Research Article

The Improved Least Square Support Vector Machine Based on Wolf Pack Algorithm and Data Inconsistency Rate for Cost Prediction of Substation Projects

Haichao Wang,[1] Yi Liang ⓘ,[2,3] Wei Ding,[1] Dongxiao Niu,[4] Si Li,[4] and Fenghua Wang[5]

[1]*Long Yuan (Beijing) Wind Power Engineering & Consulting Co., Ltd., Beijing 100034, China*
[2]*School of Management, Hebei Geo University, Shijiazhuang 050031, China*
[3]*Strategy and Management Base of Mineral Resources in Hebei Province, Hebei GEO University, Shijiazhuang 050031, China*
[4]*School of Economics and Management, North China Electric Power University, Beijing 102206, China*
[5]*State Grid Zhejiang Electric Power Company, Hangzhou 310007, China*

Correspondence should be addressed to Yi Liang; louisliang@hgu.edu.cn

Accurate and stable cost forecasting of substation projects is of great significance to ensure the economic construction and sustainable operation of power engineering projects. In this paper, a forecasting model based on the improved least squares support vector machine (ILSSVM) optimized by wolf pack algorithm (WPA) is proposed to improve the accuracy and stability of the cost forecasting of substation projects. Firstly, the optimal features are selected through the data inconsistency rate (DIR), which helps reduce redundant input vectors. Secondly, the wolf pack algorithm is used to optimize the parameters of the improved least square support vector machine. Lastly, the cost forecasting method of WPA-DIR-ILSSVM is established. In this paper, 88 substation projects in different regions from 2015 to 2017 are chosen to conduct the training tests to verify the validity of the model. The results indicate that the new hybrid WPA-DIR-ILSSVM model presents better accuracy, robustness, and generality in cost forecasting of substation projects.

1. Introduction

Poor control over the cost of substation projects easily leads to the high cost, which seriously affects the economics and sustainability of power engineering projects [1]. The forecasting of cost level is an important part of the cost control of substation projects, and it also has important guiding significance for the cost saving of substation projects. However, the attributes of historical data indicators are numerous due to the influence of factors such as the overall planning of the power grid, total capacity, topographical features, design and construction level, and the comprehensive economic level of the construction area. Simultaneously, the number of construction projects in the same period is limited, and it is impossible to collect more comparable engineering projects in a short time, which leads to less sample data and higher

difficulty for the cost forecasting of substation projects [2]. Therefore, the construction of cost forecasting model and the realization of accurate cost forecasting results of substation projects are of great significance to the sustainability of power engineering investment.

At present, few scholars have studied the cost forecasting of substation projects, but many have studied the cost forecasting of other engineering projects. The forecasting methods are mainly divided into two categories: one is the traditional forecasting method and the other is the modern intelligent forecasting method. The traditional forecasting methods mainly include time series prediction [3], regression analysis [4], Bayesian model [5], and fuzzy prediction [6]. A time series method for cost forecasting of projects based on the bill of quantities pricing model was built in [3], which could control the forecasting error within 5%. In [4], a

forecasting model based on the integral linear regression of multiple structures according to the features of building project cost was established, and the principal component factor method was introduced to solve the problem of multicollinearity among variables. In [5], a Bayesian project cost forecasting model that adaptively integrates preproject cost risk assessment and actual performance data into a range of possible project costs at a chosen confidence level is proposed. Gao et al. [6] utilized the significant cost theory to determine the significant cost factors affecting the cost and adopted the fuzzy c-means clustering method to predict the project cost. Basic theory and the verification way of this method are relatively mature and perfect, and the calculation process is also relatively simple. However, the method is often suitable for a single object, and the forecasting accuracy is not ideal.

Therefore, in the background of the rapid development of artificial intelligence technology, the use of intelligent forecasting methods to predict the cost of substation projects has a more important significance. Intelligent forecasting methods mainly include artificial neural networks (ANNs) and support vector machines (SVMs) [7]. The artificial neural networks, such as back propagation neural network (BPNN), radial basis function neural network (RBFNN), and general regression neural network (GRNN), are mainly used in the forecasting field [8]. A combination forecasting method for power grid engineering technical transformation projects was proposed in [9], which was based on Data Envelopment Analysis (DEA) and genetic BP neural network model. These factors were used as input variables of BP neural network to establish a 3-layer structure, and the forecasting model was proved by training the actual engineering data. In [10], a cost forecasting model of engineering projects based on BPNN and RBFNN was constructed, in which the radial basis function (RBF) neural network was proved to have higher forecasting accuracy on the basis of the training test with sample data from 55 engineering projects. In [11], a forecasting model for transmission line ice coating based on generalized regression neural network and fruit fly optimization algorithm is proposed and good forecasting results are achieved. However, we can find that the artificial neural network has the disadvantage of slow convergence rate and it is easy to fall into local optimization through the above research. As a result, the forecasting accuracy is greatly reduced. In [12], it is proposed that the SVM model could avoid the neural network structure selection and local minima problems. Therefore, the SVM model has been widely used in the research of projects cost forecasting. The influencing factors of substation projects cost were determined through literature review and other methods and the cost forecasting model of substation projects was established based on the theory of SVM in [13]. In [14], parameters of the SVM forecasting model were optimized by the improved firefly algorithm, which achieved a higher accuracy of the forecasting effect. Compared with the neural network, a better forecasting performance of substation projects cost can be obtained by applying the SVM model. According to the research mentioned above, it can be noted that there are some shortcomings in the

