energies

MDPI

Article

Reliability Prediction of Distribution Network Using IAO-Optimized Mixed-Kernel LSSVM

Chen Pan [1], Lijia Ren [1,*] and Junjie Wan [2]

[1] School of Electronic & Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China; pc15555396786@163.com
[2] State Grid Shanghai Electric Power Research Institute, Shanghai 200437, China; 19821653592@163.com
* Correspondence: rlijia@126.com

Abstract: For the sake of conducting distribution network reliability prediction in an accurate and efficient manner, a model for distribution network reliability prediction (IAO-LSSVM) based on an improved Aquila Optimizer (IAO) optimized mixed-kernel Least Squares Support Vector Machine (LSSVM) is thus proposed in this paper. First, the influencing factors that greatly affect the distribution network reliability are screened out through grey relational analysis. Afterwards, the radial basis kernel function and polynomial kernel function are combined and a mixed kernel LSSVM model is constructed, which has better generalization ability. However, for the AO algorithm, it is easy to fall into local extremum. In such case, the AO algorithm is innovatively improved after both the improved tent chaotic initialization strategy and adaptive t-distribution strategy are introduced. Next, the parameters of the mixed-kernel LSSVM model are optimized and the IAO-LSSVM distribution network reliability prediction model is established through using the improved AO algorithm. In the end, the prediction results and errors of the IAO-LSSVM prediction model and other models are compared in the actual distribution network applications. It is revealed that the IAO-LSSVM prediction model proposed in this paper features higher accuracy and better stability.

Keywords: reliability prediction of distribution network; mixed-kernel LSSVM; improved Aquila Optimizer; grey relational analysis

check for updates

Citation: Pan, C.; Ren, L.; Wan, J. Reliability Prediction of Distribution Network Using IAO-Optimized Mixed-Kernel LSSVM. *Energies* **2023**, *16*, 7448. https://doi.org/10.3390/en16217448

Academic Editors: Yuan Gao, Pat Wheeler, Tomislav Dragicevic and Javier Contreras

Received: 15 September 2023
Revised: 18 October 2023
Accepted: 1 November 2023
Published: 4 November 2023

Copyright: © 2023 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

As an essential part of the power system, the power distribution network features a wide range and complex structure, which have an important impact on the reliability of the distribution network. In consequence, it is of great importance to predict the power supply reliability of the distribution network accurately and effectively [1].

There are three most commonly used reliability prediction methods for distribution networks: the simulation method, analytical method, and artificial intelligence algorithm. For the simulation method, the Monte Carlo method is the most commonly used. Although it has high calculation accuracy, the Monte Carlo method has the shortcoming of long calculation time, which means it is difficult to meet the needs of online analysis [2]. The most commonly used analytical methods include the network equivalent method [3], failure mode analysis method [4], and minimum cut set method [5]. However, these three methods are complex in calculation while poor in generality. Neural networks are non-parametric data-driven methods. In reference [6], the Mind Evolutionary Algorithm (MEA) was proposed to optimize and improve the reliability prediction model of distribution networks based on Elman neural networks. Experimental results demonstrated the improved accuracy of this method. However, there may be issues such as local minima and overlearning in practical applications. In reference [7], the Improved Particle Swarm Optimization (IPSO) algorithm was suggested to optimize the reliability prediction model of distribution networks, based on the Generalized Regression Neural Network (GRNN). IPSO was utilized to

optimize the parameters of GRNN. However, the accuracy of the prediction model cannot be guaranteed when training data samples are limited. Reference [8] proposed the Support Vector Machines (SVM) as a machine learning algorithm based on small sample statistical learning theory. It overcame the drawbacks of neural networks by minimizing structural risk and achieved high prediction accuracy even when the data samples are limited. However, the standard SVM requires solving quadratic programming equations, which results in high computational complexity and slow speed. The Least Squares Support Vector Machine (LSSVM) algorithm is a machine learning algorithm specifically designed for small sample situations. Reference [9] proposed combining the structural risk minimization criterion with least squares estimation, which transformed the quadratic programming solution in standard SVM into linear equation solving, simplifying the computation process, increasing the operation speed, and solving the overfitting and underfitting problems of neural networks. This provides a new research method for reliable prediction of small sample power distribution network. Traditional LSSVM models often use a single kernel function, yet each kernel function has its own limitations that can affect prediction accuracy to some extent. Mixed-kernel LSSVM models combine different kernel functions through weighted combinations, which can balance the strengths of multiple kernel functions and achieve better prediction accuracy and generalization performance than single-kernel LSSVM models. However, compared to single-kernel LSSVM, mixed-kernel LSSVM models have more parameters and are more difficult to determine. References [9–11] have used particle swarm optimization (PSO) and the genetic algorithm (GA) to search for optimal parameters for mixed-kernel LSSVM models. Adaptive optimization (AO), as a novel swarm intelligence algorithm, has been proven to have strong global optimization ability with simple structure and better optimization efficiency than other algorithms such as GA and PSO [12].

In summary, this article proposes a global optimization method for hybrid kernel LSSVM parameters using an improved eagle-inspired optimizer (IAO). Based on this, a reliability prediction model for distribution networks using mixed kernel IAO-LSSVM is established and compared with other optimization models in terms of reliability, prediction results, and errors. The results show that the mixed kernel IAO-LSSVM model has smaller prediction errors and higher accuracy, providing certain reference significance for the future planning of distribution networks.

2. Aquila Optimizer and Its Improvements

2.1. Aquila Optimizer (AO)

As a latest novel swarm intelligence optimization algorithm, the Aquila Optimizer (AO) was inspired by the behavior of Aquila in nature. AO has two phases, namely exploration and exploitation. If $k \leq (\frac{2}{3}) * K$, the exploration steps will be executed; otherwise, the exploitation steps will be executed. The exploration steps are divided into expanded exploration and narrowed exploration while the exploitation steps are divided into expanded exploitation and narrowed exploitation. First, AO starts by determining the initial value for a set of N individuals X through the following formula:

$$X_{ij} = rand \times \left(UB_j - LB_j\right) + LB_j \tag{1}$$

In the above Equation (1), *rand* is a random number between [0,1], $i = 1, 2, \cdots N$, $j = 1, 2, \cdots d$, N is the total number of candidate solutions (population), d is the dimension of the test problem, LB_j and UB_j denote the lower bound and upper bound at dimension j, respectively.

