Article

Icing Forecasting for Power Transmission Lines Based on a Wavelet Support Vector Machine Optimized by a Quantum Fireworks Algorithm

Tiannan Ma *, Dongxiao Niu and Ming Fu

Departemnt of Economics and Management, North China Electric Power University, Beijing 102206, China; ndx@ncepu.edu.cn (D.N.); fm@ncepu.edu.cn (M.F.)
* Correspondence: matiannan_1234@ncepu.edu.cn; Tel.: +86-185-1562-1058

Abstract: Icing on power transmission lines is a serious threat to the security and stability of the power grid, and it is necessary to establish a forecasting model to make accurate predictions of icing thickness. In order to improve the forecasting accuracy with regard to icing thickness, this paper proposes a combination model based on a wavelet support vector machine (w-SVM) and a quantum fireworks algorithm (QFA) for prediction. First, this paper uses the wavelet kernel function to replace the Gaussian wavelet kernel function and improve the nonlinear mapping ability of the SVM. Second, the regular fireworks algorithm is improved by combining it with a quantum optimization algorithm to strengthen optimization performance. Lastly, the parameters of w-SVM are optimized using the QFA model, and the QFA-w-SVM icing thickness forecasting model is established. Through verification using real-world examples, the results show that the proposed method has a higher forecasting accuracy and the model is effective and feasible.

Keywords: icing forecasting; support vector machine; fireworks algorithm

1. Introduction

In recent years, a variety of extreme weather phenomena have occurred on a global scale, causing overhead power transmission line icing disasters to happen frequently, accompanied by power outages and huge economic losses due to the destruction of a large number of fixed assets. Since the first power transmission line icing accident recorded in 1932, several serious icing disasters have occurred one after another throughout the world [1], such as Canada's freezing rain disaster in 1998, which led to direct economic losses of 1 billion dollars and indirect losses of 30 billion dollars for the power system. The power system failure caused by freezing rain was also a tremendous shock to the production and life of the region. In China, the earliest recorded icing disaster occurred in 1954. After that year, several large areas of freezing rain and snow disasters have successively occurred in China, especially Southern China's freezing weather in January 2008 [2]. That disaster caused tremendous damage to China's power system, directly resulting in 8709 tower collapses, more than 27,000 line breakages, and 1497 substation outages of 110 kV and above lines. The direct property loss for the State Grid Corporation amounted to 10.45 billion RMB, and the investment in post-disaster electricity reconstruction and transformation was 39 billion RMB. According to the lessons learned from previous grid freezing accidents, and given the current development situation of China's grid and the conditions of global climate change, the grid in China will again be tested by the future large-scale icing disasters. It is always preferable to take preventative action instead of reacting to events after an accident has

already happened. Therefore, icing forecasting research with regard to overhead power transmission lines have important practical application value.

At present, there are many studies on icing forecasting for power transmission lines. According to the hydrodynamic movement law and heat transfer mechanisms, domestic and foreign scholars have established a variety of transmission line icing forecasting models that consider meteorological factors, environmental factors, and various line parameters. Generally these models can be divided into three categories: mathematical and physical models, statistical models and intelligent forecasting models. Mathematical and physical models involve simulation modeling and forecasting of the icing growth process by observing the physical process and mathematical equations of icing formation, notably including the Goodwin model [3], the Makkonen model [4] and so on. These models are established based on the experimental data, but there are differences between the experimental data and the practical data, so the forecasting results of these models are not ideal. Statistical models process historical data based on statistical theory and traditional statistical methods and do not consider the physical process of icing formation, e.g., the multiple linear regression icing model [5]. However, transmission line icing is affected by a variety of factors, and a multiple linear regression model cannot take all factors into account, meaning the icing forecasting accuracy is greatly reduced.

Intelligent forecasting models combine modern computer technology with mathematical sciences, and are able to handle high-dimensional nonlinear problems through their powerful learning and processing capabilities, which can improve the prediction accuracy. The common intelligent forecasting models are support vector machines (SVM) and Back Propagation(BP) neural network. BP neural network has a strong nonlinear fitting ability to create the non-linear relationship between the output and a variety of impact factors. With its strong learning ability and forecasting capability, the non-linear output can arbitrarily approximate the actual value. For example, paper [6] presented a short-term icing prediction model based on a three layer BP neural network, and the results showed that the BP forecasting model is accurate for transmission lines of different areas. Paper [7] presented an ice thickness prediction model based on fuzzy neural network, the tested result demonstrated its forecasting abilities of better learning and mapping. However, the single BP model easily falls into the local optimum, and cannot always reach the expected accuracy. For this problem, some scholars adopted optimization algorithms to optimize the parameters of BP neural network, thereby improving the prediction accuracy. For instance, Du and Zheng *et al.* used Genetic Algorithm (GA) to optimize BP network and built the GA-BP ice thickness prediction model [8], and this model proved that the GA-BP model was more effective than the BP model in ice thickness prediction for transmission lines. Although this combination model could improve the prediction accuracy of BP neural networks, some scholars started to use SVM to build the icing forecasting model for transmission lines due to its slow calculating speed and overall poor performance. SVM can establish the non-linear relationship between various factors and ice thickness, and has better nonlinear mapping ability and generalization ability. In addition, SVM has a strong learning ability and can quickly approach the target value through continuous repetitive learning. Therefore, theSVM model is more widely used in transmission line research. For example, paper [9–11] introduced the icing forecasting model based on the Support Vector Regression learning algorithm, and obtained the ideal results.

As is known, the standard SVM uses a Gaussian kernel function to solve the support vector by quadratic programming, and the quadratic programming will involve the calculation of an m order matrix (m is sample number). When m is bigger, the processing ability of the Gaussian function is more unsatisfactory. It will take much computing time, thereby seriously affecting the learning accuracy and predictive accuracy of the algorithm. In transmission line icing forecasting, there are many influencing factors, so a large amount of input data will make the SVM algorithm become unfeasible using the traditional Gaussian kernel for such a large-scale training sample. In view of this problem, this paper will replace the Gaussian kernel function with the wavelet kernel function, and establishes the wavelet support vector machine (w-SVM) for icing forecasting. Using wavelet kernel function in place of the Gaussian kernel is mainly based on the following considerations [12]: (1) the wavelet kernel function

has the fine characteristic of progressively describing the data, and using the SVM of the wavelet kernel function can approximate any function with high accuracy while the traditional Gaussian kernel function cannot; (2) the wavelet kernel functions are orthogonal or nearly orthogonal, and the traditional Gaussian kernel functions are relevant, even redundant; (3) the wavelet kernel function has multi-resolution analyzing ability for wavelet signals, so the nonlinear processing capacity of the wavelet kernel function is better than that of Gaussian kernel function, which can improve the generalization ability of the support vector machine regression model.

The forecasting performance of the w-SVM model largely depends on the values of its parameters; however, most researchers choose the parameters of SVM only by subjective judgment or experience. Therefore, the parameters values of SVM need to be optimized by meta-heuristic algorithms. Currently, several algorithms have been successfully applied to determine the control parameters of SVM, such as genetic algorithms [13], particle swarm optimization [14], differential evolution [15] and so on. However, those algorithms have the defects of being hard to control, achieving the global optima slowly, *etc.* In this paper, we use the fireworks optimization algorithm (FA) proposed by Tan and Zhu in 2010 [16] to determine the parameter values of w-SVM. The advantage of using FA over other techniques is that it can be easily realized and be able to reach global optimal with greater convergence speed. Besides, in order to strengthen the optimization ability and obtain better results, this paper also attempts to improve the FA model by using a quantum evolutionary algorithm.

The rest of this paper is organized as follows: In Section 2, the quantum fireworks algorithm (QFA) and w-SVM are presented in detail. Also, in this section, a hybrid icing forecasting model (QFA-w-SVM) that combines the QFA and w-SVM models is established; In Section 3, several real-world cases are selected to verify the robustness and feasibility of QFA-w-SVM, and the computation, comparison and discussion of the numerical cases are discussed in detail; Section 4 concludes this paper.

2. Experimental Section

2.1. Quantum Fireworks Algorithm

2.1.1. Fireworks Algorithm

A fireworks algorithm (FA) [16] is used to simulate the whole process of the explosion of fireworks. When the fireworks explosion generates a lot of sparks, the sparks can continue to explode to generate new sparks, resulting in beautiful and colorful patterns. In an FA, each firework can be regarded as a feasible solution of the solution space for an optimization problem, and the fireworks explosion process can be seen as a searching process for the optimal solution. In a particular optimization problem, the algorithm needs to take into account the number of sparks of each fireworks explosion, how wide the explosion radius is, and how to select an optimal set of fireworks and sparks for the next explosion (searching process).

The most important three components of FA are the explosion operator, mutation operator and selection strategy.

(1) Explosion operator. The number of sparks each fireworks explosion generates and explosion radius are calculated based on fitness value of fireworks. For the fireworks x_i $(i = 1, 2, \cdots, N)$, the calculation formula for the number of sparks S_i and explosion radius R_i are:

$$S_i = M \times \frac{y_{\max} - f(x_i) + \varepsilon}{\sum\limits_{i=1}^{N} (y_{\max} - f(x_i)) + \varepsilon} \tag{1}$$

$$R_i = \hat{R} \times \frac{f(x_i) - y_{\min} + \varepsilon}{\sum\limits_{i=1}^{N} (f(x_i) - y_{\min}) + \varepsilon} \tag{2}$$

In the above formula, $y_{\max}$, $y_{\min}$ represent the maximum and minimum fitness value of the current population, respectively; $f(x_i)$ is the fitness value of fireworks x_i; and M is a constant to adjust the number of explosive sparks. $\hat{R}$ is a constant to adjust the size of the fireworks explosion radius. ε is the minimum machine value to be used to avoid zero operation.

(2) Mutation operator. The purpose of mutation operator is to increase the diversity of the sparks population; the mutation sparks of fireworks algorithm is obtained by Gaussian mutation, namely Gaussian mutation sparks. Select fireworks x_i to make the Gaussian mutation, and the k-dimensional Gaussian mutation is:

$$\hat{x}_{ik} = x_{ik} \times e \tag{3}$$

Where $\hat{x}_{ik}$ represents k-dimensional mutation fireworks, e represents the Gaussian distribution.

In the fireworks algorithm, the new generated explosion fireworks and mutation sparks may fall out of the search space, which makes it necessary to map it to a new location, using the following formula:

$$\hat{x}_{ik} = x_{LB,k} + |\hat{x}_{ik}| \, \% \, \left(x_{UB,k} - x_{LB,k}\right) \tag{4}$$

where $x_{UB,k}$, $x_{LB,k}$ are the upper and lower search spaces, respectively; and % denotes the modulo operator for floating-point number.