traditional support vector machine. When the traditional support vector machine is used for cost forecasting, the solution process needs to be converted into a quadratic programming process through the kernel function, which reduces the efficiency and the convergence precision is not high. However, the least squares support vector machine (LSSVM) avoids the quadratic programming process by using the least squares linear system as the loss function. Meanwhile, the forecasting problem is transformed into the equation sets and the inequality constraints are transformed into equality constraints by means of using the kernel function, which increases the accuracy and speed of the forecasting [15, 16]. Therefore, some scholars have tried to use LSSVM to carry out the cost forecasting of engineering projects [17, 18]. In [17], the variable importance in projection (VIP) was used to extract the features of multiple factors that affect the project cost. And, the LSSVM was used as the nonlinear approximator to establish the cost forecasting model of engineering projects. Liu [18] proposed a cost forecasting model of projects based on chaotic theory and least squares support vector machine, which aimed at the time-varying and chaotic of the cost change.

Although LSSVM shows the better forecasting performance than SVM in cost forecasting, there is still the problem of blind selection of the penalty parameter and the kernel parameter [19]. Therefore, it is necessary to select appropriate intelligent algorithms to optimize the parameters. The common intelligent algorithms mainly include genetic algorithm [20], ant colony algorithm [21], fruit fly optimization algorithm [22], and particle swarm algorithm [23]. Although the above algorithms all have their own advantages, there are still some corresponding flaws. The genetic algorithm has the problems of easy precocity, complicated calculation, small processing scale, difficult to deal with nonlinear constraints, poor stability, and so on. The ant colony algorithm and the firefly algorithm cannot guarantee convergence to the best point, and it is easy to fall into a local optimum, leading to a decrease in the forecasting accuracy. And, the particle swarm algorithm cannot fully satisfy the optimization of parameters in the LSSVM due to the insufficient local search accuracy. Based on the above analysis, the wolf pack algorithm (WPA) was applied to optimize the parameters of LSSVM in this paper. The performance of WPA will not be affected by a small change in parameters and the selection of parameters is relatively easy, so it has a good global convergence and computational robustness, which is suitable for solving the high-dimensional, multipeak complex functions, especially [24].

Otherwise, the number of influencing factors of substation projects cost is very large. If all the influencing factors are used as the input indicators of the forecasting model, then a large amount of redundant data appears, so feature selection is also important. The feature selection refers to the identification and selection of appropriate input vectors in the forecasting model to reduce redundant data and improve the computational efficiency [25]. The DIR model refers that the feature set is divided into several subsets, and the minimum inconsistency rate is calculated by the feature

subsets to determine the optimal feature subset and complete the feature selection [26]. Both [27, 28] adopted the DIR model for feature selection and obtained effective forecasting results. The use of the DIR model for feature selection can eliminate redundant features based on the data set inconsistency, and the characteristics of the correlation between features are also taken into account at the same time. The selected optimal feature can represent all data information perfectly because the relationship between features is not ignored. Therefore, the data inconsistency rate (DIR) was chosen for the feature selection in this paper.

According to the above research, an ILSSVM model integrated DIR with WPA is proposed. It is the first time to combine these three models in cost forecasting and several comparing methods are utilized to validate the effectiveness of the proposed hybrid model. The paper is organized as follows: Section 2 introduces the implementation process of DIR and ILSSVM optimized by FOA. Section 3 provides a case to validate the proposed model. Section 4 obtains the conclusion in this paper.

There are three innovations in this paper: (1) DIR is applied to the selection of influencing factors of substation project cost prediction; (2) LSSVM model is improved by applying the wavelet kernel function to replace the traditional radial basis kernel function; (3) LSSVM parameters are optimized by WPA.

