processes

MDPI

Article

Fault Early Warning Method for Hydropower Units Based on an Improved Bees Algorithm with Two Parameters and Broad Learning System

Yuanshou Liu [1], Wei Wang [1], Jianshuang Bo [1,*], Dajun Mao [2] and Yian Wu [2]

[1] Yalong River Hydropower Development Company, Ltd., Chengdu 610051, China
[2] Faculty of Artificial Intelligence, Shanghai University of Electric Power, Shanghai 200090, China
* Correspondence: bojianshuang@ylhdc.com.cn

Abstract

To address pronounced fluctuations in monitored parameters and the difficulty of timely and accurate detection of incipient anomalies in hydropower units, this paper proposes a fault early warning method based on an improved Bees Algorithm with Two Parameters (BA2) and the broad learning system (BLS). First, while retaining the triangular distribution neighborhood search mechanism of BA2, opposition reflection search and a simulated annealing acceptance criterion are introduced to construct an improved BA2 (IBA2), and five-fold expanding window time series cross-validation is used as the fitness evaluation to automatically optimize key BLS hyperparameters. The optimized BLS is then applied to normal behavior modeling of thrust bearing temperature; operating variables under healthy conditions are used to predict thrust bearing temperature, thereby establishing a dynamic normal baseline adaptive to changes in operating conditions. On this basis, the model prediction residuals are processed using a 30-point moving average, and a fault early warning mechanism combining two-sided 3σ control limits with a persistent exceedance confirmation rule is developed to suppress false alarms caused by isolated noise and brief disturbances. Results from an actual unit show that the proposed method can identify the anomaly 381 min before the abnormality was recorded on site. Further comparisons with three prediction models show that IBA2-BLS achieves lower prediction errors and favorable stability across repeated runs. Benchmark function experiments indicate favorable optimization performance of IBA2, while ablation experiments support the effectiveness of the proposed optimization framework. These results demonstrate the feasibility of the proposed method for identifying incipient anomalies in the investigated hydropower unit and indicate its potential for condition monitoring and fault early warning applications.

Keywords: hydropower unit; fault early warning; normal behavior modeling; broad learning system; improved bees algorithm with two parameters; prediction residual

Check for updates

Academic Editor: José de Jesús Rangel Magdaleno

Received: 10 September 2026
Revised: 20 October 2026
Accepted: 22 October 2026
Published: 27 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Hydropower units are the core equipment through which hydropower plants convert hydraulic energy into electrical energy, and their operational reliability directly affects plant safety, stability, and economic performance [1]. As units age and increasingly participate in peak regulation, frequency regulation, and frequent operating condition transitions, they are subjected to prolonged coupled hydraulic, mechanical, and electrical effects, resulting in pronounced nonlinearity and uncertainty in their operating states. If incipient anomalies are not identified at an early stage, they may develop into component damage or unplanned

outages. Timely detection of abnormal changes in hydropower unit operating conditions is therefore fundamental to safe operation and predictive maintenance [2–5].

Conventional fault early warning for hydropower units mainly relies on fixed thresholds, trend analysis, and expert rules, with unit condition determined by whether monitored parameters exceed preset limits or satisfy specific rules [5]. These methods are simple and easy to implement in engineering practice, but their warning performance depends strongly on threshold settings and operating experience. Because monitored parameters vary with head, load, and operating mode, fixed thresholds cannot accurately represent the normal fluctuation ranges of state parameters under different operating conditions [1]: excessively high thresholds can delay alarms, whereas excessively low thresholds can cause false alarms during normal condition transitions. In addition, single-parameter thresholds cannot fully exploit correlations among measurement points and are therefore limited in detecting weak incipient anomalies [4,5].

With the continuous improvement of condition monitoring systems in hydropower plants, statistical analysis and normal behavior modeling have increasingly been applied to fault early warning. These methods use healthy operating data to characterize multivariable relationships under normal conditions and detect anomalies according to the deviation of actual observations from normal patterns. Previous studies have constructed composite condition indicators or normal behavior models to identify abnormal states in advance, and have demonstrated the feasibility of data-driven early warning methods using operational data from actual hydropower plants [1–5]. Compared with fixed thresholds, normal behavior modeling provides a dynamic baseline that changes with operating conditions, offering a more adaptive approach to early fault warning for hydropower units.

On this basis, machine learning and deep learning have been applied to anomaly detection and early warning for hydropower units, including deep autoencoders, hierarchical temporal memory, and recurrent neural networks [6–14]. These methods enhance the extraction of nonlinear relationships and temporal features from operational data and provide a basis for fault early warning using prediction deviations or reconstruction errors.

The broad learning system (BLS) is an efficient learning framework that does not require a deep network architecture; it forms feature representations by laterally expanding mapping and enhancement nodes and determines the output weights through a regularized closed-form solution [15,16]. Condition monitoring data for hydropower units typically consist of multiple synchronously sampled operating measurements and therefore exhibit typical multivariable tabular data characteristics. BLS can establish nonlinear mappings between operating condition variables and key state parameters and can characterize deviations from healthy operating patterns through the differences between model predictions and actual observations, making it suitable for normal behavior modeling and anomaly early warning. Moreover, its simple structure and fast training reduce the computational cost of repeated model training and parameter evaluation. Although other efficient learning models, such as the extreme learning machine, also offer fast training and nonlinear modeling capabilities [17], BLS adopts a broad architecture in which the outputs of feature mapping nodes and enhancement nodes are concatenated to form a unified feature representation. This structure allows different types of transformed features to jointly characterize the nonlinear relationships in multivariable operating data. Furthermore, the original BLS framework supports incremental updates when additional nodes or new samples are introduced [15], providing the potential for future model expansion and adaptation. These characteristics make BLS a flexible and computationally efficient framework for normal behavior modeling. However, BLS has not yet been sufficiently explored for normal behavior modeling of hydropower units or for prediction residual-based early anomaly warning. Therefore, this study first investigates the applicability of BLS to normal behavior modeling

under varying operating conditions and constructs a dynamic early warning baseline using prediction residuals from healthy state data.

The prediction performance of BLS is also affected by multiple hyperparameters, including the number of mapping nodes, the number of mapping windows, the number of enhancement nodes, the shrinkage coefficient, and the regularization coefficient. Therefore, efficiently identifying an appropriate combination of these hyperparameters is essential for constructing a stable normal behavior model. However, manual tuning relies heavily on prior experience, grid search is computationally expensive, and random search may produce unstable results. Consequently, it remains challenging to achieve a satisfactory balance between search efficiency and parameter quality under a limited computational budget. In this context, metaheuristic algorithms provide a practical alternative because they do not require gradient information and can directly search for suitable parameter combinations by minimizing the model validation error. Among these algorithms, the Bees Algorithm with Two Parameters (BA2) has a relatively simple parameter structure and combines global exploration by scout bees with intensified local search through differential worker bee allocation and triangular distribution neighborhood search [18], making it suitable for hyperparameter optimization under limited computational resources. Nevertheless, as the population gradually concentrates during the search process, population diversity may decrease, increasing the risk of local stagnation. To address these limitations, two complementary improvement strategies are introduced. Specifically, inspired by opposition-based learning [19], an opposition reflection neighborhood search mechanism is constructed to generate candidate solutions around opposite positions, thereby expanding the search coverage and mitigating the loss of population diversity. Meanwhile, the Metropolis acceptance criterion of simulated annealing [20] is incorporated into the scout bee update process, allowing inferior candidate solutions to be accepted with a certain probability to alleviate local stagnation. The two strategies operate at different but interconnected stages of the search process. Opposition reflection expands the range of candidate solutions during neighborhood search, while simulated annealing introduces flexibility into the subsequent solution acceptance process. Their interaction allows the algorithm to explore alternative search regions and potentially retain non-improving solutions that may facilitate further exploration, thereby reducing the risk of premature convergence. Through this complementary mechanism, the two strategies aim to improve the balance between exploration and exploitation. The resulting improved BA2 (IBA2) is then applied to optimize the five key hyperparameters of BLS.

Based on the above analysis, this paper proposes a fault early warning method based on IBA2-BLS for hydropower units. The main contributions of this study are summarized as follows:

(1) An IBA2-BLS fault early warning framework is developed for hydropower units. The key hyperparameters of BLS are optimized using IBA2 with five-fold expanding-window time-series cross-validation to establish a normal behavior model. Based on the prediction residuals, a fault early warning mechanism combining a 30-point moving average, two-sided 3σ control limits, and a persistent-exceedance confirmation rule is constructed to reduce the influence of isolated noise and transient disturbances.

(2) An IBA2 is developed to address potential diversity loss and local search stagnation. Opposition reflection search is introduced at the candidate generation stage to alleviate diversity loss as the scout population concentrates, while a simulated annealing acceptance criterion is incorporated into the scout bee update stage to reduce stagnation caused by greedy replacement. These two strategies complement each other by expanding the candidate search space and allowing probabilistic acceptance of inferior solutions, thereby improving the flexibility of the search process. Benchmark

function experiments show that, compared with other classical metaheuristics, IBA2 achieves better overall optimization accuracy and stability across repeated runs on most test functions.

(3) The feasibility of the proposed framework is evaluated using actual operating data from a hydropower unit. In the investigated thrust bearing temperature anomaly case, the method issued an early warning 381 min before the anomaly was recorded on site. Furthermore, comparisons with three prediction models show that IBA2-BLS achieves lower prediction errors and favorable stability across repeated runs. Additional comparisons on the same anomaly process dataset demonstrate that IBA2-BLS provides earlier warnings than the three comparison models in the investigated case.