(3) Selection strategy. In order to transmit the information to the next generation, it is necessary to select a number of individuals as the next generation.

Assume that K individuals are selected; the number of population is N and the best individual is always determined to become the fireworks of next generation. Other $N-1$ fireworks are randomly chosen using a probabilistic approach. For fireworks x_i, its probability $p(x_i)$ of being chosen is calculated as follows:

$$p(x_i) = \frac{R(x_i)}{\sum\limits_{x_j \in K} x_j} \, , \; R(x_i) = \sum\limits_{x_j \in K} d\left(x_i - x_j\right) = \sum\limits_{x_j \in K} || \, x_i - x_j \, || \tag{5}$$

where $R(x)$ is the sum of the distances between all individuals in the current candidate set. In the candidate set, if the individual has a higher density, that is, the individual is surrounded by other candidates, the probability of the individual selected will be reduced.

If $F(x)$ is the objective function of fireworks algorithm, the steps of the algorithm are shown as follows:

(1) Parameters initialized; Randomly select N fireworks and initialize their coordinates.

(2) Calculate the fitness value $f(x_i)$ of each firework, and calculate the blast radius R_i and generated sparks number of each firework. Randomly select dimension z-dimensional coordinates to update coordinates, coordinate updating formula is as follows: $\hat{x}_{ik} = x_{ik} + R_i \times U(-1,1)$, $U(-1,1)$ stands for the uniform distribution on [–1,1].

(3) Generate $\hat{M}$ Gaussian mutation sparks; randomly select sparks x_i, use the Gaussian mutation formula to obtain $\hat{M}$ Gaussian mutation sparks $\hat{x}_{ik}$, and save those sparks into the Gaussian mutation sparks population.

(4) Choose N individuals as the fireworks of next generation by using probabilistic formula from fireworks, explosion sparks and Gaussian mutation sparks population.

(5) Stop condition. If the stop condition is satisfied, then output the optimal results; if not, return step (2) and continue to cycle.

2.1.2. Quantum Evolutionary Algorithm

The development of quantum mechanics impels quantum computing to be increasingly applied in various fields. In quantum computing, the expression of a quantum state is a quantum bit, and usually quantum information is expressed by using the 0 and 1 binary method. The basic quantum states are "0" state and "1" state. In addition, the state can be an arbitrary linear superposition state

between "0" and "1". That is to say, the two states can exist at the same time, which challenges the classic bit expression method in classical mechanics to a large extent [17]. The superposition state of quantum state can be presented as shown in Equation (6).

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \quad , \quad |\alpha|^2 + |\beta|^2 = 1 \tag{6}$$

where $|0\rangle$ and $|1\rangle$ are the two quantum states α and β are the probability amplitudes. $|\alpha|^2$ represents the probability at a quantum state of $|0\rangle$ and $|\beta|^2$ represents the probability at a quantum state of $|1\rangle$.

In QFA, the updating proceeds by quantum rotating gate, and the adjustment is:

$$\begin{pmatrix} \alpha_i' \\ \beta_i' \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} \tag{7}$$

in which set $U = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$, U is quantum rotating gate, θ is quantum rotating angle, and $\theta = \arctan(\alpha/\beta)$.

A quantum evolutionary algorithm is proposed based on a probabilistic search, as the conception of qubits and quantum superposition means a quantum evolutionary algorithm has many advantages, such as better population diversity, strong global search capability, especially great robustness, and the possibility of combining with other algorithms.

2.1.3. Quantum Fireworks Algorithm

Parameters Initialized

In the solution space, randomly generate N fireworks, and initialize their coordinates. Here, we use the probability amplitude of quantum bits to encode the current position of fireworks, and the encoding method is used by the following formula:

$$P_i = \begin{bmatrix} \cos(\theta_{i1})\cos(\theta_{i2})\cdots\cos(\theta_{in}) \\ \sin(\theta_{i1})\sin(\theta_{i2})\cdots\sin(\theta_{in}) \end{bmatrix} \tag{8}$$

where $\theta_{ij} = 2\pi rand()$, $rand()$ is a random number between 0 and 1; $i = 1, 2, \cdots, m; j = 1, 2, \cdots, n; m$ is the number of fireworks, n is the number of solution space.

Therefore, the corresponding probability amplitude of individual fireworks for the quantum states $|0\rangle$ and $|1\rangle$ are as follows:

$$P_{ic} = (\cos(\theta_{i1}), \cos(\theta_{i2})\cdots\cos(\theta_{in})) \tag{9}$$

$$P_{is} = (\sin(\theta_{i1}), \sin(\theta_{i2})\cdots\sin(\theta_{in})) \tag{10}$$

Solution Space Conversion

The searching process of a fireworks optimization algorithm is carried out on the actual parameter space $[a, b]$. Due to the probability amplitude of fireworks location being in the range of $[0, 1]$, it needs to be decoded into the actual parameter space $[a, b]$ to search in the fireworks algorithm. Let the $\Delta\theta_{jd}^{k+1}$th quantum bit for the individual α_{jd}^{k+1} is $\left[\alpha_i^j, \beta_i^j\right]^T$, and its corresponding conversion equations be:

$$X_{ic}^j = \frac{1}{2}[b_i(1 + \alpha_i^j) + a_i(1 - \alpha_i^j)] \; if \; rand() < P_{id} \tag{11}$$

$$X_{is}^j = \frac{1}{2}[b_i(1 + \beta_i^j) + a_i(1 - \beta_i^j)] \; if \; rand() \geqslant P_{id} \tag{12}$$

where $rand()$ is a random number between $[0, 1]$; X_{ic}^{j} is the actual parameter value in j th dimension position when the quantum state of i th fireworks individual is $|0>$, X_{is}^{j} is the actual parameter value in j th dimension position when the quantum state of the i th fireworks individual is $|1>$. b_i and a_i are the lower and upper limits.

Assuming the FA is searching in two-dimensional space, that means $j = 1, 2$. Initialize the position of population: *InitX_ axis*; *InitY_axis*; and the position of individuals can be determined as follows:

if $rand() < P_{id}$:

$$X(i) = X\_axis + \frac{1}{2}[b_i(1 + \alpha_i^1) + a_i(1 - \alpha_i^1)] \tag{13}$$

$$Y(i) = Y\_axis + \frac{1}{2}[b_i(1 + \alpha_i^2) + a_i(1 - \alpha_i^2)] \tag{14}$$

if $rand() \geqslant P_{id}$:

$$X(i) = X\_axis + \frac{1}{2}[b_i(1 + \beta_i^2) + a_i(1 - \beta_i^2)] \tag{15}$$

$$Y(i) = Y\_axis + \frac{1}{2}[b_i(1 + \beta_i^2) + a_i(1 - \beta_i^2)] \tag{16}$$

Calculate the Fitness Value $f(x_i)$ of Each Individual, and Obtain the Explosion Radius and Generated Sparks Number S_i

$$S_i = M \times \frac{y_{\max} - f(x_i) + \varepsilon}{\sum\limits_{i=1}^{N} (y_{\max} - f(x_i)) + \varepsilon} \tag{17}$$

$$R_i = \hat{R} \times \frac{f(x_i) - y_{\min} + \varepsilon}{\sum\limits_{i=1}^{N} (f(x_i) - y_{\min}) + \varepsilon} \tag{18}$$

Individual Position Updating

The individual position update is operated by using a quantum rotating gate using the following equation:

$$\begin{bmatrix} \alpha_{jd}^{k+1} \\ \beta_{jd}^{k+1} \end{bmatrix} = \begin{bmatrix} \cos\theta_{jd}^{k+1} & -\sin\theta_{jd}^{k+1} \\ \sin\theta_{jd}^{k+1} & \cos\theta_{jd}^{k+1} \end{bmatrix} \begin{bmatrix} \alpha_{jd}^{k} \\ \beta_{jd}^{k} \end{bmatrix} \tag{19}$$

where α_{jd}^{k+1} and β_{jd}^{k+1} are the probability amplitude of j th fireworks individual in $k + 1$ th iteration for d dimension space; θ_{jd}^{k+1} is the rotating angle, which can be get from equation:

$$\theta_{jd}^{k+1} = s(\alpha_{jd}^{k}, \beta_{jd}^{k}) \Delta\theta_{jd}^{k+1} \tag{20}$$

where $s(\alpha_{jd}^{k}, \beta_{jd}^{k})$ determines the rotating angle direction and is the $\Delta\theta_{jd}^{k+1}$ rotating angle increment.

In order to adapt to operation mechanism of fireworks algorithm, we convert the updated α_{jd}^{k+1} and β_{jd}^{k+1} to a solution space.

$$X_{jc}^{d} = \frac{1}{2}[b_j(1 + \alpha_{jd}^{k+1}) + a_j(1 - \alpha_{jd}^{k+1})] \ if \ rand() < P_{id}$$
$$X_{js}^{d} = \frac{1}{2}[b_j(1 + \beta_{jd}^{k+1}) + a_j(1 - \beta_{jd}^{k+1})] \ if \ rand() \geqslant P_{id} \tag{21}$$

Then calculate the positional offset amount:

$$h_{jc}^{d} = R_i \times X_{jc}^{d} \, , \ h_{js}^{d} = R_i \times X_{js}^{d} \tag{22}$$

$if\ rand() < P_{id}$, then $d = 1$

$$X(j) = X(i) + h_{jc}^d = X(i) + R_i \times \frac{1}{2}[b_j(1 + \alpha_{jd}^{k+1}) + a_j(1 - \alpha_{jd}^{k+1})]$$

$$Y(j) = Y(i) + h_{jc}^d = Y(i) + R_i \times \frac{1}{2}[b_j(1 + \alpha_{jd}^{k+1}) + a_i(1 - \alpha_{jd}^{k+1})]$$

(23)

$if\ rand() \geqslant P_{id}$: then $d = 2$

$$X(j) = X(i) + h_{js}^d = X(i) + R_i \times \frac{1}{2}[b_j(1 + \beta_{jd}^{k+1}) + a_j(1 - \beta_{jd}^{k+1})]$$

$$Y(j) = Y(i) + h_{js}^d = Y(i) + R_i \times \frac{1}{2}[b_j(1 + \beta_{jd}^{k+1}) + a_j(1 - \beta_{jd}^{k+1})]$$

(24)

