


## Article

# Noise Elimination for Wide Field Electromagnetic Data via Improved Dung Beetle Optimized Gated Recurrent Unit

Zhongyuan Liu <sup>1</sup>, Xian Zhang <sup>2,\*</sup>, Diqian Li <sup>1</sup> , Shupeng Liu <sup>2</sup> and Ke Cao <sup>2</sup>

<sup>1</sup> Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment, Monitoring Ministry of Education, School of Geosciences and Info-Physics, Central South University, Changsha 410083, China; 205001043@csu.edu.cn (Z.L.); lidiquan@csu.edu.cn (D.L.)

<sup>2</sup> Hunan Provincial Key Laboratory of Finance & Economics Big Data Science and Technology, School of Information Technology and Management, Hunan University of Finance and Economics, Changsha 410205, China

\* Correspondence: zxian128@163.com

**Abstract:** Noise profoundly affects the quality of electromagnetic data, and selecting the appropriate hyperparameters for machine learning models poses a significant challenge. Consequently, the current machine learning denoising techniques fall short in delivering precise processing of Wide Field Electromagnetic Method (WFEM) data. To eliminate the noise, this paper presents an electromagnetic data denoising approach based on the improved dung beetle optimized (IDBO) gated recurrent unit (GRU) and its application. Firstly, Spatial Pyramid Matching (SPM) chaotic mapping, variable spiral strategy, Levy flight mechanism, and adaptive T-distribution variation perturbation strategy were utilized to enhance the DBO algorithm. Subsequently, the mean square error is employed as the fitness of the IDBO algorithm to achieve the hyperparameter optimization of the GRU algorithm. Finally, the IDBO-GRU method is applied to the denoising processing of WFEM data. Experiments demonstrate that the optimization capacity of the IDBO algorithm is conspicuously superior to other intelligent optimization algorithms, and the IDBO-GRU algorithm surpasses the probabilistic neural network (PNN) and the GRU algorithm in the denoising accuracy of WFEM data. Moreover, the time domain of the processed WFEM data is more in line with periodic signal characteristics, its overall data quality is significantly enhanced, and the electric field curve is more stable. Therefore, the IDBO-GRU is more adept at processing the time domain sequence, and the application results also validate that the proposed method can offer technical support for electromagnetic inversion interpretation.

**Keywords:** wide field electromagnetic method; noise elimination; dung beetle optimizer; gated recurrent unit; improvement strategy



Academic Editor: Dimitrios Nikolopoulos

Received: 12 November 2024

Revised: 20 December 2024

Accepted: 23 December 2024

Published: 3 January 2025

**Citation:** Liu, Z.; Zhang, X.; Li, D.; Liu, S.; Cao, K. Noise Elimination for Wide Field Electromagnetic Data via Improved Dung Beetle Optimized Gated Recurrent Unit. *Geosciences* **2025**, *15*, 8. <https://doi.org/10.3390/geosciences15010008>

**Copyright:** © 2025 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 (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In the early 21st century, Professor Jishan He proposed the wide field electromagnetic method (WFEM), which is a kind of large-depth and high-precision controlled source electromagnetic method (CSEM) [1,2]. Based on the magnetotelluric (MT) method, it was developed successively and widely used. Significant breakthroughs have been made in method theory and detection technology and equipment. The WFEM method can obtain geoelectric information of multiple frequencies by sending and receiving at one time, which improves fieldwork efficiency and anti-interference ability [3,4]. With the continuous development and expansion of industrialization, urbanization, and the scope of human

activities, electromagnetic noise has become increasingly intense. Meanwhile, noise is the main factor affecting the exploration effect of the electromagnetic method. Moreover, electromagnetic method data are the basis of inversion calculation and interpretation. To enhance the exploration result of the electromagnetic method, data denoising is significant for obtaining the high-quality of the signal [5–9]. Therefore, it is essential to address a fundamental technical challenge, how to effectively employ the new method for the elimination of electromagnetic noise.

In the frequency domain, the spectrum of the WFEM signal is known, while the spectrum of noise is unknown. The most intuitive judgment of noise is the time domain sequence, and most noise types are named according to the morphological characteristics of the time domain. Thus, a lot of research has been conducted on current denoising methods on the time domain or frequency domain of WFEM data. Such as Zhang et al. introduced a novel adaptive bidirectional mean square deviation threshold method [10]. Mo et al. developed the gray system theory alongside a robust M-estimation approach [11]. Yang et al. presented a CSEM noise evaluation technique that utilizes wavelet transform and Hilbert analytic envelope in the frequency domain [5]. Chen et al. proposed gray judgment criteria coupled with a rational function filtering method [12]. Hu et al. implemented noise separation for CSEM data based on an enhanced clustering methodology [13]. The above-mentioned frequency domain processing method can significantly improve the data quality of less interfered data. These methods rely on power spectrum selection in the frequency domain and fail when most frequency points are distorted under the influence of persistent strong noise. Therefore, for the CSEM data processing method in the time domain, Yang et al. proposed denoising CSEM data using the least squares inversion method [14]. Li et al. proposed a CSEM denoising method based on the fast Fourier transform, complementary ensemble empirical mode decomposition (CEEMD), and shift-invariant sparse coding method [6]. Yang et al. presented a subtraction and addition method for the cancellation of powerline noise [15]. Ling et al. proposed a high-frequency information extraction method based on time-domain signal reconstruction [16]. Zhang et al. introduced the signal–noise identification method for WFEM data using multi-domain features and optimized support vector machine [17]. Li et al. developed a CSEM denoising method based on a deep temporal convolutional network combined with dictionary learning [7]. Yang et al. proposed a CSEM denoising method of a same-site noise reference channel and the mixed basis of Fourier series and Legendre polynomials [18,19]. Zhang et al. proposed the WFEM signal–noise separation method based on an improved inherent time-scale decomposition and probabilistic neural network [20]. The above-mentioned CSEM denoising methods can improve the data quality and achieve good results. When machine learning and deep learning techniques are employed in electromagnetic signal processing, data quality is typically enhanced primarily by obtaining optimal parameters through extensive experimentation, owing to the challenges posed by limited data samples and the difficulty in accurately configuring parameters. Subsequently, intelligently optimized machine learning and deep learning methodologies assume a pivotal role.

With the development of computer technology, artificial intelligence technology has been applied in practical engineering. Machine learning is widely used in signal processing. Deep learning is a significant branch of the field of artificial intelligence, which realizes autonomous learning and intelligent processing by simulating neural network learning and decision making in the human brain. However, deep learning requires the construction of neural network models, and the model usually needs to set reasonable parameters to be more suitable for actual application. Therefore, many intelligent optimization algorithms have been proposed, which can find the best parameters

from different angles. In this paper, the DBO algorithm exhibits limitations in terms of its convergence speed, optimization efficiency, and global search capability. The Dung Beetle Optimizer (DBO) algorithm has been improving, and now has stronger optimization ability and faster convergence speed. The DBO algorithm is suitable for the parameter optimization of deep learning models. The gated recurrent unit (GRU) is a special type of recurrent neural network (RNN), which is designed with the idea of simplifying the complexity of long short term memory (LSTM) networks while preserving their primary function. The structure of GRU is relatively simple, the calculation cost is low, and it is more suitable for processing time-series data.

Thus, noise elimination for electromagnetic data via the improved dung beetle optimizer gated recurrent unit (IDBO-GRU) is proposed in this paper. We utilized four kinds of improvement strategies to enhance the DBO algorithm. Then, the improved DBO algorithm is used to achieve the hyperparameter optimization of the GRU algorithm. The IDBO-GRU method is applied to the denoising processing of the WFEM data. The experiments show that the solution accuracy, evaluation indexes, and convergence speed of the IDBO algorithm are superior to other intelligent optimization algorithms. The noise is effectively identified and eliminated by using the IDBO-GRU method in the simulated and measured data application, and the reconstructed data can reflect the characteristics of the periodic signals. The comparison effect of electric field curves shows that reliable WFEM data can provide technical support for electromagnetic inversion interpretation.

Note that the aim of this paper is to achieve high-precision WFEM noise elimination. The contributions of this paper are summarized as follows:

- (1) The principle of an improved dung beetle optimizer and gated recurrent unit algorithm are introduced. And the convergence of IDBO is compared with other optimization algorithms.
- (2) The four kinds of improvement strategies were utilized to enhance the DBO algorithm, and the improved DBO algorithm itself, to achieve the hyperparameter optimization of the GRU algorithm. Meanwhile, the probabilistic neural network (PNN) and the GRU algorithm in the denoising accuracy of WFEM data are compared with the proposed method.
- (3) The effectiveness of the proposed method is verified by the simulation experiments and measured WFEM data.

## 2. Method

### 2.1. Improves Dung Beetle Optimization

The Dung Beetle Optimizer (DBO) is a newly developed swarm intelligence optimization algorithm [21]. It draws inspiration from the ball-rolling, dancing, foraging, thievery, and reproduction behaviors exhibited by dung beetles. Although the DBO algorithm has achieved good optimization results in engineering applications, it still has shortcomings in convergence speed, optimization progress, and global search ability. Based on this, the DBO algorithm is improved by using the Spatial Pyramid Matching (SPM) chaotic mapping, variable spiral search strategy, Levy flight strategy, and adaptive t-distribution perturbation strategies.

#### 2.1.1. SPM Chaotic Mapping

By randomly generating initial populations, the DBO algorithm is prone to uneven population distribution, resulting in reduced population diversity and low population quality, which affects the convergence speed of the algorithm. Chaotic mapping has the characteristics of randomness, non-repeatability, and chaos ergodicity, meaning that it can make the population distribution more uniform than random generation, which depends on

probability [22,23]. Therefore, SPM chaotic mapping is used to generate initial populations to increase the diversity of potential solutions.

SPM Chaotic mapping function definition:

$$x(t+1) = \begin{cases} \text{mod}\left(\frac{x(i)}{\eta} + \mu \sin(\pi x(i)) + r, 1\right), & 0 \leq x(i) < \eta \\ \text{mod}\left(\frac{x(i)/\eta}{0.5 - \eta} + \mu \sin(\pi x(i)) + r, 1\right), & \eta \leq x(i) < 0.5 \\ \text{mod}\left(\frac{(1-x(i))/\eta}{0.5 - \eta} + \mu \sin(\pi(1-x(i))) + r, 1\right), & 0.5 \leq x(i) < 1 - \eta \\ \text{mod}\left(\frac{1-x(i)}{\eta} + \mu \sin(\pi(1-x(i))) + r, 1\right), & 1 - \eta \leq x(i) < 1 \end{cases} \quad (1)$$

where  $r$  is a random number between 0 and 1. The mod function is the remainder of two numbers divided by each other. The function is chaotic when  $\eta \in (0,1)$  and  $\mu \in (0,1)$ . Figure 1 shows the shows the distribution of particles during population initialization.

