EAI Endorsed Transactions
on Energy Web
Research Article EAI.EU

Intelligent Optimization Algorithm for Power Grid Fault Decision Based on Multi-Source Data Fusion

Kun Zhang *, Qingbiao Lin, Zhantao Fan, Shengmin Qiu, Xiaogang Wu, Zhizhong Li and Yan Guo

China Southern Power Grid Company Ltd. (CSG) Power Dispatching Control Center, Guangzhou, Guangdong 510000, China

Abstract

Power grid fault recovery scheduling often faces challenges such as strong data heterogeneity, low fault identification efficiency, and difficulty in coordinating multiple conflicting objectives. To enhance the intelligence and practicality of fault handling, a power grid fault decision-making algorithm is proposed that integrates multi-source data modeling with intelligent optimization. On the basis of unified fusion of structured and unstructured data, the algorithm incorporates power flow calculation analysis and logical reasoning to achieve rapid fault area localization and preliminary judgment, thereby significantly improving diagnostic accuracy and timeliness. In the optimization stage, an improved chimpanzee foraging optimization algorithm is developed, integrating Hammersley sequence initialization, adaptive weighting, and somersault foraging strategies to address multi-objective coordination and accelerate convergence. Experimental results demonstrate the superiority of the proposed method on typical power grid fault datasets. In the preliminary decision-making phase, the fault identification accuracy reaches 98.7%, with an area under the curve of 0.967, indicating high classification stability. In the optimization scheduling phase, the power restoration ratio peaks at 98.7%, switching operation cost is reduced to 0.15, and fault response time is shortened to 1.1 s, showcasing strong global optimization and scheduling efficiency. The above results show that the algorithm has obvious advantages in fault diagnosis accuracy, dispatch globality and optimization stability. It not only provides a deployable and scalable intelligent optimization solution for power grid fault recovery in complex scenarios, but also provides solid technical support for improving the intelligent decision-making level and operational resilience of power systems. It has important engineering and practical significance.

Keywords: Power grid fault decision, Multi-source data fusion, Intelligent optimization algorithm, Logical reasoning, Improved chimpanzee optimization algorithm

Received on 16 January 2026, accepted on 23 March 2026, published on 27 May 2026

doi: 10.4108/ew.11592

1. Introduction

With the advancement of new power system construction, the operating environment of power grids becomes increasingly complex. The system faces severe challenges in handling concurrent faults in multiple scenarios and achieving rapid decision-making and recovery [1]. Especially under the background of new energy integration, frequent power market fluctuations, and severe load variations, traditional fault handling methods relying on single data sources and static rules fail to meet high reliability and real-time requirements [2]. Existing studies mostly apply expert systems, fuzzy logic, and support vector machines for preliminary fault state judgment, achieving some results in state estimation and local fault diagnosis [3-4]. However, they generally suffer from isolated data processing, slow response, and poor adaptability. They have limited ability to

*Corresponding author. Email: zk123456202508@163.com

deeply explore semantic relationships among multi-source heterogeneous data and fault evolution paths, which restricts their application in complex dynamic environments [5]. Therefore, in the complex power grid environment where multi-source heterogeneous data are highly coupled and dynamically evolving, how to achieve precise identification of fault states and efficient restoration scheduling decisions under multiple constraints has become a key issue that urgently needs to be addressed in the current field of intelligent power grid operation and maintenance. Multi-Source Data Fusion (MSDF) integrates different structured and unstructured information to realize comprehensive perception and deep analysis of fault states, enhancing fault sensing and decision support capabilities [6-7]. The Chimpanzee Optimization Algorithm (COA) simulates collaborative foraging behavior of chimpanzees, possessing excellent global search ability and suitability for nonlinear multi-objective optimization tasks in power systems [8-9]. Meanwhile, traditional fault handling methods are unable to meet the requirements of complex power grids for real-time performance, accuracy, and global optimization capabilities. Therefore, it is urgent to develop a power grid fault decision-making method that integrates multi-source data and intelligent optimization mechanisms. Therefore, multi-source real-time data are combined with power grid operations to propose an intelligent optimization algorithm for power grid fault decision that fuses MSDF mechanism with an improved COA. The aim is to improve fault recovery scheduling decision accuracy and resource allocation efficiency under multi-objective constraints. The innovation lies in realizing fusion modeling of structured and unstructured information, combining logical reasoning and power flow simulation to strengthen fault localization. This provides a deployable technical path and theoretical support for autonomous perception, intelligent judgment, and precise scheduling of power grids in complex environments.

2. Related works

MSDF is a key technology that integrates information from different sources. It fully mines semantic correlations and dynamic evolution patterns among diverse data sources, enhancing model interpretability and generalization ability. Scholars worldwide have conducted in-depth studies on this topic. For example, He et al. raised a cross-network node classification framework based on multi-source data selection to address the negative transfer caused by improper source network choice in multi-source cross-network node classification tasks. Their method fused multi-source transfer learning and network embedding strategies, effectively avoiding degradation of optimal solutions [10]. He et al. proposed a diagnostic method based on multi-sensor data fusion and a dual-scale residual network to improve the accuracy of train bearing fault diagnosis. They built a multi-sensor fusion mechanism that extracted high-dimensional nonlinear signal features into three-dimensional pixel matrices, achieving efficient feature extraction and discrimination [11]. Guo et al. developed a deep learning-based online real-time ship data fusion method to tackle the inefficiency in fusing automatic identification system and video data in inland waterway ship traffic monitoring [12]. Yang et al. proposed a dynamic implicit modeling method based on MSDF to solve issues of high heterogeneity and uncertainty in tunnel adverse geological modeling. Their method combined globally supported radial basis functions and Boolean implicit computation, realizing local adaptive updating of the model [13]. Existing research on MSDF mainly focuses on cross-network classification, industrial fault diagnosis, and multimodal perception. It has achieved good results in feature fusion and representation learning, but it focuses on single-task or static-scene modeling. It lacks the ability to model multi-source data in a unified manner and provide real-time decision support for complex power grid operating environments. In particular, it is still insufficient in the coupling analysis of multi-source information and power grid physical mechanisms.

Regarding power grid fault decision methods, some theories and practical applications have become relatively mature, and extensive research has been conducted by scholars worldwide. Atrigna et al. proposed a machine learning-based prediction method for power distribution network outages to address the impact of frequent heat waves caused by climate change. Their method combined historical fault data and meteorological information to model and predict power grid fault risks under heat wave conditions [14]. Branco et al. developed a hybrid time series prediction method that fused random walk filters and ensemble aggregation models to improve the prediction of insulator faults in power grids. The integrated learning model forecasted leakage current trends of insulators [15]. Dwivedi and Tajer raised a graph-structured data-driven framework based on graph recurrent neural networks and Markov decision process modeling to address modeling complexity in cascade fault chain identification and real-time search in power systems [16]. Huang et al. proposed a fault prediction model based on particle swarm optimization and support vector machines to solve low operational maintenance efficiency of relay protection devices in substations. Their model globally optimized support vector machine parameters via the optimization algorithm, improving prediction accuracy [17]. In contrast, while existing research on power grid fault decision-making has made some progress in fault prediction, state assessment, and local optimization, most methods are still based on a single data source and lack the ability to integrate multi-source data fusion, fault identification, and recovery scheduling into a unified model. Furthermore, they still fall short in global optimization and real-time response under multi-objective constraints. In recent years, there have been studies that have applied blockchain and artificial intelligence technologies to the operation and security of smart grids. For instance, Khan et al. proposed a distributed power system operation and control framework based on blockchain to enhance node privacy and security [18]. Mamoon Rashid et al. conducted a systematic review of the integration application of artificial intelligence and blockchain in the automation of smart grids, highlighting their potential advantages in data security and decision

reliability [19]. Although these methods have made progress in grid operation security and data protection, they mainly focus on the design of system security mechanisms and rarely address the integrated optimization problem of fault identification and recovery scheduling driven by multi-source data fusion.

