059-4069.
- [26] Sun, G., Zhang, Z., Zheng, B., & Li, Y. (2019). Multi-sensor data fusion algorithm based on trust degree and improved genetics. *Sensors*, 19(9), 2139.
- [27] Li, T., Li, S., Xu, Z., Alkayem, N. F., Bao, Q., & Wang, Q. (2025). A Multi-Sensor Fusion Approach Combined with RandLA-Net for Large-Scale Point Cloud Segmentation in Power Grid Scenario. *Sensors*, 25(11), 3350.