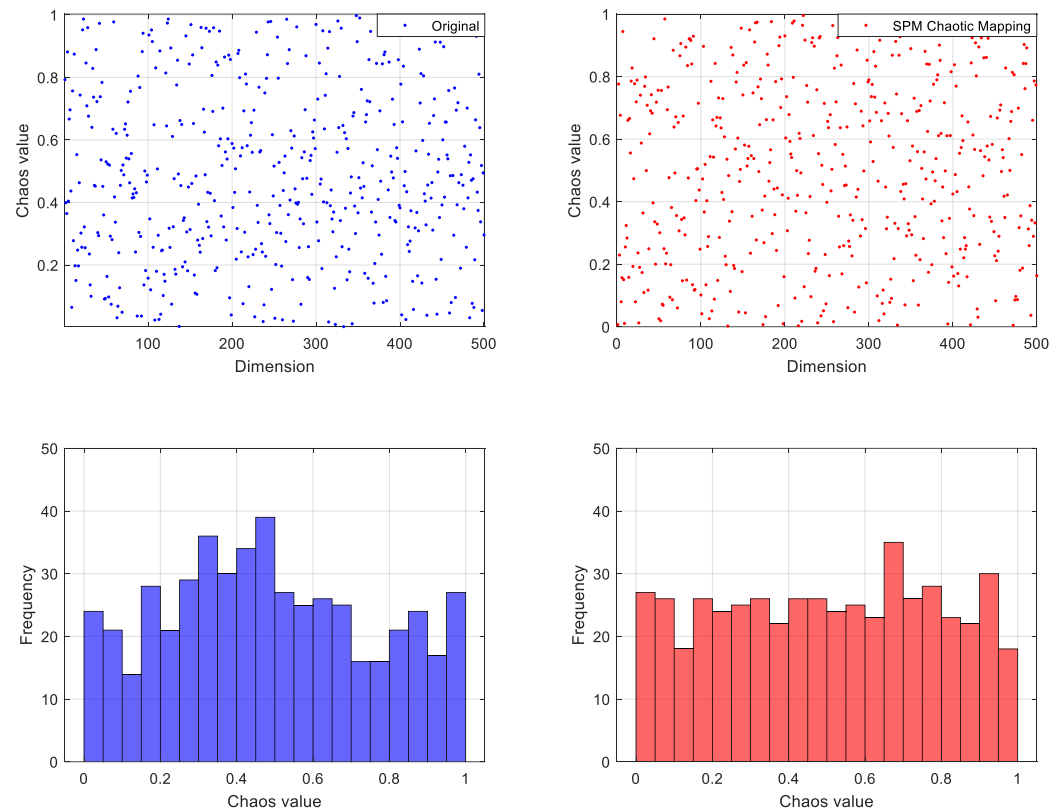


Figure 1. Distribution of particles.

As shown in Figure 1, the original algorithm employs random generation for initializing the population, and the randomly generated population exhibits non-uniformity across the entire space. Based on the distribution of particles (red bar graph), the incorporation of SPM chaotic mapping mitigates the uncertainties associated with random initialization and addresses the issue of uneven population distribution in space.

### 2.1.2. Variable Spiral Search Strategy

Inspired by the whale optimization algorithm, the variable spiral position update strategy is introduced to make the follower’s position update more flexible, and

various position update search paths are developed [24]. In the process of the follower position update, spiral parameter  $z$  is not fixed. As a result, the search method is monotonous and falls easily into the local optimal, thus weakening the searchability of the algorithm. Therefore, we designed an adaptive  $z$  variable to dynamically adjust the spiral shape of the follower search, broadening the follower’s ability to explore the unknown region, and improving the search efficiency and global search ability of the algorithm. The variable spiral search strategy formula is as follows:

$$x_{i,j}^{t+1} = \begin{cases} e^{zl} \cdot \cos(2\pi l) \cdot Q \cdot \exp\left(\frac{x_{worst}^t - x_{i,j}^t}{i^2}\right), & \text{if } i > \frac{n}{2} \\ x_p^{t+1} + |x_{i,j}^t - x_p^{t+1}| \cdot A^+ \cdot L \cdot e^{zl} \cdot \cos(2\pi l), & \text{otherwise} \end{cases} \quad (2)$$

$$z = e^{k \cos(\pi(1-(i/i_{max})))}$$

where  $k = 5$ ,  $L \in [-1, 1]$  represents uniformly distributed random numbers.  $x_p^{t+1}$  represents the optimal position of individuals in the  $t + 1$  iteration.  $x_{i,j}^t$  represents the  $j$  dimensional position of the  $i$  individual in the  $t$  iteration.  $x_{worst}^t$  represents the position of individuals with the worst fitness in the  $t$  iteration.  $Q$  represents random numbers subject to normal distribution.  $A^+ = A^T(AA^T)^{-1}$  is the matrix of  $1 \times \text{dim}$ , where  $\text{dim}$  is the dimension.

### 2.1.3. Levy Flight Strategy

Levy flight strategy can improve the randomness of algorithm solutions, enrich the diversity of population locations, and improve operation efficiency [25]. The location update formula is as follows:

$$x_i(t + 1) = x_i(t) + l \oplus \text{levy}(\lambda) \quad (3)$$

where  $x_i(t)$  represents the position of the  $i$  individual in the  $t$  iterative,  $\oplus$  represents the arithmetic symbol of point-to-point multiplication, and  $l$  represents the step control parameter.  $\text{levy}(\lambda)$  is the path obeying the levy distribution, indicating the introduction of the Levy flight strategy, and satisfies  $\text{levy} \sim \mu = t^{-\lambda}, 1 < \lambda \leq 3$ . Among them, the Levy distribution is usually simulated using the Monte Carlo algorithm [26].

Thus, the variable spiral search strategy and levy flight strategy can enhance the global optimization ability of the algorithm, which not only ensures the convergence speed of the algorithm but also increases the diversity of individuals.

### 2.1.4. Adaptive T-Distribution Perturbation Strategy

T-distribution combines the advantages of Cauchy distribution and Gaussian distribution, uses T-distribution perturbation to improve global and local search capabilities, and avoids falling into the local optimal value [27]. The improved strategy uses the freedom parameter  $iter$  to optimize the search direction and distance of individuals, and the step size can adaptively change with the increase in the number of iterations.

The perturbation strategy is defined as follows:

$$x_i^t = x_i + x_i \cdot t(iter) \quad (4)$$

where  $x_i^t$  represents the position of the mutant dung beetle,  $x_i$  represents the position of the  $i$  individual dung beetle, and  $t(iter)$  represents the T-distribution of the number of iterations  $iter$ , which is the degree of freedom.

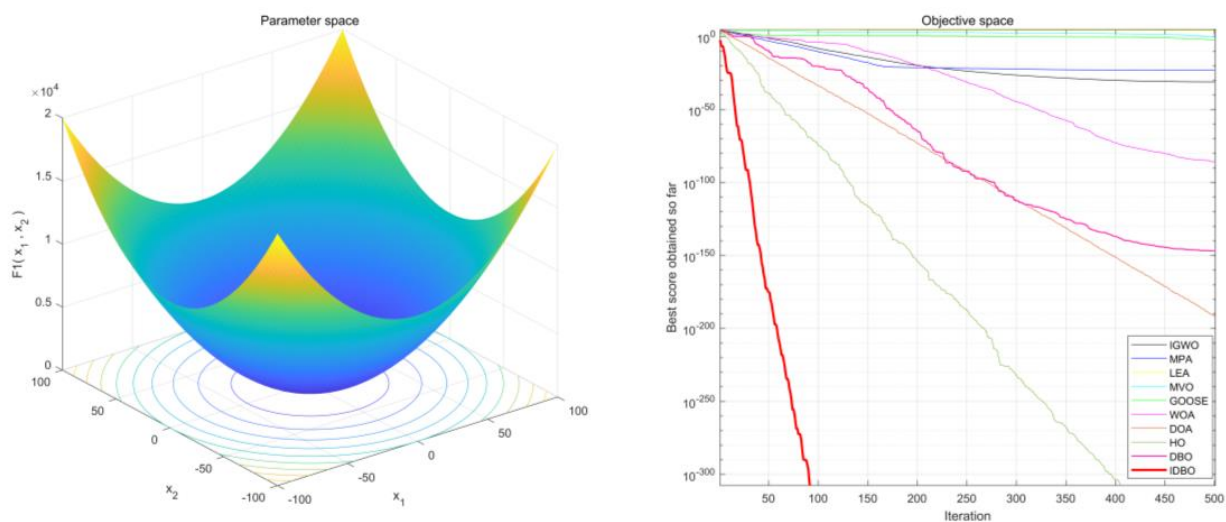
By using the adaptive T-distribution perturbation strategy, the greedy rule is added after the variation perturbation is updated, and the fitness values of the old and new positions are compared to determine whether to update the positions.

$$x_g^{new}(t) = \begin{cases} x_g^l(t), & f(x_g^l(t)) < f(x_g(t)) \\ x_g(t), & f(x_g^l(t)) > f(x_g(t)) \end{cases} \quad (5)$$

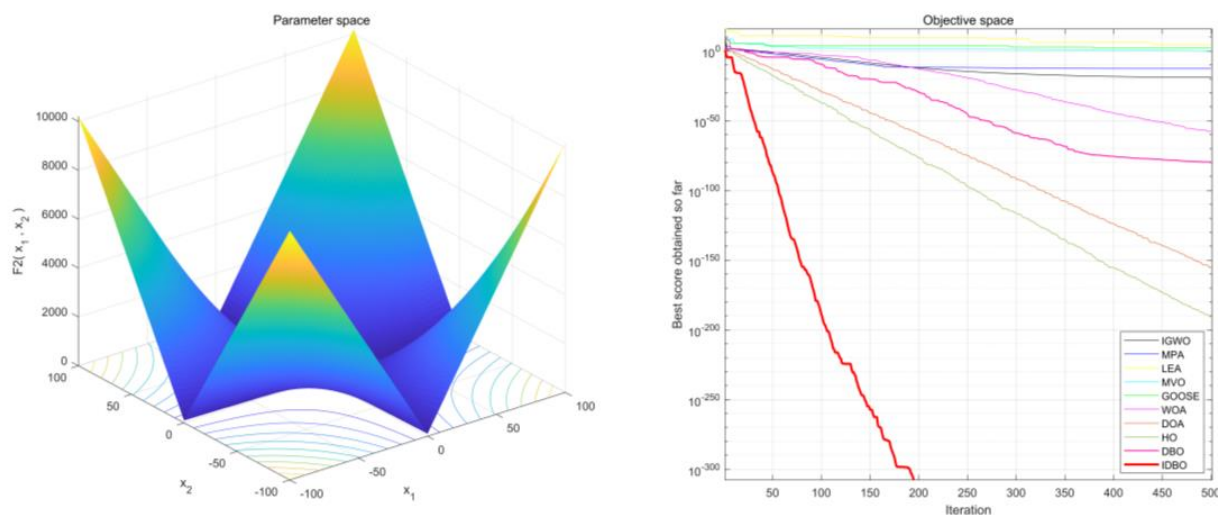
where  $x_g^{new}(t)$  represents the position of the individual after the greedy rule update,  $x_g^l(t)$  represents the individual after the disturbance,  $x_g(t)$  represents the individual before the disturbance, and  $f()$  represents the fitness function.