In summary, existing studies have achieved certain results in MSDF and intelligent decision-making for power grid faults. Some methods showed good recognition accuracy and decision efficiency in specific scenarios. However, current algorithms still have problems such as coarse fusion granularity, poor dynamic adaptability, and insufficient semantic understanding of heterogeneous data. These issues limited their ability to meet higher requirements of real-time performance, accuracy, and multi-objective coordination in complex power grids. Therefore, this study aims to construct a unified data fusion framework based on MSDF and introduce intelligent optimization algorithms to achieve multi-objective recovery strategy optimization. The goal is to improve fault response efficiency and scheduling practicality of power grids in dynamic and complex environments. The main contributions of this study are as follows: (1) A multi-source data fusion framework for power grid fault decision-making is constructed, enabling unified modeling and efficient collaborative processing of heterogeneous data from Supervisory Control and Data Acquisition (SCADA) systems, meteorological sources, and other inputs, thereby enhancing the completeness and consistency of fault information perception. (2) A fault identification method integrating power flow calculation and logical reasoning is proposed, which achieves precise localization of fault sections and enhanced risk perception through the synergistic interaction of physical mechanism analysis and rule-based inference. (3) A preliminary fault decision model combining Decision Tree (DT) and Random Forest (RF) is developed, which improves the accuracy of fault type identification and the generalization capability of the model based on multi-source fused features. (4) COA is introduced and improved by incorporating the Hammersley sequence, adaptive weight factors, and somersault foraging strategy, thereby constructing a multi-objective optimization model to realize global optimization of fault recovery paths and scheduling strategies. (5) Extensive experiments with multiple evaluation metrics demonstrate the superior performance of the proposed method in fault identification and recovery scheduling, highlighting its strong engineering applicability and practical potential.

3. Intelligent algorithm for power grid fault decision based on multi-source data fusion

3.1. Power grid fault decision algorithm based on MSDF and logical reasoning

In complex operational scenarios, power grid fault decision tasks often need to process multi-source heterogeneous data from dispatch control systems, relay protection devices, Geographic Information System (GIS), and other sources [20-21]. Data from different systems vary significantly in structure, frequency, and expression formats. If not effectively fused and uniformly processed, these differences limit the accuracy and response efficiency of fault identification [22]. According to the structural characteristics of the power grid, data are categorized into structured types, such as measurement values from Supervisory Control and Data Acquisition (SCADA) systems and protection action records, and unstructured types, such as weather forecast data [23]. This study integrates MSDF with the Hadoop distributed platform to build a global indexing framework named MSDF-H, which achieves unified and efficient retrieval of multi-source measurement data. The distributed framework of MSDF-H is shown in Figure 1.

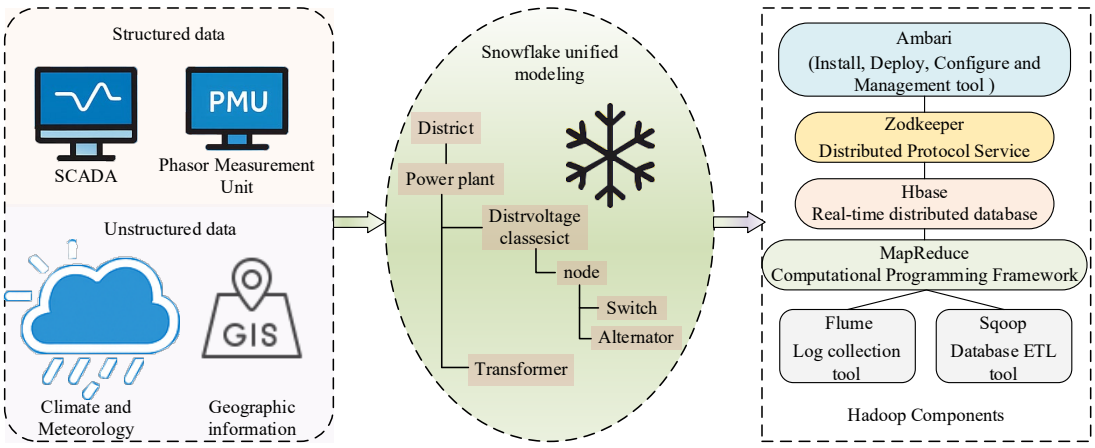


Figure 1. MSDF-H distributed framework diagram

In Figure 1, this framework is built on the Hadoop distributed computing system. It integrates various structured and unstructured data through a snowflake architecture that enables hierarchical and unified modeling from regions and power sources to faulty devices. Hadoop components perform efficient data storage and parallel computation scheduling, providing data support for subsequent fault identification. Based on this, the study further constructs a power grid fault identification framework using Power Flow Calculation (PFC) and logical reasoning for the fault perception and diagnosis layer. The one-dimensional array of the node hierarchy structure in PFC is shown in Equation (1) [24].

$$\begin{cases} LayerM(i) = l_i, \ (i = 1, 2, \ldots, N) \\ NU(i) = j \end{cases} \tag{1}$$

In Equation (1), $Layer(\cdot)$、$NU(\cdot)$represent the layer number of the node and the parent node respectively, N represents the number of nodes in the power grid, l_i is the layer number of the i-th node, and j is the parent node of node i. The reverse current in power flow calculation is shown in Equation (2).

$$I_{ij} = \sum_{k \in Child(j)} I_{jk} + \frac{S_j^*}{V_j^*} \tag{2}$$

In Equation (2), I_{ij} represents the branch current from node j to its parent node i, I_{jk} represents the branch current flowing from node j to its downstream child node k, S_j^* is the complex power load of node j, V_j^* is the voltage of node j, and $Child(j)$ is the set of downstream child nodes of node j. The forward voltage calculation is shown in Equation (3).

$$V_j = V_i - Z_{ij} \cdot I_{ij} \tag{3}$$

In Equation (3), V_i is the voltage of the parent node i, and Z_{ij} represents the line resistance from i to j. Based on the above MSDF and PFC analysis, the study designs the full implementation process of MSDF-HP to achieve efficient fault identification. The process is shown in Figure 2.

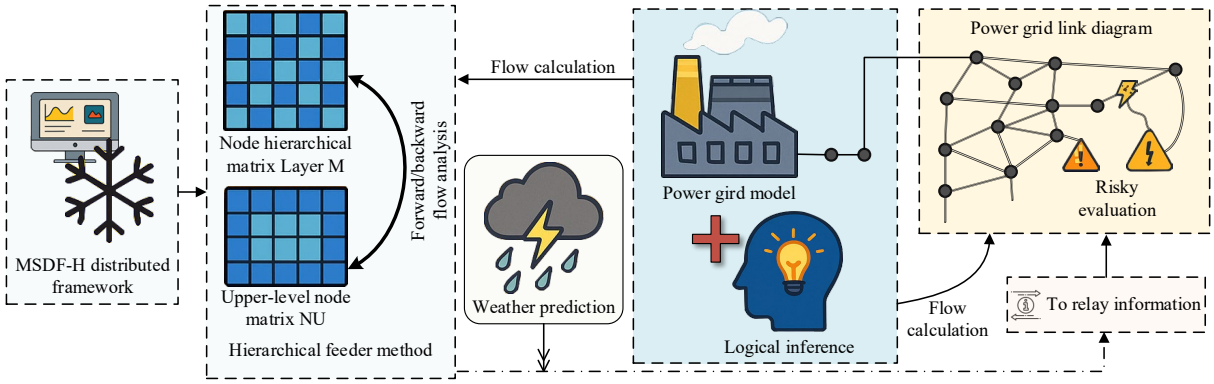


Figure 2. Complete operation flow chart of MSDF-HP

As shown in Figure 2, this method builds a layered feeder structure using the node hierarchy matrix and upper-level node matrix. It performs PFC analysis efficiently through forward and backward substitution and obtains key electrical parameters [25]. Based on this, the method uses weather forecast data and relay protection action information to construct reasoning rules based on the physical structure of the power grid, forming a fault judgment mechanism for risk perception. Through the collaboration of PFC analysis and logical reasoning, it quickly identifies fault areas and marks risk regions visually. The fault judgment logic combining the power grid model and logical reasoning is shown in Equation (4).