Detection of cross-border. If the generated explosion sparks exceed the possible domain boundary, the position of sparks can be updated with the following equations:

$$X(j) = X_{LB,k} + |X(j)|\,\%(X_{UB,k} - X_{LB,k})$$

$$Y(j) = Y_{LB,k} + |Y(j)|\,\%(Y_{UB,k} - Y_{LB,k})$$

(25)

Individual Mutation Operation

The main reason for the premature convergence and local optimum of the fireworks group is that the diversity of the population is lost in the process of population search. In the quantum fireworks algorithm, in order to increase the diversity of the population, the Gauss mutation in the original algorithm is replaced by a quantum mutation. Randomly select fireworks x_i, and generate $\hat{M}$ quantum mutation sparks, and its operation formula is shown as follows:

$$\begin{bmatrix} 01 \\ 10 \end{bmatrix} \begin{bmatrix} \cos(\theta_{ij}) \\ \sin(\theta_{ij}) \end{bmatrix} = \begin{bmatrix} \sin(\theta_{ij}) \\ \cos(\theta_{ij}) \end{bmatrix} = \begin{bmatrix} \cos(\frac{\pi}{2} - \theta_{ij}) \\ \sin(\frac{\pi}{2} - \theta_{ij}) \end{bmatrix}$$

(26)

Let the probability of individual be P_m, and $rand()$ be a random number between $[0, 1]$; if $rand() < P_m$, the mutation can be operated with the above formula and the probability amplitude in quantum bit is changed; finally, the mutated individual can be converted into the solution space and save it to the mutation sparks population.

$$X_{jc}^d = \frac{1}{2}[b_j(1 + \cos(\frac{\pi}{2} - \theta_{ij})) + a_j(1 - \cos(\frac{\pi}{2} - \theta_{ij}))]\ if\ rand() < P_{id}$$

$$X_{js}^d = \frac{1}{2}[b_j(1 + \sin(\frac{\pi}{2} - \theta_{ij})) + a_j(1 - \sin(\frac{\pi}{2} - \theta_{ij}))]\ if\ rand() \geqslant P_{id}$$

(27)

$if\ rand() < P_{id}$, then $d = 1$

$\hat{X}(\hat{M}) = X(\hat{M}) \times X_{jc}^d$　　$\hat{Y}(\hat{M}) = Y(\hat{M}) \times X_{jc}^d$

$if\ rand() \geqslant P_{id}$: then $d = 2$

$\hat{X}(\hat{M}) = X(\hat{M}) \times X_{js}^d$　　$\hat{Y}(\hat{M}) = Y(\hat{M}) \times X_{js}^d$

Detection of cross-border. If the generated explosion sparks exceed the possible domain boundary, the position of sparks can be updated by the following equations:

$$\hat{X}(\hat{M}) = X_{LB,k} + |\hat{X}(\hat{M})|\,\%(X_{UB,k} - X_{LB,k})$$

$$\hat{Y}(\hat{M}) = Y_{LB,k} + |\hat{Y}(\hat{M})|\,\%(Y_{UB,k} - Y_{LB,k})$$

(28)

(6) Choose N individuals as the fireworks of the next generation by using probabilistic formula $p(x_i)$ from fireworks, explosion sparks and Gaussian mutation sparks population.

(7) Stop condition. If the stop condition is satisfied, then output the optimal results; if not, return to step (2) and continue to cycle.

2.2. Wavelet Support Vector Machine

2.2.1. Basic Theory of Support Vector Machine (SVM)

A support vector machine, proposed by Vapnik, is a kind of feed forward network [18]; its main purpose is to establish a hyper-plane to make the input vector project into another high-dimensional space. Given a set of data $\{(x_i, d_i)\}_{i=1}^{N}$, where x_i is the input vector; d is the expected output; it is further assumed that the estimate value of d is y, which is obtained by the projection of a set of nonlinear functions:

$$y = \sum_{j=0}^{m} w_j \phi_j(x) = w^T \phi(x) \tag{29}$$

where $\phi(x) = [\phi_0(x), \phi_1(x), \cdots, \phi_m(x)]^T$, $w = [w_0, w_1, \cdots, w_m]^T$; Let $\phi_0(x) = 1$; w_0 represent the bias b, and the minimization risk function can be described as follows:

$$\Phi(w, \xi, \xi') = \frac{1}{2} w^T w + C \sum_{i=1}^{N} (\xi_i + \xi_i') \tag{30}$$

The minimization risk function must satisfy the conditions:

$$\begin{cases} d_i - w^T \phi(x_i) \leqslant \varepsilon + \xi_i \\ w^T \phi(x_i) - d_i \leqslant \varepsilon + \xi_i' \\ \qquad \xi_i \geqslant 0 \\ \qquad \xi_i' \geqslant 0 \end{cases} \tag{31}$$

where $i = 1, 2, \cdots, N$, ξ_i' and ξ_i are slack variables; loss function is a ε-insensitive loss function, C is a constant.

Establish the Lagrange function and obtain:

$$\begin{aligned} J(w, \xi, \xi', \alpha, \alpha', \gamma, \gamma') \quad &= C \sum_{i=1}^{N} (\xi_i + \xi_i') + \frac{1}{2} w^T w - \sum_{i=1}^{N} \alpha_i [w^T \phi(x_i) - d_i + \varepsilon + \xi_i] \\ &- \sum_{i=1}^{N} a_i' [d_i - w^T \phi(x_i) + \varepsilon + \xi_i'] - \sum_{i=1}^{N} (\gamma_i \xi_i + \gamma_i' \xi_i') \end{aligned} \tag{32}$$

where α_i and α_i' are Lagrange multiplier; take the partial derivative of variables $w, \xi, \xi', \alpha, \alpha', \gamma, \gamma'$ and obtain:

$$\begin{aligned} w &= \sum_{i=1}^{N} (\alpha_i - \alpha_i') \phi(x_i) \\ \gamma_i &= C - \alpha_i; \ \gamma_i' = C - \alpha_i' \end{aligned} \tag{33}$$

The above problem can be converted into a dual problem:

$$\max Q\left(\alpha_i, \alpha_i'\right) = \sum_{i=1}^{N} d_i(\alpha_i - \alpha_i') - \varepsilon \sum_{i=1}^{N} (\alpha_i + \alpha_i') - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\alpha_i - \alpha_i')(\alpha_j - \alpha_j') \phi^T(x_i) \phi(x_j)$$

$$\begin{cases} \sum_{i=1}^{N} (\alpha_i - \alpha_i') = 0 \\ 0 \leqslant \alpha_i \leqslant C \\ 0 \leqslant \alpha_i' \leqslant C \\ \qquad\qquad i = 1, 2, \cdots, N \end{cases} \tag{34}$$

Solve the equation and obtain: $w = \sum\limits_{i=1}^{N} (\alpha_i - \alpha_i')\phi(x)$

Then $F(x, w) = w^T x = \sum\limits_{i=1}^{N} (\alpha_i - \alpha_i')\phi^T(x)\phi(x)$, let $K(x^T, x) = \phi^T(x)\phi(x)$ be the kernel function.

In this paper, we choose the wavelet kernel function to replace the Gaussian kernel function, and the construction of the wavelet kernel function will be introduced in detail in Section 2.2.2.

2.2.2. Construction of Wavelet Kernel Function

The kernel function $k(x, x')$ to SVM is the inner product of the image of two input space data points in the spatial characteristic. It has two important features: first, the symmetric function to inner product kernel variables is $k(x, x') = k(x', x)$; second, the sum of the kernel function on the same plane is a constant. In general, only if the kernel function satisfies the following two theorems, can it become the kernel of support vector machine [19].

Mercer Lemma

$k(x, x')$ represents a continuous symmetric kernel, which can be expanded into a series as:

$$k(x, x') = \sum_{i=1}^{\infty} \lambda_i g_i(x) g_i(x') \tag{35}$$

where λ_i is positive, in order to ensure the above expansion is absolutely uniform convergence, the sufficient and necessary condition is:

$$\iint k(x, x')g(x)g(x)dxdx' \geqslant 0, x, x' \in R^n \tag{36}$$

For all $g()$ needs to satisfy the condition: $\begin{cases} g() \neq 0 \\ \int_{R^n} g^2(\xi)d\xi < \infty \end{cases}$, $g(x_i)$ stands for the expansion of the characteristic function, λ_i stands for eigenvalue and all are positive, thus the $k(x, x')$ is positive definite.

Smola and Scholkopf Lemma

If the support vector machine's kernel function has meet the Mercer Lemma, then it only needs to prove $k(x, x')$ to satisfy the follow formula:

$$F(x)(\omega) = (2\pi)^{-n/2} \int_{R^n} \exp(-J(\omega \cdot x))k(x)dx \geqslant 0, x \in R^n \tag{37}$$

Construction of Wavelet Kernel

If the wavelet kernel satisfies the condition: $\psi(x) \in L^2(R) \cap L'(R)$ and $\hat{\psi}(x) = 0, \hat{\psi}(x)$ is the Fourier transform of $\psi(x)$, the $\psi(x)$ can be defined as [20]:

$$\psi_{\sigma,m}(x) = (\sigma)^{1/2}\psi(\frac{x-m}{\sigma}), x \in R \tag{38}$$

where σ means contraction-expansion factor, and m is horizontal floating coefficient, $\sigma > 0, m \in R$.

When the function $f(x), f(x) \in L^2(R)$, the wavelet transform to $f(x)$ can be defined as:

$$W(\sigma, m) = \sigma^{1/2} \int_{-\infty}^{+\infty} f(x)\psi^*(\frac{x-m}{\sigma})dx \tag{39}$$

where $\psi^*(x)$ is the complex conjugation of $\psi(x)$. The wavelet transform $W(\sigma, m)$ is reversible and also can be used to reconstruct the original signal, so:

$$f(x) = C_\varphi^{-1} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W(\sigma, m)\psi_{\sigma,m}(x) \frac{d\sigma}{\sigma^2} dm \tag{40}$$

Among them $C \underline{\underline{\varphi}} \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(w)|^2}{|w|} dw < \infty$, $\hat{\psi}(w) = \int \psi(x)\exp(-Jwx)dx$.

For the above equation, is a constant. The theory of wavelet decomposition is to approximate the function group by the linear combination of the wavelet function.