The remainder of this paper is organized as follows: Section 2 reviews related research on condition monitoring and fault early warning for hydropower units; Section 3 describes the BLS model, the IBA2 search mechanism, and the BLS hyperparameter optimization method; Section 4 presents the case data, input variable selection, and experimental setup; Section 5 analyzes the hyperparameter optimization, healthy state prediction, and anomaly early warning results; Section 6 evaluates the proposed method through benchmark function tests, prediction model comparisons, and ablation experiments; and Section 7 concludes the paper and outlines future work.

2. Literature Review

Hydropower units operate for long periods under complex and varying conditions; state parameters such as temperature, vibration, and shaft swing are influenced not only by equipment health but also by load, head, and operating mode. Conventional fixed threshold alarms cannot reliably distinguish normal operating condition fluctuations from equipment anomalies, which can lead to delayed or false alarms. Establishing normal state baselines from historical operating data and identifying anomalies in advance have therefore become important research topics for safe hydropower unit operation and predictive maintenance.

Betti et al. [1] used self-organizing maps to learn normal operating patterns of hydropower plants and constructed key performance indicators from deviations between observed and normal states, enabling early detection of potential faults. Calvo-Bascones et al. [2] described normal behavior patterns using K-means clustering, self-organizing maps, and local probability density distributions, and applied them to abnormal bearing temperature monitoring in a hydropower plant. De Santis and Costa [3] applied extended isolation forests to early fault detection in small hydropower plants and validated their anomaly detection capability using one year of actual operating data.

Zemouri et al. [4] combined sparse dictionary learning with a variational autoencoder, trained a reconstruction model on healthy state data, and identified early anomalies in hydrogenerator units from reconstruction errors. Oviedo et al. [5] integrated time series clustering, autoencoders, and adaptive anomaly thresholds to achieve real-time anomaly monitoring of multivariable hydropower system data without fault labels. Wang et al. [6] combined variational mode decomposition with a deep autoencoder and identified abnormal hydropower unit states from autoencoder reconstruction residuals. Ma et al. [7] used variational mode decomposition to process complex monitoring signals and hierarchical temporal memory for anomaly detection. Xiong et al. [8] developed a recurrent neural network based on conditional quantile regression for real-time anomaly identification in hydropower units.

Baron et al. [9] incorporated vibration monitoring into predictive maintenance for turbine generator units in small hydropower plants to identify abnormal operating behavior and support maintenance decisions. Bütüner et al. [10] established a normal behavior

model of a governor hydraulic system using actual supervisory control and data acquisition (SCADA) data from a hydropower unit and used the model output as a dynamic threshold that varied with operating conditions to identify anomalies in advance. Jian and Chen [11] combined Internet of Things monitoring with a long short-term memory (LSTM) model to construct a real-time early warning system for hydropower plants through state parameter prediction, anomaly detection, and dynamic threshold adjustment. Yi et al. [12] integrated isolation forest, local outlier factor, and density estimation models for anomaly discrimination in hydropower units.

Liu et al. [13] combined boosted trees, autoencoders, and LSTM models to perform anomaly detection at a single measurement point, health state prediction at multiple measurement points, and trend analysis, and constructed a real-time integrated early warning system by setting alarm thresholds. Wu et al. [14] integrated generalized machine acoustic signatures, an auxiliary classifier generative adversarial network, and an autoassociative kernel regression model to extract early fault features of turbine generator units from prediction residuals, and validated the early warning capability using actual turbine operating data.

In summary, existing studies have employed self-organizing maps, isolation forests, autoencoders, recurrent neural networks, and multimodel fusion for anomaly detection and fault early warning in hydropower units, providing effective tools for condition monitoring under complex operating conditions. Nevertheless, several limitations remain.

On the one hand, deep learning models generally require repeated iterative training, and their performance is sensitive to hyperparameters such as network structure, learning rate, window length, and alarm threshold, increasing the computational cost of repeated training and parameter adjustment; on the other hand, BLS, which features lateral expansion and fast solution of output weights, has rarely been applied to normal behavior modeling and fault early warning for hydropower units.

Moreover, hyperparameters such as the number of mapping windows, number of mapping nodes, number of enhancement nodes, and regularization parameter are mutually coupled, and efficient determination of a suitable parameter combination requires further investigation.

Therefore, this study uses BLS to establish a normal behavior model for hydropower units and employs an improved BA2 to automatically optimize its key hyperparameters, with the aim of improving state parameter prediction accuracy and the reliability of fault early warning.

3. Proposed Fault Early Warning Method

3.1. BLS Model

The BLS, proposed by Chen and Liu [15], is a learning model that enables rapid modeling by laterally expanding feature mapping nodes and enhancement nodes.

Unlike deep learning models that rely on multilayer network structures and backpropagation training, BLS uses a flat network architecture consisting mainly of a feature mapping layer, an enhancement layer, and an output layer; its output weights can be obtained directly using a generalized inverse or regularized least-squares method.

Let the input matrix containing N samples be $\boldsymbol{X} \in \mathbb{R}^{N \times D}$, and let the output matrix be $\boldsymbol{Y} \in \mathbb{R}^{N \times Q}$, where D and Q denote the dimensions of the input and output variables, respectively.

BLS uses N_w feature windows, each containing N_f mapping nodes.

The output of the i-th feature window is expressed as

$$\boldsymbol{Z}_i = \phi_i\left(\boldsymbol{X}\boldsymbol{W}_{e_i} + \boldsymbol{\beta}_{e_i}\right) \tag{1}$$

where $i = 1, 2, \ldots, N_w$; $\boldsymbol{W}_{e_i}$ and $\boldsymbol{\beta}_{e_i}$ denote the mapping weight matrix and bias vector of the ith feature window, respectively, and $\phi_i(\cdot)$ is the feature mapping function.

Each feature window generates N_f mapping nodes; hence, $\boldsymbol{Z}_i \in \mathbb{R}^{N \times N_f}$.

Horizontally concatenating all feature windows yields the feature mapping matrix

$$\boldsymbol{Z} = \left[\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_{N_w}\right] \tag{2}$$

where $\boldsymbol{Z} \in \mathbb{R}^{N \times (N_w N_f)}$.

Next, the feature mapping matrix $\boldsymbol{Z}$ is used as the input to generate N_e enhancement nodes.

The linear input to the enhancement layer is

$$\tilde{\boldsymbol{H}} = \boldsymbol{Z}\boldsymbol{W}_h + \boldsymbol{\beta}_h \tag{3}$$

where $\boldsymbol{W}_h$ and $\boldsymbol{\beta}_h$ are the enhancement layer weight matrix and bias vector, respectively.

To control the input range of the enhancement node activation function, the shrinkage coefficient α is introduced, and its scaling factor is defined as

$$s = \frac{\alpha}{\max\left|\tilde{\boldsymbol{H}}\right|} \tag{4}$$

The enhancement node matrix is expressed as

$$\boldsymbol{H} = \xi\left(s\tilde{\boldsymbol{H}}\right) \tag{5}$$

where $\xi(\cdot)$ is the nonlinear activation function and $\boldsymbol{H} \in \mathbb{R}^{N \times N_e}$.

The feature mapping nodes and enhancement nodes are horizontally concatenated to form the BLS broad feature matrix

$$\boldsymbol{A} = [\boldsymbol{Z}, \boldsymbol{H}] \tag{6}$$

The relationship between the broad feature matrix and the output matrix is expressed as:

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{W}_o \tag{7}$$

where $\boldsymbol{W}_o$ is the output layer weight matrix.

The output weights are obtained by solving the regularized least-squares problem:

$$\boldsymbol{W}_o = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} + \lambda_o \boldsymbol{I}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y} \tag{8}$$

where λ_o is the output layer regularization coefficient and $\boldsymbol{I}$ is the identity matrix.

For a new input sample $\boldsymbol{X}_{\text{new}}$, the mapping weights, enhancement layer weights, and output layer weights determined during training are used to sequentially generate the feature mapping matrix $\boldsymbol{Z}_{\text{new}}$, the enhancement node matrix $\boldsymbol{H}_{\text{new}}$, and the broad feature matrix

$$\boldsymbol{A}_{\text{new}} = [\boldsymbol{Z}_{\text{new}}, \boldsymbol{H}_{\text{new}}] \tag{9}$$

The model prediction is then

$$\hat{\boldsymbol{Y}} = \boldsymbol{A}_{\text{new}}\boldsymbol{W}_o \tag{10}$$

BLS obtains a broad feature representation through random feature mapping and nonlinear enhancement and directly solves the output layer weights without iterative back-propagation. It therefore features a simple network structure, fast training, and convenient lateral expansion.

3.2. Improved BA2

BA2 is a variant of the BA that uses two parameters [18]. This study retains its scout bee ranking, differential worker bee allocation, and triangular distribution neighborhood search mechanisms; based on the concept of opposition-based learning [19], an opposition reflection neighborhood search is constructed to expand the coverage of candidate solutions in the BLS hyperparameter space; in addition, the Metropolis acceptance criterion of simulated annealing [20] is introduced into the scout bee update stage, allowing temporarily inferior candidate solutions to be accepted with a certain probability and thereby alleviating local stagnation.

3.2.1. Scout Bee Initialization

Suppose the optimization problem contains D_h hyperparameters; the lower and upper bounds of the j-th dimension are L_j and U_j, respectively. The algorithm randomly generates N_s scout bees in the search space.

At iteration $k = 0$, the initial position of the rth scout bee in dimension j is

$$z_{(r,j)}^{(0)} = L_j + rand(0,\ 1)\left(U_j - L_j\right) \tag{11}$$

The components of all dimensions form the scout bee position vector $\mathbf{z}_r^{(0)}$. If a candidate position exceeds the search bounds, the corresponding component is clipped to the interval $[L_j,U_j]$.