$$F_j = \begin{cases} 1, \ if \ P_j = 1 \wedge T_j > t_{\text{th}} \wedge W_j > \theta \\ 0, \ otherwise \end{cases} \tag{4}$$

In Equation (4), P_j represents whether the protection device of node j has acted (where 1 indicates an active

state), T_j is the action duration of the protection device, T_j is the meteorological intensity, θ is the threshold for severe weather identification, F_j represents whether node j is inferred as a fault point, and t_{th} is the delay threshold of the protection device [26]. Relying only on the physical structure and logical reasoning still leads to limited coverage in complex environments or large-scale networks [27]. Therefore, this study introduces DT and RF to construct a preliminary fault decision algorithm framework, named MSHP-DRF. This framework improves automatic classification of fault types and locations and enhances the level of intelligent response. The schematic of MSHP-DRF is shown in Figure 3.

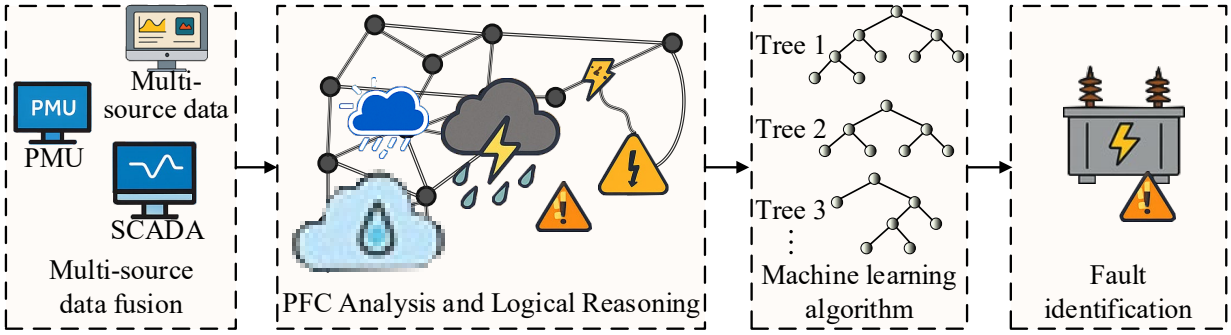


Figure 3. Framework diagram of the MSHP-DRF algorithm

As shown in Figure 3, the algorithm first uses the MSDF module to integrate real-time and historical data from SCADA, Phasor Measurement Units (PMU), and other sources, providing complete information support for fault decisions. Then, it combines PFC analysis and logical reasoning to quickly identify fault areas and visually mark risk regions. Based on this, the method builds multiple decision trees using fused features. It splits nodes according to key features, such as voltage and current, and outputs the fault type and location using a voting mechanism. The decision tree selects the feature based on the minimum Gini index, as shown in Equation (5) [28].

$$\begin{cases} Gain(D,A) = H(D) - \sum_{v \in A} \frac{|D^v|}{|D|} H(D^v) \\ Gini(D) = 1 - \sum_{k=1}^{K} p_k^2 \end{cases} \tag{5}$$

In Equation (5), D is the current MSDF dataset, A is the candidate feature, $H(D)$ is the information entropy of D, D^v represents the subset under feature $A = v$, $|D^v|$ and $|D|$ are the sizes of the subset and dataset respectively, $H(D^v)$ is the entropy of the subset, p_k is the proportion of class k in D, and K is the total number of classes. Overall, this method enables fast fault identification and intelligent prediction through the collaborative drive of MSDF, logical reasoning, and machine learning. It provides decision support for emergency handling of key loads and coverage areas.

3.2 Power grid fault decision design combining MSDF and improved COA

Under the support of MSDF and logical reasoning, power grid fault identification achieves an effective transformation from data-driven approaches to intelligent diagnosis. However, based on fault classification results, how to efficiently formulate repair and reconfiguration strategies remains a key issue in the distribution network restoration process [29-31]. Therefore, to further support rapid response, decision-making, and optimized handling after faults, this study introduces the COA optimization strategy to intelligently optimize recovery paths, reconfiguration switch operations, and load redistribution strategies in faulty areas. The algorithm simulates the collaborative behavior of chimpanzee groups during hunting. It guides fault recovery plans toward optimal solutions under multi-objective constraints, such as power supply recovery rate and operational cost, through a role-driven search mechanism and discrete update strategy [32]. The COA process is shown in Figure 4.

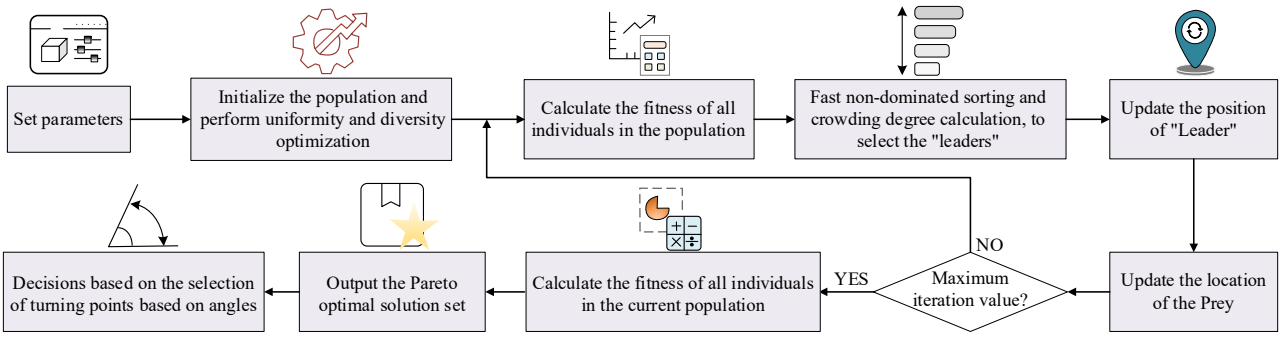


Figure 4. COA operation flow chart

As shown in Figure 4, the algorithm first sets parameters and initializes population positions. It introduces a uniform distribution mechanism to enhance population diversity. Then it calculates the fitness of each individual, selects "leader" individuals through sorting and competition, and updates all individuals' states based on their positions [33]. A disturbance mechanism is added to avoid falling into local optima, and the algorithm checks whether the maximum number of iterations has been reached. If not, the iteration continues; otherwise, the current optimal solution set is output as the final result. The hunting behavior of chimpanzee groups is divided into the exploration phase and the hunting phase. The mathematical model of the exploration phase is shown in Equation (6) [34].

$$\begin{cases} d = \left| c \cdot X^t_{prey} - m \cdot X^t_{chimp} \right| \\ X^{t+1}_{chimp} = X^t_{prey} - a \cdot d \end{cases} \tag{6}$$

In Equation (6), d and t represent the distance between the current individual and the prey and the iteration number, respectively. X^t_{prey} and X^t_{chimp} are the position vectors of the prey and the current chimpanzee. a, m, and c are coefficient vectors, which are calculated using Equation (7).