### 2.2. Validation and Comparison of the Optimization Algorithm

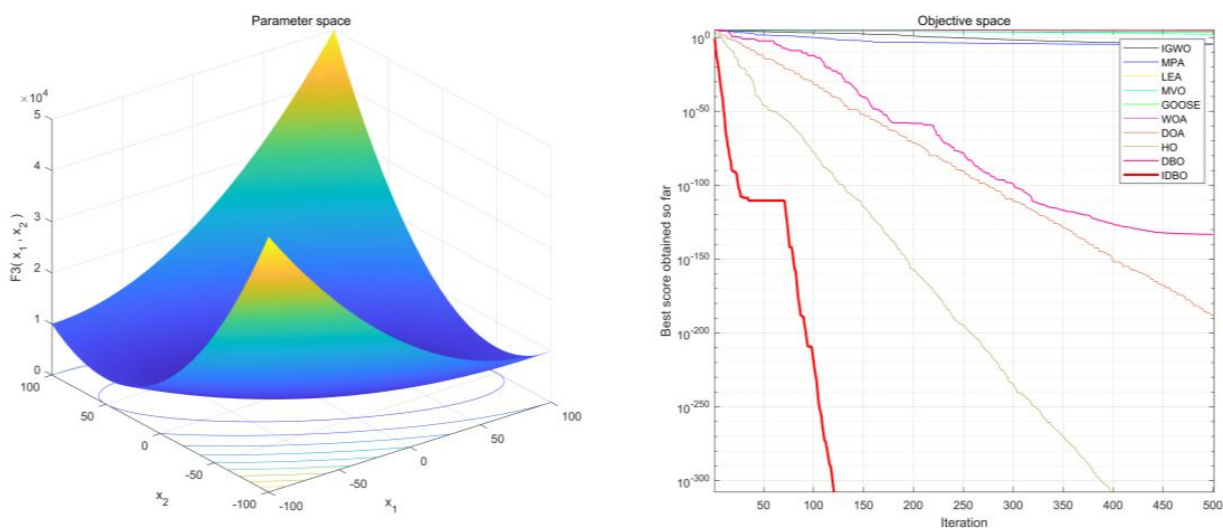
To verify the improved strategies and the convergence performance of the IDBO algorithm, the eight benchmark functions given by MATLAB were used to compare the convergence of multiple intelligent optimization algorithms, and the IDBO algorithm was comprehensively analyzed. Among them, a variety of intelligent optimization algorithms such as improved grey wolf optimizer algorithm (IGWO) [17], marine predators algorithm (MPA) [28], love evolutionary algorithm (LEA) [29], multi-verse optimizer algorithm (MVO) [30], goose optimization algorithm (GOOSE) [31], whale optimization algorithm (WOA) [32], dingo optimization algorithm (DOA) [33], hippopotamus optimization algorithm (HO) [34], and dung beetle optimizer (DBO) and its improved algorithm. Figures 2–9 show the convergence comparison effect of the eight benchmark functions. The eight benchmark functions are F1 as Sphere function, F2 as Schwefel’s Problem2.22, F3 as Schwefel’s Problem1.2, F4 as Schwefel’s Problem2.21, F9 as Generalized Rastrigin’s Function, F10 as Ackley’s Function, F11 as Generalized Griewank’s Function, and F12 as Generalized Penalized Function. Among them, the population size is 50 and maximum number of cycles is 500.



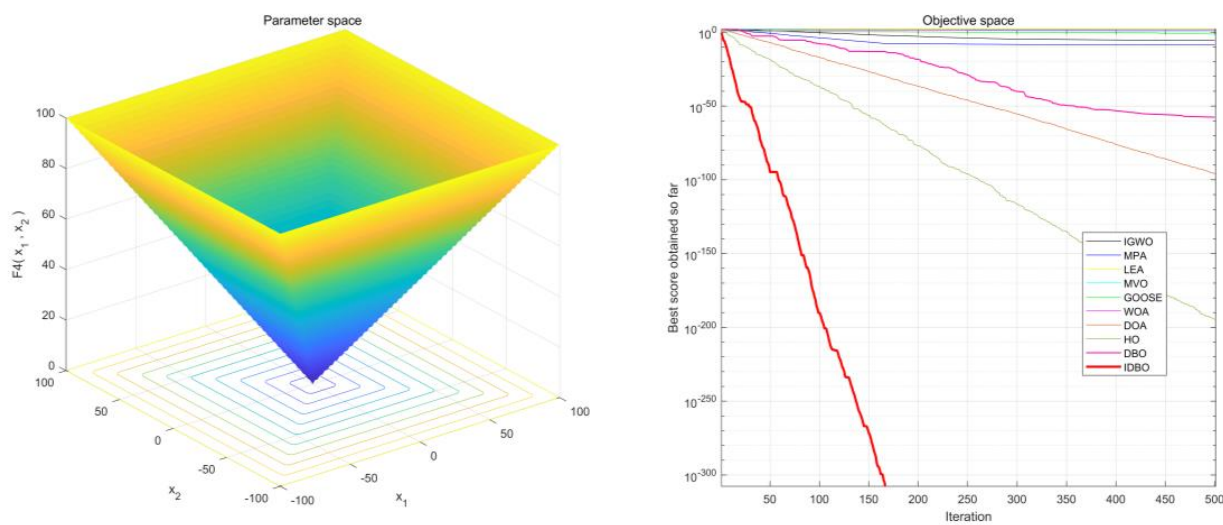
**Figure 2.** The comparison effect of benchmark function F1: (Left): parameter space; (Right): convergence curve.



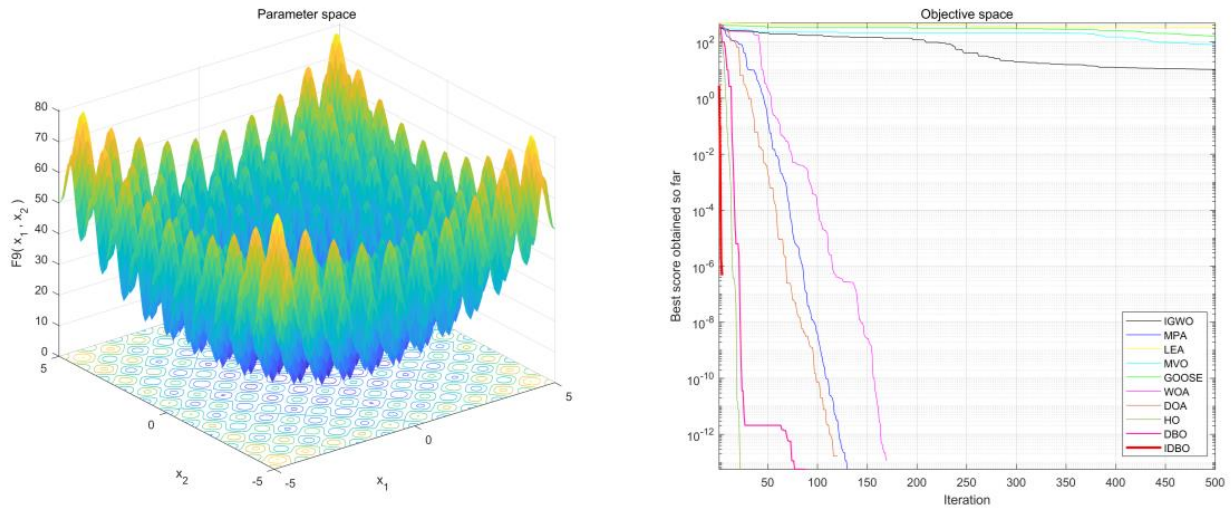
**Figure 3.** The comparison effect of benchmark function F2: (Left): parameter space; (Right): convergence curve.



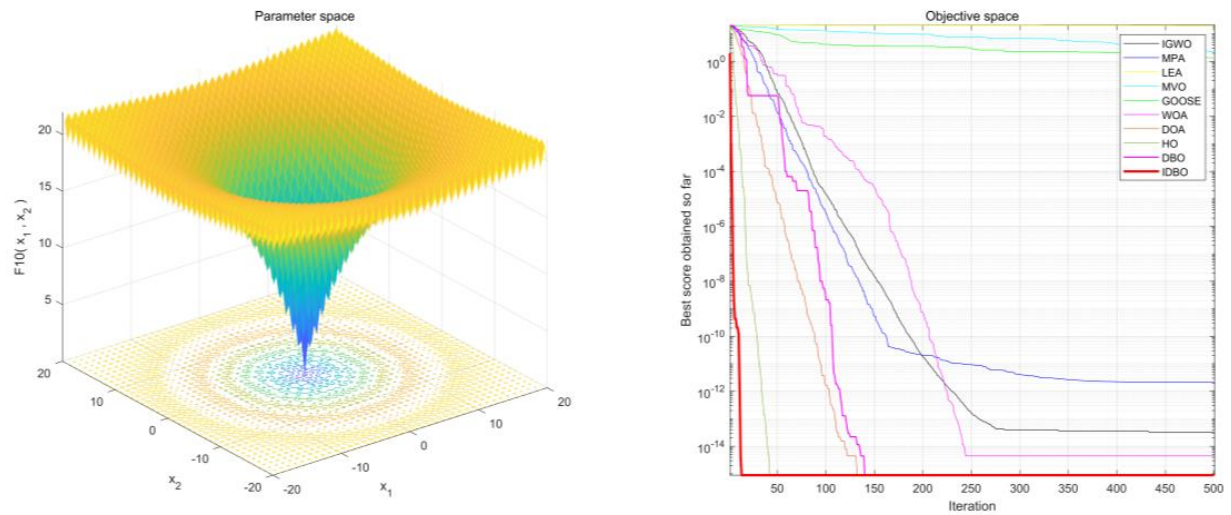
**Figure 4.** The comparison effect of benchmark function F3: (Left): parameter space; (Right): convergence curve.



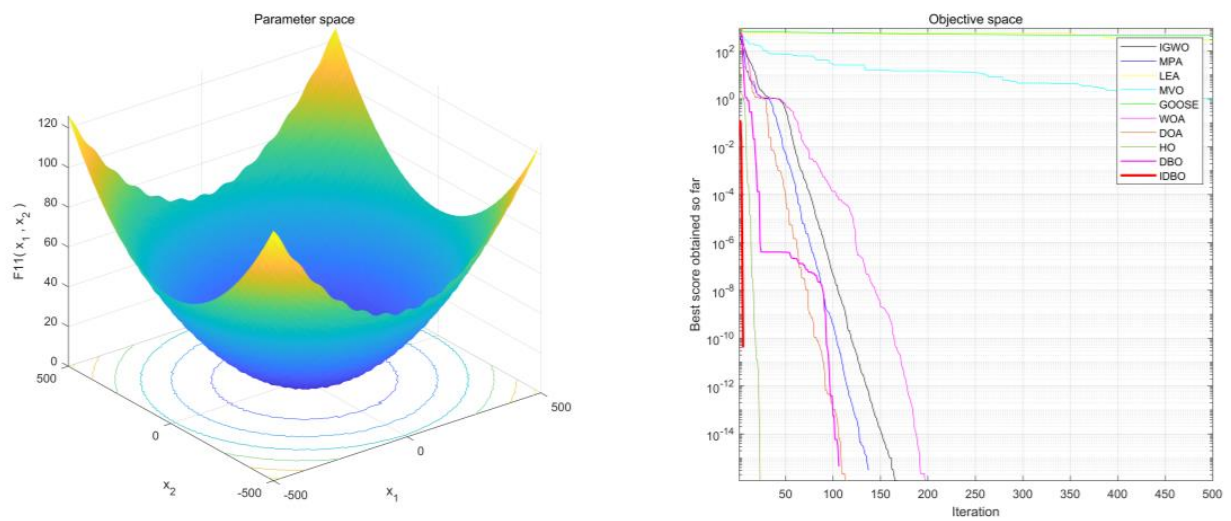
**Figure 5.** The comparison effect of benchmark function F4: (Left): parameter space; (Right): convergence curve.