2. Basic Theory

2.1. Wolf Pack Algorithm. Wolf pack algorithm (WPA) is derived from the study of wolf hunting behaviors. During the hunting process, different divisions of labor are assigned to wolves according to their respective roles. Wolves in the pack are divided into three types: the leader wolf, the scout wolf, and the ferocious wolf. And, these three wolves work together to complete the hunting process. The wandering behavior, summoning behavior and siege behavior of WPA are the three main behaviors, which are bionic based on the behavioral features of the wolves in the hunting process. Simultaneously, the generation of leader wolf and the replacement of wolf pack are followed by the rules of "survival of the fittest" and "winner-take-all," respectively [29]. The optimal solution is obtained through continuous iterations of the bionic behaviors. The bionic model of wolf pack algorithm is as shown in Figure 1.

The principle and steps of wolf pack algorithm are as follows:

(1) *Initialization of the Wolf Pack.* Suppose that there are N artificial wolves in the D-dimensional space and the location of the ith wolf is shown as follows:

$$X_i = (x_{i1}, x_{i2}, \ldots, x_{id}), \quad 1 \leq i \leq N, 1 \leq d \leq D. \quad (1)$$

The initial position is generated by the following equation:

$$x_{id} = x_{\min} + \text{rand} \times (x_{\max} - x_{\min}), \quad (2)$$

where rand is a uniformly distributed random number in [0, 1] and $x_{\max}$ and $x_{\min}$ are the upper and lower limits of the search space, respectively.

(2) *Generation of the Leader Wolf.* The wolf at the optimal position of the target function value is chosen as the leader wolf. The leader wolf neither updates the position of the hunting process nor participates in hunting activities, but directly iterates. If $Y_{\text{lead}} < Y_i$, then $Y_{\text{lead}} = Y_i$, namely, the scout wolf becomes a leader wolf at this time. Conversely, if the scout wolf i migrates towards h directions until the maximum H is obtained or location cannot be further optimized, the search stops at this moment. Among the h positions searched by the wolf i, the position of the jth point at the dth dimension is shown as follows:

$$y_{ijd} = y_{id} + \text{rand} \times \text{step}_a. \quad (3)$$

(3) *Close to the Prey.* The location update of the wolf pack is promoted by the summoning behavior of the leader wolf. Driven by the summoning behavior, the new location is obtained by wolves based on the summons of the leader wolf. The updated position of wolf i at the dth dimension is shown as follows:

$$z_{id} = x_{id} + \text{rand} \times \text{step}_b \times (x_{id} - x_{lid}), \quad (4)$$

where step_a is the length of the wolf's stride in the search process, step_b is the length of stride when the wolf moves to the target, x_{id} is the current position of the wolf i at the dth dimension, and x_{lid} is the position of the leader wolf at the dth dimension.

(4) *Encirclement of the Prey.* After discovering the prey, the surrounding wolves will complete the reclamation of the target prey according to the signals issued by the leader wolf. Equations of the reclamation and the reclamation steps are as follows:

$$X_i^{t+1} = \begin{cases} X_i^t, & r_m < \theta, \\ X_i + \text{rand} \times ra, & r_m > \theta, \end{cases} \quad \text{ra}(t) = \text{ra}_{\min}$$
$$\times (x_{\max} - x_{\min}) \times e^{\left((\ln(\text{ra}_{\min}/\text{ra}_{\max}))/\max t\right)}, \quad (5)$$

where t is the number of iterations, ra is the length of the wolf's stride during the reclamation, X_i is the position of the leader that issued the signals, and X_i^t is the position of the wolf i in the tth iteration.

(5) *Update Mechanism of Wolves Competition.* Wolves will be eliminated without food in the process of reclamation hunting; instead, the wolves that survived are the first to obtain food. Weak wolves that without food will be eliminated while an equal
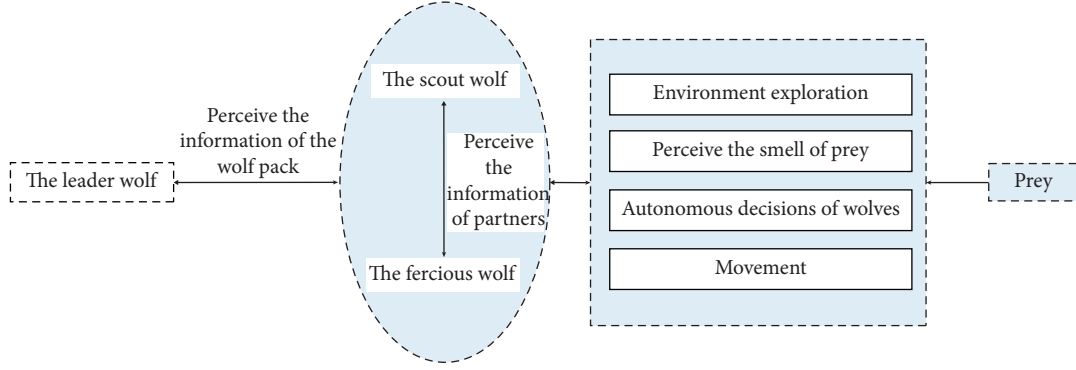
FIGURE 1: Bionic graph of Wolf pack algorithm.

number of new wolves that are randomly generated will be added into the wolf pack.