2.1.1. Expanded Exploration

In this phase, the Aquila flies high over the ground and explores the search space widely, and then a vertical dive will be taken once the area of the prey is determined. The mathematical representation of this behavior is written as follows:

$$X_1(k+1) = X_{best}(k) \times (\frac{1-k}{K}) + (X_M(k) - X_{best}(k) * rand) \quad (2)$$

In the above Equation (2), *rand* is a random value between 0 and 1, K is the number of total iterations, $X_{best}(k)$ is the best-obtained solution until kth iteration, $X_M(k)$ denotes the locations mean value of the current solutions connected at kth iteration, and it is computed as follows:

$$X_M(k) = \frac{1}{N}\sum_{i=1}^{N} X(k),\ j= 1, 2, \cdots d \quad (3)$$

wherein, d is the dimension size of the optimization variable.

2.1.2. Narrowed Exploration

The Aquila circles above the target prey when the prey area is found from a high soar, and then attacks. This behavior is mathematically presented as in Equation (4) below.

$$X_2(k+1) = X_{best}(k) \times Levy(d) + X_R(k) + (y - x) * rand \quad (4)$$

wherein, X_R denotes a random chosen individual. $Levy(d)$ is the levy flight distribution function, and it is calculated by Equation (5). d is the dimension size. y and x represent the spiral shape in the search. The calculations are as follows in Formulas (6) to (9):

$$\text{Levy}(d) = s \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}}, \sigma = \left(\frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma(\frac{1+\beta}{2}) \times \beta \times 2^{(\frac{\beta-1}{2})}} \right) \quad (5)$$

$$y = r \times cos(\theta),\ x = r \times sin(\theta) \quad (6)$$

$$r = r_1 + U \times d_1 \quad (7)$$

$$\theta = -\omega \times d_1 + \theta_1 \quad (8)$$

$$\theta_1 = \frac{3 \times \pi}{2} \quad (9)$$

wherein, s and β are equal to 0.01 and 1.5, respectively, while u and v are random numbers between 0 and 1. As a random value, r_1 is between 0 and 20. d_1 is integer numbers from 1 to the length of the search space (d). U = 0.00565 and $\omega = 0.005$ are small values.

2.1.3. Expanded Exploitation

When the prey area is specified accurately, the Aquila descends vertically for a preliminary attack to discover the prey reaction. It is mathematically represented as in Equation (10).

$$\begin{gathered} X_3(k+1) = (X_{best}(k) - X_M(k)) \times \alpha - rand + \\ ((UB - LB) \times rand + LB) \times \delta \end{gathered} \quad (10)$$

wherein, *rand* is between 0 and 1; α and δ are 0.1. LB denotes the lower bound and UB denotes the upper bound of the search space.

2.1.4. Narrowed Exploitation

When the Aquila draws close to the prey, it attacks the prey over the land in light of its stochastic movements. The mathematical representation of this behavior is written as in Equation (11).

$$X_4(k+1) = QF \times X_{best}(k) - (G_1 \times X(k) \times rand) - G_2 \times Levy(d) + rand \times G_1 \quad (11)$$

wherein, QF is used to balance the search strategy, see Formula (12) for the calculation. G_1 and G_2 refer to Formulas (13) and (14), respectively.

$$QF(k) = k^{\frac{2\times rand()-1}{(1-K)^2}} \quad (12)$$

$$G_1 = 2 \times rand() - 1 \quad (13)$$

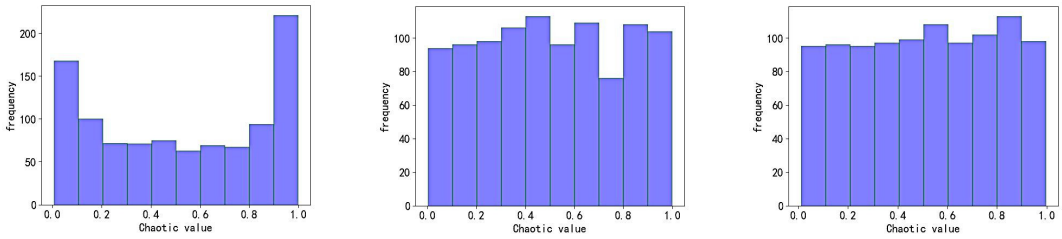
$$G_2 = 2 \times (1 - \frac{k}{K}) \quad (14)$$

2.2. Improved Aquila Optimizer (IAO)

Compared with other swarm intelligence optimization algorithms, AO not only has high optimization accuracy, but also fast convergence speed [12]. Nevertheless, just like other optimization algorithms, the population diversity of AO will also decrease at the later iteration, and local optimal problems will appear. Therefore, an improved AO is proposed in this paper. First of all, the improved tent chaotic mapping is utilized to initialize the population so as to improve the global search ability of the algorithm. Then, the adaptive t-distribution strategy is introduced for the position update, which effectively improves the defect that it is easy to fall into local extremum. The improved strategy is as below:

1. Improved tent chaos mapping

Random generation is used in AO to initialize the population, leading to uneven distribution of the Aquila population. Chaotic mapping is characterized by regularity and ergodicity [13], so in this paper, chaotic sequencing is used to initialize the position of the Aquila. The most commonly used chaotic mappings are logistic mapping and tent mapping. As can be seen from Figure 1a, the distribution probability of logistic mapping values is higher between [0,0.1] and [0.9,1], and this inhomogeneity seriously affects the optimization efficiency of the algorithm. It can be observed from Figure 1b that it is more uniform in tent mapping distribution.



(**a**) (**b**) (**c**)

Figure 1. Distribution histogram of different chaotic sequences. (**a**) Logistic chaotic sequence distribution histogram. (**b**) Tent chaotic sequence distribution histogram. (**c**) Improved tent chaotic sequence distribution histogram.

The tent mapping function is calculated as follows: Formula (15)

$$x_{i+1} = \begin{cases} 2x_i \ , \ 0 \le x \le \frac{1}{2} \\ 2(1-x_i) \ , \ \frac{1}{2} \le x \le 1 \end{cases} \quad (15)$$

The expression after the Bernoulli transformation is as shown below:

$$x_{i+1} = (2x_i) mod 1 \tag{16}$$

Tent chaotic sequencing has the problem that it will fall into small periodic points and unstable periodic points during iteration [14]. Hence, a random variable rand(0,1)×(1/*NT*) is added to the original, and the expression is as follows:

$$x_{i+1} = \begin{cases} 2x_i + \text{rand}(0,1) \times \frac{1}{NT}, 0 \le x \le \frac{1}{2} \\ 2(1-x_i) + \text{rand}(0,1) \times \frac{1}{NT}, \frac{1}{2} \le x \le 1 \end{cases} \tag{17}$$

The expression after the Bernoulli transformation is as shown below:

$$x_{i+1} = (2x_i) mod 1 + \text{rand}(0,1) \times \frac{1}{NT} \tag{18}$$

wherein, *NT* represents the number of particles in the chaotic sequence, while rand (0,1) represents the random number between (0,1).