3.2.2. Worker Bee Allocation

At the beginning of each iteration, the scout bees are sorted in ascending order of objective function value. A smaller rank r indicates better fitness and therefore a larger worker bee allocation. Let $n_{\max}$ denote the maximum number of worker bees assigned to the highest ranked scout bee; the number of worker bees assigned to the r-th scout bee is:

$$n_r = \text{round}(n_{\max} - (r-1)(n_{\max} - 1)/(N_s - 1)) \tag{12}$$

This linear allocation assigns more search resources to promising regions while ensuring that lower ranked scout bees receive at least one worker bee.

When the scout bees are ranked again, each bee's current position and search state are moved together as a single unit.

3.2.3. Search Scale Based on the Triangular Distribution

In the original BA2 search mechanism, the mode of the triangular distribution is determined by the rank of the scout bee. Let m_r denote the triangular-distribution mode associated with the *r-th* ranked scout bee, the mode values are defined as:

$$m_r = \frac{r-1}{N_s - 1} \tag{13}$$

For each worker bee q assigned to scout bee r at iteration k, a fresh random variable u is independently sampled from $\mathrm{U}(0, 1)$, and the triangular search-scale coefficient $\tau_{r,q}^k$ is then generated from $T(0,m_r,1)$ by inverse-transform sampling:

$$\tau_{r,q}^k = \sqrt{u\, m_r}, \qquad 0 < u < m_r \tag{14}$$

$$\tau_{r,q}^k = 1 - \sqrt{(1-u)(1-m_r)}, \qquad m_r \le u < 1 \tag{15}$$

The actual search radius of worker bee q in dimension j at iteration k is obtained by scaling the current patch size of its scout bee:

$$R_{r,q,j}^k = \tau_{r,q}^k S_{r,j}^k \tag{16}$$

where $R^k_{r,q,j}$ denotes the search radius, $S^k_{r,j}$ denotes the current patch size, which is initialized as $U_j - L_j$, and $\tau^k_{r,q}$ is the triangular random coefficient.

3.2.4. Dual Operator Worker Bee Foraging

To balance local exploitation and global exploration, each worker bee selects either the original neighborhood search or the opposition reflection search with equal probability.

Let $\varepsilon_{r,q,j}$ follow a uniform distribution over $\left[-R^{(k)}_{r,q,j}, R^{(k)}_{r,q,j}\right]$.

If the random number $b_{r,q} < 0.5$, the original neighborhood search is performed:

$$y^{(N,k)}_{(r,q,j)} = z^{(k)}_{(r,j)} + \varepsilon_{(r,q,j)} \tag{17}$$

If $b_{r,q} \geq 0.5$, the opposite position of the current scout bee is first calculated as

$$\overline{z}^{(k)}_{(r,j)} = L_j + U_j - z^{(k)}_{(r,j)} \tag{18}$$

A reflection search is then performed around the opposite position:

$$y^{(OR,k)}_{(r,q,j)} = \overline{z}^{(k)}_{(r,j)} + \varepsilon_{(r,q,j)} \tag{19}$$

Accordingly, the final candidate position of the worker bee can be written as:

$$y^{(k)}_{(r,q,j)} = y^{(N,k)}_{(r,q,j)}, b_{(r,q)} < 0.5 \tag{20}$$

$$y^{(k)}_{(r,q,j)} = y^{(OR,k)}_{(r,q,j)}, b_{(r,q)} >= 0.5 \tag{21}$$

All candidate components are clipped to the interval $[L_j, U_j]$.

After all worker bees assigned to the r-th scout bee complete foraging, the worker bee with the minimum objective function value is selected as the candidate update solution, denoted by $\mathbf{y}^{*(k)}_r$.

3.2.5. Standard Simulated Annealing Update

The scout bee update follows the standard simulated annealing criterion.

Let the objective function difference between the candidate update solution and the current scout bee be

$$\Delta J^{(k)}_r = J(y^*_r(k)) - J\left(z^{(k)}_r\right) \tag{22}$$

If $\Delta J^{(k)}_r \leq 0$, the candidate solution is better than or equal to the current solution and is accepted directly; if $\Delta J^{(k)}_r > 0$, the candidate solution is accepted according to the standard Metropolis probability:

$$P_{acc} = exp\left(-\Delta J^{(k)}_r / \Theta^{(k)}\right) \tag{23}$$

A uniformly distributed random number is generated. If the acceptance condition is satisfied, the current scout bee is replaced by the candidate solution; otherwise, its current position is retained.

The simulated annealing temperature follows a geometric cooling schedule, where η is the cooling coefficient:

$$\Theta^{(k+1)} = \eta\Theta^{(k)}, 0 < \eta < 1 \tag{24}$$

3.2.6. Search Range Update and Overall Procedure

After each iteration, the patch size of each scout bee is updated as:

$$S^{k+1}_{r,j} = \begin{cases} S^k_{r,j}, & \Delta J^k_r < 0, \\ \rho S^k_{r,j}, & \Delta J^k_r \geq 0, \end{cases} \qquad 0 < \rho < 1 \tag{25}$$

The objective function values of all updated scout bees are then recalculated and the bees are ranked again; meanwhile, the historical global best position is recorded independently to prevent the loss of the best solution found when simulated annealing accepts an inferior candidate.

The algorithm repeatedly performs worker bee allocation, triangular distribution search scale generation, dual operator foraging, best worker selection, simulated annealing update, and patch-size update until the maximum number of iterations or the runtime limit is reached; the historical global best position and its objective function value are then returned.

3.3. BLS Hyperparameter Optimization

The five key BLS hyperparameters are encoded as a scout bee position vector: the number of mapping nodes per window, the number of mapping windows, the number of enhancement nodes, the shrinkage coefficient, and the output layer regularization coefficient.

Each candidate position is first decoded into a valid hyperparameter set and then evaluated using five-fold expanding window time series cross-validation [21]. In each fold, BLS is trained using all samples preceding the validation period, and the root mean square error (RMSE) is calculated on the subsequent contiguous validation samples.

IBA2 uses the mean validation RMSE across the five folds as the fitness function and minimizes it:

$$f(\mathbf{v}) = \frac{1}{5}\sum_{k=1}^{5}\sqrt{\frac{1}{N_k}\sum_{i=1}^{N_k}\left[y_{k,i} - \hat{y}_{k,i}(\mathbf{v})\right]^2} \tag{26}$$

where v is a candidate BLS hyperparameter vector; N_k is the number of validation samples in fold k; and $y_{k,i}$ and $\hat{y}_{k,i}(\mathbf{v})$ are the actual and predicted values, respectively, of sample i in fold k.

The BLS hyperparameter optimization procedure is illustrated in Figure 1.

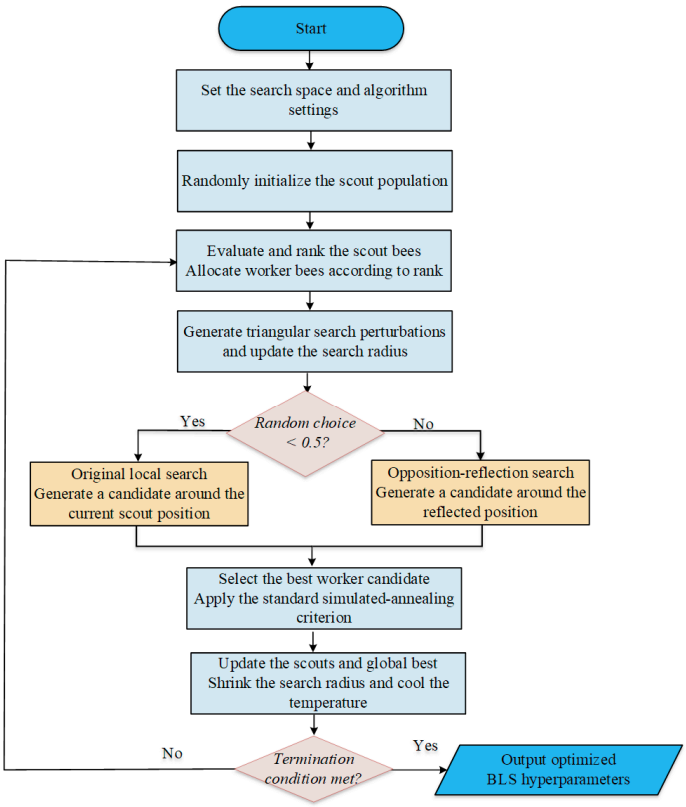


Figure 1. Flowchart of the Improved BA2-Based BLS Hyperparameter Optimization.

4. Case Study and Experimental Setup

4.1. Data Source and Task Definition

To verify the feasibility of the proposed IBA2-BLS fault early warning method, a case study was conducted using actual operating data from a hydropower unit at a hydropower plant. The investigated unit is equipped with a suspended-type hydrogenerator (model SF60-28/6500), with a rated power output of 60 MW, a rated voltage of 10.5 kV, a rated frequency of 50 Hz, and a rated rotational speed of 214.3 r/min. The generator adopts a closed-loop self-circulating air-cooling system. The monitoring data were collected at 1-min intervals. During operation, the unit experienced an abnormal increase in thrust bearing temperature, and the anomaly was recorded on site at 22:03 on 19 November 2025. The healthy operating samples were selected based on the plant's operation records and included normal power-generation operation, normal cessation of power generation, and subsequent cooling processes.