Assuming $\psi(x)$ is a one-dimensional function, based on tensor theory, the multi-wavelet function can be defined as:

$$\psi_l(x) = \prod_{i=1}^{l} \psi(x_i), x \in R^{l \times d}, x_i \in R^d \tag{41}$$

The horizontal floating kernel function can be built as:

$$k(x, x') = \prod_{i=1}^{l} \psi(\frac{x_i - x_i'}{\sigma_i}), \sigma_i > 0 \tag{42}$$

In support vector machines, the kernel function should satisfy the Fourier transform; therefore, only when the wavelet kernel function satisfies the Fourier transform can it be used for support vector machines. Thus, the following formula needs to be proved.

$$F[k](w) = (2\pi)^{-l/2} \int_{R^l} \exp(-J(wx))k(x)dx \geqslant 0 \tag{43}$$

In order to keep the generality of wavelet kernel function, choose the Morlet mother wavelet as follows:

$$\psi(x) = \cos(1.75x)\exp(-x^2/2) \tag{44}$$

σ can be figured out. Where $x \in R^N$, $\sigma, x_i \in R^N$, the multi-dimensional wavelet function is an allowable multidimensional support vector machine functions.

$$
\begin{aligned}
k(x, x') &= \prod_{i=1}^{l} \left\{ \cos\left[1.75(\frac{||x_i - x_i'||^2}{2\sigma^2}) \right] \exp\left[-\frac{||x_i - x_i'||^2}{2\sigma^2} \right] \right\} \\
&= \prod_{i=1}^{l} \cos(\frac{1.75X_i}{\sigma})\exp(-||x_i||^2/2\sigma^2)
\end{aligned}
\tag{45}
$$

Proof: $\int_{R^l} \exp(-Jwx)k(x)dx$

$$= \int_{R^l} \exp(-Jwx) \cdot (\prod_{i=1}^{l} \left\{ \cos[1.75(\frac{x_i - x_i'}{\sigma})]\exp(-\frac{||x_i - x_i'||^2}{2\sigma^2}) \right\})dx$$

$$= \prod_{i=1}^{l} \int_{R^l} \exp(-Jw_i x_i) \cdot (\frac{\exp(j1.75x_i/\sigma) + \exp(-1.75x_i/\sigma)}{2}) \cdot \exp(-||x_i||^2/2\sigma^2)dx_i$$

$$= \prod_{i=1}^{l} \frac{1}{2} \int_{R^l} \left\{ \exp\left[-\frac{||x_i||^2}{2\sigma^2} + (\frac{1.75j}{\sigma} - jw_i\sigma) \cdot x_i \right] + \exp\left[-\frac{||x_i||^2}{2\sigma^2} - (\frac{1.75j}{\sigma} - jw_i\sigma) \cdot x_i \right] \right\} dx_i$$

$$= \prod_{i=1}^{l} \frac{|\sigma| \sqrt{2\pi}}{2} \left\{ \exp(-\frac{(1.75 - w_i\sigma)^2}{2}) + \exp(-\frac{(1.75 + w_i\sigma)^2}{2}) \right\}$$

Then $F(x)(w) = \prod_{i=1}^{l} (\frac{|\sigma|}{2})(\exp(-\frac{(1.75 - w_i\sigma)^2}{2}) + \exp(-\frac{(1.75 + w_i\sigma)^2}{2}))$

$|\sigma| \neq 0 \Rightarrow F(x)(w) \geqslant 0 \Rightarrow$ wavelet kernel function is an allowable support vector kernel function.

2.3. Quantum Fireworks Algorithm for Parameters Selection of Wavelet Support Vector Machine (w-SVM) Model

It is extremely important to select the parameters of w-SVM which can affect the fitting and learning ability. In this paper, the constructed quantum fireworks algorithm (QFA) is used for selecting the appropriate parameters of the w-SVM model in order to improve the icing forecasting accuracy. The flowchart of QFA for parameter selection of the fireworks the w-SVM model is shown in Figure 1 and the details of the QFA-w-SVM model are shown as follows:

(1)   Initialize the parameters. Initialize the number of fireworks N, the explosion radius $A1$, the number of explosive sparks M, the mutation probability P_{id}, the maximum number of iterations *Maxgen*, the upper and lower bound of the solution space V_{up} and V_{down} respectively, and so on.

(2)   Separate the sample date as training samples and test samples, then normalize the sample data.

(3)   Initialize the solution space.

   a   In the solution space, randomly initialize N positions, that is, N fireworks. Each of the fireworks has two dimensions, that is, C and σ.

   b   Use the probability amplitude of quantum bits to encode current position of fireworks according to Equations (9)–(10).

   c   Converse the Solution space according to Equations (11)–(12).

   d   Input the training samples, use w-SVM to carry out a training simulation for each fireworks, and calculate the value of the fitness function corresponding to each of the fireworks.

(4)   Initialize the global optimal solution by using the above initialized solution space, including the global optimal phase, the global optimal position quantization of fireworks, the global best fireworks, and the global best fitness value.

(5)   Start iteration and stop the cycle when the maximum number of iterations *Maxgen* is achieved.

   a   According to the fitness value of each firework, calculate the corresponding explosion radius $R(i)$ and the number of sparks $S(i)$ generated by each explosion. The purpose of calculating $R(i)$ and $S(i)$ is to obtain the optimal fitness values, which means that if the fitness value is smaller, the explosion radius is larger, and the number of sparks generated by each explosion is bigger. Therefore, more excellent fireworks can be retained as much as possible.

   b   Generate the explosive sparks. When each fireworks explodes, carry out the solution space conversion for each explosion spark and control their spatial positions through cross border detection.

   c   Generate $\hat{M}$ Gaussian mutation sparks. Carry out a fireworks mutation operation according to Equations (26)–(27), resulting in $\hat{M}$ mutation sparks.

   d   Use the initial fireworks, explosive sparks and Gauss mutation sparks to establish a group, and the fitness value of each firework in the new group is calculated by using w-SVM.

   e   Update the global optimal phase, the global optimal position quantization of fireworks, the global best fireworks, and the global best fitness value.

   f   Use roulette algorithm to choose the next generation of fireworks, and obtain the next generation of N fireworks.

   g   Use w-SVM to carry out a training simulation for each fireworks, and update the value of the fitness function corresponding to each of the fireworks.

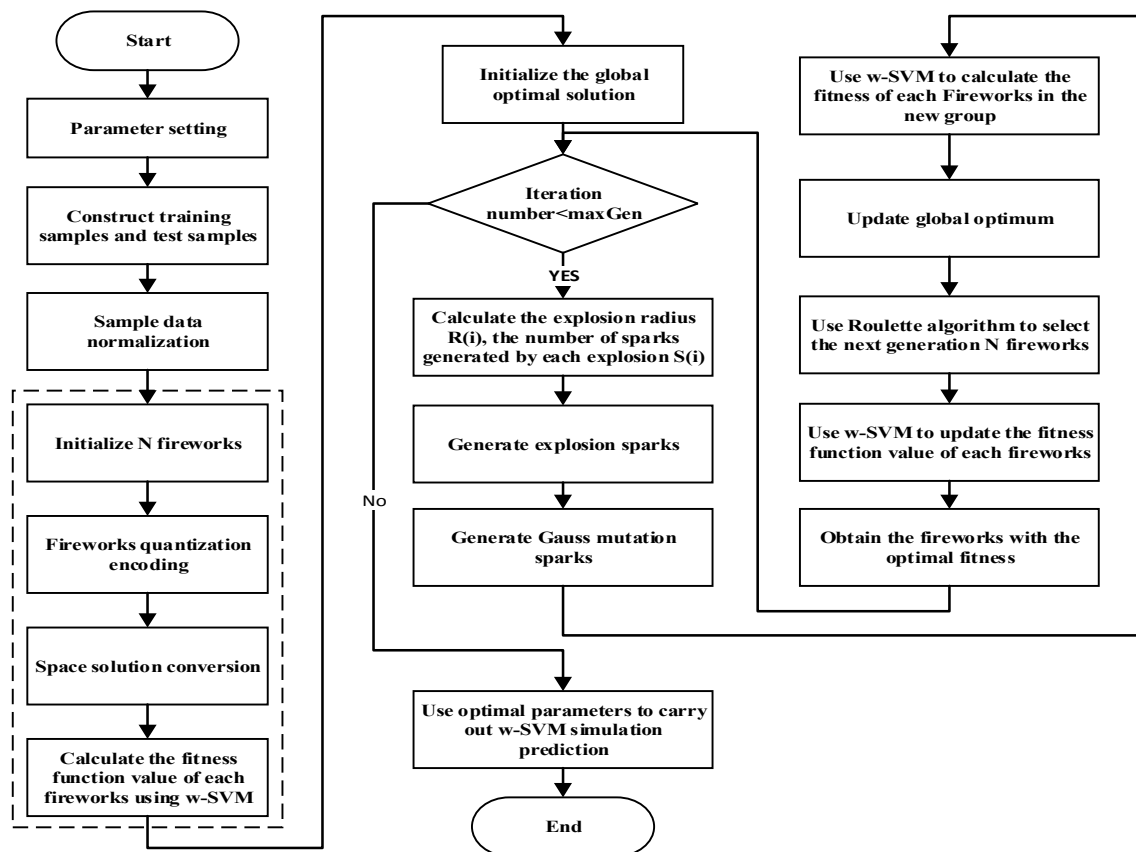(6)   After all iterations, the best fireworks can be obtained, which correspond to the best parameters of C and σ.

Figure 1. The flowchart of the quantum fireworks algorithm (QFA)-w-SVM (wavelet Support Vector Machine) model.

3. Case Study and Results Analysis

3.1. Data Selection

Transmission line icing is affected by many factors, which mainly include wind direction, light intensity, air pressure, altitude, condensation level, terrain, alignments, wires hanging height, wire stiffness, wire diameter, load current and so on. However, the necessary meteorological conditions are: (1) the relative air humidity must be above 85%; (2) the wind speed should be greater than 1 m/s; (3) the temperature needs to reach 0 °C and below. In addition, the impact factors, which have the greater correlations with the line icing, mainly include: wind direction, light intensity and air pressure, *etc*. In general, when the wind direction is parallel to the wire or the angle between the wire and the wire is less than 45, the extent of line icing is lighter; when the wind direction is vertical to the wire or the angle between the wire and the wire is more than 45, the extent of line icing is more severe. Similarly, the lower the light intensity is, the more severe the line icing is.

In this paper, three power transmission lines, named "Qianpingxian-95", "Fusha-I-xian" and "Yangtongxian" in PingXi, ChangSha and ZhaoYang of Hunan province, respectively, are selected as the case studies to demonstrate the effectiveness, feasibility and robustness of the proposed method. The data from the above mentioned three transmission lines are provided by Key Laboratory of Disaster Prevention and Mitigation of Power Transmission and Transformation Equipment (Changsha, China).