In this paper, an improved tent chaotic mapping is adopted to initialize the population, rendering the position distribution of the initial population more uniform. The steps are as below:

- Generate a random number x_0 between (0,1), i = 0.
- As in Equation (18), an x sequence is generated, and i increases by 1
- When i reaches the maximum number of iterations, the x sequence is saved
- As shown in Formula (19), the elements of x sequence are mapped to the Aquila individual so as to obtain x_i^{new}
- Wherein, *UB* and *LB* represent the upper and lower bounds of the search space, respectively.

$$x_i^{\text{new}} = LB + (UB - LB)x_i \tag{19}$$

2. Adaptive t-distribution strategy

T-distribution is also known as student distribution [15]. The shape of the distribution function curve is up to the value of the parameter degree of freedom n. As indicated in Figure 2, the smaller the degree of freedom n is, the flatter the shape of the curve is. If the degree of freedom n = 1, the t-distribution is similar to Cauchy distribution, expressed as $t(n = 1) \rightarrow C(0,1)$. The larger the degree of freedom n is, the higher the curve shape is, and the curve is close to the standard normal distribution curve. When the degree of freedom n is infinite, the t-distribution is like the Gaussian distribution, expressed as $t(n = \infty) \rightarrow N(0,1)$.

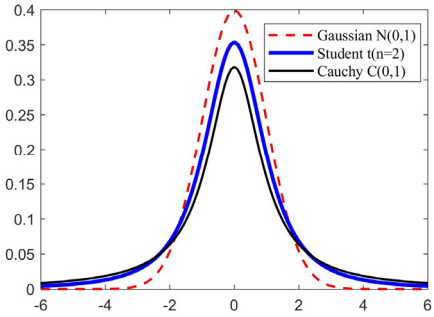


Figure 2. Distribution map of t distribution, Gaussian distribution, and Cauchy distribution function.

The adaptive t-distribution is used for the Aquila's position update, as shown in the following formula:

$$x_i^t = x_i + x_i * t(k) \tag{20}$$

wherein, x_i^t represents the Aquila's position after t variation, x_i is the position of the ith Aquila individual, and $t(k)$ indicates that in the t-distribution the current iteration number is selected as the parameter freedom. When the current iteration number k becomes smaller, the t-distribution is alike to the Cauchy distribution variation. In this case, the t-distribution operator obtains a larger value, and the step size of the position variation becomes larger, which means the algorithm has a better global search ability. During the middle of iteration, in t-distribution, it changes from Cauchy distribution variation to Gaussian distribution variation, and the t-distribution operator's value is relatively in the middle. In the algorithm combining the advantages of the two, it has both global and local search ability. When the number of iterations k is larger during the later iteration, the t-distribution is similar to Gaussian distribution variation. At this point, the t-distribution operator's value and the variation step are both smaller, which means the algorithm has better local search ability. This strategy is beneficial to finding the global optimal point in the algorithm.

2.3. IAO Performance Evaluation

2.3.1. Parameter Setting and Benchmarking Function Selection

In a bid to verify the performance of the IAO algorithm, GA [16], PSO [17], GWO [18], SSA [19], AO, and IAO are employed to conduct simulation experiments on the test function. The parameter settings of each algorithm are outlined in Table 1, and each algorithm has a population size of 30. The multi-dimensional single-peak function Sphere and multi-dimensional multi-peak function Ackley are selected for the benchmarking function, and the dimensions d of each function are 10, 30, and 100 dimensions, respectively. The value range and optimal solution information of the test function are shown in Table 2.

Table 1. Parameter setting table.

Algorithm	Parameter Settings
GA	Crossover probability is 0.7, and Mutation probability is 0.2
PSO	Learning factor: $c_1 = c_2 = 1.5$, Initial weight: $W = 0.729$
GWO	a linearly decreases from 2 to 0, $r_1, r_2 \in [0,1]$
SSA	Safety threshold: $ST = 0.6$, Finder ratio: $PD = 0.2$, Investigator ratio: $SD = 0.1$
AO	$U = 0.00565$, $\omega = 0.005$, $s = 0.01$, $\beta = 1.5$, $\alpha = 0.1$, $\delta = 0.1$, $r_1 \in [0,20]$, $u,v \in [0,1]$
IAO	Same as AO

Table 2. Benchmarking function.

Function	Range of Values	Optimal Value
Sphere	[100,100]	0
Ackley	[32.768,32.768]	0

2.3.2. Comparative Analysis of Algorithm Performance Results

The Sphere and Ackley benchmark test functions were optimized using GA, PSO, GWO, SSA, AO, and IAO algorithms. Table 3 shows the results obtained by independently running these four algorithms twenty times, yielding the average and standard deviation for each. By comparing the average and standard deviation, the precision and stability of each algorithm can be reflected. The closer the value is to 0, the higher the algorithm's stability.

In the multi-dimensional single-peak function Sphere, as dimensionality increases, the optimization accuracy of GA, PSO, GWO, SSA, and AO noticeably decreases because these algorithms tend to get stuck in local optima, making them unsuitable for high-dimensional problems. In contrast, the improved IAO algorithm, which incorporates tent map chaos initialization and adaptive t-distribution strategy, enhances its ability to escape local optima and perform global optimization. As a result, the optimization accuracy of the IAO algorithm is further improved with increasing functional dimensionality. Compared to the other five algorithms, the IAO algorithm exhibits extremely high optimization precision

and greater stability, with both mean and standard deviation being the smallest. This observation holds true for the multi-dimensional multi-peak function Ackley as well, where the optimization accuracy of GA, PSO, GWO, SSA, and AO still decreases with dimensionality, but the mean and standard deviation of IAO hardly change and remain the smallest, which further demonstrates its superiority over the other five algorithms in terms of optimization accuracy and stability.

Table 3. Comparison of test function optimization results.