First, the model inputs and prediction output were determined according to the unit's measurement point configuration and the correlations between monitored variables and thrust bearing temperature; next, 8000 consecutive healthy operation samples were divided chronologically at a 7:3 ratio; the first 5600 samples were used as the training set for IBA2-BLS model training, and the remaining 2400 samples were used as the test set to evaluate healthy state prediction performance and establish the early warning thresholds; finally, 800 samples before the recorded anomaly, the anomaly time sample, and 800 samples after it were extracted to form a 1601-sample anomaly process dataset for evaluating the model's advance warning capability. The symmetric observation window was selected to capture the temporal evolution of the monitored variables before and after the recorded anomaly. At a 1-min sampling interval, the 800 samples preceding the anomaly (approximately 13.3 h) allow the residual behavior before the early warning and the subsequent emergence of persistent deviations to be examined, while the 800 samples after the anomaly provide an observation period for examining the subsequent development of abnormal behavior.

4.2. Input Variable Selection

Thrust bearing temperature is jointly affected by unit load, rotational speed state, and the thermal states of adjacent bearings. Based on the unit's measurement point configuration, Z9, which reflects thrust bearing temperature, was selected as the prediction output of the normal behavior model. Candidate temperature variables included seven upper guide measurement points (Z17, Z19, Z20, Z21, Z23, Z25, and Z26) and eight lower guide measurement points (Z27, Z29, Z31, Z32, Z33, Z35, Z37, and Z38). Active power and frequency were also retained as basic input variables characterizing the unit's operating condition.

Combining physical considerations with Pearson correlation analysis provides a practical approach to preliminary input variable selection in equipment condition monitoring [22]. Accordingly, Pearson correlation was used to identify representative temperature measurement points from the upper and lower guide bearings, as shown in Table 1. Among the upper guide temperature points, Z26 had the highest correlation with Z9 (0.9267); among the lower guide temperature points, Z38 had the highest correlation (0.9395). Therefore, Z26 and Z38 were selected to represent the upper guide and lower guide thermal states, respectively, and were combined with active power and frequency as model inputs. Selecting one representative point per group reduced the number of temperature-related inputs from 15 to two while retaining measurements from both bearing locations. The selected inputs were subsequently used by BLS to model their nonlinear relationships with Z9 through its feature mapping and enhancement nodes.

Table 1. Pearson Correlation Coefficients between Candidate Temperature Measurement Points and Z9.

Upper-Guide Measurement Point	Pearson Coefficient	Lower-Guide Measurement Point	Pearson Coefficient
Z17	0.8552	Z27	0.8701
Z19	0.8776	Z29	0.8850
Z20	0.9004	Z31	0.7641
Z21	0.9015	Z32	0.7392
Z23	0.8721	Z33	0.9343
Z25	0.8461	Z35	0.8335
Z26	0.9267	Z37	0.8791
-	-	Z38	0.9395

The bearing arrangement and the locations of the selected temperature measurement points are illustrated in Figure 2. Z9 is located at the thrust bearing and serves as the prediction output, while Z26 and Z38 are located at the upper and lower guide bearings, respectively, and are used as model inputs. In addition, active power and frequency are obtained from the unit's electrical metering system and are used as input variables to characterize the operating conditions.

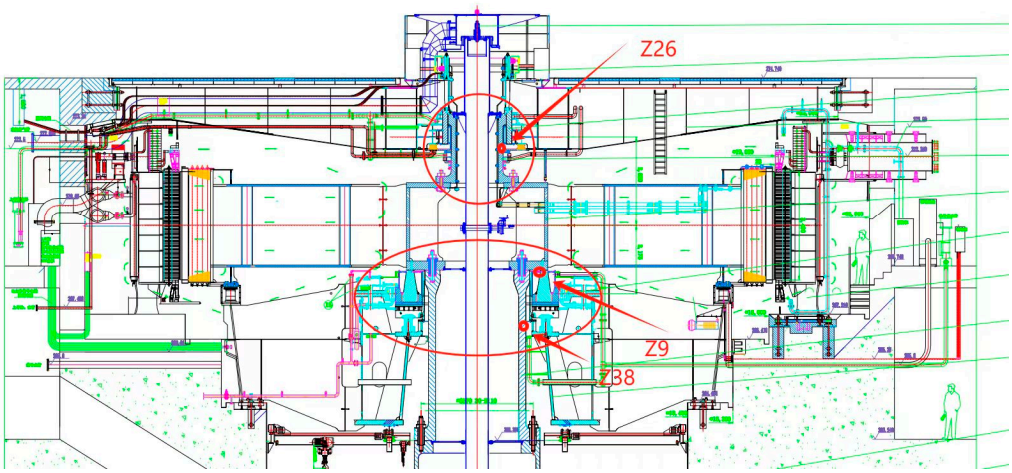


Figure 2. Bearing arrangement and locations of the temperature measurement points in the hydrogenerator.

4.3. Experimental Setup

A total of 8000 consecutive healthy operation samples were used in the experiments, of which 5600 were allocated for model training and 2400 for independent testing. Active power, frequency, Z26, and Z38 were used as inputs, and Z9 was used as the prediction output. The experiments were implemented in Python 3.11.4. For IBA2, N_s was set to 30, $n_{\max}$ to 3, Θ_0 to 0.25, η to 0.95; and ρ to 0.99. Regarding the termination criterion, a fixed maximum runtime of 20 min was adopted for IBA2-based hyperparameter optimization, consistent with the maximum computing time rule commonly used in metaheuristic optimization studies [23]. This criterion provides a predefined computational time budget without requiring additional stagnation-detection parameters.

The BLS used tanh as the activation function, and the remaining hyperparameters and their search ranges are listed in Table 2.

Table 2. BLS Hyperparameter Search Space.

Hyperparameter	Description	Search Range
Nf	Number of mapping nodes per window	[5, 15]
Nw	Number of mapping windows	[5, 15]
Ne	Number of enhancement nodes	[20, 200]
α	Enhancement-layer shrinkage coefficient	[0.1, 1.5]
λ_o	Output-layer regularization coefficient	$[10^{-8}, 1]$

5. Results and Analysis

5.1. Hyperparameter Optimization Results

Figure 3 illustrates the convergence of the five-fold mean validation RMSE during IBA2-based BLS hyperparameter optimization, while Table 3 lists the optimal hyperparameter combination obtained by IBA2. Based on these results, the final BLS model was trained using all 5600 training samples. The remaining 2400 independent healthy test samples were then used to evaluate the prediction performance of the optimized model and to establish the normal-state residual control limits for subsequent anomaly early warning.

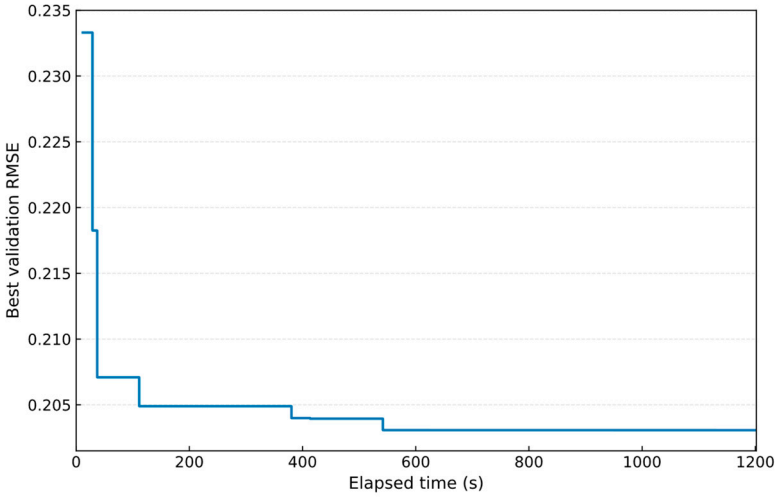


Figure 3. Convergence of the Five-Fold Mean RMSE during IBA2 Optimization of BLS.

Table 3. Optimal BLS Hyperparameters Obtained by IBA2.

Parameter	Optimal Value
Nf	6
Nw	7
Ne	159
α	1.0710
λ_o	1.2548×10^{-8}

5.2. Healthy State Prediction and Threshold Determination

As shown in Figure 4, the predicted thrust bearing temperature generally follows the variation trend of the measured temperature on the independent healthy test set. This indicates that the optimized BLS model can characterize the relationship between the operating

variables and thrust bearing temperature under the investigated healthy operating conditions, thereby providing a normal behavior baseline. However, the comparison between measured and predicted values alone cannot provide a quantitative criterion for anomaly identification. Therefore, the prediction residuals are further analyzed to determine the early warning thresholds. In addition, it should be noted that a pronounced decrease in thrust bearing temperature is observed between sample indices 1000 and 1300 in Figure 4. During this period, the unit exited power generation and entered the subsequent cooling stage. Therefore, the observed temperature decrease was associated with a normal operating condition transition rather than an equipment fault.

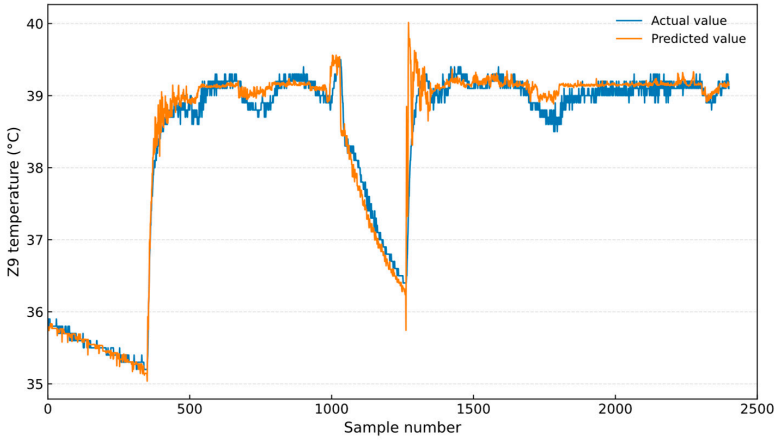


Figure 4. Measured and Predicted Z9 Values on the Independent Healthy Test Set.