**Figure 6.** The comparison effect of benchmark function F9: (Left): parameter space; (Right): convergence curve.

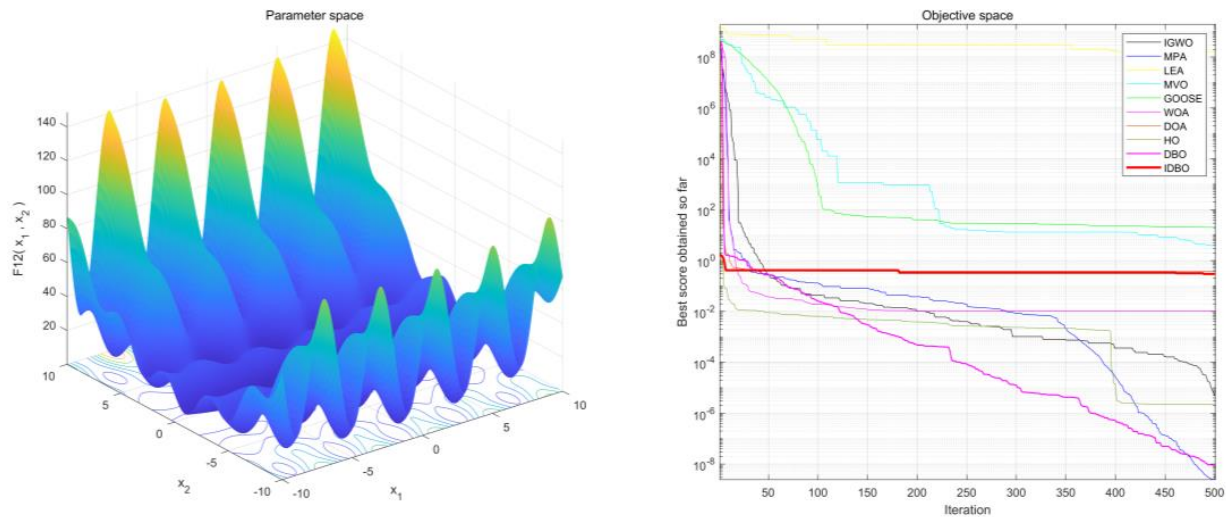


**Figure 7.** The comparison effect of benchmark function F10: (Left): parameter space; (Right): convergence curve.



**Figure 8.** The comparison effect of benchmark function F11: (Left): parameter space; (Right): convergence curve.





**Figure 9.** The comparison effect of benchmark function F12: (Left): parameter space; (Right): convergence curve.

Table 1 shows the comparison effect of evaluation indexes in different optimization algorithms. According to the analysis of Table 1, the IDBO algorithm can obtain the optimal evaluation indexes in the given eight benchmark functions, and the best value is 0, and the average value and standard deviation are the smallest compared with other intelligent optimization algorithms. Through the analysis of convergence, the four strategies are proven to improve the DBO algorithm effectively.

**Table 1.** The comparison effect of evaluation indexes in different optimization algorithms.

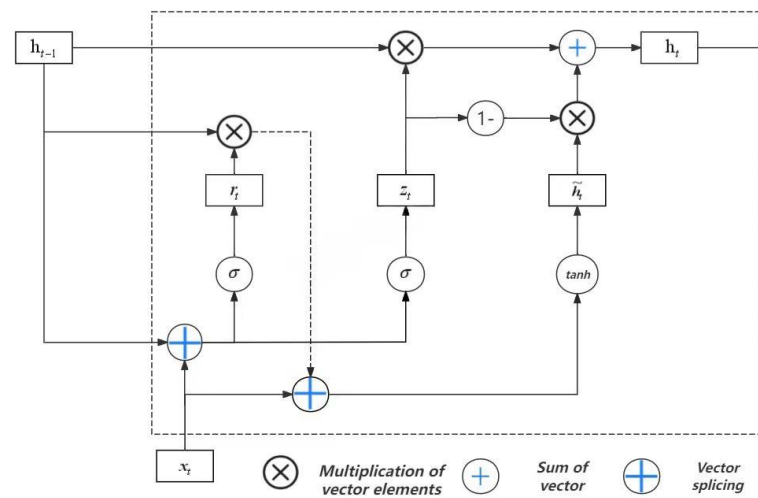
Function	Index	IGWO	MPA	LEA	MVO	GOOSE	WOA	DOA	HO	DBO	IDBO
F1	Best	$6 \times 10^{-34}$	$1.4 \times 10^{-25}$	39,633	0.65	0.005	$1.2 \times 10^{-87}$	$2 \times 10^{-138}$	0	$4 \times 10^{-125}$	0
	Mean	417	366	56,869	3421	210	600	164	130	235	0.0005
	Std	3383	3241	12,323	8369	3269	5131	2784	2896	2864	0.01
F2	Best	$1 \times 10^{-20}$	$1.3 \times 10^{-13}$	$1.2 \times 10^5$	0.64	0.47	$5.1 \times 10^{-59}$	$3.4 \times 10^{-77}$	$7 \times 10^{-198}$	$1.2 \times 10^{-74}$	0
	Mean	$3.4 \times 10^9$	2304	$6 \times 10^{16}$	$2.1 \times 10^9$	$1.8 \times 10^6$	$3.4 \times 10^9$	$8.7 \times 10^7$	$9.9 \times 10^8$	0.8	0.0008
	Std	$7.7 \times 10^{10}$	$5.1 \times 10^4$	$1 \times 10^{18}$	$3 \times 10^{10}$	$4.1 \times 10^7$	$7.6 \times 10^{10}$	$1.9 \times 10^9$	$2.2 \times 10^{10}$	7.8	0.018
F3	Best	$1.2 \times 10^{-6}$	$7.9 \times 10^{-5}$	62,467	77	4501	9448	$1 \times 10^{-316}$	0	$2 \times 10^{-118}$	0
	Mean	1906	869	90,381	9977	22,727	41,744	334	130	11,053	0.0002
	Std	7104	6380	66,885	15,694	31,560	26,320	5473	2865	21,156	0.006
F4	Best	$4.5 \times 10^{-7}$	$2.9 \times 10^{-9}$	61.51	0.57	41.71	0.24	$2.2 \times 10^{-76}$	$7 \times 10^{-201}$	$2.1 \times 10^{-67}$	0
	Mean	3.15	1.32	69.64	19.42	52.14	13.88	0.55	0.21	2.73	0.028
	Std	10.9	7.20	8.04	19.63	10.99	22.96	5.43	3.82	13.64	0.61
F9	Best	10.5	0	332	81.03	164.49	0	0	0	0	0
	Mean	89.9	8.93	378.9	198.4	295.2	20.99	9.06	1.72	2.89	0.005
	Std	85.1	42.78	37.9	55.03	54.71	71.71	50.12	22.01	28.3	0.123
F10	Best	$3 \times 10^{-14}$	$2 \times 10^{-12}$	19.82	2.23	1.35	$4.4 \times 10^{-15}$	$8 \times 10^{-16}$	$8 \times 10^{-16}$	$8 \times 10^{-16}$	$8 \times 10^{-16}$
	Mean	0.64	0.45	20.39	8.71	3.96	0.55	0.25	0.06	0.42	0.003
	Std	2.81	2.32	0.27	4.53	3.29	2.53	1.82	0.95	2.57	0.088
F11	Best	0	0	285.6	0.88	451.6	0	0	0	0	0
	Mean	4.8	3.18	513.3	29.23	514.03	5.56	2.15	1.17	4.33	0.0004
	Std	38.5	29.64	111.7	61.16	60.86	46.98	25.99	26.01	47.15	0.007
F12	Best	$4 \times 10^{-6}$	$2.4 \times 10^{-9}$	$1.5 \times 10^8$	3.82	20.38	0.01	0.37	$2.2 \times 10^{-6}$	$8.7 \times 10^{-9}$	0.29
	Mean	$1.5 \times 10^6$	$8.6 \times 10^5$	$3.5 \times 10^8$	$1.5 \times 10^7$	$1.6 \times 10^7$	$3.5 \times 10^6$	$9.3 \times 10^5$	$1.5 \times 10^6$	$2.1 \times 10^6$	0.37
	Std	$2.3 \times 10^7$	$1.1 \times 10^7$	$2.2 \times 10^8$	$7.1 \times 10^7$	$6.6 \times 10^7$	$3.4 \times 10^7$	$2 \times 10^7$	$3.3 \times 10^7$	$2.7 \times 10^7$	0.11

According to the analysis of Figures 2–9 and Table 1, under the same population size and iteration times, the solution accuracy, evaluation indexes, and convergence speed of the IDBO algorithm are superior to other intelligent optimization algorithms. By analyzing the performance of the algorithm with the given benchmark function, the IDBO algorithm

can jump out of the local optimal better and obtain a more stable global optimization ability. Thus, the IDBO algorithm is more suitable for engineering optimization problems.

### 2.3. Gated Recurrent Unit

In the processing of time-series data, gated recurrent units (GRU) are utilized for capturing comprehensive data information in order to acquire temporal dependencies and sequential patterns. The structure of GRU is shown in Figure 10. The GRU network structure streamlines the internal architecture and enhances computational efficiency by consolidating the input gate, forget gate, and output gate structures from LSTM into an update gate and reset gate. Additionally, a single update gate is employed to achieve network forgetting and selective memory, resulting in a substantial reduction in parameters.



**Figure 10.** The structure of GRU.

The update gate  $z_t$  decides how much historical and current information to use to update the current latent state at time  $t$ .

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \tag{6}$$

where  $z_t$  is the gating update for the signal, and the mean determines the degree of memory of the candidate’s implicit state.  $h_{t-1}$  is the historical implied state.  $x_t$  is the input data at time  $t$ .  $W_z$  is weight matrix.  $\sigma$  is the sigmoid function.

The reset gate  $r_t$  determines how much past information can be forgotten.

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \tag{7}$$

where  $r_t$  is reset signal, and  $W_r$  is weight matrix.