$$\begin{cases} a = 2f \cdot r_1 - f \\ c = 2r_2 \end{cases} \tag{7}$$

In Equation (7), r_1 and r_2 are random vectors within [0,1], f is a convergence factor, a is a random disturbance in the distance between the chimpanzee and the prey, randomly generated within the range of $[-f, f]$. c is an adjustment factor controlling the driving behavior of chimpanzees, with a range of [0,2]. In the hunting phase, other individuals in the population update their positions based on the four leaders. The model is shown in Equation (8) [35].

$$\begin{cases} X^{t+1}_1 = X^t_{attacker} - a_1 \cdot \left| c_1 X^t_{attacker} - m_1 \cdot X^t \right| \\ X^{t+1}_2 = X^t_{barrier} - a_2 \cdot \left| c_2 X^t_{barrier} - m_2 \cdot X^t \right| \\ X^{t+1}_3 = X^t_{chaser} - a_3 \cdot \left| c_3 X^t_{chaser} - m_3 \cdot X^t \right| \\ X^{t+1}_4 = X^t_{driver} - a_4 \cdot \left| c_4 X^t_{driver} - m_4 \cdot X^t \right| \\ X^{t+1} = \left(X^{t+1}_1 + X^{t+1}_2 + X^{t+1}_3 + X^{t+1}_4 \right) / 4 \end{cases} \tag{8}$$

In Equation (8), $X^t_{attacker}$ and $X^t_{barrier}$ are the position vectors of the current attacker and encircler, X^t_{chaser} and X^t_{driver} are the position vectors of the chaser and driver, and X^{t+1} is the updated position vector of the current chimpanzee. To enhance the intelligence level of fault handling strategies and adapt to discrete multi-objective optimization requirements in the fault recovery process, this study improves the COA by integrating Hammersley Sequence (HS), adaptive weight factors, and a somersault foraging strategy. The proposed method is named MSHP-DRF-ICOA. The overall structure of MSHP-DRF-ICOA is shown in Figure 5.

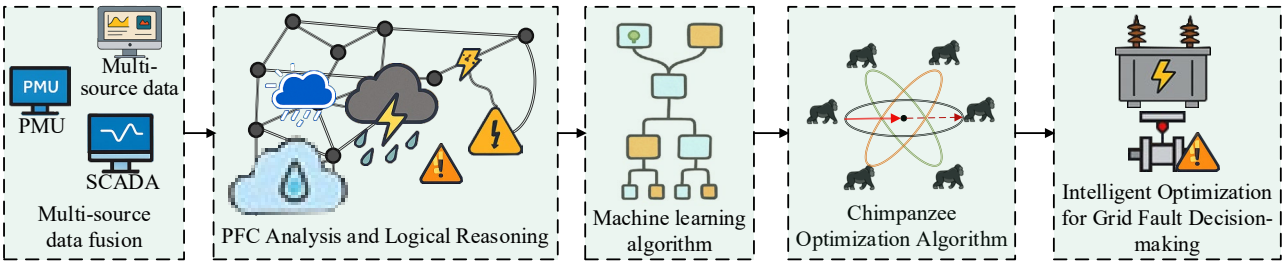


Figure 5. MSHP-DRF-ICOA overall structure diagram

As shown in Figure 5, the algorithm framework first fuses multi-source data such as SCADA and PMU to construct unified data features. Then it performs PFC analysis and logical reasoning to identify fault areas and perceive risks. Based on this, machine learning algorithms classify fault types, and the improved COA performs multi-objective optimization for recovery paths and regulation decisions, finally achieving intelligent decision-making and rapid response for power grid faults. The HS initialization equation is shown in Equation (9) [36].

$$x_{\beta}^{\alpha} = lb^{\alpha} + \left(ub^{\alpha} - lb^{\alpha}\right) \cdot h_{\beta,\alpha} \tag{9}$$

In Equation (9), x_{β}^{α} is the initial value of the β-th individual on the α-th dimension. lb^{α} and ub^{α} are the upper and lower bounds of the α-th variable. $h_{\alpha,\beta}$ is the value of the β-th individual on the α-th dimension. The generation method of HS is shown in Equation (10).

$$h_{\beta,\alpha} = \Phi_{b_\alpha}(\beta) = \sum_{g=1}^{L} a_g b_\alpha^{-g} \tag{10}$$

In Equation (10), $\Phi_{b_\alpha}(\beta)$ is the low-discrepancy sequence constructed based on prime number b_α. a_g is the g-th digit of β in the b_α base. b_α is the prime base used in the α-th dimension, and L is the number of expansion terms. The somersault foraging update equation is shown in Equation (11) [37].

$$X^{t+1} = X^{t} + S\left(r_1 \cdot X_{best}^{t} - r_2 \cdot X^{t}\right),\ if\ \ i/N > r,\ r \in [0,1] \tag{11}$$

In Equation (11), S is the somersault factor, X_{best}^{t} is the prey position, i is the current iteration number, N is the maximum number of iterations. Based on the aforementioned multi-source data fusion, fault identification, and optimization decision-making method construction, the complete pseudo-code description of the MSHP-DRF-ICOA algorithm is presented below to clearly present the overall process and key steps:

Algorithm Basic procedure of the proposed MSHP-DRF-ICOA

Input: Multi-source data D_{ms}, grid topology G(V, E), protection records Y, weather threshold δ, protection delay threshold τ, number of RF trees N_{tree}, population size N_{pop}, maximum iterations T_{max}, variable bounds lb and ub, somersault factor S.
Output: Fault diagnosis result R_{fault} and optimal restoration decision X_{best}

```
Preprocess D_ms and construct the MSDF-H framework
Generate node-layer and parent-node arrays by Eq. (1)
Compute branch current and node voltage by Eqs. (2) and (3)
Perform logical inference by Eq. (4) and form fused diagnostic features
Train the MSHP-DRF model
for k = 1 to Ntree do
Build a decision tree using the split criterion in Eq. (5)
end for
Obtain R_fault by random-forest voting
Encode restoration variables according to R_fault
Initialize population P by Hammersley sequence using Eqs. (9) and (10)
Evaluate P and determine the initial best solution X_best
Set t = 1
    while t ≤ T_max do
        Identify four leader chimpanzees according to fitness
        for i = 1 to N_pop do
            Update X_i by Eqs. (6)~(8)
            if somersault search is activated then
                Refine X_i by Eq. (11)
            end if
            if X_i exceeds bounds then
                Repair X_i within [lb, ub]
            end if
            Evaluate X_i and update X_best if improved
        end for
    Update adaptive parameters and set t ← t + 1
end while
Return R_fault and X_best
```

4. Algorithm performance validation for intelligent optimization of power grid fault decision

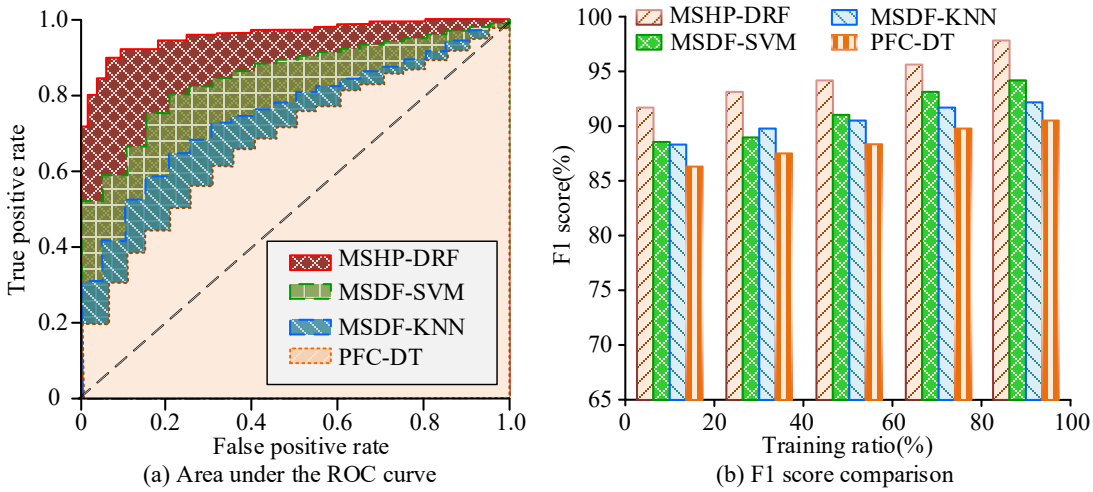
4.1 Performance validation of power grid fault decision algorithm based on MSHP-DRF