The prediction residuals of the test set are shown in Figure 5. The residuals were then used to determine the early warning thresholds; to reduce the influence of isolated noise and brief operating condition fluctuations on threshold determination, a 30-point moving average was applied to the prediction residuals. Based on the mean and standard deviation of the moving average residuals from healthy data, two-sided 3σ limits [24] were used to determine the lower and upper control limits as −0.5170 and 0.4133, respectively, which were adopted as the anomaly early warning thresholds.

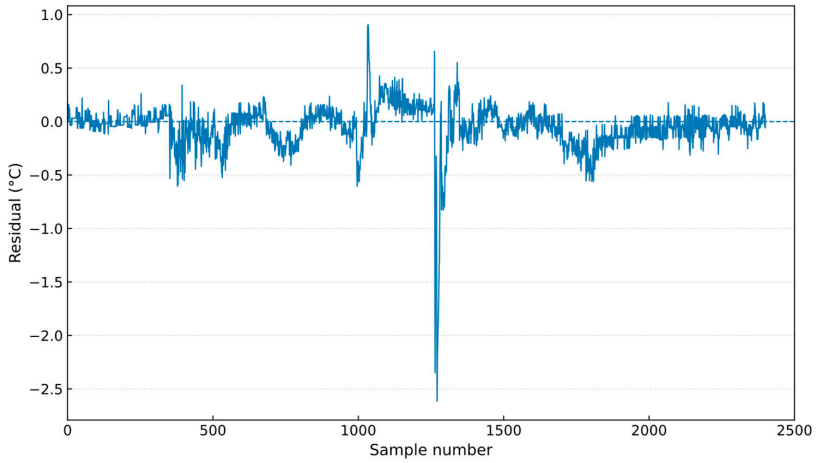


Figure 5. Prediction Residuals on the Independent Healthy Test Set.

5.3. Early Warning Results for the Anomaly Process

After the early warning thresholds were determined, the 1601 consecutive samples comprising 800 records before and 800 records after the site recorded anomaly time, together with the anomaly time record itself, were input to the trained IBA2-BLS model, and

prediction residuals were calculated for each sampling instant. To reduce the effects of isolated noise and brief operating disturbances, a 30-point moving average was applied to the prediction residuals, and the resulting moving average residuals were compared with the two-sided 3σ control limits established from the independent healthy data.

To avoid false alarms caused by transient exceedances, an anomaly was confirmed and an alarm was issued only when the moving average residual remained above the upper control limit or below the lower control limit for 20 consecutive moving windows; the results are shown in Figure 6.

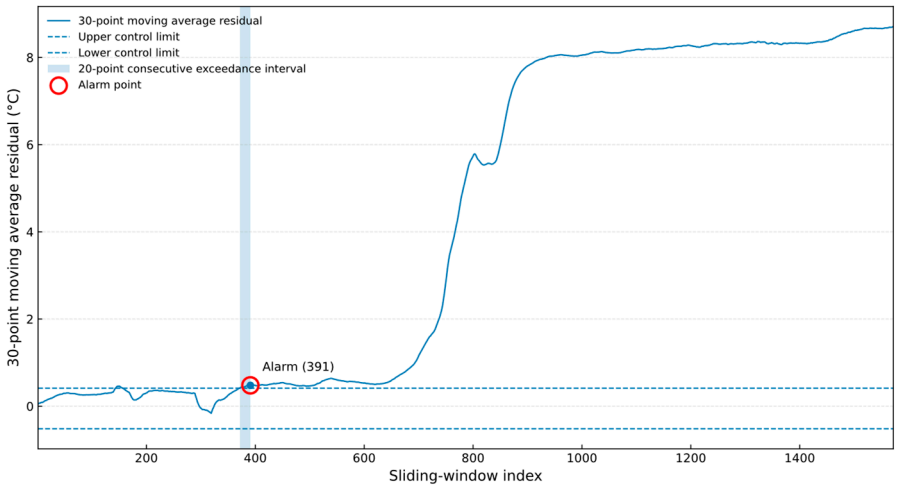


Figure 6. Moving Average Residuals and Persistent Exceedance Early Warning Result for the Anomaly Process.

As shown in Figure 6, the moving average residual remained above the upper control limit from moving window index 372 (15:23 on 19 November 2025) onward and satisfied the anomaly confirmation condition of 20 consecutive exceedance windows at index 391 (15:42). Compared with the site recorded anomaly time of 22:03, the proposed method issued the early warning 381 min, i.e., 6 h 21 min, in advance.

6. Algorithm Performance Analysis

6.1. IBA2 Performance Analysis

To evaluate the global optimization performance of IBA2, particle swarm optimization (PSO) [25], differential evolution (DE) [26], and the grey wolf optimizer (GWO) [27] were selected as comparison algorithms. Experiments were conducted on eight CEC 2017-based benchmark functions, namely F1, F3, F5, F10, F12, F17, F22, and F29, implemented using the Python OPFUNU library (version 1.0.4) [28]. The function labels and definitions followed those provided by OPFUNU 1.0.4. All benchmark functions were configured with 30 dimensions and a search space of $[-100, 100]^{30}$, and were treated as minimization problems.

To reduce the influence of manually selected hyperparameters on the experimental results, Optuna 5.0.0 [29] was employed to tune the hyperparameters of each algorithm separately for each benchmark function. For each algorithm–function combination, 30 Optuna trials were conducted. Each candidate hyperparameter configuration was evaluated through 10 independent runs, with a maximum runtime of 10 s per run.

For IBA2, PSO, and DE, Optuna's Tree-structured Parzen Estimator sampler was used for hyperparameter optimization. For all four algorithms, the population size was searched within the range of 20–100. For IBA2, the search ranges of the remaining hyperparameters were defined as follows: the maximum number of worker bees per scout bee ($n_{\max} \in [2, 8]$),

the initial temperature ($\Theta_0 \in [0.05, 0.5]$), the cooling coefficient ($\eta \in [0.90, 0.99]$), and the search-radius decay coefficient ($\rho \in [0.95, 0.999]$). For PSO, the search ranges of the inertia weight, cognitive coefficient (c_1), and social coefficient (c_2) were set to [0.4, 0.9], [1.0, 2.5], and [1.0, 2.5], respectively. For DE, the scaling factor (F) and crossover probability (CR) were searched within [0.1, 1.5] and [0.1, 1.0], respectively. For GWO, Optuna's GridSampler was employed to perform a grid search over 30 predefined integer candidate population sizes within the range of 20–100.

After hyperparameter optimization, the best hyperparameter configuration obtained for each algorithm on each benchmark function was fixed, and 10 independent runs were performed. The best (Best), mean (Mean), and worst (Worst) final fitness values obtained across the 10 runs were recorded to evaluate the optimization accuracy and performance of each algorithm. The experimental results are presented in Table 4.

Table 4. Optimization Results of IBA2 and Comparison Algorithms on the Test Functions.

Test Function	Algorithm	Best	Mean	Worst
F1	IBA2	1.0000×10^{2}	1.0000×10^{2}	1.0000×10^{2}
	PSO	6.5223×10^{8}	6.5434×10^{9}	1.7213×10^{10}
	DE	3.0634×10^{2}	3.5237×10^{3}	1.2171×10^{4}
	GWO	6.0412×10^{9}	7.8941×10^{9}	8.8016×10^{9}
F3	IBA2	3.0267×10^{2}	3.2481×10^{2}	3.3181×10^{2}
	PSO	3.8128×10^{2}	7.9496×10^{2}	1.8662×10^{3}
	DE	3.3181×10^{2}	3.3187×10^{2}	3.3243×10^{2}
	GWO	6.6529×10^{2}	8.0484×10^{2}	9.0220×10^{2}
F5	IBA2	5.0000×10^{2}	5.0000×10^{2}	5.0000×10^{2}
	PSO	5.0000×10^{2}	5.0000×10^{2}	5.0001×10^{2}
	DE	5.0000×10^{2}	5.0001×10^{2}	5.0002×10^{2}
	GWO	5.0000×10^{2}	5.0000×10^{2}	5.0000×10^{2}
F10	IBA2	1.2040×10^{3}	1.4430×10^{3}	1.8803×10^{3}
	PSO	5.8764×10^{4}	6.9139×10^{7}	6.8325×10^{8}
	DE	3.1426×10^{3}	1.0766×10^{4}	2.6783×10^{4}
	GWO	7.6933×10^{4}	3.8457×10^{6}	6.1992×10^{6}
F12	IBA2	1.2666×10^{3}	2.1121×10^{3}	3.6491×10^{3}
	PSO	2.9800×10^{5}	5.9386×10^{8}	3.1593×10^{9}
	DE	2.6617×10^{3}	1.1057×10^{4}	2.8666×10^{4}
	GWO	5.4458×10^{7}	1.7555×10^{8}	2.5180×10^{8}
F17	IBA2	2.7087×10^{3}	3.8731×10^{3}	6.3655×10^{3}
	PSO	3.2978×10^{4}	1.8312×10^{5}	1.1041×10^{6}
	DE	5.6676×10^{3}	1.7163×10^{4}	4.0391×10^{4}
	GWO	5.8556×10^{5}	1.4133×10^{6}	2.0923×10^{6}
F22	IBA2	2.4000×10^{3}	2.4000×10^{3}	2.4000×10^{3}
	PSO	3.1707×10^{3}	1.6507×10^{4}	2.8899×10^{4}
	DE	2.9373×10^{3}	4.1719×10^{3}	7.9546×10^{3}
	GWO	6.3565×10^{3}	8.9425×10^{3}	1.0553×10^{4}
F29	IBA2	4.8641×10^{3}	9.8571×10^{3}	1.5309×10^{4}
	PSO	2.1070×10^{5}	2.2931×10^{9}	8.7814×10^{9}
	DE	8.1662×10^{6}	9.7034×10^{7}	2.4710×10^{8}
	GWO	6.1053×10^{7}	6.7601×10^{8}	9.4640×10^{8}

As shown in Table 4, IBA2 achieved the lowest or joint-lowest Best, Mean, and Worst values across the eight selected test functions. It obtained lower values than PSO, DE, and GWO on F1, F3, F10, F12, F17, F22, and F29. On F5, the results of IBA2 and GWO were identical at the reported precision, while PSO and DE also yielded similar values. These

findings indicate favorable optimization accuracy and run-to-run consistency for IBA2 on the selected benchmark functions.