Under the action of the update gate and reset gate, the current moment candidate’s hidden state and implied output state can be updated as:

$$h_t = (1 - z_t) * h_{t-1} + z_t * \hat{h}_t \tag{8}$$

Among them, the candidate implied states  $\hat{h}_t$ :

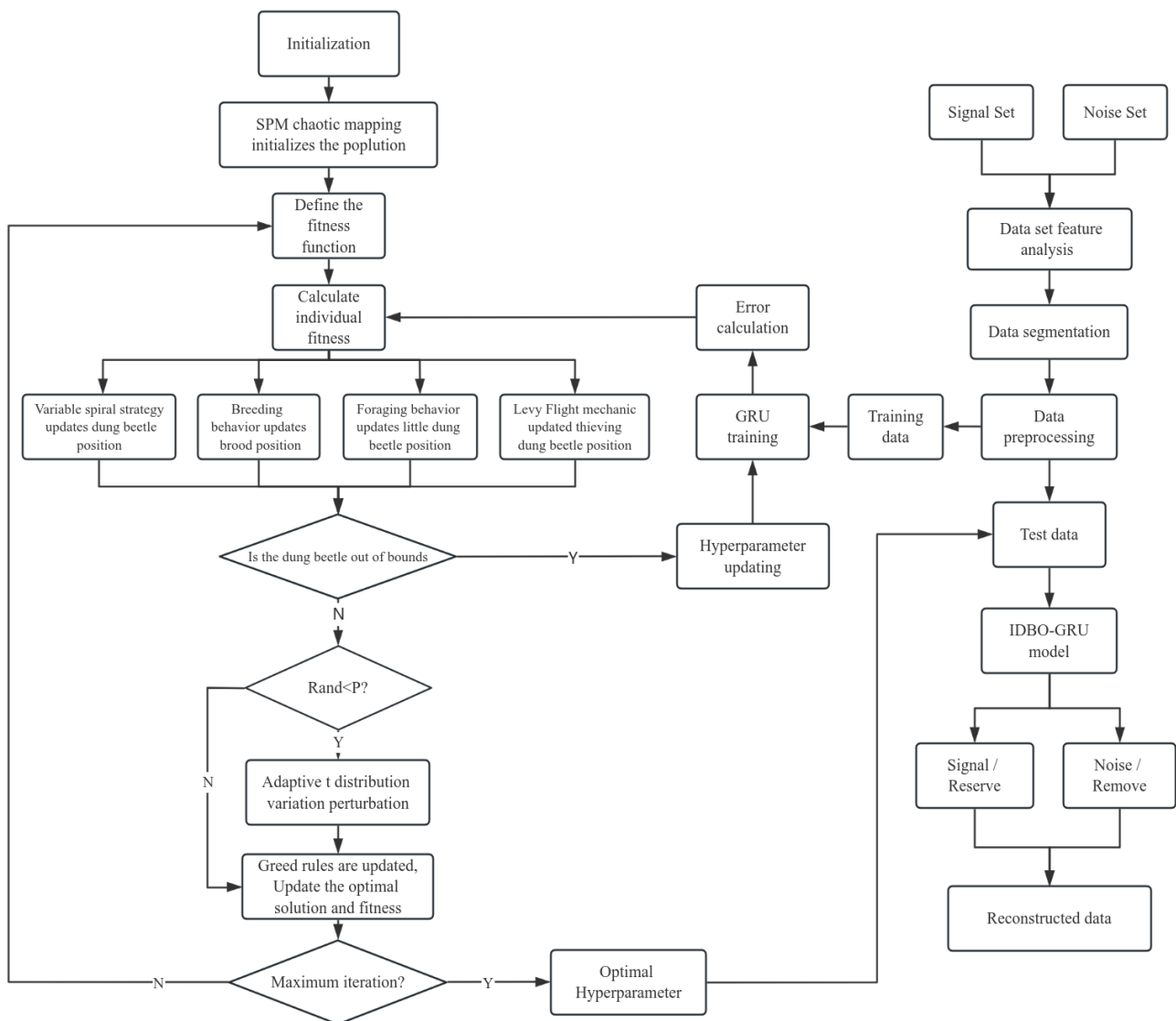
$$\hat{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \tag{9}$$

where the candidate’s implied state is integrated with merging the informative features of the input and historical data, and the operation is related to the reset signal obtained by the reset gate. However,  $h_t$  represents the final state of the current moment, which includes the process of forgetting and memory. The product of  $1 - z_t$  and the implied state  $h_{t-1}$  at the upper time represents the forgetting process. The closer it is to 1, the more information will be forgotten at the upper moment. The size of  $z_t$  determines the memory degree of the candidate’s hidden state, that is, how much of the hidden state is retained before.

### 3. Experiments and Applications

#### 3.1. The Flow of Algorithm

The algorithm flow of noise elimination for WFEM via IDBO-GRU, as shown in Figure 11; the specific steps are as follows:



**Figure 11.** The algorithm flow of noise elimination for WFEM via IDBO-GRU.

- (1) The WFEM data are divided into signal sets and noise sets, and data set features are analyzed.
- (2) Date segmentation, preprocessing, training data, and test data are divided.
- (3) Initialize the IDBO algorithm parameters.
- (4) SPM chaotic mapping is used to initialize dung beetle populations.

(5) The fitness function is defined, and the fitness value of the dung beetle positions is calculated according to the fitness function; the training data are entered into the GRU network for training, and the error is calculated.

(6) Update the position of all dung beetles, by using the variable spiral strategy to update the position of rolling dung beetles, reproduction behavior to update the position of nursery dung beetles, foraging behavior to update the position of small dung beetles, and Levy flight mechanism to update the position of thieving dung beetles.

(7) Determine whether the updated dung beetle is out of bounds. The hyperparameter of the GRU is updated.

(8) If the random number is less than the given value, perturbation is performed according to the adaptive T-distribution variation strategy to generate a new solution.

(9) Determine whether to update according to the greed rule, and update the current optimal solution and fitness.

(10) Repeat the above steps until the maximum number of iterations, and output the global optimal value and its optimal solution; these are the optimal hyperparameters.

(11) The optimal hyperparameters are verified and the IDBO-GRU model is obtained.

(12) IDBO-GRU is used for signal–noise identification, and the identified signal is reserved, and the identified noise is directly removed. The identified signal is used to reconstruct effective data.

### 3.2. The Applications of Simulation Analysis

To analyze the signal and noise types in WFEM data, we established a comprehensive library of signal–noise samples [35]; among them, 30 samples are signal and 120 samples are noise, which includes effective signal, pulse noise, attenuation noise, triangle noise, and square wave noise, respectively. Figure 12 shows various time domain data types alongside their corresponding frequency spectrum and time–frequency spectrum (Short Time Fourier Transform spectrum, STFT) graphs. Upon scrutiny of the sample database, it becomes apparent that the effective signal displays periodicity with a stable amplitude, as evidenced by its consistent time–frequency spectral characteristics. Conversely, diverse forms of noise appear as anomalies in the time-domain sequences, and waveforms with distinct morphological features, leading to an increase in effective signal amplitude and gradual disordering of the spectrum, partially obscured by noise. Consequently, the corresponding time–frequency spectrum fails to accurately depict the periodicity and frequency characteristics of the effective signal.

Thus, the data from the known sample library are analyzed, and the sample library is utilized for the optimized GRU training and test verification. Figure 13 illustrates the prediction results and confusion matrix of the sample library.

Through training on the sample library data and testing them, it can be observed that the predicted signal and noise align perfectly with the categories of actual sample database data (as shown in Figure 13 on the left). Additionally, upon examining each column of the confusion matrix, it becomes evident that 30 samples are accurately classified into Category 1 while 120 samples fall into Category 2 based on their target class labels. Similarly, these numbers match those obtained from the output class (that is, predicted class labels). The confusion matrix consistently reflects the class of both real and predicted data. However, the presence of noise severely impacts the training performance and recognition accuracy of deep learning models. By integrating an intelligently optimized GRU model, we aim to enhance the training efficiency and signal-to-noise identification accuracy. By using the proposed method, the robustness of the model against various complex noises, following data preprocessing, enables the training set, when integrated with the optimized GRU model, to achieve high-precision noise recognition and denoising.

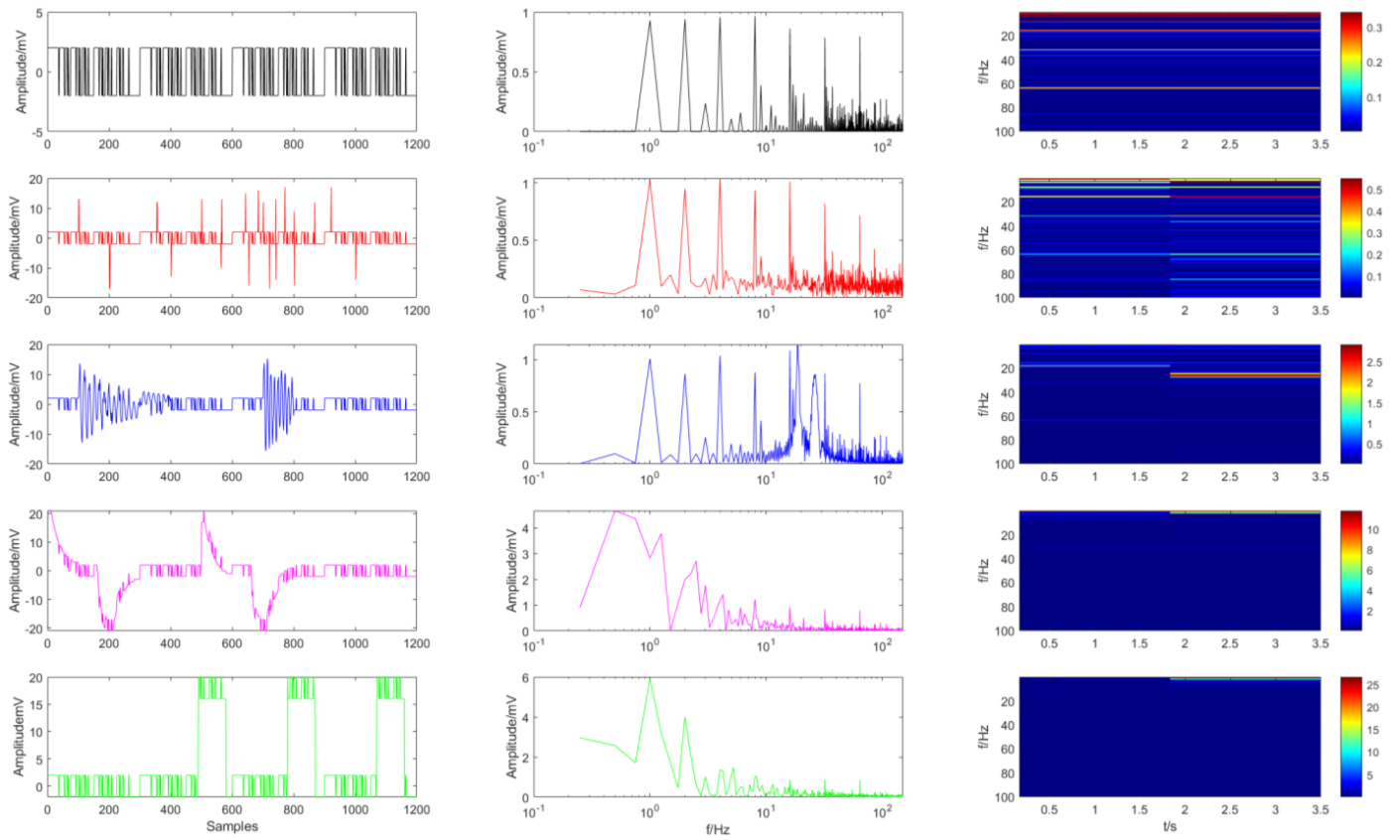


Figure 12. A set of the sample library data: (Left): time domain; (Middle): frequency spectrum; (Right): STFT spectrum.