Function	Algorithm	d = 10		d = 30		d = 50	
		Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation
Sphere	GA	1.64×10^{-1}	7.14×10^{-1}	3.14×10^{0}	4.16×10^{-1}	4.535×10^{1}	1.356×10^{0}
	PSO	3.02×10^{-2}	3.55×10^{-2}	5.13×10^{-1}	3.56×10^{-1}	2.877×10^{1}	3.935×10^{0}
	GWO	4.14×10^{-4}	3.76×10^{-4}	7.54×10^{-2}	5.16×10^{-2}	3.854×10^{-1}	2.857×10^{0}
	SSA	2.63×10^{-8}	1.14×10^{-8}	3.18×10^{-6}	4.14×10^{-6}	5.455×10^{-7}	6.732×10^{-7}
	AO	6.14×10^{-17}	5.33×10^{-17}	5.48×10^{-15}	6.19×10^{-15}	7.55×10^{-12}	3.55×10^{-12}
	IAO	7.14×10^{-130}	0	7.36×10^{-159}	0	3.55×10^{-190}	0
Ackley	GA	3.64×10^{0}	4.56×10^{0}	8.18×10^{1}	6.14×10^{-1}	7.593×10^{1}	8.479×10^{1}
	PSO	5.02×10^{-1}	7.57×10^{-1}	1.55×10^{0}	5.38×10^{-1}	9.258×10^{1}	3.398×10^{0}
	GWO	5.37×10^{-2}	7.86×10^{-2}	6.26×10^{-1}	7.41×10^{-1}	7.398×10^{0}	6.985×10^{0}
	SSA	2.83×10^{-8}	4.14×10^{-8}	9.14×10^{-7}	6.29×10^{-7}	4.698×10^{-7}	9.874×10^{-7}
	AO	3.67×10^{-17}	8.35×10^{-17}	5.33×10^{-16}	3.84×10^{-16}	9.683×10^{-16}	8.489×10^{-16}
	IAO	7.17×10^{-68}	0	7.22×10^{-68}	0	7.19×10^{-68}	0

2.3.3. Comparative Analysis of Algorithm Convergence Curves

Under the same operating conditions, each algorithm had a population size of 30 and a maximum iteration of 1000, and each algorithm runs independently 20 times. The convergence curves of GA, PSO, GWO, and SSA, AO, IAO on two test functions (d = 30) are shown in Figures 3 and 4, respectively. The *x*-axis represents the iteration number, while the *y*-axis represents the logarithm of the fitness value (i.e., lg(fitness value)). The fitness value here refers to the value of the test function. The benchmark test function curves can clearly demonstrate the algorithm's convergence speed, accuracy, and ability to escape local spaces. The earlier the curve appears to change direction, the faster the convergence speed. The lower the lg(fitness value), the higher the optimization accuracy.

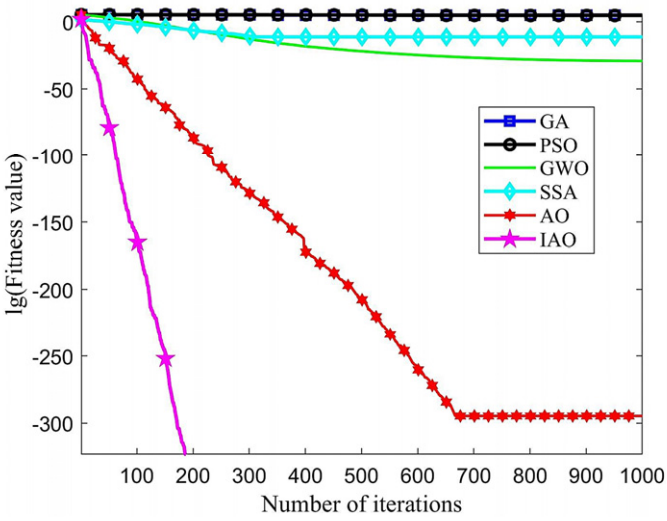


Figure 3. Sphere function optimization curve.

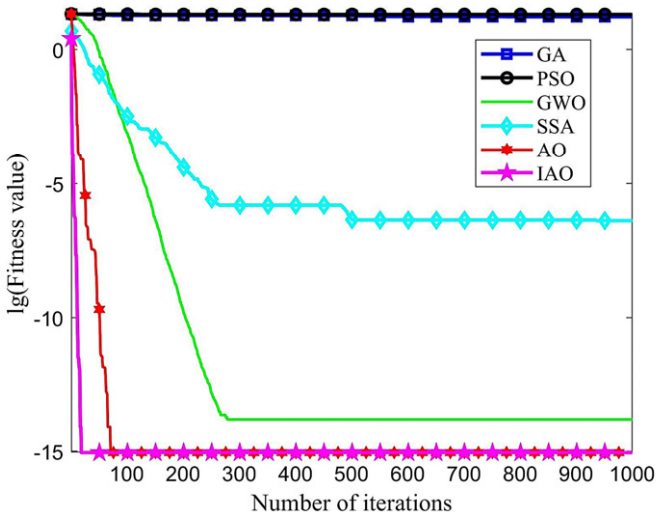


Figure 4. Ackley function optimization curve.

The convergent curves of GA, PSO, GWO, and SSA in the optimization process are relatively gentle, with slow inflection point occurrence. Therefore, these four algorithms tend to become stuck in local optima early on in the iteration process and cannot escape with increasing iterations, resulting in poor convergence effectiveness and optimization accuracy. Compared to AO and other algorithms, IAO has a faster downward curve and inflection point appearance. Thus, IAO outperforms the aforementioned algorithms in terms of optimization accuracy and convergence speed.

3. Mixed Kernel IAO-LSSVM Prediction Model

3.1. Mixed Kernel LSSVM Model

The distribution network reliability data are based on each year and have a small quantity of data; LSSVM has advantages in dealing with regression problem of small samples [20]. So, in this paper LSSVM is used to achieve reliability prediction of the distribution network. First proposed by Suykens et al. [9]. LSSVM is an improved SVM featuring simpler and faster computation.

It is assumed that the sample set (x_i, y_i) of the reliability training data of the distribution network is obtained, where $i = 1, 2 \ldots l$, $x_i \in R^n$ is the n-dimensional input, $y_i \in R$ is the one-dimensional output, x_i represents the influence factor data on distribution network reliability [21], and y_i indicates the reliability index data of the distribution network. When the Least Square Support Vector Machine (LSSVM) is used to process such regression problems as the reliability prediction of distribution networks, the nonlinear problems in the original low-dimensional space are transformed into linear regression in the high-dimensional space by means of nonlinear mapping function $\varphi(\cdot)$ [22], that is, in the high-dimensional space, the input and output of samples are fitted:

$$y(x) = w^T \varphi(x) + b \tag{21}$$

wherein w is the weight matrix and b is the offset quantity.