To further examine the convergence behavior of the algorithms, the run with the lowest final objective function value among the 10 independent runs of each algorithm on each test function was selected. The corresponding current best objective function value was plotted against runtime, as shown in Figure 7. On F1, F3, F10, F12, F17, F22, and F29, IBA2 reached lower final objective function values than the comparison algorithms. On F5, multiple algorithms approached similar final values. These convergence results are consistent with the optimization results presented in Table 4 and further demonstrate that IBA2 can effectively explore promising regions of the search space while maintaining strong optimization performance.

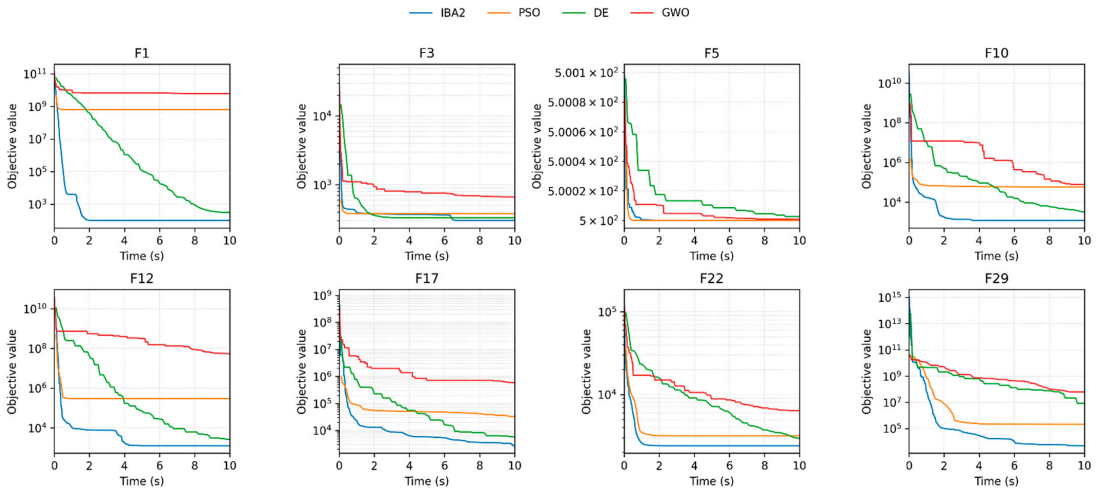


Figure 7. Best Run Convergence Curves of IBA2 and Comparison Algorithms on Eight Test Functions.

6.2. Prediction Model Performance Comparison

To evaluate the prediction performance of the proposed IBA2-BLS, it was compared with three prediction models: an adaptive genetic algorithm optimized back propagation neural network (AGABP) [30]; a deep belief network optimized by an improved particle swarm optimization algorithm (IPSO-DBN) [31]; and a gated recurrent unit model whose hyperparameters are optimized by an improved ant lion optimizer (IALO-GRU) [32].

Because model training is stochastic, each model was independently run 10 times and evaluated using RMSE, mean absolute error (MAE), and the coefficient of determination (R^2). The mean values over the 10 runs are reported in Table 5.

Table 5. Performance Comparison of Different Prediction Models.

Model	RMSE	MAE	R^2
IBA2-BLS	0.2180	0.1368	0.9724
AGABP	0.2391	0.1660	0.9668
IPSO-DBN	0.2350	0.1520	0.9679
IALO-GRU	0.2589	0.1702	0.9611

As shown in Table 5, IBA2-BLS achieved the best prediction performance across all three evaluation metrics. Compared with AGABP, IBA2-BLS reduced RMSE and MAE by 8.82% and 17.59%, respectively, while increasing R^2 by 0.0056. Compared with IPSO-DBN, RMSE and MAE were reduced by 7.23% and 10%, respectively, with an absolute increase of 0.0045 in R^2. Compared with IALO-GRU, IBA2-BLS achieved reductions of 15.80%

and 19.62% in RMSE and MAE, respectively, while R^2 increased by 0.0113. These results indicate that IBA2-BLS provides more accurate predictions and more effectively captures the nonlinear mapping between the operating variables and the target state parameter.

To jointly evaluate prediction accuracy and stability across repeated runs, the relative deviation index (RDI) was used to normalize the results of the 10 independent runs. RMSE and MAE are minimization metrics, whereas R^2 is a maximization metric. Accordingly, the RDI is uniformly defined as

$$RDI_{ir}^{g} = \begin{cases} \frac{x_{ir}^{g} - x_{\min}}{x_{\max} - x_{\min}}, & g \in \{\mathrm{RMSE}, \mathrm{MAE}\}, \\ \frac{x_{\max} - x_{ir}^{g}}{x_{\max} - x_{\min}}, & g = R^2, \end{cases} \tag{27}$$

where x_{ir}^{g} denotes the value of metric g for the *r-th* run of the *i-th* method, and $x_{\min}$ and $x_{\max}$ denote, respectively, the minimum and maximum values of that metric across all runs in the current comparison set. A smaller RDI indicates better relative performance on the corresponding metric.

The mean RDI values and 95% confidence intervals for RMSE, MAE, and R^2 of the different prediction models are shown in Figures 8–10. IBA2-BLS achieved the lowest mean RDI on all three metrics; its confidence intervals were also generally narrow, indicating good stability across repeated runs while maintaining high prediction accuracy.

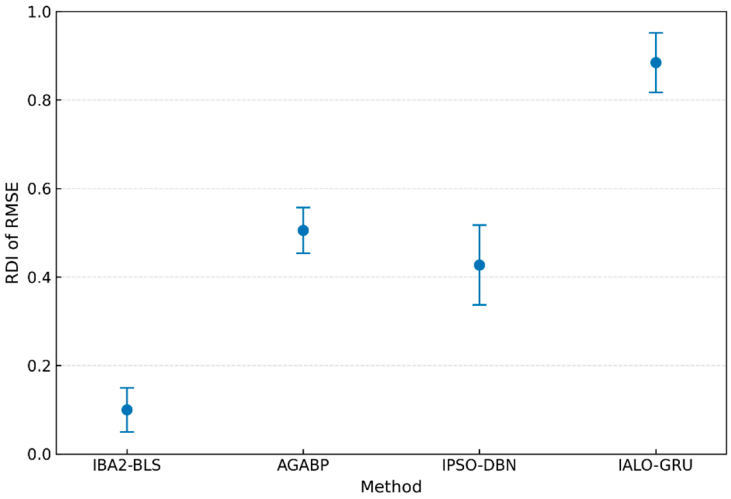


Figure 8. Mean RDI and 95% Confidence Intervals of RMSE for Different Prediction Models.

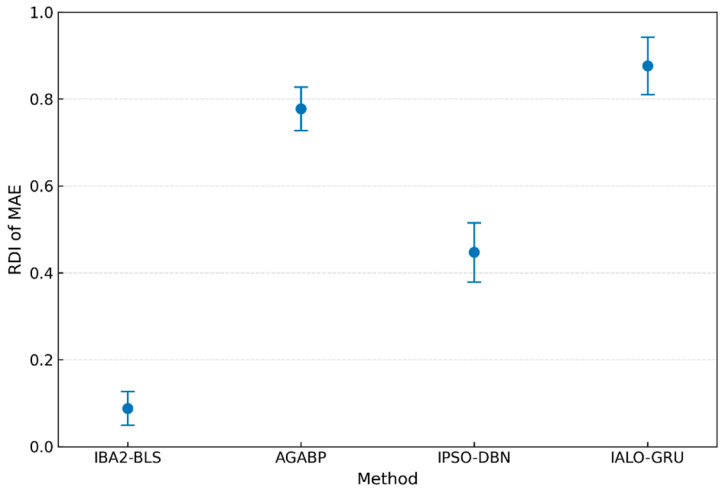


Figure 9. Mean RDI and 95% Confidence Intervals of MAE for Different Prediction Models.