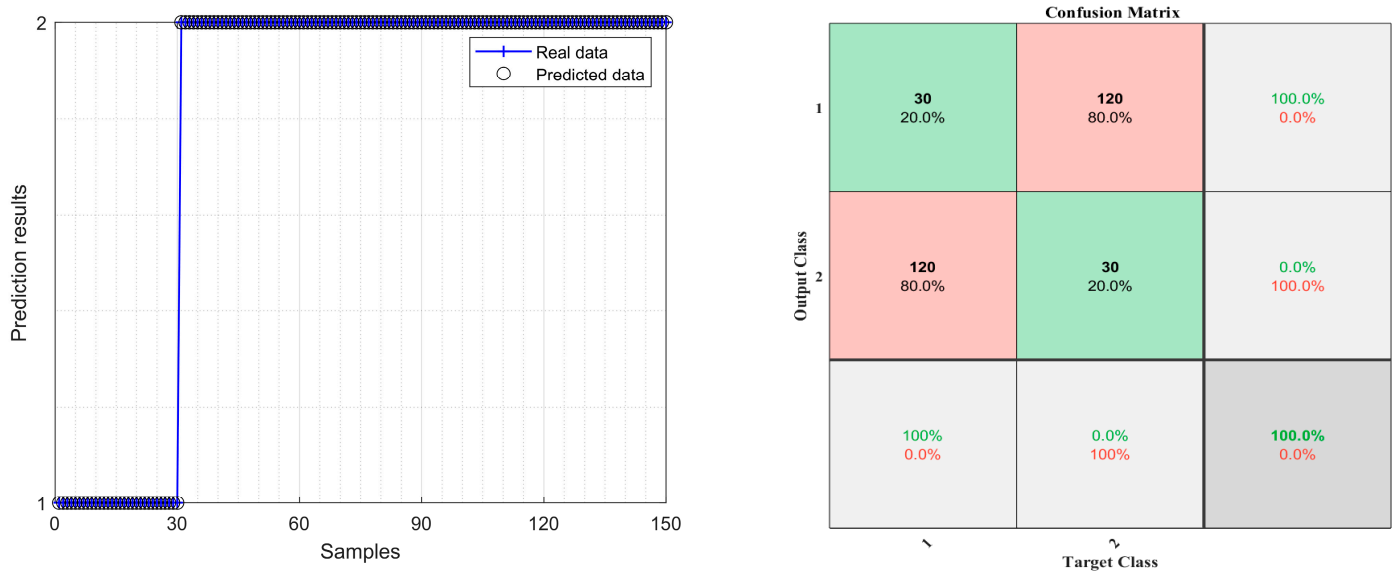
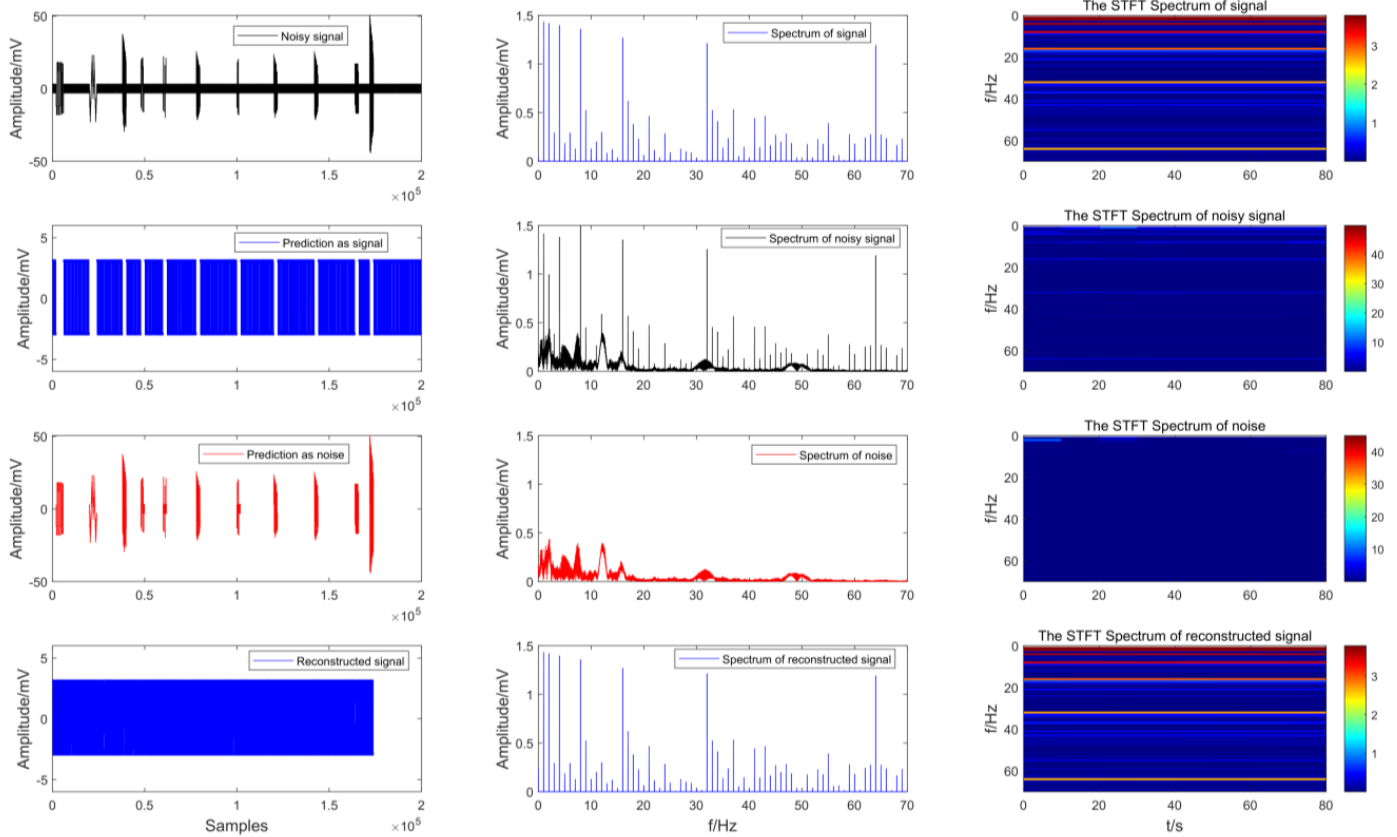


Figure 13. Prediction results (Left) and confusion matrix (Right) of the sample library.

In order to verify the processing effect of the proposed method, the simulated pseudo-random periodic signal is denoised and analyzed. Figure 14 shows the processing effect of signal noise prediction, spectrum, and short-time Fourier spectrum of the simulated data.

As can be seen from Figure 14, the analog noisy data contain a variety of noise types that lead to an increase in the amplitude of the effective signal, resulting in a large number of noise effects in the corresponding spectrum, and the amplitude of the effective signal in

different frequency segments increases or decreases accordingly. The effective signal in the STFT spectrum is overwhelmed by noise, and the information on each frequency cannot be displayed. After processing by the method proposed in this paper, the signal and noise can be effectively predicted and the noise can be eliminated, the noise waveform can be completely eliminated in the reconstructed data, and the spectrum and STFT spectrum can be effectively restored.



**Figure 14.** The signal–noise prediction effect of simulated data: (Left): time domain; (Middle): frequency spectrum; (Right): STFT spectrum.

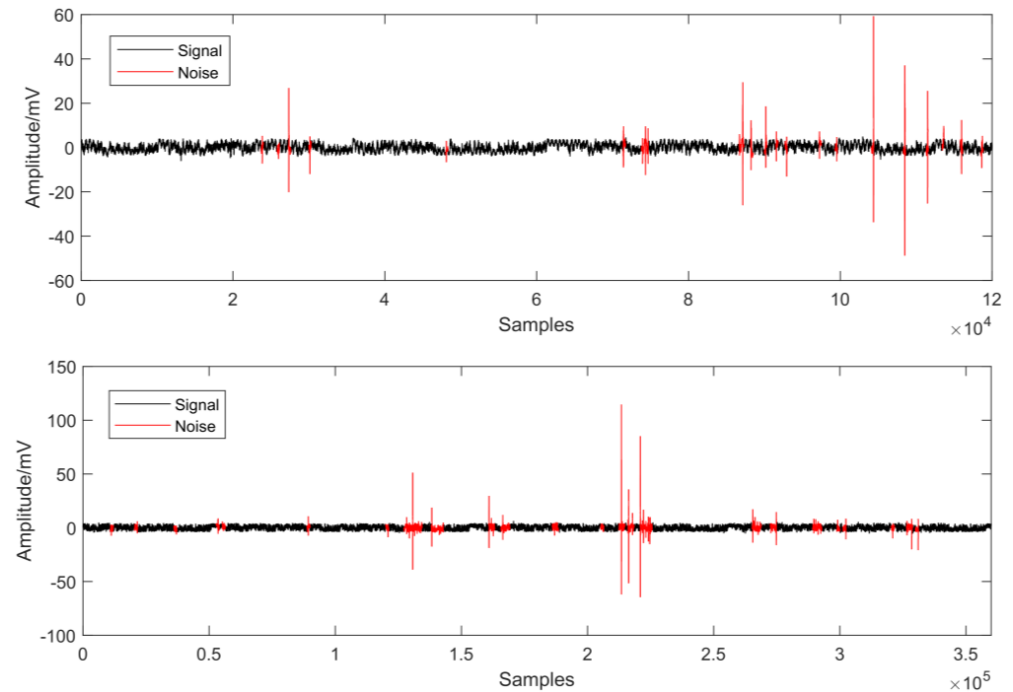
For the effectiveness of the quantitative analysis method, the amplitude of the electric field before and after denoising was calculated, as shown in Table 2. From the analysis of Table 2, it can be seen that the electric field amplitude of the original data is stable. Due to the addition of noise at different times, the electric field amplitude of the noisy data seriously decreases at 2 Hz and increases at 8 Hz, resulting in the instability of the overall frequency amplitude. After processing by this method, the electric field amplitude is closer to the original electric field value.