The principle of structural risk minimization is adopted in defining optimization problems:

$$\begin{cases} minJ(w, b, e) = \frac{1}{2} w^T w + \frac{1}{2} C \sum_{i=1}^{l} e_i^2 \\ s.t. y_i = w^T \varphi(x_i) + b + e_i \end{cases} \tag{22}$$

In the formula above, C is the penalty parameter. The higher the value of C is, the more likely it is to be over-fitted; the lower it is, the more prone it is to be under-fitted. Therefore, it is necessary to make a reasonable choice of C. e_i is the fitting error.

Then, use the Lagrange function to solve the optimization problem as below:

$$L(w,b,e,\alpha) = J(w,e) - \sum_{i=1}^{l} \alpha_i \left[w^T \varphi(x_i) + b + e_i - y_i \right] \tag{23}$$

wherein, α_i is the Lagrange multiplier.

Based on the condition of KKT, the partial derivative of Equation (23) is equal to 0, and then

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \rightarrow w = \sum\limits_{i=1}^{l} \alpha_i \varphi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum\limits_{i=1}^{l} \alpha_i = 0 \\ \frac{\partial L}{\partial e} = 0 \rightarrow \alpha_i = Ce_i \\ \frac{\partial L}{\partial \alpha} = 0 \rightarrow w^T \varphi(x_i) + b + e_i - y_i = 0 \end{cases} \tag{24}$$

Next, define the kernel function $K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j)$ which satisfies the Mercer principle, the necessary and sufficient conditions of the kernel function, and then the prediction model can be expressed as follows:

$$y(x) = \sum_{i=1}^{l} \alpha_i K(x, x_i) + b \tag{25}$$

The performance of LSSVM is determined by kernel function. As for the traditional kernel functions, as a global kernel function, the polynomial kernel function has poor local learning ability. As a local kernel function, the radial basis kernel function (RBF) has good local learning ability. Therefore, the advantages of the two kernel functions are combined to construct a mixed kernel function:

$$K = \lambda(xx_i + 1)^q + (1 - \lambda) exp\left(\frac{-\|x - x_i\|}{2\sigma^2} \right) \tag{26}$$

wherein, λ is the weight coefficient and $\lambda \in [0,1]$, q represents the parameters of the polynomial kernel function and σ^2 means the parameters of the radial basis kernel function.

3.2. Establishment of Mixed Kernel IAO-LSSVM Model

For the establishment of the mixed kernel LSSVM distribution network reliability prediction model, the grey relational analysis method [23] is used to analyze the correlation degree of the influencing factors. The influencing factors with a correlation degree higher than a set threshold value are selected as the input of the model. This reduces the complexity of model training and the calculation time. However, for the prediction accuracy of the mixed kernel LSSVM model, parameters C, λ, q and σ^2 are the determining factors, so parameter optimization is particularly important.

In this paper, in order to optimize the parameters, the improved Aquila Optimizer is used, where each Aquila represents a set of parameters (C,λ,q,σ^2). The specific modeling steps are as follows:

1. Initialize the parameters. Set the optimization intervals of C, λ, q, and σ^2, as well as relevant parameter values of the IAO algorithm. Formulas (18) and (19) are adopted to initialize the Aquila population.
2. Select the fitness function. The MSE of the predicted value y_i' and the true value y_i of the training sample are taken as the fitness value of the Aquila individual, where y_i' is calculated by Formula (25). The calculation formula of MSE is as shown below:

$$\text{MSE} = \frac{1}{l} \sum_{i=1}^{l} \left(y_i' - y_i \right)^2 \tag{27}$$

3. The fitness values of Aquila individuals are calculated, based on what current best position $X_{best}(k)$ is determined. $X_M(k)$, G_1, G_2, Levy(d) are updated
4. When $k \leq (2/3)*K$ and $rand \leq 0.5$, refer to the Formula (2) for the location update of the Aquila individual; otherwise, refer to the Formula (4).
5. When $(2/3)*K<k \leq K$, and $rand \leq 0.5$, the Formula (10) is used to update the location of the Aquila individual; otherwise, the Formula (11) will be used.
6. In accordance with Equation (20), the t-distribution mutation operation is conducted on the Aquila position.
7. Calculate the fitness value of each Aquila after t-distribution variation. Update individual fitness values as well as global optimal information.
8. Termination condition. The parameters corresponding to the optimal Aquila individual position (C,λ,q,σ^2) are output if the maximum number of iterations K has been reached. Thus, the mixed kernel IAO-LSSVM distribution network reliability prediction model is established. Otherwise, return to Step 3
9. The grey relational analysis method is used for analyzing the data. Firstly, the distribution network data are processed in a dimensionless manner. Secondly, the entropy weight correlation degree of influencing factors of power supply reliability is calculated. Finally, the influencing factors are sorted in accordance with the grey correlation degree, and the entropy weight correlation degrees higher than the set threshold are selected as the input of the mixed kernel IAO-LSSVM model.
10. Data set division. The influencing factors selected in step 9 are taken as the input of the model, while the distribution network reliability index is taken as the output of the model. Then, the distribution network reliability data are divided into training set and test set, and are normalized.
11. The established mixed kernel IAO-LSSVM model is used for the distribution network reliability prediction.

3.3. Model Evaluation Index

Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE) are selected as the evaluation indexes of the model so as to compare the performance of each model. The calculation formulas of each index refer to Formulas (28), (29), and (30), respectively.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{S}\left(\hat{Y}_i - Y_i\right)^2}{S}} \tag{28}$$

$$\text{MAPE} = \frac{1}{S}\sum_{i=1}^{S}\frac{\left|\hat{Y}_i - Y_i\right|}{Y_i} * 100\% \tag{29}$$

$$\text{MAE} = \frac{1}{S}\sum_{i=1}^{S}\left|\hat{Y}_i - Y_i\right| \tag{30}$$

wherein, $\hat{Y}_i$ represents the predicted value, Y_i indicates the real value, and S is the total number of data samples for model performance testing.

4. Case Analysis

The distribution network reliability data of Hubei Province from 2006 to 2015 [6] are selected for the experiment, and the distribution network reliability index involves power supply reliability. Refer to Table 4 for 27 power supply reliability rates and influence factor data of the Hubei power distribution network from 2006 to 2015, and Table 5 for the result table of the entropy weight correlation degree of influencing factors of the power supply reliability rate of the distribution network in Hubei Province.