To systematically evaluate the performance of the proposed MSHP-DRF algorithm in power grid fault identification tasks, three baseline algorithms (PFC-DT, MSDF-KNN, and MSDF-SVM) were selected as benchmarks. The experimental data were constructed based on fault records, SCADA/PMU monitoring values, GIS topology, and concurrent meteorological information from a typical 10 kV urban distribution network. Specifically, the dataset includes voltage, current, phase angle, and power flow measurements from SCADA/PMU (sampling interval: 1 s), network topology and node connectivity from GIS, and meteorological variables such as wind speed (0 m/s–25 m/s), temperature (-10 °C to 40 °C), and precipitation (0 mm to 50 mm). Fault records cover typical events including line faults, transformer faults, and transient disturbances. To ensure data quality and consistency, all data were temporally aligned based on unified timestamps, followed by noise filtering, missing value interpolation (linear interpolation), and min–max normalization to [0,1]. Multi-source features were then fused into a unified feature vector with a dimension of 18–24 features per sample. Each sample corresponds to a specific network state and is labeled with fault type and location, enabling supervised learning. The dataset included 35 regional nodes and 480 sets of fault and non-fault samples. After cleaning and normalization, a structured multidimensional fusion input was formed. All experiments were conducted in a Python environment. The algorithms were trained in a supervised manner using 80% of the samples, and the remaining 20% were used for testing and evaluation. The hardware configuration and parameter settings are shown in Table 1.

Table 1. Hardware configuration and parameter setting

Item	Configuration
Dataset	Self-constructed dataset
The number of regions	35
The number of samples	480
Operating system	Windows 10 64-bit
Experiment condition	Intel Core i7-7500UCPU@16G, NVIDIA GeForce RTX 3060X
Simulation tool	MATLAB R2019b

First, the study compared the Area Under the Curve (AUC) and F1 score of the four algorithms. The experimental results are shown in Figure 6.

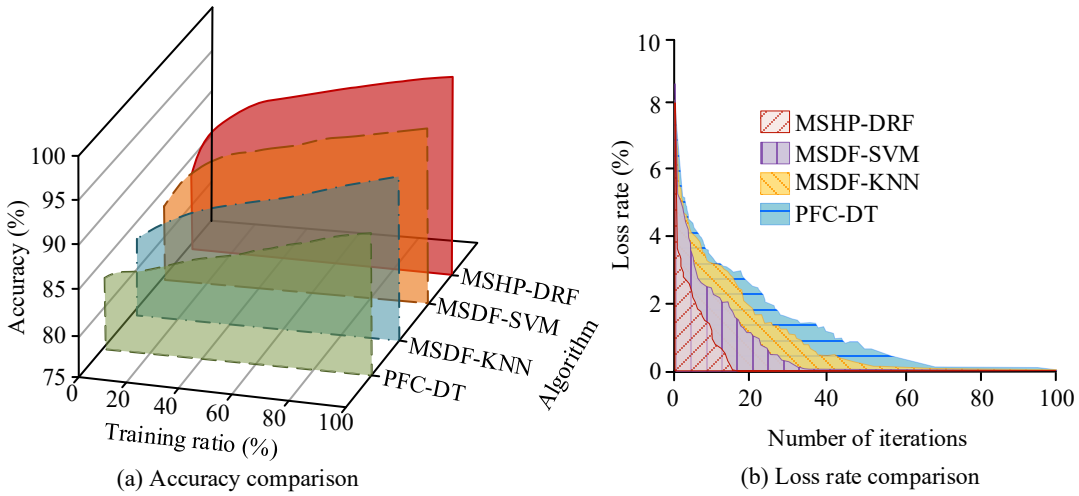


(a) Area under the ROC curve (b) F1 score comparison

Figure 6. Comparison of AUC values and F1 scores

As shown in Figure 6(a), in the Receiver Operating Characteristic (ROC) curve comparison, the MSHP-DRF algorithm had the curve closest to the upper left corner. Its AUC reached 0.967, significantly higher than the comparison algorithms, which scored 0.852, 0.793, and 0.728. This result indicated that MSHP-DRF had stronger overall discriminative ability in fault classification and more reliable prediction outcomes. Further observations from Figure 6(b) showed that the F1 score of MSHP-DRF consistently remained above 92.5%, ultimately reaching 97.9%, far exceeding those of the comparison algorithms. This demonstrated an excellent trade-off between classification precision and recall, as well as stable performance. Overall, MSHP-DRF outperformed the other models in terms of fault identification accuracy, generalization ability, and robustness. To further evaluate the comprehensive performance of each algorithm in power grid fault identification, the study conducted a quantitative comparison

based on accuracy and loss rate. The results are shown in Figure 7.



(a) Accuracy comparison (b) Loss rate comparison

Figure 7. Comparison of accuracy and loss rate

As shown in Figure 7(a), as the proportion of training data increased, the prediction accuracy of MSHP-DRF in fault identification remained the highest. When using the full training set, it reached 98.73%, significantly higher than the 85.62% of MSDF-SVM, 88.32% of MSDF-KNN, and 87.44% of PFC-DT. According to Figure 7(b), as the number of iterations increased, MSHP-DRF converged the fastest, reaching a loss rate of 1.2% after only 18 iterations, which is close to zero, and remained stable afterward. Throughout the iteration process, its convergence speed consistently outpaced the three comparison algorithms. In summary, the proposed MSHP-DRF algorithm demonstrated superior prediction performance and convergence characteristics in intelligent diagnosis and rapid response for power grid faults. To further validate the superiority of the proposed method, statistical analysis was conducted over 10 independent runs, and the results are summarized in Table 2.

Table 2. Statistical validation of MSHP-DRF performance (mean ± std, n=10 runs)

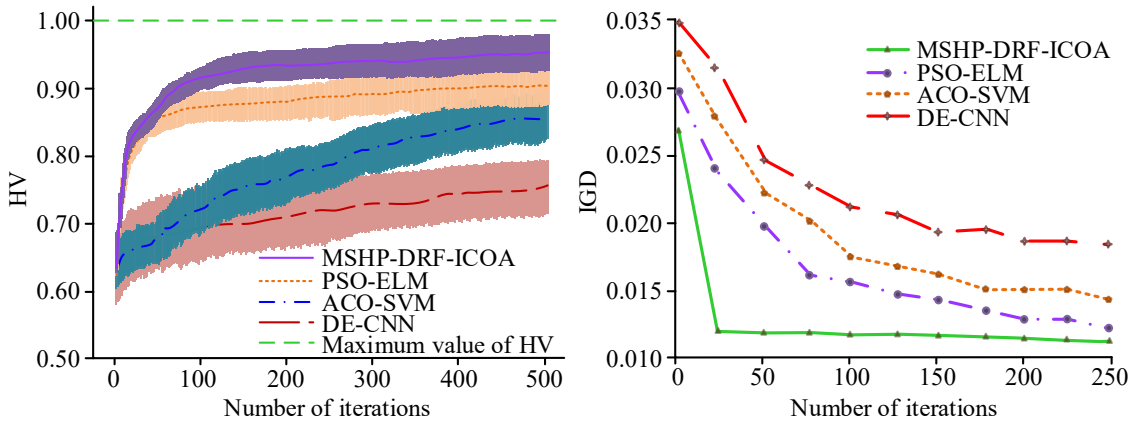
Algorithm	Accuracy (%)	AUC	F1-score (%)	Loss (%)	*p*-value
MSHP-DRF	98.73±0.31	0.967±0.006	97.9±0.42	1.2±0.08	—
MSDF-SVM	85.62±0.88	0.852±0.014	88.6±0.91	3.8±0.21	<0.01
MSDF-KNN	88.32±0.79	0.793±0.018	90.2±0.85	3.1±0.18	<0.01
PFC-DT	87.44±0.83	0.728±0.021	89.1±0.88	3.5±0.20	<0.01