2.2. Data Inconsistency Rate (DIR).

The feature selection of many historical cost data of substation projects takes aim at distinguishing the most relevant data features, which makes the input vector of the power projects cost forecasting model with a strong pertinence and reduces the redundancy of the input information in order to improve the accuracy of the cost forecasting of the substation projects. The inconsistency rate of data can accurately describe the discrete characteristics of the input features. Different feature characteristics can be obtained by different division modes and different division modes can obtain different frequency distributions. The discrimination ability of data categories can be distinguished by the inconsistency rate of calculation. The higher the inconsistency rate of data, the stronger the classification ability of feature vectors.

It is necessary to know the specific calculation formula of the inconsistency rate for selecting the features by means of the inconsistency rate method. Therefore, it is assumed that the collected cost data of substation projects have g features, such as main transformer capacity, floor space, and main transformer unit price. which are, respectively, represented by the values of $G1, G2, \ldots, Gg$. Γ is the feature set and L is a subset of the feature set. Next, set the standard class M with c categories and N data instances. The feature value corresponding to the feature Fi is represented as z_{ji}. λ_i is the class value of M, then the data instance can be represented as $[z_j, \lambda_i]$, where $z_j = [z_{j1}, z_{j2}, z_{j3}, \ldots, z_{jg}]$. The calculation formula of the inconsistency rate is shown as follows:

$$\tau = \frac{\sum_{k=1}^{p} \left(\sum_{l=1}^{c} f_{kl} - \max_l \{f_{kl}\} \right)}{N}, \tag{6}$$

where f_{kl} is the number of data instances that centrally belongs to the feature subset of the x_k mode. x_k is the data set with a total of P feature partition modes, where $(k = 1, 2, \ldots, p$ and $p \leq N)$. The feature selection on the basis of the inconsistency rate is shown as follows:

(1) Initialize the optimal feature subset Γ to an empty set.

(2) Calculate the inconsistency rate of datasets $G1, G2, \ldots, Gg$ which belongs to the subset mode that consists of the subset Γ and each remaining feature.

(3) Calculate the inconsistency rate statistical table of the feature subsets and rank the inconsistent data from small to large.

(4) Select the feature subset L with the smallest number of features. If $\tau_L \approx \tau_\Gamma$ or $\tau_{L'}/\tau_L$ is the minimum of all adjacent inconsistency rate, then the optimal feature subset is L, where L' is the last feature subset adjacent to L.

By calculating the inconsistency rate, on the one hand, redundant features can be eliminated based on the inconsistency of data set. On the other hand, it can also take into account the correlation between features in the selection process, which has a better presentation of all data information through the selected optimal features.

2.3. Improved Least Squares Support Vector Machine (LSSVM)

2.3.1. Least Squares Support Vector Machine.

Least squares support vector machine (LSSVM) is an extension of support vector machine (SVM). It constructs the optimal decision surface by projecting the input vector into the high-dimensional space nonlinearly. Next, the inequalities of SVM are inverted into the equation sets by applying the risk minimization principle, which reduces complexity and speeds up the rate of calculation [30].

Suppose that $T = \{(x_i, y_i)\}_{i=1}^{N}$ is the given sample set, N is the total number of samples. The regression model of the sample is shown as follows:

$$y(x) = w^T \cdot \phi(x) + b, \tag{7}$$

where $\phi(*)$ is the function that projects training samples into a high-dimensional space, w is the weighted vector, and b is the offset parameter.

The optimization problem of LSSVM can be converted into the following function to solve [31]:

$$\min \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^{N} \xi_i^2 \tag{8}$$

$$\text{s.t.} \quad y_i = w^T \phi(x_i) + b + \xi_i, \quad i = 1, 2, 3, \ldots, N,$$

where γ is the penalty coefficient, which is used to balance the complexity and accuracy of the model, and ξ_i is the estimation error. The above equations can be solved by converting them into the Lagrangian function, which is detailed in 2.3.3.