Table 4. Power supply reliability rate and influence factor data of the Hubei power distribution network from 2006 to 2015.

	2006	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**
X_1	1.56	1.63	1.72	1.8	1.95	2.04	2.82	3.01	3.35	3.68
X_2	38.55	40.93	42.7	45.1	48.5	50.8	53.6	56.9	59.2	61.6
X_3	1.95	2.03	2.1 2	2.1 5	2.25	2.37	2.06	2.1 2	2.35	2.61
X_4	37.98	40.79	43.52	46.11	48.97	51.13	54.22	57.63	61.4	64.35
X_5	51.03	54.52	57.86	61.03	65.22	69.09	73.76	78.57	84.4	87.78
X_6	36.75	38.97	42.02	40.55	45.6	44.7	43.5	44.2	43.27	42.3
X_7	41.25	39.87	37.99	36.4	34.3	33.7	32.8	29.7	28.9	27.8
X_8	10.02	8.62	7.93	7.5	6.92	6.87	6.66	6.38	6.01	5.83
X_9	0.1 1	0.14	0.1 7	0.22	0.35	0.51	0.62	0.73	0.79	0.82
X_{10}	32.61	35.72	37.85	39.87	44.76	46.29	49.88	52.37	54.87	57.99
X_{11}	1.83	1.92	1.98	2.03	2.06	2.08	1.84	1.95	2.33	2.61
X_{12}	39.05	41.06	43.27	45.2	47.3	49.85	51.03	53.29	55.35	59.81
X_{13}	38.98	40.97	42.65	43.07	48.7	53.6	55.46	57.12	59.85	62.53
X_{14}	45.09	44.72	44.56	43.13	42.89	41.87	42.24	41.95	41.66	40.51
X_{15}	33.87	32.95	32.06	31.68	31.05	30.27	30.73	30.44	30.15	29.93
X_{16}	10.36	9.79	9.05	8.1 9	7.83	7.55	6.92	6.61	6.33	6.02
X_{17}	17.25	18.02	18.98	19.31	19.77	20.33	20.82	21.36	21.89	22.3
X_{18}	2.17	2.36	2.75	3.01	3.55	4.84	5.68	6.33	7.1 8	8.67
X_{19}	29.76	31.87	33.87	35.76	40.65	42.18	45.77	48.26	50.76	53.88
X_{20}	38.75	39.86	41.27	43.76	46.98	48.25	49.78	55.98	58.98	61.25
X_{21}	15.98	15.27	14.98	14.36	13.65	12.67	11.81	10.08	8.96	7.6
X_{22}	0	0	0	0	0.03	0.12	0.36	0.57	2.13	4.36
X_{23}	37.021	35.098	33.54	31.52	29.87	27.986	25.40	23.016	20.825	18.007
X_{24}	11.57	12.93	14.75	16.23	17.55	19.86	22.13	26.36	30.01	36.99
X_{25}	14.19	13.28	12.05	11.68	11.03	10.75	9.21	7.98	5.39	5.02
X_{26}	59.02	57.98	55.76	53.73	49.85	45.9	42.13	37.31	32.29	30.13
X_{27}	6.9	7.3	9.8	11.3	20.5	31.6	42.3	48.5	57.1	65.3
X_0	0.99823	0.99842	0.99863	0.99897	0.99918	0.99927	0.99938	0.99947	0.99955	0.99938

Table 5. Result table of entropy weight correlation degree of influencing factors of power supply reliability rate of the distribution network in Hubei Province.

110 KV	**Influencing Factor**	**35 KV**	**Ifluencing Factor**	**10 KV**	**Influencing Factor**
Cabling rate X_1	0.9011	Cabling rate X_9	0.6540	Insulation rate of overhead lines X_{17}	0.9642
Main structure ratio X_2	0.9374	Main structure ratio X_{10}	0.9191	Cabling rate X_{18}	0.8125
Capacity–load ratio X_3	0.9711	Capacity–load ratio X_{11}	0.9717	Main structure ratio of the overhead network X_{19}	0.9194
N − 1 security main transformer ratio X_4	0.9294	N − 1 security main transformer ratio X_{12}	0.9471	Main structure ratio of cable network X_{20}	0.9458
N − 1 security line ratio X_5	0.9284	N − 1 security line ratio X_{13}	0.9379	Average power supply radius X_{21}	0.9527
Mean value of maximum load rate of Main transformer X_6	0.9666	Mean value of maximum load rate of the Main transformer X_{14}	0.9883	Effective coverage of distribution automation X_{22}	0.9642
Mean value of maximum line load rate X_7	0.9619	Mean value of the maximum line load rate X_{15}	0.9823	High-loss distribution transformers X_{23}	0.9479
Heavy-load line ratio X_8	0.9397	Heavy-load line ratio X_{16}	0.9468	N − 1 security line ratio X_{24}	0.8547
–		–		Heavy-load line ratio X_{25}	0.9374
–		–		Mean value of maximum line load rate X_{26}	0.9531
–		–		Non-power-cut maintenance working ratio X_{27}	0.6555

4.1. Analyze Influencing Factors

The entropy weight-grey correlation analysis method is used to calculate the entropy weight correlation degree of each influencing factor X_1–X_{27} relative to the power supply reliability indicator X_0. The size of the entropy weight correlation degree reflects the degree of influence of each factor on the reliability indicator. The larger the entropy weight correlation degree, the closer the correlation between the two. Table 5 presents the entropy weight correlation degrees of each influencing factor.

Among them, correlation degrees of entropy weight greater than 0.7 are called a strong correlation factor, while those below 0.3 are weak. Therefore, in this paper, the influencing factors with a correlation degree of no less than 0.7 are selected as the input of the model. The influencing factors X_9 and X_{27} are rejected, and the remaining 25 influencing factors are chosen as the input of the model, while the power supply reliability X_0 is taken as the output of the model.