As is known, a large freezing disaster occurred in the south of China, which caused huge damage to the power grid system. Hunan province, located in the southern part of China, was one of the most serious areas affected by this disaster. During the disaster period, one third of 500 and 200 kV substations were out of action in the Hunan power grid. According to the statistics, there were 481 line breakages of 500 kV transmission lines, 673 line breakages of 200 kV transmission lines, 142 tower

collapses of 500 kV AC and DC transmission lines, 633 tower collapses of 220 kV transmission lines, and 1203 tower collapses of 110 kV transmission lines. Moreover, in the Hunan area, which suffered from the influence of topography, it is easy to form the stationary front due to the mountains block, when cold air enters the area. Therefore, it is due to their certain typicality that we select those three transmission lines in Hunan province as cases to verify the validity and robustness of the proposed method.

Case 1: the data on "Qianpingxian-95" are from 10 January 2008 to 15 February 2008, which has 221 data groups for the training and testing in total. The former 180 data groups are regarded as a training set and the last 41 groups of data are a testing set. The input vectors of SVM are average temperature, relative humidity, wind speed, wind direction, and sunlight intensity, and the output vector is ice thickness. The sample data are shown in Figure 2.

Case 2: the data of "Fusha-I-xian" are from 12 January 2008 to 25 February 2008, which has 287 data groups. The former 240 groups are taken as a training set and the remaining 47 groups as a testing set. The input vectors are same as that of *Case 1.* The sample data are shown in Figure 3.

Case 3: the data of "Yangtongxian" are from 8 January 2008 to 24 February 2008, whose data groups total up to 329. The former 289 groups are taken as a training set and the remaining 40 groups as a testing set. The input vectors are still same as that of Case 1. The sample data are shown in Figure 4.
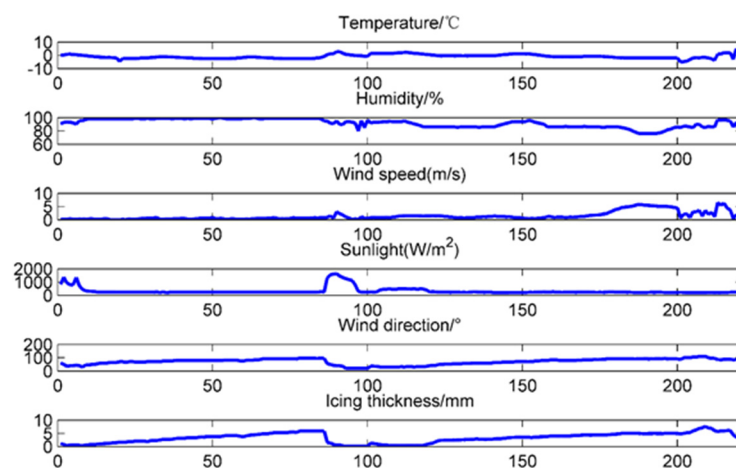
Figure 2. The original sample data charts.
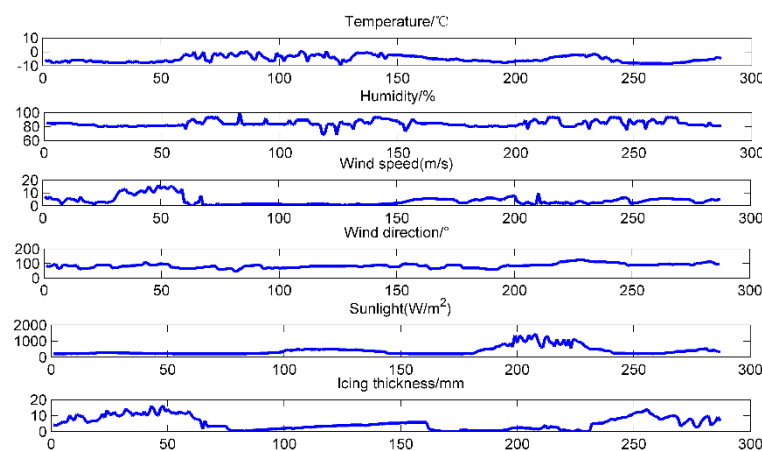
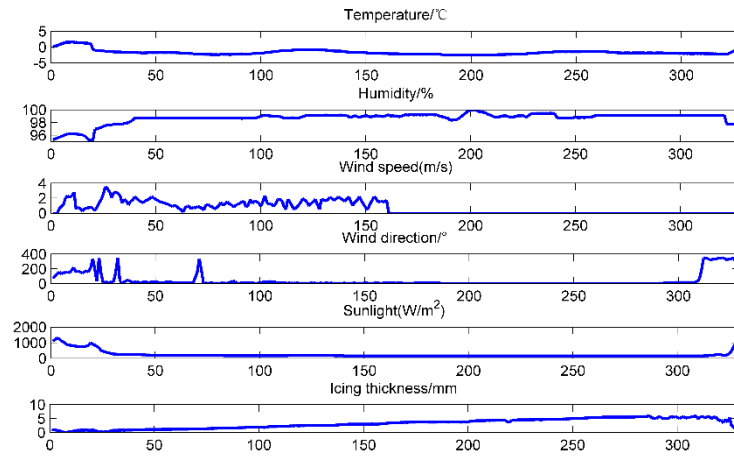**Figure 3.** Original data of "Fusha-I-xian".

Figure 4. Original data charts of "Yangtongxian".

3.2. Data Pre-Treatment

Before the calculation, the data must be screened and normalized to put them in the range of 0 to 1 using the following formula:

$$Z = \{z_i\} = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \qquad i = 1, 2, 3, ..., n \tag{46}$$

where $y_{\max}$ and $y_{\min}$ are the maximum and minimum value of sample data, respectively. The values of each data are in the range [0,1] for eliminating the dimension influence.

Furthermore, this paper will use a quantum fireworks algorithm (QFA) to optimize the parameters C, σ of the wavelet support vector machine to find optimal parameters to improve the prediction accuracy. In the parameters optimization, we adopt the mean square error (MSE) as the fitness function to realize the process of QFA, and the formula of MSE is as follows:

$$function(i) = \frac{\sqrt{\sum\limits_{i=1}^{n} (y_i - y_i')^2}}{n} \tag{47}$$

where y_i is the actual value; y_i' is the prediction value; and n is the sample number.

3.3. Case Study 1

In this case, "Qianpingxian-95", which is a 220kV high voltage line, is selected to performed the simulation. After the above preparation, apply the constructed model to verify its feasibility and robustness.

Firstly, initialize the parameters of QFA. Let the maximum iteration number $Maxgen = 200$, population number $PopNum = 40$, Sparks number determination constant $M = 100$, explosion radius determination constant $\hat{R} = 150$, the border of parameter C be $[2^{-5}\ 2^{10}]$, the border of parameter σ be $[2^{-5}\ 2^{5}]$, the upper and lower limits of searching space of fireworks individual be $V_{up} = 512$ and $V_{down} = -512$, respectively, and mutation rate $P_{id} = 0.05$. Then use the steps of QFA to optimize the parameters of w-SVM and obtain $C = 18.3516$, $\sigma = 0.031402$. Finally, we predict the icing thickness of the testing sample after putting the optimal parameters into the w-SVM regression model.

Figure 5 shows the optimization process of the quantum fireworks algorithm (QFA). As we can see, the proposed model obtains the optimal value when the iteration number is 35, and the optimal value is 0.11; this illustrates that the proposed algorithm can obtain the global optimum with a fast convergence speed. Figure 6 shows the forecasting results of the proposed method. This paper also

selects w-SVM optimized by a regular fireworks algorithm (FA-w-SVM), SVM optimized by particle swarm optimization algorithm (PSO-SVM), SVM model, and multiple linear regression model (MLR) to make a comparison, and the convergence curves of FA-w-SVM and PSO-SVM are also shown in Figure 5. The forecasting results of four models are shown in Figure 7.
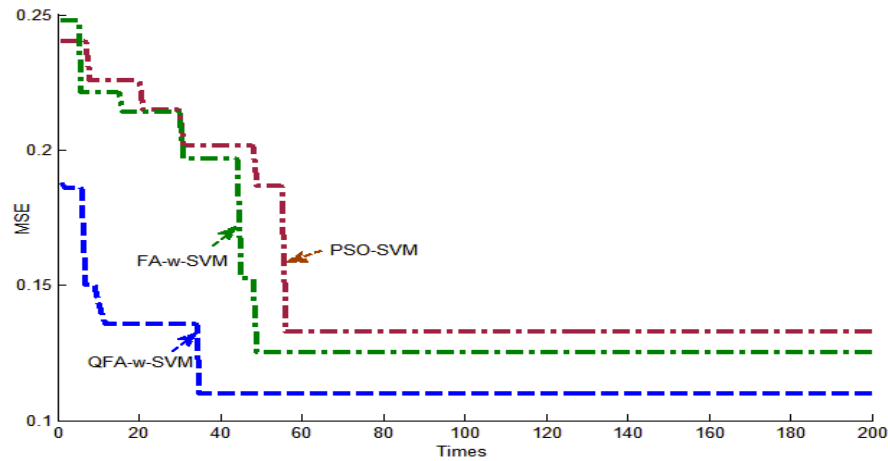

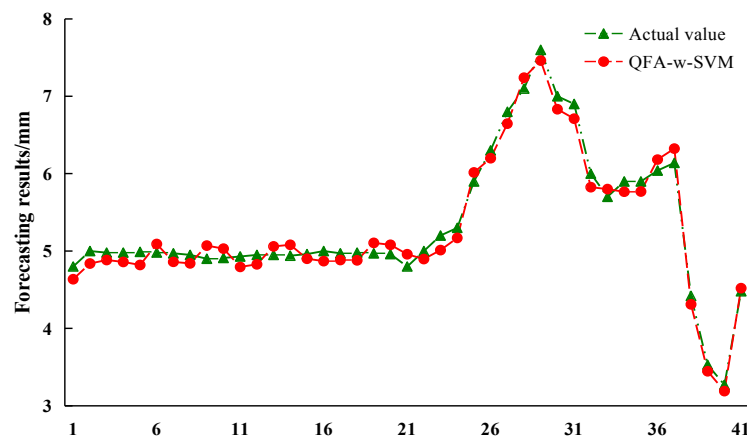**Figure 5.** The iteration process of QFA-w-SVM. PSO-SVM: particle swarm optimization algorithm.
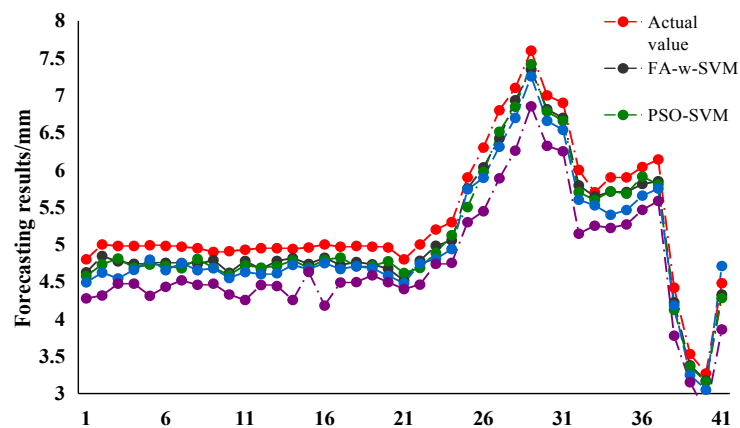

Figure 6. Forecasting value of QFA-w-SVM.