**Table 2.** The comparison effect of the electric field amplitude.

Frequency/Hz	Original/mV	Noisy/mV	Processed/mV
1	1.4345	1.4155	1.4346
2	1.4212	0.9970	1.4214
4	1.3972	1.3820	1.3970
8	1.3608	1.4993	1.3606
16	1.2694	1.3527	1.2695
32	1.2149	1.2572	1.2152
64	1.1933	1.1946	1.1932

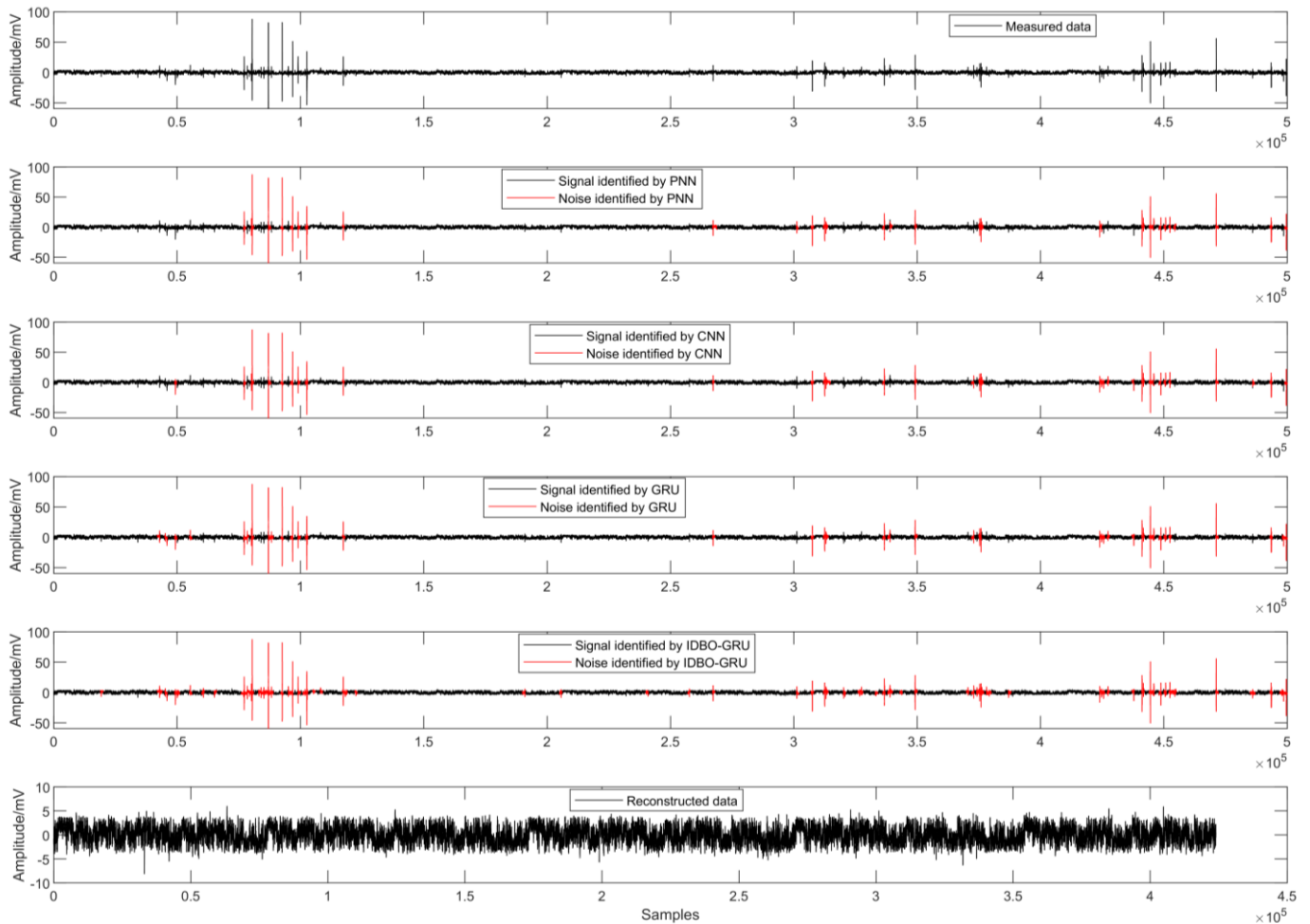
### 3.3. The Applications of Measured Data

The measured data were collected by WFEM detection, and were compared with the time-domain data for signal–noise identification and denoising processing. Figure 15 shows the signal–noise classification effect of the measured WFEM data. When analyzing the different lengths of the time-series sequence, the noise and signal can be identified effectively by using the proposed method.



**Figure 15.** The signal–noise classification effect of measured WFEM data.

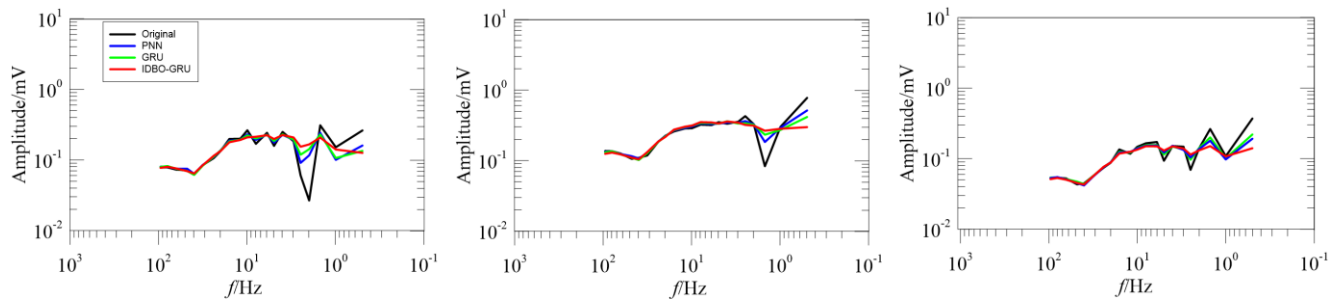
In order to further validate the algorithm’s effectiveness, we applied it to measured WFEM data. The time-domain sequence usually contains typical noise such as abnormal waveforms and abrupt interference, leading to an increase in the amplitude of the original effective signal and obscuring the pseudo-random waveform and periodic characteristics of WFEM data. Figure 16 shows the denoising comparison effect of the measured WFEM data. Analysis of Figure 16 reveals several abnormal abrupt signals in the time series of measured WFEM data, resulting in increased signal amplitudes. The most abrupt or abnormal signals can be identified by the PNN method, CNN method, and GRU method. Meanwhile, there are also some misidentifications, where effective signals are recognized as noise, or some noise is inaccurately identified. In particular, it is difficult to choose the hyperparameters of CNN and GRU, which leads to the unsatisfactory recognition effect. In cost and time, the PNN method took 3.13 s, CNN method took 24.08 s, and GRU method took 14.54 s; the proposed method took 93 s. Although the efficiency improved in the PNN method, CNN method, and GRU method, these methods cannot obtain accurate identification results. It can be seen that the PNN method, CNN method, and GRU method also need an artificial parameter setting, which leads easily to an undesirability effect in WFEM data processing. In this paper, we employed the IDBO-GRU method for WFEM signal–noise identification, significantly improving recognition accuracy. The processing results demonstrate complete identification of abrupt signals and abnormal noise in the measured data. The entire time series of noise is effectively identified and eliminated, while retaining the effective signals without any influence from noise or abnormal waveforms in the reconstructed data. These processing results align with the periodic characteristics of WFEM time-domain data.



**Figure 16.** The denoising comparison effects of measured WFEM data.

Figure 17 shows the comparison effect of the WFEM electric field curve. Among them, the electric field curve is also known as the curve of the electric field that changes with frequency. The three measured points are gathered along the same measurement line. The electric field curve is plotted by extracting the amplitude value of the required frequency from the original data spectrum. Upon analysis of the electric field curve (Figure 17), it becomes evident that typical noise types or significant large-scale noise interference appear in the original time domain, impacting data quality in the low-frequency band. This results in erratic behavior of the electric field curve at different frequency points and sharp fluctuations in amplitude within the low-frequency band (below 10 Hz). Notably, at 2 Hz, there is considerable fluctuation and overall instability in the curve. The application of a PNN method does not effectively enhance this effect as its smoothing factor requires manual adjustment, often leading to unsatisfactory outcomes in post-processing. While employing a GRU method improves overall trends significantly, there remains a slight increasing and unstable trend within the low-frequency band. Therefore, both PNN and GRU methods necessitate optimal parameter selection to achieve a more stable electric field curve effect. Through using the proposed method, no abnormal jumps are observed in the overall trend of the electric field curve, and amplitude stability is enhanced. These results indicate the effective elimination of noise from the original data and a substantial improvement in data quality, thereby providing technical support for electromagnetic method inversion interpretation.





**Figure 17.** The comparison effect of WFEM electric field curve.

#### 4. Discussions

Strong electromagnetic interference represents a significant technical challenge in the application of electromagnetic exploration, making the distinction between signal and noise a critical research focus in the field. Electromagnetic noise in practical scenarios exhibits a high degree of complexity and unpredictability, often manifesting as the superimposition of various types of noise and useful signals, which significantly complicates the process of signal–noise separation. Consequently, this paper introduces a wide field electromagnetic denoising method based on an IDBO-GRU algorithm. By employing multiple strategies to enhance the DBO algorithm and optimizing the hyperparameters of the GRU model, the proposed approach aims to achieve accurate signal–noise identification and separation for WFEM data through the application of intelligently optimized deep learning techniques.

With the rapid advancement of machine learning and artificial intelligence methodologies, deep learning models have increasingly been utilized in signal processing and pattern recognition. In this paper, we have chosen the Improved Dung Beetle Optimization (IDBO) algorithm for multi-strategy enhancement (Figures 1–9) and compared its convergence with those of several other intelligent optimization algorithms. The results clearly demonstrate that the IDBO exhibits superior convergence efficiency and precision, as well as enhanced global search capabilities. Additionally, the convergence performance metrics of IDBO are quantitatively evaluated in Table 1, contributing to further studies on parameter optimization. Subsequently, we detail the principles, optimization parameters, and procedures of the Gated Recurrent Unit (GRU) model, applying it to the signal–noise identification and separation processes of WFEM data (Figures 10 and 11). Within this study, a data sample library was constructed, dividing the data set into training and testing sets, with the outcomes analyzed based on these samples (Figures 12 and 13). Experimental findings, derived from both simulated and real-world data, indicate that noisy data can be effectively filtered out, allowing for the direct removal of identified noise while preserving the recognized signal components to reconstruct useful data. Meanwhile, compared with the PNN method, CNN method, and GRU method, the proposed method can obtain a more stable reconstructed signal. Consequently, the reconstructed data exhibit a reduced length compared to the original, in line with the periodic and pseudo-random characteristics of the effective WFEM signal (Figures 14–16). In addition, we only use the electric field component for data denoising and calculating the electric field curve to assess the effectiveness of the denoising process, and calculations that do not take into account the phase. The efficacy of the proposed method is substantiated through electric field curve analysis (Figure 17).

Furthermore, the presence of noise severely impacts the training performance and recognition accuracy of deep learning models. By integrating an intelligently optimized GRU model, we aim to enhance the training efficiency and signal-to-noise identification accuracy. In the future, there is a need to further enhance the efficiency, cost-effectiveness, and robustness of intelligently optimized deep learning models.