4.2. Model Parameter Determination

The eight groups of data from 2006 to 2013 shown in Table 3 are used as training samples for power supply reliability prediction in 2014 and 2015 in order to verify the predictive performance of the IAO-LSSVM model. The accuracy of the same test set (all the combination models use the mixed-core LSSVM) is compared for the models including single-kernel RBF-LSSVM, mixed-core LSSVM, GA-LSSVM, PSO-LSSVM, GWO-LSSVM, SSA-LSSVM, AO-LSSVM, and IAO-LSSVM. For each optimization algorithm, the parameter settings are identical to those in the IAO performance evaluation, with dimensions d = 4 and K = 100. In the RBF-LSSVM model, the radial basis kernel function is selected, with parameters C = 100, σ^2 = 5. For the mixed kernel LSSVM model, the mixed kernel function combining radial basis kernel function and polynomial kernel function is used, with parameters C = 100, λ = 0.5, q = 1, = 5. As for other model parameters C, λ, q, and σ^2, the lower and upper bounds are [100, 0, 0.7, 3] and [300, 1, 1.3, 5], respectively.

The fitness curves of IAO and AO optimized LSSVM refer to Figure 5. The ISSA-LSSVM curve shows a faster descent rate and a turning point, indicating that the model has faster convergence speed and better convergence accuracy, leading to better performance. Table 6 presents the optimization results of hybrid core LSSVM parameters.

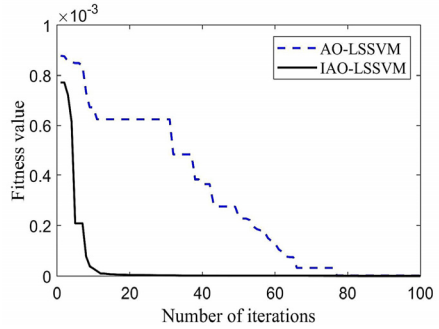


Figure 5. AO-LSSVM and IAO-LSSVM model fitness curves.

Table 6. Optimization results of hybrid core LSSVM parameters.

Optimization Algorithm	Parameter C	Parameter λ	Parameter q	Parameter σ^2
GA	216.74	0.39	1.21	4.17
PSO	291.81	0.49	0.98	3.95
GWO	247.90	0.37	0.99	4.70
SSA	244.65	0.19	1.07	4.24
AO	223.38	0.14	1.02	3.34
IAO	286.72	0.29	1.10	4.37

5. Discussion

In experiment 1, a random selection of 9 years of samples from 2006 to 2015 was chosen for training in order to predict the reliability of power supply for a year outside of these samples. Figure 6 displays the comparison between the prediction results of the IAO-LSSVM model and the actual values of the training samples, which shows a high degree of agreement between the predicted and actual values. This proves that the model is capable of accurately predicting the reliability of distribution networks, and demonstrates the algorithm's wide applicability for analyzing trends in reliability changes.

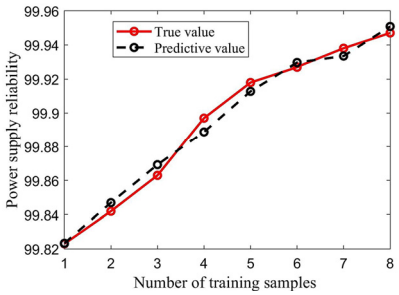


Figure 6. Comparison of prediction results of training samples.

In experiment 2, eight samples from 2006 to 2013 were selected as training samples to predict power supply reliability in 2014 and 2015. The predictive effects of the RBF-LSSVM model, hybrid kernel LSSVM model, GA-LSSVM model, PSO-LSSVM model, GWO-LSSVM model, SSA-LSSVM model, AO-LSSVM model, and IAO-LSSVM model were compared with actual values. To provide a more intuitive display of the results, bar graphs of the predictive results and actual values for each model were created in the testing sample (as shown in Figure 7). Additionally, Table 7 lists three types of errors for the eight models when applied to the testing sample and time required to run the model.

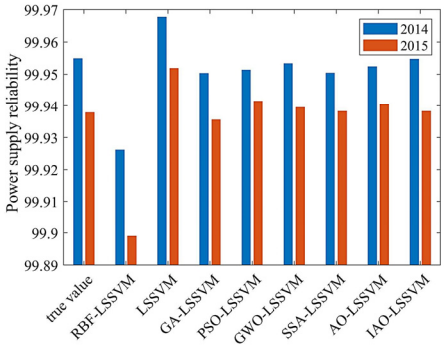


Figure 7. Histogram of sample prediction results of each model.

Table 7. Error comparison of test data and time required to run the model.

Prediction Model	RMSE	MAPE/%	MAE	Running Time (s)
RBF-LSSVM	0.0494	0.0487	0.0486	2.539
LSSVM	0.0241	0.0182	0.0184	3.305
GA-LSSVM	1.0305×10^{-3}	9.8324×10^{-4}	9.8629×10^{-4}	2.243
PSO-LSSVM	7.9773×10^{-4}	7.6437×10^{-4}	7.6398×10^{-4}	1.450
GWO-LSSVM	2.6909×10^{-4}	2.1875×10^{-4}	2.1864×10^{-4}	0.8050
SSA-LSSVM	1.7178×10^{-4}	1.7032×10^{-4}	1.7023×10^{-4}	0540
AO-LSSVM	2.8671×10^{-4}	2.8317×10^{-4}	2.8301×10^{-4}	0.445
IAO-LSSVM	5.8567×10^{-5}	4.8856×10^{-5}	4.8832×10^{-5}	0.118

As can be seen from Figure 7, IAO-LSSVM has a good fitting degree for the training samples. The prediction accuracy of the IAO-LSSVM model for test samples is much higher than that of other models. For the error comparison in Table 5, it shows that the prediction error of IAO-LSSVM is the minimum among the eight models. The root mean square error (RMSE) of the three evaluation indexes are chosen to analyze the eight models. The RBF-LSSVM model with single kernel function has a maximum error of 0.0494. Compared with RBF-LSSVM, the mixed kernel LSSVM has a smaller error of 0.0241. As the parameters for these two models are obtained based on previous experience, certain errors still exist comparing with other optimized models. The IAO-LSSVM model among all the combination models has the minimum prediction error of 5.8567×10^{-5}, which is one order of magnitude higher than others.

Therefore, it can be judged that the IAO-LSSVM prediction model demonstrates smaller error and higher accuracy than other common models in use, and it has achieved good results on power supply reliability prediction of the distribution network. The required running times for each algorithm's prediction models are shown in Table 5, with the IAO algorithm requiring the shortest time compared to other algorithms. Therefore, it can be concluded that the IAO algorithm has a higher optimization speed compared to other algorithms and proves that the tent chaos initialization and adaptive t-distribution strategy have optimization effects on the convergence speed of the AO algorithm.