Figure 7. Forecasting value of each method. MLR: multiple linear regression model.

For further analysis, this paper will use the relative error (RE), the mean absolute percentage error (MAPE), mean square error (MSE) and average absolute error (AAE) to evaluate the results of prediction.

$$RE = \left| \frac{Y_i - Y'_i}{Y_i} \right| \times 100\% \quad MAPE = \frac{1}{n} \left| \frac{Y_i - Y'_i}{Y_i} \right| \times 100\%$$

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (Y_i - Y'_i)^2 \quad AAE = \frac{1}{n} (\sum_{i=1}^{n} |Y_i - Y'_i|) / (\frac{1}{n} \sum_{i=1}^{n} Y_i) \tag{48}$$

where Y_i is the actual value of power load; Y'_i is the forecasting value of power load; $i = 1, 2, \cdots, n$.

The relative errors (RE) value of the QFA-w-SVM, FA-w-SVM, PSO-SVM, SVM and MLR models are shown in Figure 8. It can be clearly seen that the RE curve of QFA-w-SVM is the lowest among the other four models, which demonstrates that the accuracy of the proposed algorithm is much higher than other mentioned algorithms. The RE curve of the MLR model is the highest, and this demonstrates the forecasting results based on the MLR model are not satisfied here.
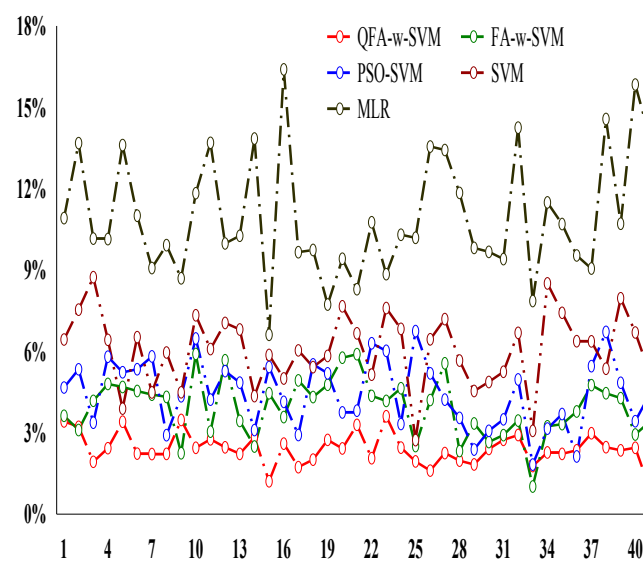
Figure 8. Forecasting errors of each model.

The relative error range [−3%,+3%] is always regarded as a standard to evaluate the performance of a forecasting model. As we can see from Figure 7, 35 points of QFA-w-SVM means 85% forcasting points are in the range [−3%,+3%], and only 6 forecasting points are out of this scope. In the model of FA-w-SVM, only 8 out of 41 points are in the range [−3%,+3%], which means that 78% of forecasting points are out of the scope. In addition, only 5 forecasting points of PSO-SVM and one forecasting point of SVM are in the scope of [−3%,+3%]; none of the forecasting points of the MLR model are in the scope. These results demonstrate that the QFA-w-SVM model has a better performance in icing forecasting when compared with other models. In addition, the maximum and minimum relative errors (*Max*RE and *Min*RE) can also reflect the forecasting accuracy of icing forecasting models. Firstly, both the *Max*RE and *Min*RE values of combined models are smaller than those of single models, which proves the optimization algorithm can facilitate improving accuracy of w-SVM through finding optimal parameters. Secondly, the *Max*RE and *Min*RE values of QFA-w-SVM model are 3.61% and 0.89%, respectively, both of which are the smallest among the five icing forecasting models, which illustrates that QFA-w-SVM has a better nonlinear fitting ability in sample training. Thirdly, in the FA-w-SVM model, the *Max*RE and *Min*RE values are 5.91% and 1.02%, respectively, both of which are smaller than that of PSO-SVM model(6.75% and 1.82%), which means FA-w-SVM has better forecasting accuracy. Finally, both the *Max*RE and *Min*RE values of SVM are 8.73% and 2.73%, respectively, both of

which are smaller than that of MLR, that mean the nonlinear processing ability is better than that of MLR; in other words, the robustness of SVM is much stronger than that of the MLR model.

The MAPE, MSE and AAE values of the above mentioned five models are shown in Figure 9. The experimental results show that the forecasting effect of the SVM model obviously precedes the multiple linear regression model; this illustrates that the intelligent forecasting model has strong learning ability and nonlinear mapping ability which the multiple linear regression model cannot match. In addition, the MAPE value of single SVM is 4.988%, which is much higher than that obtained by hybrid QFA-w-SVM, FA-w-SVM and PSO-SVM models (which are 1.56%, 2.83% and 2.776%, respectively); this result proves that the optimization algorithm can help single SVM to find the optimal parameters, and heighten the learning capacity and forecasting accuracy. Furthermore, the MAPEs of QFA-w-SVM and FA-w-SVM are 1.56% and 2.83%, respectively; this shows that the QFA model can greatly improve the optimization performance of regular FA model through the combination of quantum optimization algorithm, which makes the algorithm can easily get the optimal results. Meanwhile, the MAPE value of PSO-SVM model is 2.776%, which is higher than that of QFA-w-SVM; this illustrates that QFA-w-SVM model has a better global convergence performance compared with PSO-SVM, and also proves that w-SVM has better nonlinear mapping capability compared with SVM.
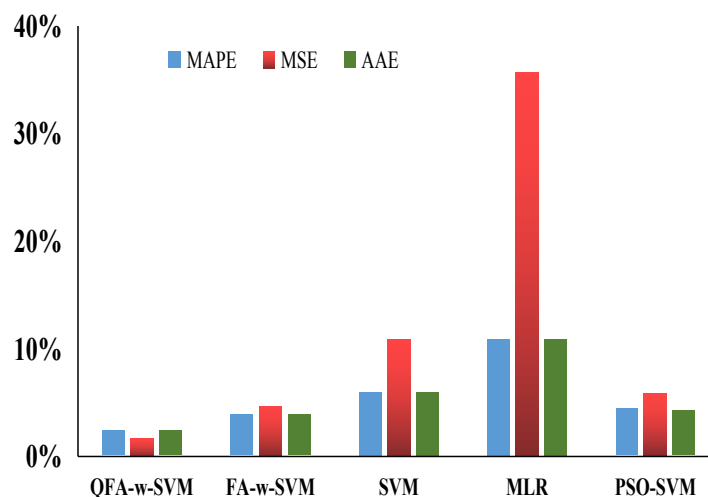

Figure 9. Values of mean absolute percentage error (MAPE), mean square error (MSE) and average absolute error (AAE).

Furthermore, the MSE values of QFA-w-SVM, FA-w-SVM, PSO-SVM, SVM and MLR models are 0.0174, 0.0466, 0.0590, 0.1094 and 0.3581, respectively, and the AAE values of those five models are 0.024, 0.0389, 0.0439, 0.0602 and 0.1092, respectively. As we know, when the value of MSE and AAE is smaller, the prediction effect of model is more ideal. It can be clearly seen that the MSE value and AAE value of QFA-w-SVM is the smallest among other models, which directly demonstrates the feasibility and effectiveness of QFA-w-SVM model, and QFA can improve the optimization performance of the regular FA to help SVM find the optimal parameters to improve the forecasting accuracy.

3.4. Case Study 2

"Fusha-I-xian", which is a representative 500 kV transmission line between Fuxing and Shapin of Hunan province, is chosen to prove the robustness and stability of the proposed QFA-w-SVM icing forecasting model. The sample data of "Fusha-I-xian" are also predicted by the five models and the results of five models are also used to make a comparison.

Figure 10 shows the iteration trend of the QFA-w-SVM searching of optimization parameters. As can be seen from Figure 3, the convergence can be seen in generation 39 with the optimal MSE value of 0.1124, and the parameters of w-SVM are obtained with $C = 23.68$, $\sigma = 0.1526$, respectively.

In FA-w-SVM model, the convergence can be seen in generation 53 with the optimal MSE value of 0.1307, and the parameters of w-SVM are obtained with $C = 31.58$, $\sigma = 0.0496$. In the PSO-SVM model, the convergence can be seen in generation 58 with the optimal MSE value of 0.1324, and the parameters of SVM are obtained with $C = 19.48$, $\sigma = 0.5962$. This proves that the proposed QFA-w-SVM model can find the global optimal value with the faster convergence compared with FA-w-SVM and PSO-SVM models. This also validates the stability of the proposed icing forecasting model.
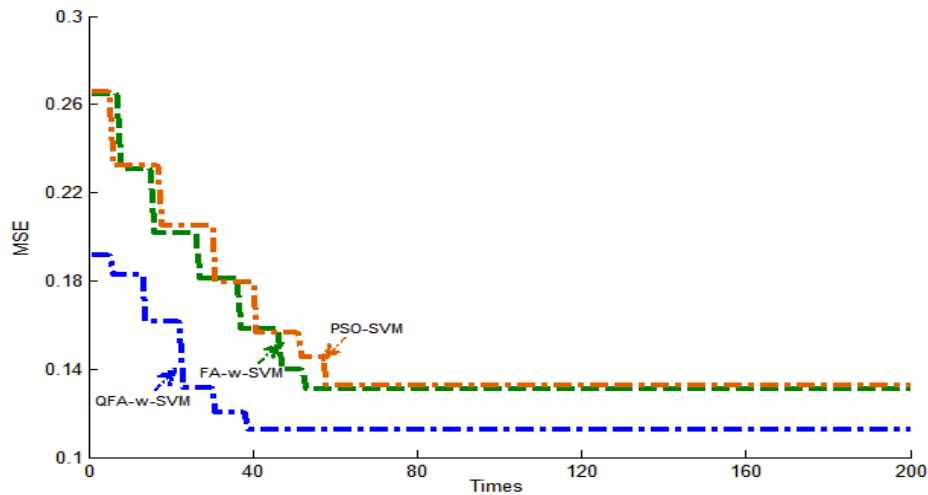
Figure 10. The convergence curves of algorithms.