As shown in Table 2, MSHP-DRF achieves the highest mean performance across all evaluation metrics, with an accuracy of 98.73±0.31%, AUC of 0.967±0.006, and F1-score of 97.9±0.42%, while maintaining the lowest loss (1.2 ±0.08%). In comparison, the baseline models exhibit lower mean performance and larger standard deviations, indicating reduced stability. Furthermore, the p-values obtained from two-tailed t-tests are all below 0.01, demonstrating that the performance improvements of MSHP-DRF over MSDF-SVM, MSDF-KNN, and PFC-DT are statistically significant. These results confirm that the proposed algorithm not only outperforms existing methods in terms of average performance but also provides more consistent and reliable results, thereby offering stronger validation for its superiority in power grid fault decision tasks.

4.2 Performance of fault decision algorithm integrating MSHP-DRF and improved COA

To verify the performance advantages of MSHP-DRF-ICOA in intelligent optimization for power grid fault decisions, the study further compared it with three combined algorithms: Ant Colony Optimization and Support Vector Machine (ACO-SVM), Particle Swarm Optimization and Extreme Learning Machine (PSO-ELM), and Differential Evolution and Convolutional Neural Network (DE-CNN)[38-40]. All

experiments were conducted on a Windows 10 64-bit system with an Intel Core i7-7500U processor and 16 GB RAM. The models were trained in a Python environment using both multi-objective test functions and a self-constructed multi-source power grid fault dataset. First, the study compared the Hypervolume (HV) and Inverted Generational Distance (IGD) metrics on the test set. The experimental results are shown in Figure 8.

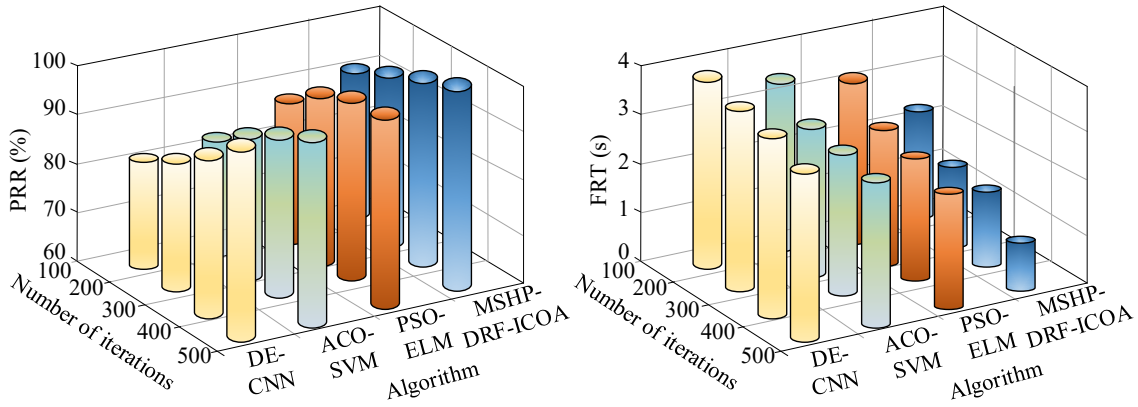


(a) The HV indicators of the four algorithm

(b) The IGD evolution trajectories of four algorithms on the DTLZ4 problem

Figure 8. Comparison of HVs and IGD of four algorithms

As shown in Figure 8(a), all four algorithms showed increasing HV curves with more iterations. The HV value of MSHP-DRF-ICOA consistently led and eventually stabilized at 0.951 after 500 iterations, indicating excellent solution set coverage and convergence performance. As shown in Figure 8(b), MSHP-DRF-ICOA converged to a low IGD value of approximately 0.0112 at the 23rd iteration. This result indicates that the solution set is closer to the Pareto optimal front, with better distribution quality and diversity. These results demonstrated that the algorithm had outstanding multi-objective decision-making capability and global optimization stability in power grid fault scenarios. Finally, the study compared the Power Restoration Ratio (PRR) and Fault Response Time (FRT) of the four algorithms using real power grid fault decision data. The experimental results are shown in Figure 9.



(b) Four algorithms of PRR

(b) Four algorithms of SOC

Figure 9. Comparison chart of PRR and FRT

As shown in Figure 9(a), the PRR of MSHP-DRF-ICOA remained the highest across all iteration stages and ultimately reached 98.71%. This result significantly surpassed ACO-SVM (91.61%), PSO-ELM (87.12%), and DE-CNN (87.54%), demonstrating its advantages in restoration coverage and scheduling accuracy. Figure 9(b) showed that MSHP-DRF-ICOA also converged faster. It successfully reduced the FRT to 1.1 seconds, which is notably better than those of the comparison algorithms. These findings confirmed that MSHP-DRF-ICOA not only

balanced recovery efficiency and resource constraints but also demonstrated strong engineering practicality and adaptability in global scheduling for multi-objective power grid fault decision tasks. Considering that in actual power grid operation there may be situations such as missing monitoring data, measurement noise interference, communication delays, and switch malfunctions, this study conducted supplementary experiments to evaluate robustness to abnormal disturbances. In addition, considering that different load scenarios have different requirements for recovery strategies, further experiments were carried out to assess practical application effects. It selected the power supply restoration rate, fault response time, and restoration success rate as comprehensive evaluation indicators. The results are shown in Table 3.

Table 3. Comparative results of robustness and practical application performance for four algorithms

Test type	Scenario setting	MSHP-DRF-ICOA	ACO-SVM	PSO-ELM	DE-CNN
Standard condition	No abnormal disturbance	98.7±0.32/1.1±0.05 /98.5±0.28	91.6±0.74/1.8±0.09 /91.2±0.71	87.1±0.91/2.1±0.12 /86.5±0.88	87.5±0.85/2.0±0.11 /86.9±0.82
Abnormal data	10% missing data	96.8±0.41/1.3±0.06 /96.2±0.39	89.4±0.82/2.0±0.11 /88.7±0.79	84.6±0.96/2.3±0.14 /83.9±0.92	85.1±0.91/2.2±0.13 /84.3±0.87
	15% Gaussian noise	95.9±0.48/1.4±0.07 /95.4±0.45	88.6±0.89/2.1±0.12 /87.9±0.85	83.8±1.02/2.5±0.15 /83.1±0.98	84.2±0.97/2.4±0.14 /83.6±0.94
	200 ms delay	96.2±0.44/1.5±0.08 /95.7±0.41	89.1±0.85/2.2±0.13 /88.4±0.81	84.2±0.98/2.6±0.16 /83.5±0.94	84.8±0.93/2.5±0.15 /84.0±0.90
System fault	Switch failure	95.6±0.52/1.6±0.09 /94.9±0.49	87.8±0.94/2.3±0.14 /86.8±0.91	82.5±1.05/2.7±0.18 /81.6±1.01	83.1±1.01/2.6±0.17 /82.2±0.98
	Communication interruption	95.1±0.55/1.7±0.10 /94.3±0.52	87.2±0.97/2.4±0.15 /86.0±0.93	81.9±1.09/2.8±0.19 /80.8±1.05	82.4±1.04/2.7±0.18 /81.4±1.00
Practical case	Residential feeder	98.1±0.36/1.2±0.05 /97.6±0.33	91.0±0.79/1.9±0.10 /90.3±0.76	86.4±0.93/2.2±0.12 /85.6±0.90	86.9±0.88/2.1±0.11 /86.1±0.85
	Industrial park	97.5±0.39/1.3±0.06 /97.0±0.36	90.2±0.81/2.0±0.11 /89.4±0.78	85.7±0.95/2.3±0.13 /84.8±0.91	86.1±0.90/2.2±0.12 /85.2±0.87
	Extreme weather	96.9±0.43/1.5±0.08 /96.1±0.40	88.9±0.87/2.2±0.13 /87.8±0.84	84.0±1.01/2.6±0.16 /82.9±0.97	84.5±0.96/2.5±0.15 /83.4±0.92

Note: each value is presented as PRR/FRT/RSR, corresponding to power restoration ratio (%)/fault response time (s)/restoration success rate (%).