2.3.2. Improved Method of LSSVM

(1) Horizontal Weighting of Input Vectors. The cost forecasting of substation projects is mostly a multiinput and single-output model. The values of the input vectors are horizontally distributed with the item serial number. And, the influence of the actual value of the substation projects cost influencing factors on the final forecasting value can be reflected by means of the weighted processing. Therefore, the weighted processing of input vectors is shown as follows:

$$\widehat{x}_i = x_{ki} \cdot \delta (1 - \delta)^{n-i}, \quad k = 1, 2, \ldots, l, \tag{9}$$

where $\widehat{x}_i$ is the weighted input vector, x_{ki} is the original input vector, k is the dimension number of input vector, and δ is a constant.

(2) Longitudinal Weighting of the Training Sample Sets. The cost forecasting value is not only related to the elements in the input vectors but also has a certain correlation with the sample groups, which means that the close sample has a greater influence on the forecasting value and the long-range sample has less influence on the forecasting value. Therefore, it is necessary to reduce the influence of close samples on the forecasting model by assigning different degrees of subordination to the influencing factors of the current substation projects cost and also enhance the influence of the long-range samples on the forecasting model at the same time. The linear membership degree μ_i is used to calculate the degree of the given membership and the equation is shown as follows [32]:

$$\mu_i = \frac{\beta + i(1 - \beta)}{N}, \quad 0 \leq \mu_i \leq 1, \tag{10}$$

where μ_i is the degree of membership and β is a constant between [0, 1] and $i = 1, 2, \ldots, N$.

Then the input sample set can be changed as follows:

$$T = \{(x_1, y_1, \mu_1) \ (x_2, y_2, \mu_2) \ \cdots \ (x_N, y_N, \mu_N)\}. \tag{11}$$

The determination of β affects the fitting performance of LSSVM directly [33]. The value of β can be obtained through the gray correlation coefficient, and the calculation formulas are shown as follows:

$$r(x_0(k), x_i(k)) = \frac{\Delta_{ki}(\min) + \rho \Delta_{ki}(\max)}{\Delta_{ki} + \rho \Delta_{ki}(\max)},$$

$$\Delta_{ki} = |x_0(k) - x_i(k)|, \quad \rho \in [0, 1], \tag{12}$$

$$\beta_i = \sum_{k=1}^{N} r(x_0(k), x_i(k)).$$

In this paper, $x_0 = Y, Y = \{y_1, y_2, \ldots, y_N\}$ since the forecasting model for cost prediction of substation projects is usually a multiinput and single-output model.

2.3.3. Weighted Least Squares Support Vector Machine.
The improvement in the above section is applied to the LSSVM to get the weighted least squares support vector machine. The objective function is described as follows:

$$\min \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^{N} \mu_i \xi_i^2 \tag{13}$$

$$\text{s.t.} \quad y_i = w^T \phi(x_i) + b + \xi_i, \quad i = 1, 2, 3, \ldots, N.$$

To solve the above problem, the Lagrangian function is established as follows:

$$L(w, b, \xi_i, \alpha_i) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^{N} \mu_i \xi_i^2$$

$$- \sum_{i=1}^{N} \alpha_i [w^T \phi(x_i) + b + \xi_i - y_i], \tag{14}$$

where α_i is the Lagrange multiplier. The variables of the function are deduced and the derivation is equivalent to zero. The specific calculation is shown as follows:

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \longrightarrow w = \sum_{i=1}^{N} \alpha_i \phi(x_i), \\[2mm] \dfrac{\partial L}{\partial b} = 0 \longrightarrow \sum_{i=1}^{N} \alpha_i = 0, \\[2mm] \dfrac{\partial L}{\partial \xi} = 0 \longrightarrow \alpha_i = \gamma \mu_i \xi_i, \\[2mm] \dfrac{\partial L}{\partial \alpha} = 0 \longrightarrow w^T + b + \xi_i - y_i = 0. \end{cases} \tag{15}$$

Convert the equations to the following problem by eliminating w and ξ_i:

$$\begin{bmatrix} 0 & e_n^T \\ e_n & \Omega + \gamma^{-1} \mu^{-1} \cdot I \end{bmatrix} \cdot \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}, \tag{16}$$

where $\Omega = \phi^T(x_i)\phi(x_i)$, $e_n = [1, 1, \ldots, 1]^T$, $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_n]$, and $y = [y_1, y_2, \ldots, y_n]^T$.
The equation (17) is obtained as follows:

$$y(x) = \sum_{i=1}^{N} \alpha_i K(x_i, x) + b, \tag{17}$$

where $K(x_i, x)$ is the kernel function. The wavelet kernel function $K(x_i, x) = \prod_{i=1}^{N} \Psi((x_i - x'_i)/\sigma_i)$ is selected to replace the Gaussian kernel function in the standard least squares support vector machine and the construction of the wavelet kernel function will be detailed in the next section. The wavelet kernel function is brought into $y(x)$ as follows:

$$y(x) = \sum_{i=1}^{N} \alpha_i \prod_{i=1}^{N} \Psi\left(\frac{x_i - x'_i}{\sigma_i}\right) + b,$$

$$\psi(x) = \cos(1.75x) \cdot \exp\left(\frac{-x^2}{2}\right). \tag{18}$$

Finally, the regression model of the weighted least squares support vector machine is shown as follows:

$$y(x) = \sum_{i=1}^{N} \alpha_i \prod_{i=1}^{N} \left\{ \cos\left[\frac{1.75(x_i - x'_i)}{\sigma_i}\right] \cdot \exp\left[\frac{-(x_i - x'_i)^2}{2}\right] \right\} + b. \tag{19}$$

In this paper, the wavelet kernel function was used to replace the traditional radial basis kernel function, which is mainly based on the following considerations. (a) The wavelet kernel function has the excellent specialty of describing the data information step by step, and the LSSVM that uses the wavelet kernel function as the kernel function can simulate any function with high precision. However, the traditional Gaussian function is relatively less effective. (b) The wavelet kernel function is orthogonal or approximately orthogonal, while the traditional Gaussian kernel function is related or even redundant. (c) The wavelet kernel function can analyze and process the multiresolution of wavelet signals. Therefore, the nonlinear processing ability of the wavelet kernel function is better than the Gaussian kernel function, which can improve the generalization ability and robustness of the LSSVM regression model.

2.3.4. Construction of the Wavelet Kernel Function. The kernel function of LSSVM is the inner product of two input spatial data points in a spatial feature, and it has two obvious features. Firstly, $k(x, x') = k(x', x)$ is the symmetric function of inner product kernel variables. Secondly, the sum of the kernel functions on the same plane is a constant. In short, only when the kernel function satisfies the following two theorems can it become the kernel function of the least squares support vector machine [34, 35].

(1) Mercer's Theorem. $k(x, x')$ is the continuous symmetric core that can be extended to the following form:

$$k(x, x') = \sum_{i=1}^{\infty} \lambda_i g_i(x) g_i(x'), \tag{20}$$

where λ_i is a positive value. The following necessary and sufficient conditions need to be met to ensure complete convergence of the above extensions.

$$\iint k(x, x') g(x) g(x) \mathrm{d}x \mathrm{d}x' \geq 0, \quad x, x' \in R^n, \tag{21}$$

For all the functions of $g(*)$, the condition that
$$\begin{cases} g(*) \neq 0 \\ \int_{R^n} g^2(\xi) \mathrm{d}\xi < \infty \end{cases}$$ needs to be satisfied.

Where $g(x_i)$ is the feature function, λ_i is the eigenvalue, and all of them are positive. Therefore, it is known that the kernel function $k(x, x')$ is a positive definite function.

(2) Smola and Scholkopf Theorem. When the kernel function satisfies Mercer's theorem, $k(x, x')$ can be used as the kernel equation of the least square support vector machine when it is proved to satisfy the following equation (22):

$$F(x)(\omega) = (2\pi)^{-(n/2)} \int_{R^n} \exp(-J(\omega \cdot x)) k(x) \mathrm{d}x \geq 0, \quad x \in R^n. \tag{22}$$

(3) Construction of the Wavelet Kernel Function. When the wavelet kernel equation satisfies conditions that $\psi(x) \in L^2(R) \cap L'(R)$ and $\widehat{\psi}(x) = 0$, $\widehat{\psi}(x)$ is the Fourier transform of $\psi(x)$. Then, $\psi(x)$ can be defined as follows [36]:

$$\psi_{\sigma,m}(x) = (\sigma)^{(1/2)} \psi\left(\frac{x - m}{\sigma}\right), \quad x \in R, \tag{23}$$

where σ is the shrinkage coefficient, m is the horizontal float coefficient, and $\sigma > 0, m \in R$.

When $f(x)$ satisfies $f(x) \in L^2(R)$, $f(x)$ can be wavelet transformed as follows:

$$W(\sigma, m) = \sigma^{(1/2)} \int_{-\infty}^{+\infty} f(x) \psi^*\left(\frac{x - m}{\sigma}\right) \mathrm{d}x, \tag{24}$$

where $\psi^*(x)$ is the complex conjugate function of $\psi(x)$. The wavelet transform function $W(\sigma, m)$ is reversible and can be used for reconstructing the original signals, and then the following formula can be obtained:

$$f(x) = C_\varphi^{-1} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W(\sigma, m) \psi_{\sigma,m}(x) \frac{\mathrm{d}\sigma}{\sigma^2} \mathrm{d}m, \tag{25}$$