## 5. Conclusions

To enhance the optimization and convergence performance of the DBO algorithm, as well as to broaden the application scope of the GRU model, in this paper, an improved dung beetle optimized gated recurrent unit algorithm is developed, and applied to WFEM noise elimination. On the one hand, four kinds of improvement strategies were used for improving the DBO algorithm, thus enhancing the optimization ability. On the other hand, GRU is a simple deep learning algorithm whose hyperparameters determine the classification and prediction effect of the network. Meanwhile, the IDBO-GRU algorithm can realize WFEM signal–noise identification and denoising. The experiments and applications show that the solution accuracy, evaluation indexes, and convergence speed of the IDBO algorithm are superior to other intelligent optimization algorithms. The proposed method can accurately identify and eliminate noise in the simulation and measured data. Its frequency domain and time–frequency domain effect also indicate that the reconstructed data are line in with the periodic signal feature by using the proposed method. The processed electric field curve exhibits greater stability and has no abnormal fluctuation compared to alternative methods. Thus, the satisfactory performance and application in the results verify the effectiveness of the design and the optimization method. In the future, we will develop a novel intelligent optimization algorithm characterized by enhanced optimization capabilities and improved convergence accuracy, enabling adaptive selection of optimal parameters. Additionally, an intelligent optimization network model will be established to decrease the algorithm’s complexity, while data-driven signal processing in intelligent deep learning will be implemented. And we will conduct a more explicit analysis to achieve WFEM noise classification and predictive denoising.

**Author Contributions:** Z.L.: Writing—original draft, Visualization, Software. X.Z.: Writing—review, Conceptualization, Project administration, Methodology. D.L.: Resources, Project administration. S.L.: Writing—original draft, Investigation, Formal analysis. K.C.: Resources, Supervision. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the National Natural Science Foundation of China (No. 42474170), the Key R&D Program of Hunan Province (No. 2024AQ2002), Hunan Province talent lifting project (2024TJ-X47), the Science and Technology Innovation Program of Hunan Province (No. 2023RC1014), the Open Research Fund Program of Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment Monitoring (Central South University), Ministry of Education (No. 2024YSJS12), and the Research Foundation of Education Bureau of Hunan Province, China (No. 24B0927).

**Data Availability Statement:** The data cannot be made publicly available upon publication because they are not available in a format that is sufficiently accessible or reusable by other researchers. The data that support the findings of this study are available upon reasonable request from the authors.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. He, J.S. On the closed addition in a three element set and the  $2n$  sequence pseudo-random signal coding. *J. Cent. South Univ. Sci. Technol.* **2010**, *41*, 632–637.
2. He, J.S. Wide field electromagnetic sounding methods. *J. Cent. South Univ. Sci. Technol.* **2010**, *41*, 1065–1072.
3. He, J.S.; Tong, T.G.; Liu, J.X. Mathematical analysis and realization of an sequence pseudo-random multi-frequencies signal. *J. Cent. South Univ. Sci. Technol.* **2009**, *40*, 1666–1671.
4. Li, D.Q.; Xie, W.; Chen, D.X. Three-dimensional modeling for E-Ex wide field electromagnetic methods. *Trans. Nonferrous Met. Soc. China* **2013**, *23*, 2459–2470.
5. Yang, Y.; Li, D.Q.; Tong, T.G.; Zhang, D.; Zhou, Y.T.; Chen, Y.K. Denoising controlled-source electromagnetic data using least-squares inversion. *Geophysics* **2018**, *83*, E229–E244. [[CrossRef](#)]

6. Li, G.; He, Z.S.; Tang, J.T.; Deng, J.Z.; Liu, X.Q.; Zhu, H.J. Dictionary learning and shift-invariant sparse coding denoising for controlled-source electromagnetic data combined with complementary ensemble empirical mode decomposition. *Geophysics* **2021**, *86*, E185–E198. [[CrossRef](#)]
7. Li, G.; Wu, S.L.; Cai, H.Z.; He, Z.S.; Liu, X.Q.; Zhou, C.; Tang, J.T. IncepTCN: A new deep temporal convolutional network combined with dictionary learning for strong cultural noise elimination of controlled-source electromagnetic data. *Geophysics* **2023**, *88*, E107–E122. [[CrossRef](#)]
8. Li, J.; Liu, Y.C.; Tang, J.T.; Peng, Y.Q.; Zhang, X.; Li, Y. Magnetotelluric data denoising method combining two deep-learning-based models. *Geophysics* **2023**, *88*, E13–E28. [[CrossRef](#)]
9. Li, J.; Luo, Y.C.; Li, G.; Liu, Y.C.; Tang, J.T. Atom-profile updating dictionary learning with nucleus sampling attention mechanism sparse coding for audio magnetotelluric denoising. *Geophysics* **2024**, *89*, E73–E85. [[CrossRef](#)]
10. Zhang, B.M.; Jiang, Q.Y.; Mo, D.; Xiao, L.Y. A new method for handling gross errors in electromagnetic prospecting data. *Chin. J. Geophys.* **2015**, *58*, 2087–2102.
11. Mo, D.; Jiang, Q.Y.; Li, D.Q.; Chen, C.J.; Zhang, B.M.; Liu, J.W. Controlled-source electromagnetic data processing based on gray system theory and robust estimation. *Appl. Geophys.* **2017**, *14*, 570–580. [[CrossRef](#)]
12. Chen, C.J.; Jiang, Q.Y.; Mo, D.; Li, G.; Zhou, F. De-noising pseudo-random electromagnetic data using gray judgment criterion and rational function filtering. *Chin. J. Geophys.* **2019**, *62*, 3854–3865.
13. Hu, Y.F.; Liu, Z.J.; Li, D.Q.; Zhang, X.; Suo, G.Y. Noise separation of CSEM data based on improved clustering method. *Chin. J. Geophys.* **2024**, *67*, 394–408.
14. Yang, Y.; He, J.S.; Li, D.Q. A noise evaluation method for CSEM in the frequency domain based on wavelet transform and analytic envelope. *Chin. J. Geophys.* **2018**, *61*, 344–357.
15. Yang, Z.; Tang, J.T.; Xiao, X.; Jiang, Q.Y.; Huang, X.Y.; Hu, S.G. Application of powerline noise cancellation method in correlation identification of controlled source electromagnetic method. *J. Geophys. Eng.* **2021**, *18*, 339–354. [[CrossRef](#)]
16. Ling, F.; Yang, Y.; Li, G.; Zhou, C.Y.; Huang, M.; Wang, X.; Zhang, H.; Zhu, Y.Z.; Sun, H.F. Extracting useful high-frequency information from wide-field electromagnetic data using time-domain signal reconstruction. *J. Cent. South Univ.* **2022**, *29*, 2150–2163. [[CrossRef](#)]
17. Zhang, X.; Li, D.Q.; Li, J.; Liu, B.; Jiang, Q.Y.; Wang, J.H. Signal-noise identification for wide field electromagnetic method data using multi-domain features and IGWO-SVM. *Fractal. Fract.* **2022**, *6*, 80. [[CrossRef](#)]
18. Yang, Y.; Zhang, H.; Zhu, Y.Z.; Zhou, C.Y.; Sun, H.F. Denoising land-based controlled-source electromagnetic data based on a same-site noise reference channel. *Geophys. J. Int.* **2023**, *235*, 2285–2304. [[CrossRef](#)]
19. Yang, Y.; Zhou, C.Y.; Zhang, H.; Peng, Y.H.; Sun, H.F. De-noising CSEM data using least-squares method based on mixed basis of Fourier series and Legendre polynomials. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5921812. [[CrossRef](#)]
20. Zhang, X.; Li, D.Q.; Hu, Y.F.; Zhu, Y.Q.; Li, F. WFEM signal-noise separation method based on improved inherent time-scale decomposition and probabilistic neural network. *Prog. Geophys.* **2024**, *39*, 0241–0252. (In Chinese)
21. Xue, J.K.; Shen, B. Dung beetle optimizer: A new metaheuristic algorithm for global optimization. *J. Supercomput.* **2023**, *79*, 7305–7336. [[CrossRef](#)]
22. Khayam, K.N.; Mehmood, Z.; Chaudhry, H.N.; Ashraf, M.U.; Tariq, U.; Altouri, M.N.; Alsubhi, K. Local-Tetra-Patterns for Face Recognition Encoded on Spatial Pyramid Matching. *Comput. Mater. Contin.* **2022**, *70*, 5039–5058.
23. Wang, Y.; Wang, N.; Gao, T.; Bu, F.Y.; Zhu, X.Q. An improved prairie dog optimization algorithm integrating multiple strategies and its application. *Eng. Res. Express* **2024**, *6*, 035224. [[CrossRef](#)]
24. Anitha, J.; Pandian, S.I.A.; Agens, S.A. An efficient multilevel color image thresholding based on modified whale optimization algorithm. *Expert Syst. Appl.* **2021**, *178*, 115003. [[CrossRef](#)]
25. Rhee, I.; Shin, M.; Hong, S.; Lee, K.; Kim, S.J.; Chong, S. On the Levy-walk Nature of Human Mobility. *IEEE/ACM Trans. Netw.* **2011**, *19*, 630–643. [[CrossRef](#)]
26. Brockmann, D.; Hufnagel, L.; Geisel, T. The scaling laws of human travel. *Nature* **2006**, *439*, 462–465. [[CrossRef](#)]
27. Zhu, X.M.; Liu, S.; Zhu, X.L.; You, X.M. Enhancing sparrow search algorithm with hybrid multi-strategy and its engineering applications. *J. Intell. Fuzzy Syst.* **2023**, *45*, 5601–5632. [[CrossRef](#)]
28. Faramarzi, A.; Heidarinejad, M.; Mirjalili, S.; Gandomi, A.H. Marine Predators Algorithm: A nature-inspired metaheuristic. *Expert Syst. Appl.* **2020**, *152*, 113377. [[CrossRef](#)]
29. Gao, Y.S.; Zhang, J.H.; Wang, Y.L.; Wang, J.P.; Qin, L. Love Evolution Algorithm: A stimulus-value-role theory-inspired evolutionary algorithm for global optimization. *J. Supercomput.* **2024**, *80*, 12346–12407. [[CrossRef](#)]
30. Mirjalili, S.; Mirjalili, S.M.; Hatamlou, A. Multi-verse optimizer: A nature-inspired algorithm for global optimization. *Neural Comput. Appl.* **2016**, *27*, 495–513. [[CrossRef](#)]
31. Hamad, R.K.; Rashid, T.A. GOOSE algorithm: A powerful optimization tool for real-world engineering challenges and beyond. *Evol. Syst.* **2024**, *15*, 1249–1274. [[CrossRef](#)]
32. Mirjalili, S.; Lewis, A. The whale optimization algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [[CrossRef](#)]

33. Hernan, P.V.; Adrian, F.P.D.; Gustavo, E.C.; Ana, B.M.C.; Jonas, V.A.; Fernando, R.P. A Bio-Inspired Method for Engineering Design Optimization Inspired by Dingoes Hunting Strategies. *Math. Probl. Eng.* **2021**, *2021*, 9107547.
34. Amiri, M.H.; Mehrabi Hashjin, N.; Montazeri, M.; Mirjalili, S.; Khodadadi, N. Hippopotamus optimization algorithm: A novel nature-inspired optimization algorithm. *Sci. Rep.* **2024**, *14*, 5032. [[CrossRef](#)]
35. Zhang, X.; Li, D.Q.; Liu, B.; Hu, Y.F.; Mo, Y. Intelligent processing of electromagnetic data using detrended and identification. *Mach. Learn. Sci. Technol.* **2023**, *4*, 045041. [[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.