As shown in Table 3, MSHP-DRF-ICOA achieves the best performance and stability across all scenarios. Under standard conditions, it reaches 98.7±0.32% PRR, 1.1±0.05 s FRT, and 98.5±0.28% RSR, outperforming ACO-SVM (91.6±0.74/1.8±0.09/91.2±0.71), PSO-ELM (87.1±0.91/2.1±0.12/86.5±0.88), and DE-CNN (87.5±0.85/2.0±0.11/86.9±0.82) with lower variance.Under abnormal data conditions, it maintains 95.9%–96.8% PRR with standard deviations below 0.5, while FRT remains within 1.5±0.08 s, showing stronger robustness than comparison models (typically std > 0.8). Under system faults, PRR remains around 95% (±0.52) with smaller fluctuations than other methods (>±1.0), indicating better fault tolerance.In practical scenarios, the method consistently achieves ≥96.9% PRR and ≤1.5 s FRT, with lower variance (e.g., 96.9±0.43%) than ACO-SVM (88.9±0.87%) and PSO-ELM (84.0±1.01%). In summary, MSHP-DRF-ICOA maintains high performance under abnormal data and system fault interference, and demonstrates good application results in various real-world scenarios, verifying its robustness and engineering application value in complex power grid fault decision-making.

5. Conclusion

To address the problems of strong data heterogeneity, insufficient decision accuracy, and conflicting optimization objectives in traditional power grid fault decision-making, this study proposed a fault decision algorithm for power grids that integrates multi-source data modeling with an intelligent optimization mechanism, named MSHP-DRF-ICOA. Based on unified modeling of multi-source structured and unstructured data, the algorithm combines power flow calculation with logical reasoning to achieve rapid fault area identification and intelligent classification. At the optimization decision stage, the study introduced an improved chimpanzee optimization algorithm, integrating the Hammersley sequence, adaptive weight mechanism, and somersault strategy to construct an efficient model for multi-objective restoration scheduling. Experimental results showed that the proposed MSHP-DRF algorithm achieved an AUC of 0.967, an F1-score of 97.9%, and a prediction accuracy of 98.7%, demonstrating outstanding performance in fault identification. In terms of multi-objective optimization, MSHP-DRF-ICOA maintained an HV of 0.951 and reduced the IGD to 0.023, showing excellent convergence

to the Pareto front and a well-distributed solution set. Meanwhile, it achieved a maximum power restoration ratio of 98.7% and reduced the fault response time to 1.1 seconds, significantly outperforming the comparison algorithms. These results demonstrated strong global search capability and practical effectiveness in power grid restoration. In summary, MSHP-DRF-ICOA showed significant advantages in fault localization accuracy, decision-making efficiency, and multi-objective optimization performance, with good engineering applicability and deployment potential. However, the current algorithm did not fully consider factors such as load priority and time window constraints in the restoration strategy. In future work, the method could be further extended to incorporate real-world scenario constraints and multi-energy system collaborative restoration to enhance its generalization ability and intelligent decision-making performance in complex power grid environments.

Use of AI tools declaration

This paper has not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization: Zhang Kun and Lin Qingbiao; methodology: Fan Zhantao; software: Qiu Shengmin; validation: Zhang Kun: Wu Xiaogang and Li Zhizhong; formal analysis: Zhang Kun; investigation: Guo Yan; resources: Qiu Shengmin; data curation: Li Zhizhong; writing—original draft preparation: Qiu Shengmin; writing—review and editing: Fan Zhantao; visualization: Lin Qingbiao; supervision: Zhang Kun; project administration: Zhang Kun; funding acquisition: Guo Yan. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

This work was supported by the Research and Application of Power Grid Abnormal Scheduling Decision Technology Driven by Data and Knowledge Collaboration, with the project number 000005KC24010014 (Science and Technology Project number: ZDKJXM20240015).

References

[1] Zhang L, Wang J, Liu Z. (2023) Power grid operation optimization and forecasting using a combined forecasting system. J Journal of Forecasting 42(1): 124-153. https://doi.org/10.1002/for.2888

[2] Khaleel M, Abulifa S A, Abulifa A A. (2023) Artificial intelligent techniques for identifying the cause of disturbances in the power grid. J Brilliance: Research of Artificial Intelligence 3(1): 19-31. https://doi.org/10.47709/brilliance.v3i1.2165

[3] Wang J, Gao D, Zhu S, et al. (2023) Fault diagnosis method of photovoltaic array based on support vector machine. J Energy sources, part a: recovery, utilization, and environmental effects 45(2): 5380-5395. https://doi.org/10.1080/15567036.2019.1671557

[4] Rana S. (2025) AI-driven fault detection and predictive maintenance in electrical power systems: A systematic review of data-driven approaches, digital twins, and self-healing grids. J American Journal of Advanced Technology and Engineering Solutions 1(1): 258-289. https://doi.org/10.63125/4p25x993

[5] Liu Z, Qian R, Jin X, et al. (2024) Multi-source heterogeneous data fusion technology for electric power based on big data mining. J Journal of Computational Methods in Sciences and Engineering 24(6): 3366-3380. https://doi.org/10.1177/147279782412938

[6] Wang Z, Wang S, Wang S, et al. (2025) An intelligent prediction method of surface residual stresses based on multi-source heterogeneous data. J Journal of Intelligent Manufacturing 36(1): 441-457. https://doi.org/10.1007/s10845-023-02238-6

[7] Cheng S, Yu R, Shen A, et al. (2025) MSDF-TL: A transfer learning and physics-constrained multi-source data fusion framework for weld state prediction. J Journal of Manufacturing Processes 154(5): 504-523. https://doi.org/10.1016/j.jmapro.2025.09.070

[8] Abedi F, Ghanimi H M A, Algarni A D, et al. (2023) Chimp Optimization Algorithm Based Feature Selection with Machine Learning for Medical Data Classification. J Comput. Syst. Sci. Eng. 47(3): 2791-2814. https://doi.org/10.32604/csse.2023.038762

[9] Sarihaddu C K, Raaza A, Akre S. (2025) Cognitive workload detection via binary chimp optimization algorithm and machine learning. J Arabian Journal for Science and Engineering 50(21): 17605-17616. https://doi.org/10.1007/s13369-025-10034-y

[10] He H, Yang H, Zhang W, et al. (2023) Msds: A novel framework for multi-source data selection based cross-network node classification. J IEEE Transactions on Knowledge and Data Engineering 35(12): 12799-12813. https://doi.org/10.1109/TKDE.2023.3277957

[11] He D, Lao Z, Jin Z, et al. (2023) Train bearing fault diagnosis based on multi-sensor data fusion and dual-scale residual network. J Nonlinear Dynamics 111(16): 14901-14924. https://doi.org/10.1007/s11071-023-08638-w