where $C_\varphi = \int_{-\infty}^{+\infty} (|\widehat{\psi}(w)|^2/|w|) \mathrm{d}w < \infty$ and $\widehat{\psi}(w) = \int \psi(x) \exp(-Jwx) \mathrm{d}x$. where C_φ is a constant. The wavelet decomposition theory is an infinite approximation to a set of functions based on linear combination of the wavelet functions. Suppose that $\psi(x)$ is a one-dimensional function, then the multidimensional wavelet equation can be described according to the tensor theory as follows:

$$\psi_l(x) = \prod_{i=1}^{l} \psi(x_i), x \in R^{l \times d}, \quad x_i \in R^d. \tag{26}$$

In this way, the horizontal floating kernel function is constructed as follows:

$$k(x, x') = \prod_{i=1}^{l} \psi\left(\frac{x_i - x_i'}{\sigma_i}\right), \quad \sigma_i > 0. \tag{27}$$

The kernel equation needs to satisfy the Fourier transform in the least squares support vector machine. Therefore, the wavelet kernel function can be used as the kernel function of LSSVM when it satisfies the Fourier transform. It is necessary to prove that the following equation is established:

$$F(k)(w) = (2\pi)^{(-l/2)} \int_{R^l} \exp(-J(wx))k(x)dx \geq 0. \tag{28}$$

In this paper, the Morlet wavelet mother function is chosen to prove the above equation, which ensures the generalization of the wavelet kernel function, namely,

$$\psi(x) = \cos(1.75x)\exp\left(-\frac{x^2}{2}\right). \tag{29}$$

And, $k(x, x')$ can be expressed as follows:

$$
\begin{aligned}
k(x, x') &= \prod_{i=1}^{l} \left\{ \cos\left[1.75\left(\frac{\|x_i - x_i'\|^2}{2\sigma^2}\right)\right]\exp\left[-\frac{\|x_i - x_i'\|^2}{2\sigma^2}\right]\right\} \\
&= \prod_{i=1}^{l} \cos\left(\frac{1.75X_i}{\sigma}\right)\exp\left(-\frac{\|x_i\|^2}{2\sigma^2}\right),
\end{aligned}
\tag{30}
$$

where σ can be obtained through the sample fitting. $x \in R^N$ and $\sigma, x_i \in R^N$. It can be seen from the above equation that the multidimensional wavelet function can be used as a kernel function of the multidimensional least squares support vector machine. And, the mathematical proof process is shown as follows:

$$
\begin{aligned}
&\int_{R^l} \exp(-Jwx)k(x)dx \\
&= \int_{R^l} \exp(-Jwx) \cdot \left(\prod_{i=1}^{l}\left\{\cos\left[1.75\left(\frac{x_i - x_i'}{\sigma}\right)\right]\exp\left(-\frac{\|x_i - x_i'\|^2}{2\sigma^2}\right)\right\}dx\right) \\
&= \prod_{i=1}^{l} \int_{R^l} \exp(-Jw_i x_i) \cdot \left(\frac{\exp(j1.75x_i/\sigma) + \exp(-1.75x_i/\sigma)}{2}\right) \cdot \exp\left(-\frac{\|x_i\|^2}{2\sigma^2}\right)dx_i \\
&= \prod_{i=1}^{l} \frac{1}{2}\int_{R^l}\left\{\exp\left[-\frac{\|x_i\|^2}{2\sigma^2} + \left(\frac{1.75j}{\sigma} - jw_i\sigma\right)\cdot x_i\right] + \exp\left[-\frac{\|x_i\|^2}{2\sigma^2} - \left(\frac{1.75j}{\sigma} - jw_i\sigma\right)\cdot x_i\right]\right\}dx_i \\
&= \prod_{i=1}^{l} \frac{|\sigma|\sqrt{2\pi}}{2}\left\{\exp\left(-\frac{(1.75 - w_i\sigma)^2}{2}\right) + \exp\left(-\frac{(1.75 + w_i\sigma)^2}{2}\right)\right\}.
\end{aligned}
\tag{31}
$$

Finally, the equation is obtained as follows:

$$
\begin{aligned}
F(x)(w) = \prod_{i=1}^{l}\left(\frac{|\sigma|}{2}\right)&\left(\exp\left(-\frac{(1.75 - w_i\sigma)^2}{2}\right)\right. \\
&\left.+\exp\left(-\frac{(1.75 + w_i\sigma)^2}{2}\right)\right),
\end{aligned}
\tag{32}
$$

where $|\sigma| \neq 0$, so $F(x)(w) \geq 0$. In addition, it can be concluded that the wavelet kernel function can be used as the kernel function of least squares support vector machine.