6. Conclusions and Future Work

In terms of problems needing to be solved, in the paper, the grey relational analysis method is firstly used to analyze the reliability data of distribution network. Then, the improved tent chaos initialization strategy and the adaptive t-distribution strategy are introduced to improve the AO algorithm, which enhances the searchability of the algorithm. Meanwhile, the parameters of the mixed kernel LSSVM model are optimized through the improved AO algorithm, and it effectively avoids blind parameter selection. In the paper, the distribution network reliability data come from fewer samples; however, their advantages are maximized due to the LSSVM model being used for reliability prediction. In terms of the prediction results, in this paper, the power supply reliability prediction results and errors of the RBF-LSSVM model, mixed kernel LSSVM model, GA-LSSVM model, PSO-LSSVM model, GMO-LSSVM model, SSA-LSSVM model, AO-LSSVM, and IAO-LSSVM are compared and analyzed. The findings show that the IAO-LSSVM prediction model features higher accuracy and also better stability.

Regarding future work, the LSSVM model adopted in this paper adopts a single radial basis kernel function, which has certain limitations, and the linear combination of the two kernel functions can be considered to form a mixed-kernel LSSVM model. By improving the model, the reliability of the distribution network can be accurately predicted, and the future development trend of the distribution network can be accurately judged, which provides certain reference significance for the future planning of distribution networks.

Author Contributions: This paper was a collaborative effort among all authors. All authors participated in the analysis, discussed the results, and wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, M.; Luan, Y.; Lu, X.; Hu, J.L.; Wu, D. *Summary of Power Supply Reliability Forecast of Distribution Network*, Technology ed.; China Electric Power: Beijing, China, 2012; pp. 40–42.
2. Munoz-Delgado, G.; Contreras, J.; Arroyo, J.M. Distribution Network Expansion Planning with an Explicit Formulation for Reliability Assessment. *Trans. Power Syst.* **2017**, *99*, 95–99.

3. Billinton, R.; Wang, P. Reliability network equivalent approach to Distribution System Reliability Evaluation. *IEEE Proc.-Gener. Transm. Distrib.* **1998**, *145*, 149–153. [CrossRef]
4. Billinton, R.; Billinton, J.E. Distribution System Reliability Indices. *IEEE Trans. Deliv.* **1989**, *4*, 561–568. [CrossRef]
5. Yan, Z.; Zhangchao, C. Evaluation of a 35/10 kV Electric Distribution System Reliability Level. *J. Shanghai Jiaotong Univ.* **1995**, *29*, 198–201.
6. Wang, Y.X.; Pan, X. Reality evaluation of distribution network based on improved-elman feedback dynamic neural network. *Water Resour. Power* **2019**, *37*, 177–180.
7. Zhan, C.Y. *Research on Model Optimization of Distribution Network Reliability Prediction*; Anhui University of Science & Technology: Huainan, China, 2020.
8. Liu, Q.; Lu, H.; Wang, Q.; Zhang, H.L. A regional traffic volume rolling prediction model based on support vector machine. *J. Harbin Inst. Technol.* **2011**, *43*, 138–143.
9. Suykens, J.; Vandewalle, J. Least Squares Support Vector Machine Classifiers. *Neural Process. Lett.* **1999**, *9*, 293. [CrossRef]
10. Dong, H.; Shi, L.S.; Zhao, P.C. Reliability Prediction of Urban Power Network Based on PSO-LSSVM Model. *Proc. CSU-EPSA* **2014**, *26*, 82–86.
11. Deng, H.; Zhu, X.; Zhang, Q.; Zhao, J. Prediction of Short-term Pubic Transportation Flow Based on Multiple-kernel Least Square Support Vector Machine. *J. Transp. Eng. Inf.* **2012**, *10*, 84–88.
12. Abualigah, L.; Yousri, D.; Abd Elaziz, M.; Ewees, A.A.; Al-Qaness, M.A.; Gandomi, A.H. Aquila Optimizer: A novel meta-heuristic optimization algorithm. *Comput. Ind. Eng.* **2021**, *157*, 107250. [CrossRef]
13. Liu, L.; Sun, S.Z.; Yu, H.; Yue, X.; Zhang, D. A Modified Fuzzy C-means (FCM) Clustering Algorithm and its Application on Carbonate Fluid Identification. *J. Appl. Geophys.* **2016**, *129*, 28–35. [CrossRef]
14. Zhang, N.; Zhao, Z.D.; Bao, X.A.; Qian, J.; Wu, B. Gravitational Search Algorithm Based on Improved Tent Chaos. *Control. Decis.* **2020**, *35*, 893–900.
15. Chen, X.; Zhao, S.; Liu, F. Robust identification of Linear ARX Models with Recursive EM Algorithm Based on Student's T-Distribution. *J. Frankl. Inst.* **2021**, *358*, 1103–1121. [CrossRef]
16. Li, H.; Peng, Y.; Deng, C.; G, D.Q. Review of Hybrids of GA and PSO. *Comput. Eng. Appl.* **2018**, *54*, 20–28+39.
17. Li, X.; Deng, D.; Liao, D.; Chen, J.H. A Review on Application of Particle Swarm Optimization in Electric Power Systems. *Power Syst. Prot. Control.* **2007**, *23*, 77–84.
18. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey Wolf Optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]
19. Xue, J.; Shen, B. A Novel Swarm Intelligence Optimization Approach: Sparrow Search Algorithm. *Syst. Sci. Control. Eng. Open Access J.* **2020**, *8*, 22–34. [CrossRef]
20. Zhang, J.; Tang, Z.H.; Gui, W.H.; Chen, Q.; Liu, J.P. Interactive image segmentation with a regression based ensemble learning paradigm. *Front. Inf. Technol. Electron. Eng.* **2017**, *18*, 1002–1021. [CrossRef]
21. Allan, R.N.; Billinton, R. Bibliography on the application of probability methods in power system reliability evaluation. *IEEE Trans. Power Syst. Publ. Power Eng. Soc.* **1999**, *14*, 51–57. [CrossRef]
22. Xiong, Y. Improved FCM-LSSVM Prediction Model for Penicillin Fed-batch Fermentation. *Control. Eng. China* **2017**, *24*, 2237–2242.
23. Yuan, K.; Luo, P.; Wang, G.; Jin, D.M.; Yao, L.Y. A New Method of Power System False Data Attack Detection Based on Grey Analysis. *Adv. Technol. Electr. Eng. Energy* **2019**, *38*, 17–23.

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.