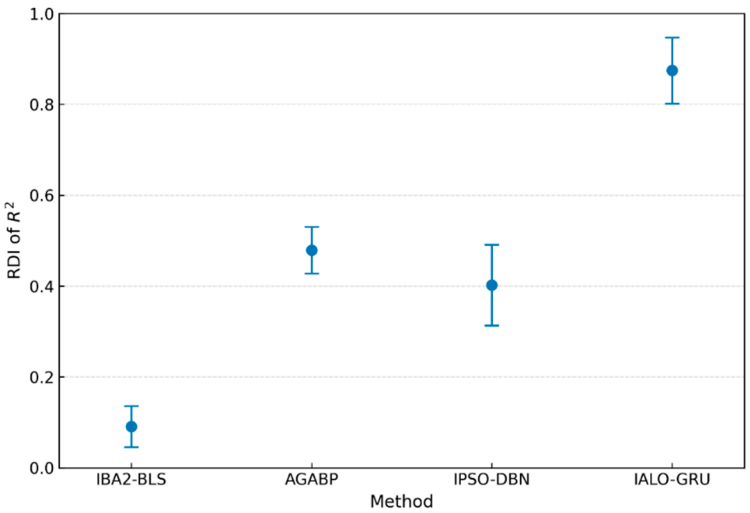


Figure 10. Mean RDI and 95% Confidence Intervals of R^2 for Different Prediction Models.

Taken together, Table 5 and Figures 8–10 show that IBA2-BLS achieves both high prediction accuracy and good stability across repeated runs under the same data split and evaluation procedure, further confirming the effectiveness of the proposed method for target state parameter prediction.

6.3. Ablation Study

To examine the contributions of BLS hyperparameter optimization and the IBA2 search mechanism to model performance, three models were considered in the ablation study: Basic BLS, BA2-BLS, and IBA2-BLS. Basic BLS does not use metaheuristic hyperparameter optimization, BA2-BLS uses BA2 to search the key BLS hyperparameters, and IBA2-BLS uses the IBA2 proposed in this study. Under the same data split and experimental conditions, each model was independently run 10 times; the mean values of the three evaluation metrics are reported in Table 6.

As shown in Table 6, compared with Basic BLS, BA2-BLS reduced RMSE and MAE by approximately 4.08% and 10.47%, respectively, and increased R^2 by 0.0026, indicating that automatically searching the key BLS hyperparameters with BA2 improves prediction performance. On this basis, IBA2-BLS further reduced RMSE and MAE by 4.34% and 8.56%, respectively, relative to BA2-BLS, and increased R^2 by 0.0026; these results indicate that the combined effects of the introduced opposition reflection search and simulated annealing acceptance criterion improve the BLS hyperparameter search results.

Table 6. Ablation Results for Different Models.

Model	RMSE	MAE	R^2
Basic BLS	0.2376	0.1671	0.9672
BA2-BLS	0.2279	0.1496	0.9698
IBA2-BLS	0.2180	0.1368	0.9724

The mean RDI values and 95% confidence intervals of the three models on RMSE, MAE, and R^2 are shown in Figures 11–13. IBA2-BLS has the lowest mean RDI on all three metrics, BA2-BLS is intermediate, and Basic BLS has the highest values.

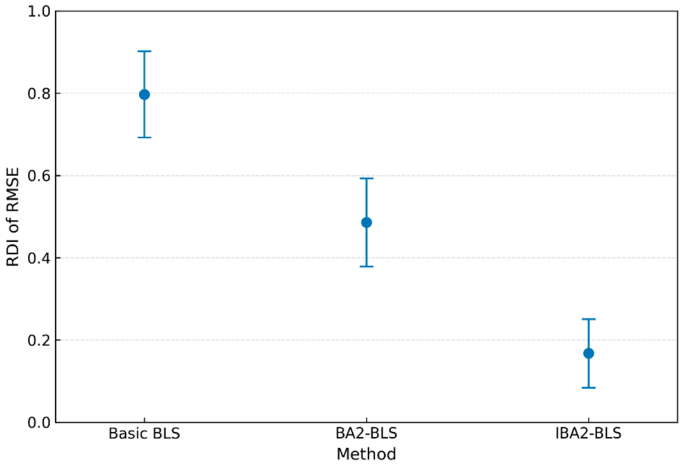


Figure 11. Mean RDI and 95% Confidence Intervals of RMSE in the Ablation Study.

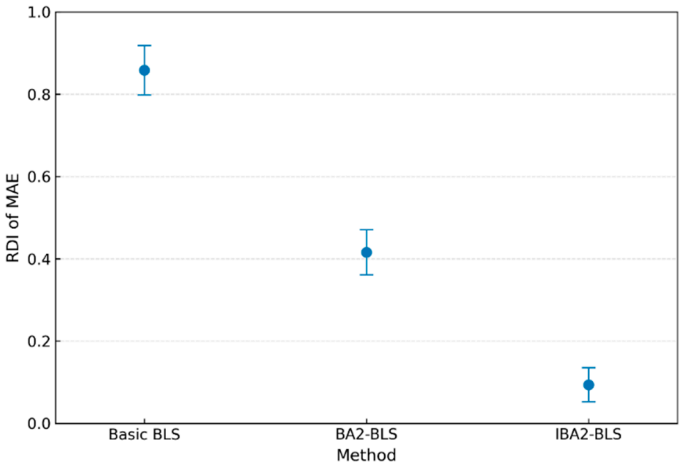


Figure 12. Mean RDI and 95% Confidence Intervals of MAE in the Ablation Study.

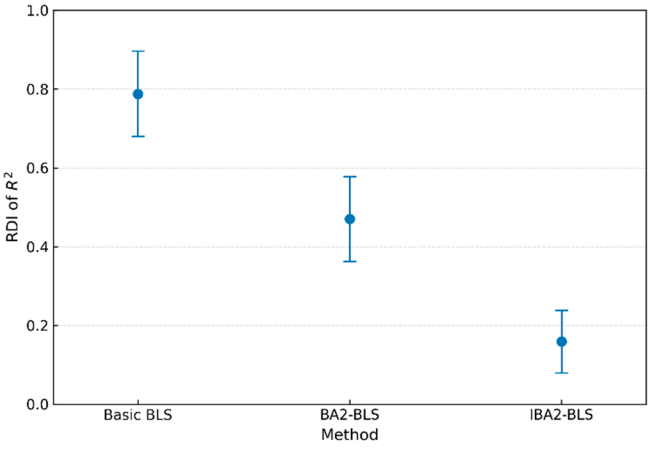


Figure 13. Mean RDI and 95% Confidence Intervals of R^2 in the Ablation Study.

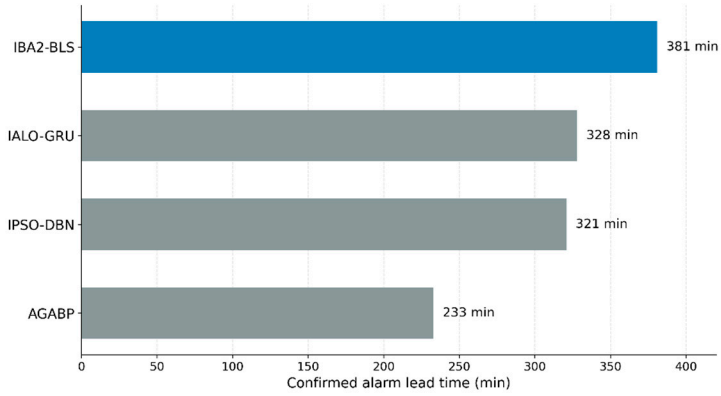
These results indicate that both BLS hyperparameter optimization and the IBA2 search mechanism contribute to performance gains.

6.4. Early Warning Performance Comparison

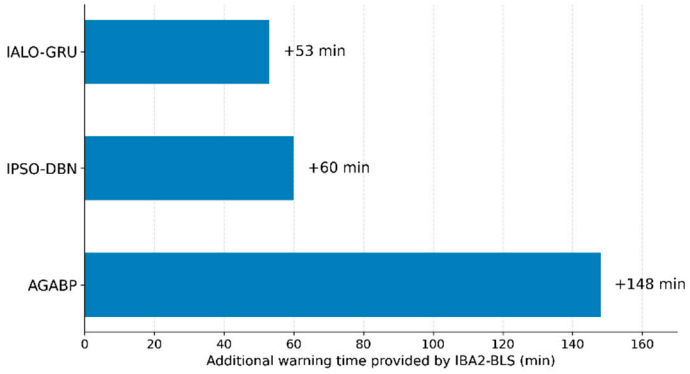
To further evaluate the effectiveness of the proposed IBA2-BLS method in anomaly early warning, three prediction models, AGABP, IPSO-DBN, and IALO-GRU, were selected for comparison using the same anomaly process dataset. Based on the results of 10 independent

runs of each comparison model, a comprehensive ranking method was employed to select a representative run with relatively favorable prediction performance. Specifically, the 10 runs were ranked in ascending order of RMSE and MAE and descending order of R^2. The rankings of the three metrics were then summed to obtain a comprehensive ranking score for each run. The run with the lowest comprehensive ranking score was selected as the representative run; in the event of a tie, the run with the lower RMSE was preferred. The selected models were subsequently applied to the same anomaly process dataset, using consistent residual processing and anomaly confirmation rules to compare their anomaly detection times and early warning capabilities.

Figure 14 compares the early warning lead times of different models on the anomaly process dataset. As shown in the figure, all four models issued warnings before the anomaly time recorded on site, although their warning lead times differed. Among them, IBA2-BLS achieved the longest warning lead time of 381 min, whereas IALO-GRU, IPSO-DBN, and AGABP achieved lead times of 328 min, 321 min, and 233 min, respectively. Compared with these three models, IBA2-BLS issued warnings 53 min, 60 min, and 148 min earlier, respectively, demonstrating better early anomaly detection performance in the investigated case. These results indicate that although different normal behavior prediction models can identify equipment anomalies through prediction residuals, the proposed IBA2-BLS method can detect deviations from normal operating behavior earlier and provides more timely warnings in this particular anomaly case.