Figure 11 shows the forecasting results of the five models. The forecasting errors of the five models are shown in Figure 12. It can be seen from the comparison of the forecasting curve and the actual value that the forecasting results of ice thickness of those five models have approximation to actual curve. Among them, the proposed QFA-w-SVM has the best fitting learning and fitting ability, and 44 forecasting points mean that almost 94% of points are in the scope of [−3,+3]; only three forecasting points are out of this scope. In FA-w-SVM model, 27 forecasting points means that nearly 57% points are in the range [−3,+3], and 20 points fall out this scope. In PSO-SVM model, 20 forecasting points means that nearly 43% of points are in the scope, and 27 points are out of the range. In SVM model, only 2 points are in this range, and the rest of the forecasting points are not in this scope. In MLR model, none of points are in this scope. This reveals that QFA-w-SVM model has a higher forecasting accuracy than other mentioned models, and it also has stronger robustness and nonlinear fitting ability.
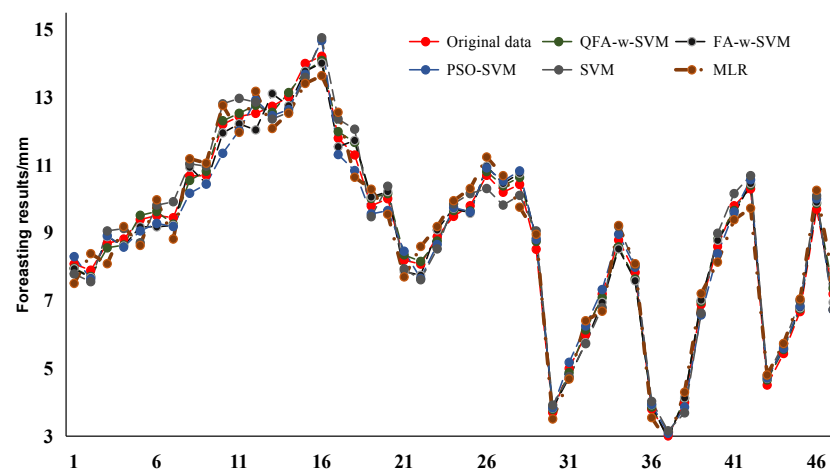

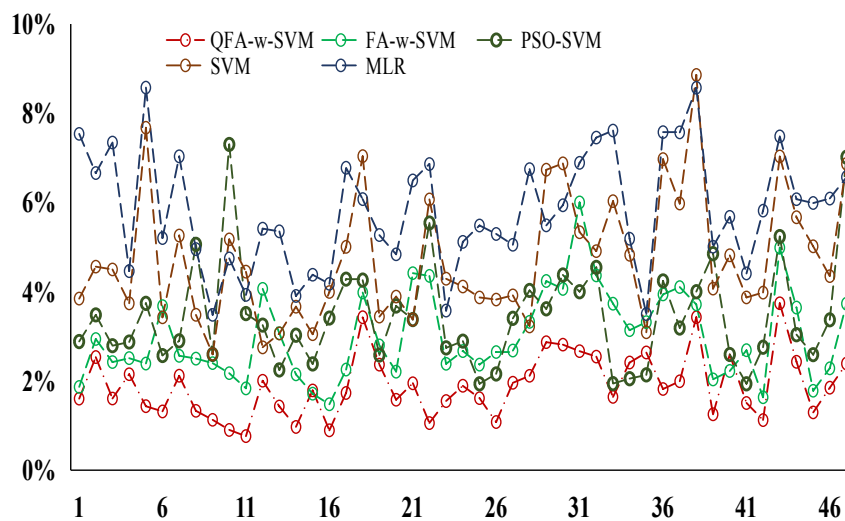**Figure 11.** Forecasting results of the five models.

Figure 12. Forecasting errors of each model.

The values of MAPE, MSE, AAE of the five models in icing forecasting of "Fusha-I-xian" are shown in Figure 13. It can be seen that the proposed QFA-w-SVM model still has the smallest MAPE, MSE and AAE values, which are 1.9%, 0.026 and 0.018, respectively. This again reveals that the proposed QFA-w-SVM model has the best performance in the icing forecasting results. In addition, the MAPE, MSE and AAE values of FA-w-SVM are 3.01%, 0.063 and 0.028, respectively, which are both larger than that of QFA-w-SVM, but smaller than the rest of three models. In PSO-SVM, the MAPE, MSE and AAE values are 3.45%, 0.099 and 0.034, respectively, which are smaller than those of SVM and MLR models. That proves that the combined algorithms have better forecasting performance and the optimization algorithms can help single regression model to achieve a better accuracy through finding better parameters. This result agrees with the one presented in Section 3.3.
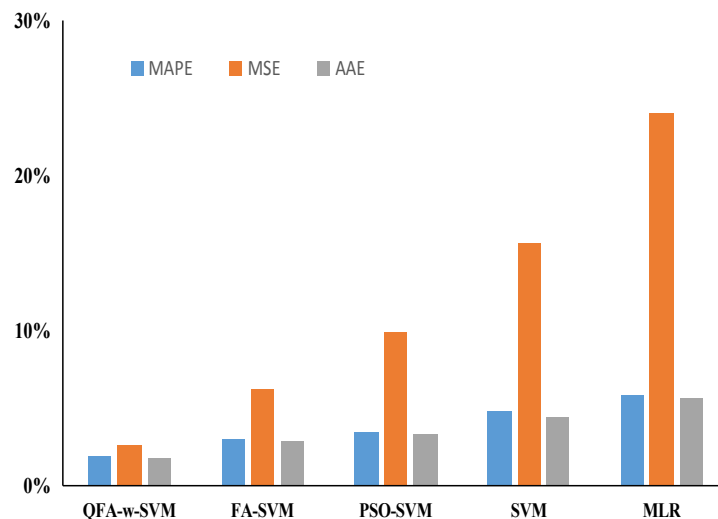

Figure 13. Values of MAPE, MSE, AAE.

3.5. Case Study 3

In this case, the icing data from a 110 kV transmission line called "Yangtongxian" is selected to validate the robustness and stability of the proposed QFA-w-SVM icing forecasting model. Similarly, the five models are still used to make a comparison for prediction results.

Figure 14 shows the convergence plots of QFA-w-SVM, FA-w-SVM and PSO-SVM. As shown in the figure, QFA-w-SVM obtains the optimal value after 40 iterations while FA-w-SVM and PSO-SVM models converge to global solutions after 53 and 57 iterations, respectively. In other words, the QFA-w-SVM still has the fastest convergence speed and stronger global optimization ability compared with other models.
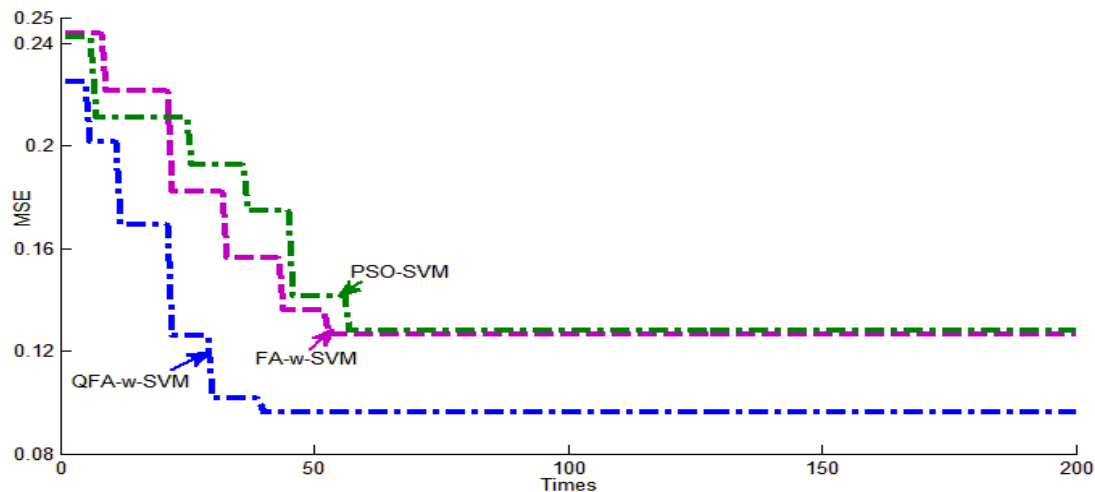
Figure 14. The convergence curves of algorithms.