[12] Guo Y, Liu R W, Qu J, et al. (2023) Asynchronous trajectory matching-based multimodal maritime data fusion for vessel traffic surveillance in inland waterways. J IEEE Transactions on Intelligent Transportation Systems 24(11): 12779-12792. https://doi.org/10.1109/TITS.2023.3285415

[13] Yang B, Ding Y, Zhu Q, et al. (2024) Implicit modelling and dynamic update of tunnel unfavourable geology based on multi-source data fusion using support vector machine. J Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards 18(1): 257-274. https://doi.org/10.1080/17499518.2023.2239778

[14] Atrigna M, Buonanno A, Carli R, et al. (2023) A machine learning approach to fault prediction of power distribution grids under heatwaves. J IEEE Transactions on Industry Applications 59(4): 4835-4845. https://doi.org/10.1109/TIA.2023.3262230

[15] Branco N W, Cavalca M S M, Ovejero R G. (2024) Bootstrap aggregation with Christiano–Fitzgerald random walk filter for fault prediction in power systems. J Electrical Engineering, 106(3): 3657-3670. https://doi.org/10.1007/s00202-023-02146-1

[16] Dwivedi A, Tajer A. (2023) GRNN-based real-time fault chain prediction. J IEEE Transactions on Power Systems

39(1): 934-946. https://doi.org/10.1109/TPWRS.2023.3258740
[17] Yuming H, Jiaohong L, Zhenguo M, et al. (2023) On combined PSO-SVM models in fault prediction of relay protection equipment. J Circuits, Systems, and Signal Processing 42(2): 875-891. https://doi.org/10.1007/s00034-022-02056-w
[18] Khan A A, Liu D S, Shaikh A A, et al. (2021) EPS-Ledger: Blockchain hyperledger sawtooth-enabled distributed power systems chain of operation and control node privacy and security. J Electronics 10(19): 2395. https://doi.org/10.3390/electronics10192395
[19] Rashid M, Li H, Javed A R, et al. (2023) Artificial intelligence and blockchain technology for secure smart grid and power distribution automation: A state-of-the-art review. J Sustainable Energy Technologies and Assessments 57: 103282. https://doi.org/10.1016/j.seta.2023.103282
[20] Ghadi Y Y, Mazhar T, Aurangzeb K, et al. (2024) Security risk models against attacks in smart grid using big data and artificial intelligence. J PeerJ Computer Science 10: e1840. https://doi.org/10.7717/peerj-cs.1840
[21] Guo P, Xiao K, Wang X, et al. (2024) Multi-source heterogeneous data access management framework and key technologies for electric power Internet of Things. J Global energy interconnection 7(1): 94-105. https://doi.org/10.1016/j.gloei.2024.01.009
[22] Ganjkhani M, Gholami A, Giraldo J, et al. (2023) Multi-source data aggregation and real-time anomaly classification and localization in power distribution systems. J IEEE Transactions on Smart Grid 15(2): 2191-2202. https://doi.org/10.1109/TSG.2023.3316548
[23] Jia Y, Sheng W, Cai Z, et al. (2025) The application of big data and artificial intelligence technology in power system state monitoring and fault diagnosis: A review. J Advances in Resources Research 5(3): 1318-1334. https://doi.org/10.50908/arr.5.3_1318
[24] Colak A M. (2024) Parameters optimized with the pattern search algorithm for fuzzy control based active power factor correction. J Electric Power Components and Systems 52(7): 1115-1128. https://doi.org/10.1080/15325008.2023.2297999
[25] Soares M B, Guimarães É C, Rodrigues D B, et al. (2025) Control strategy for voltage disturbance compensation and power factor correction with on-site short-circuit level calculation using grid-following inverter. J IET Power Electronics 18(1): e70123. https://doi.org/10.1049/pel2.70123
[26] Su T, Zhao J, Gomez-Exposito A, et al. (2025) Grid-enhancing technologies for clean energy systems. J Nature Reviews Clean Technology 1(1): 16-31. https://doi.org/10.1038/s44359-024-00001-5
[27] Ghazal T M, Hasan M K, Mokhtar U A, et al. (2025) Machine learning-based real-time outage fault detection for distribution networks in smart grid. J Energy Reports 14: 3739-3752. https://doi.org/10.1016/j.egyr.2025.10.045
[28] Li Z. (2025) Fault traceability algorithm for communication networks based on fault tree and decision tree. J International Journal of Network Security 27(1): 152-162. https://doi.org/10.6633/IJNS.202501_27(1).17
[29] Wang C, Yan M, Pang K, et al. (2023) Cyber-physical interdependent restoration scheduling for active distribution network via ad hoc wireless communication. J IEEE Transactions on Smart Grid 14(5): 3413-3426. https://doi.org/10.1109/TSG.2023.3244785
[30] Li Y, Su S, Hu F, et al. (2025) A novel fault location method for distribution networks with distributed generators based on improved seagull optimization algorithm. J Energy Reports 13(2): 3237-3245. https://doi.org/10.1016/j.egyr.2025.02.058
[31] Alhanaf A S, Farsadi M, Balik H H. (2024) Fault detection and classification in ring power system with DG penetration using hybrid CNN-LSTM. J IEEE Access 12: 59953-59975. https://doi.org/10.1109/ACCESS.2024.3394166
[32] Arunachalam S, Kumar A K V, Reddy D N, et al. (2025) Modeling of chimp optimization algorithm node localization scheme in wireless sensor networks. J Int J Reconfigurable & Embedded Syst 14(1): 221-230. https://doi.org/10.11591/ijres.v14.i1.pp221-230
[33] Chen L F, Zhang S P, Qin Y Y, et al. (2025) Software defect prediction based on support vector machine optimized by reverse differential chimp optimization algorithm. J International Journal of Data Science and Analytics 20(5): 4423-4448. https://doi.org/10.1007/s41060-025-00726-x
[34] Saadeh M, Koutsougeras C. (2025) Analysis of decision-making algorithm to operate an industrial grid conveying system. J Journal of Intelligent Manufacturing and Special Equipment 6(2): 198-209. https://doi.org/10.1108/JIMSE-05-2025-0009
[35] Kumar V R, Jeyanthy P A. (2025) Fault Classification and Detection in Transmission Lines by Chimp optimization algorithm (ChOA) Associated support Vector Machine. J JITSI: Jurnal Ilmiah Teknologi Sistem Informasi 6(1): 86-101. https://doi.org/10.62527/jitsi.6.1.284
[36] Anka F. (2025) A multi-strategy chimp optimization algorithm for solving global and constraint engineering problems. J Knowledge and Information Systems 67(8): 6753-6802. https://doi.org/10.1007/s10115-025-02422-5
[37] Saifullah S, Yudhana A, Suryotomo A P. (2025) Automatic brain tumor segmentation: advancing U-Net with ResNet50 encoder for precise medical image analysis. J IEEE Access 13(3): 43473-43489. https://doi.org/10.1109/ACCESS.2025.3547430
[38] Manakkadu S, Dutta S. (2024) Ant colony optimization based support vector machine for improved classification of unbalanced datasets. J Procedia Computer Science 237: 586-593. https://doi.org/10.1016/j.procs.2024.05.143
[39] Deevi D P, Kodadi S, Allur N S, et al. (2025) Improvement and application of particle swarm optimization algorithm. J Intelligent Decision Technologies 19(4): 2347-2367. https://doi.org/10.1177/18724981251331781
[40] Martini Silva R, de Oliveira Vitor A L, Favoretto Castoldi M, et al. (2026) Optimized time–frequency analysis for induction motor fault detection using hybrid differential evolution and deep learning techniques. J International Journal of Adaptive Control and Signal Processing 40(1): 2-26. https://doi.org/10.1002/acs.4075