(**a**) Confirmed Early Warning Lead Times of Different Models



(**b**) Additional Warning Lead Time Achieved by IBA2-BLS

Figure 14. Comparison of early warning performance among different prediction models.

7. Conclusions

This paper addresses the fluctuations of monitored parameters associated with operating condition changes and the difficulty of timely identification of incipient anomalies in hydropower units, as well as the challenge of selecting appropriate BLS hyperparameters,

by proposing a fault early warning method based on IBA2-BLS. Healthy operating data are used to establish a BLS normal behavior model between operating variables and thrust bearing temperature. IBA2 is constructed by incorporating opposition reflection search and a simulated annealing acceptance criterion into BA2, and five-fold expanding-window time-series cross-validation is used to optimize the key BLS hyperparameters. Two-sided 3σ control limits are then established from moving average prediction residuals of independent healthy data, together with an alarm confirmation rule requiring 20 consecutive moving-window exceedances. Benchmark function comparisons show that IBA2 obtains favorable Best, Mean, and Worst results on most of the eight selected test functions, indicating good optimization performance and stability under the experimental settings considered.

Results from the investigated hydropower unit show that IBA2-BLS can track the thrust bearing temperature trend under healthy operating conditions. Comparisons with three prediction models indicate that IBA2-BLS achieves lower prediction errors and favorable stability across repeated runs on the investigated dataset, while the ablation experiments support the effectiveness of the overall optimization framework. During anomaly process validation, the model confirmed the alarm at 15:42 on 19 November 2025, which was 381 min earlier than the anomaly time recorded on site. Further comparisons on the same anomaly process dataset show that IBA2-BLS provides earlier warnings than the three comparison models in the investigated case. These results suggest that the proposed method can identify deviations from normal operating behavior before the site-recorded anomaly in this specific case.

Owing to data limitations, this study validates the method using only one thrust bearing temperature anomaly process from a single unit, and its applicability to other units and fault scenarios remains to be further investigated. In addition, the input variable selection procedure, which combined physical considerations with Pearson correlation analysis, did not explicitly evaluate nonlinear dependencies or potential multicollinearity among the selected variables. Future work will extend the dataset to different units, operating conditions, and fault types, and further explore feature selection methods that account for nonlinear dependencies and multicollinearity, as well as online model updating, adaptive control limits, and cross-unit transfer learning.

Author Contributions: Conceptualization, Y.L.; methodology, Y.L. and W.W.; formal analysis, W.W. and D.M.; investigation, Y.W.; validation, Y.W.; data curation, Y.L. and J.B.; supervision, D.M.; writing—original draft preparation, Y.L., W.W. and J.B.; writing—review and editing, W.W., J.B., D.M. and Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data are not publicly available due to confidentiality restrictions related to industrial equipment operating data.

Conflicts of Interest: The authors Yuanshou Liu, Wei Wang, Jianshuang Bo were employed by the company Yalong River Hydropower Development Company, Ltd., Chengdu 610051, Sichuan, China. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Betti, A.; Crisostomi, E.; Paolinelli, G.; Piazzi, A.; Ruffini, F.; Tucci, M. Condition monitoring and predictive maintenance methodologies for hydropower plants equipment. *Renew. Energy* **2021**, *171*, 246–253. [CrossRef]
2. Calvo-Bascones, P.; Sanz-Bobi, M.A.; Welte, T.M. Anomaly detection method based on the deep knowledge behind behavior patterns in industrial components. Application to a hydropower plant. *Comput. Ind.* **2021**, *125*, 103376. [CrossRef]
3. de Santis, R.B.; Costa, M.A. Extended isolation forests for fault detection in small hydroelectric plants. *Sustainability* **2020**, *12*, 6421. [CrossRef]

4. Zemouri, R.; Ibrahim, R.; Tahan, A. Hydrogenerator early fault detection: Sparse dictionary learning jointly with the variational autoencoder. *Eng. Appl. Artif. Intell.* **2023**, *120*, 105859. [CrossRef]
5. Oviedo, A.I.; Vargas, J.F.; Hincapie, R.C.; Molina, A.F.; Vergara, E.A.; Tello, D.M. Unsupervised Real-Time Anomaly Detection in Hydropower Systems via Time Series Clustering and Autoencoders. *Technologies* **2025**, *13*, 534. [CrossRef]
6. Wang, H.; Liu, X.; Ma, L.; Zhang, Y. Anomaly detection for hydropower turbine unit based on variational modal decomposition and deep autoencoder. *Energy Rep.* **2021**, *7*, 938–946. [CrossRef]
7. Ma, T.; Zhu, Z.; Wang, L.; Wang, H.; Ma, L. Anomaly detection for hydropower turbine based on variational modal decomposition and hierarchical temporal memory. *Energy Rep.* **2022**, *8*, 1546–1551. [CrossRef]
8. Xiong, L.; Liu, J.; Yang, F.; Zhang, G.; Dang, J. Anomaly detection of hydropower units based on recurrent neural network. *Front. Energy Res.* **2022**, *10*, 856635. [CrossRef]
9. Baron, P.; Kočiško, M.; Hlavatá, S.; Franas, E. Vibrodiagnostics as a predictive maintenance tool in the operation of turbo generators of a small hydropower plant. *Adv. Mech. Eng.* **2022**, *14*, 16878132221101023. [CrossRef]
10. Bütüner, M.A.; Koşalay, İ.; Gezer, D. Machine-learning-based modeling of a hydraulic speed governor for anomaly detection in hydropower plants. *Energies* **2022**, *15*, 7974. [CrossRef]
11. Jian, X.; Chen, J. Real-time monitoring and early warning system for hydropower safety based on the internet of things. *Intell. Decis. Technol.* **2025**, *19*, 2295–2309. [CrossRef]
12. Yi, T.; Zhao, X.; Shi, Y.; Jing, X.; Lei, W.; Guo, J.; Meng, Y.; Zhang, Z. Anomaly Detection Method for Hydropower Units Based on KSQDC-ADEAD Under Complex Operating Conditions. *Sensors* **2025**, *25*, 4093. [CrossRef] [PubMed]
13. Liu, D.; Kong, L.; Ren, W.; Guo, Y. Real-time early warning of multi-measurement point time series of hydropower generating units based on an integrated machine learning model. *Neural Comput. Appl.* **2026**, *38*, 557. [CrossRef]
14. Wu, L.; Tian, J.; Tang, J.; Liu, D.; Cao, H. Early fault feature extraction and early warning algorithm for hydraulic turbine generator units based on generalised machine acoustic signatures. *Discov. Artif. Intell.* 2026, *in press*. [CrossRef]
15. Chen, C.P.; Liu, Z. Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *29*, 10–24. [CrossRef] [PubMed]
16. Zhao, H.; Zheng, J.; Xu, J.; Deng, W. Fault diagnosis method based on principal component analysis and broad learning system. *IEEE Access* **2019**, *7*, 99263–99272. [CrossRef]
17. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]
18. Ismail, A.H.; Ruslan, W.; Pham, D.T. A user-friendly Bees Algorithm for continuous and combinatorial optimisation. *Cogent Eng.* **2023**, *10*, 2278257. [CrossRef]
19. Tizhoosh, H.R. Opposition-based learning: A new scheme for machine intelligence. In *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*; IEEE: New York, NY, USA, 2005; Volume 1, pp. 695–701.
20. Kirkpatrick, S.; Gelatt, C.D., Jr.; Vecchi, M.P. Optimization by simulated annealing. *Science* **1983**, *220*, 671–680. [CrossRef] [PubMed]
21. Tashman, L.J. Out-of-sample tests of forecasting accuracy: An analysis and review. *Int. J. Forecast.* **2000**, *16*, 437–450. [CrossRef]
22. Liu, J.H.; Corbita, N.T., Jr.; Paclijan, G.B. Predictive framework for detecting main bearing faults in direct-drive horizontal-axis wind turbines using supervisory control and data acquisition. *Adv. Mech. Eng.* **2026**, *18*, 16878132261464392. [CrossRef]
23. Corominas, A. On deciding when to stop metaheuristics: Properties, rules and termination conditions. *Oper. Res. Perspect.* **2023**, *10*, 100283. [CrossRef]
24. Yang, B.; Zheng, R.; Han, Y.; Huang, J.; Li, M.; Shu, H.; Su, S.; Guo, Z. Recent advances in fault diagnosis techniques for photovoltaic systems: A critical review. *Prot. Control Mod. Power Syst.* **2024**, *9*, 36–59. [CrossRef]
25. Shi, Y.; Eberhart, R. A modified particle swarm optimizer. *Evol. Comput. Proc.* **1998**, *890*, 69–73.
26. Storn, R.; Price, K. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **1997**, *11*, 341–359. [CrossRef]
27. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]
28. Van Thieu, N. Opfunu: An open-source python library for optimization benchmark functions. *J. Open Res. Softw.* **2024**, *12*, 8. [CrossRef]
29. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, July 2019*; ACM: New York, NY, USA, 2019; pp. 2623–2631.
30. Yu, H.; Wang, T. A method for real-time fault detection of liquid rocket engine based on adaptive genetic algorithm optimizing back propagation neural network. *Sensors* **2021**, *21*, 5026. [CrossRef] [PubMed]

31. Zhang, Z.; Wang, S.; Wang, P.; Jiang, P.; Zhou, H. Research on fault early warning of wind turbine based on IPSO-DBN. *Energies* **2022**, *15*, 9072. [CrossRef]
32. Wang, Y.; Sun, W.; Liu, H.; Wang, S.; Zhou, Q. Study on gearbox fault warning based on the improved M-IALO-GRU model. *Appl. Sci.* **2025**, *15*, 3175. [CrossRef]

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