The forecasting results and relative errors are shown in Figures 15 and 16 respectively. From the figures, it is obvious that the proposed QFA-w-SVM model still has better accuracy in icing forecasting, which demonstrates that the proposed model is superior for solving the icing forecasting problem compared to other algorithms, because it can achieve global optima in fewer iterations and is a critical factor in convergence process of algorithms.
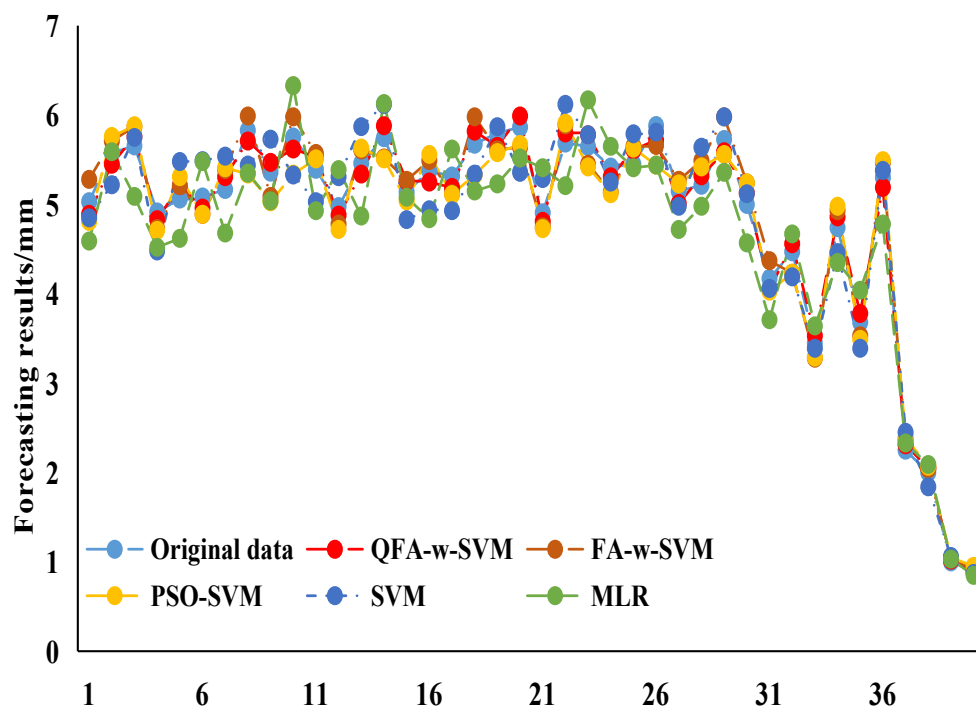
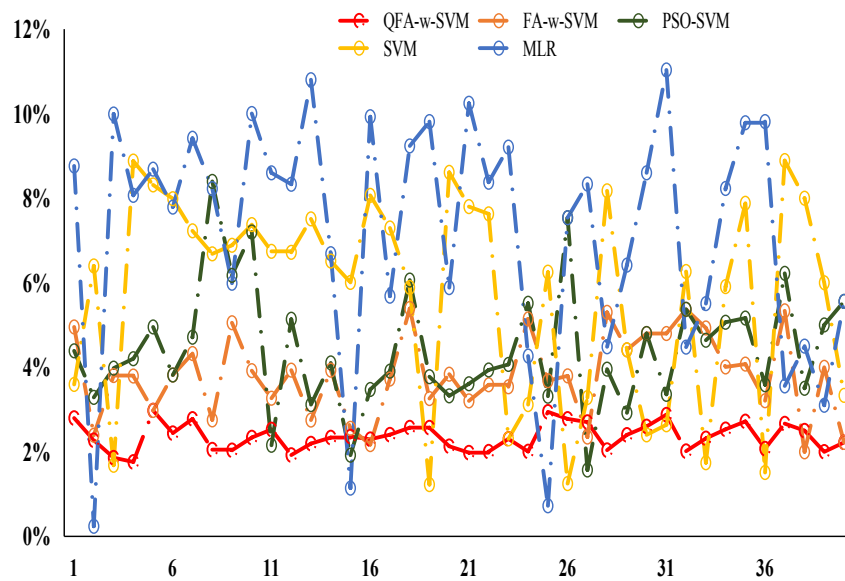**Figure 15.** Forecasting results of each model.

Figure 16. Forecasting errors of each model.

Figure 17 shows the MAPE, MSE, and AAE values of the five models; it is clear that the values of MAPE, MSE, AAE of the proposed QFA-w-SVM model are the smallest among the five models, which again demonstrates that the proposed model performs with great robustness and stability.
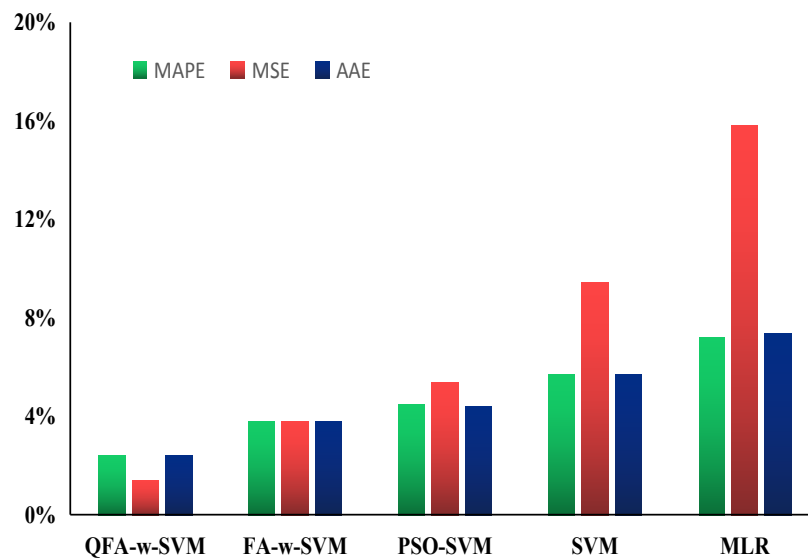

Figure 17. MAPE, MSE, AAE values of each model.

In summary, the proposed QFA-w-SVM model can greatly close the gap between forecasting values and original data, which means it outperforms the FA-w-SVM, PSO-SVM, SVM, and MLR models in icing forecasting with different voltage transmission lines. Moreover, the QFA-w-SVM which uses QFA to select the parameters of the SVM model, can effectively make a great improvement for icing forecasting accuracy. It might be a promising alternative for icing thickness forecasting.

4. Conclusions

For better forecasting accuracy for icing thickness, this paper proposes an intelligent method for a wavelet support vector machine (w-SVM) based on a quantum fireworks optimization algorithm

(QFA). Firstly, the regular fireworks optimization algorithm is improved through combination with a quantum optimization algorithm and proposes the quantum fireworks optimization algorithm (QFA); the steps of QFA are listed in detail. Secondly, in order to make use of the advantages of kernel function of SVM, the wavelet kernel function is applied to SVM instead of a Gaussian kernel function. Finally, this paper uses the proposed QFA model to optimize the parameters of w-SVM, and builds the icing thickness forecasting model of QFA-w-SVM. During the application of icing thickness forecasting, this paper also gives full consideration to the impact factors and selects the temperature, humidity, wind speed, wind direction and sunlight as the main impact factors. Through the numerical calculations of a self-written program, the empirical results show that the proposed forecasting model of QFA-w-SVM has great robustness and stability in icing thickness forecasting and is effective and feasible.

Author Contributions: Drafting of the manuscript: Tiannan Ma and Dongxiao Niu; Implementation of numerical simulations and preparations of figures: Tiannan Ma and Ming Fu; Finalizing the manuscript: Tiannan Ma. Planning and supervision of the study: Dongxiao Niu and Tiannan Ma.

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

Reference

1. Yin, S.; Lu, Y.; Wang, Z.; Li, P.; Xu, K.J. Icing Thickness forecasting of overhead transmission line under rough weather based on CACA-WNN. *Electr. Power Sci. Eng.* **2012**, *28*, 28–31.
2. Jia-Zheng, L.U.; Peng, J.W.; Zhang, H.X.; Li, B.; Fang, Z. Icing meteorological genetic analysis of hunan power grid in 2008. *Electr. Power Constr.* **2009**, *30*, 29–32.
3. Goodwin, E.J.I.; Mozer, J.D.; Digioia, A.M.J.; Power, B.A. Predicting ice and snow loads for transmission line design. **1983**, *83*, 267–276.
4. Makkonen, L. Modeling power line icing in freezing precipitation. *Atmos. Res.* **1998**, *46*, 131–142. [CrossRef]
5. Liao, Y.F.; Duan, L.J. Study on estimation model of wire icing thickness in hunan province. *Trans. Atmos. Sci.* **2010**, *33*, 395–400.
6. Sheng, C.; Dong, D.; Xiaotin, H.; Muxia, S. Short-term Prediction for Transmission Lines Icing Based on BP Neural Network. In Proceedings of the IEEE 2012 Asia-Pacific Power and Energy Engineering Conference (APPEEC), Sanya, China, 31 December 2012–2 January 2013; pp. 1–5.
7. Liu, J.; Li, A.-J.; Zhao, L.-P. A prediction model of ice thickness based on T-S fuzzy neural networks. *Hunan Electr. Power* **2012**, *32*, 1–4.
8. Du, X.; Zheng, Z.; Tan, S.; Wang, J. The study on the prediction method of ice thickness of transmission line based on the combination of GA and BP neural network. In Proceedings of the 2010 International Conference on IEEE E-Product E-Service and E-Entertainment (ICEEE), Henan, China, 7–9 November 2010; pp. 1–4.
9. Zarnani, A.; Musilek, P.; Shi, X.; Ke, X.D.; He, H.; Greiner, R. Learning to predict ice accretion on electric power lines. *Eng. Appl. Artif. Intell.* **2012**, *25*, 609–617. [CrossRef]
10. Huang, X.-T.; Xu, J.-H.; Yang, C.-S.; Wang, J.; Xie, J.-J. Transmission line icing prediction based on data driven algorithm and LS-SVM. *Autom. Electr. Power Syst.* **2014**, *38*, 81–86.
11. Li, Q.; Li, P.; Zhang, Q.; Ren, W.P.; Cao, M.; Gao, S.F. Icing load prediction for overhead power lines based on SVM. In Proceedings of the 2011 International Conference on IEEE Modelling, Identification and Control (ICMIC), Shanghai, China, 26–29 June 2011; pp. 104–108.
12. Li, J.; Dong, H. Modeling of chaotic systems using wavelet kernel partial least squares regression method. *Acta Phys. Sin.* **2008**, *57*, 4756–4765.
13. Wu, Q. Hybrid model based on wavelet support vector machine and modified genetic algorithm penalizing Gaussian noises for power load forecasts. *Expert Syst. Appl.* **2011**, *38*, 379–385. [CrossRef]
14. Liao, R.J.; Zheng, H.B.; Grzybowski, S.; Yang, L.J. Particle swarm optimization-least squares support vector regression based forecasting model on dissolved gases in oil-filled power transformers. *Electr. Power Syst. Res.* **2011**, *81*, 2074–2080. [CrossRef]
15. Dos Santosa, G.S.; Justi Luvizottob, L.G.; Marianib, V.C.; Dos Santos, L.C. Least squares support vector machines with tuning based on chaotic differential evolution approach applied to the identification of a thermal process. *Expert Syst. Appl.* **2012**, *39*, 4805–4812. [CrossRef]

16. Tan, Y.; YU, C.; Zheng, S.; Ding, K. Introduction to fireworks algorithm. *Int. J. Swarm Intell. Res.* **2013**, *4*, 39–70. [CrossRef]

17. Gao, J.; Wang, J. A hybrid quantum-inspired immune algorithm for multiobjective optimization. *Appl. Math. Comput.* **2011**, *217*, 4754–4770. [CrossRef]

18. Wang, J.; Li, L.; Niu, D.; Tan, Z. An annual load forecasting model based on support vector regression with differential evolution algorithm. *Appl. Energy* **2012**, *94*, 65–70. [CrossRef]

19. Dai, H.; Zhang, B.; Wang, W. A multiwavelet support vector regression method for efficient reliability assessment. *Reliab. Eng. Syst. Saf.* **2015**, *136*, 132–139. [CrossRef]

20. Kang, J.; Tang, L.W.; Zuo, X.Z.; Li, H.; Zhang, X.H. Data prediction and fusion in a sensor network based on grey wavelet kernel partial least squares. *J. Vib. Shock* **2011**, *30*, 144–149.