3. DIR-ILSSVM Optimized by the WPA Algorithm

The cost forecasting model of substation projects combining WPA, DIR, and ILSSVM is constructed as illustrated in Figure 2. As is shown from the figure, the cost forecasting model of substation projects proposed in this paper mainly includes three sections. The first section is the feature selection based on the inconsistency rate. The second section is the sample training based on the ILSSVM model. The third section is the cost forecasting based on the ILSSVM model. When the established feature subset L cannot satisfy stopping conditions of the algorithm, the program will continue to loop until it reaches the desired precision, and an optimal feature subset is output. Therefore, in the cost forecasting model of the substation projects constructed in this paper, the purpose of the first section is to find the optimal feature subset and the optimal regression model parameters by iterative calculations. The purpose of the second section is to calculate the forecasting accuracy of the training sample in each iteration, so that the fitness function value can be calculated. The third section will make use of the optimal feature subsets and parameters obtained in the above two sections. And, finally, the substation projects cost of the test sample can be forecasted through retraining of the ILSSVM regression model.
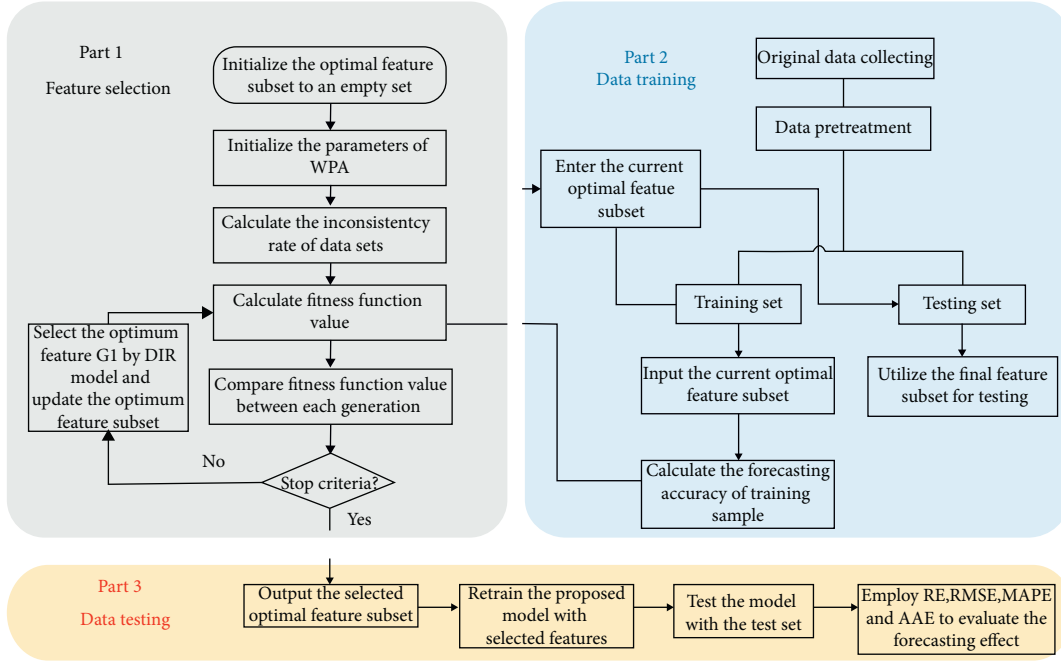
FIGURE 2: The flow chart of WPA-DIR-ILSSVM.

The specific steps for cost forecasting of substation projects are listed as follows:

(1) *Determine the Initial Candidate Feature Values.* By combing the related references [37–39], in this paper, the candidate features of the influencing factors of substation projects cost are selected as follows, including the floor space, construction properties, substation voltage level, main transformer capacity, number of outlets on the high-voltage side, number of outlets on the low-voltage side, topography, duration, substation type, number of transformers, economic development level of the construction area, inflation rate, transformer single unit price, high-side circuit breaker unit price, high-side circuit breaker unit, low-voltage capacitor quantity, high-voltage fuse price, current transformer price, power capacitor price, reactor price, power bus price, arrester price, measuring instrument price, relay protection device price, signal system price, automatic device price, site leveling fee, foundation treatment fee, designer's skill level, number of accidents, deviation rate of project volume, construction progress level, and days of rain and snow. In the IR algorithm, the optimal feature subset needs to be initialized to an empty set, namely, $\Gamma =$.

(2) *Initialize the Parameters of WPA.* Set the total number of wolf $N = 50$, the number of iterations $t = 100$, $step_a = 1.5$, $step_b = 0.8$, $q = 6$, and $h = 5$.

(3) *Calculate the Inconsistency Rate.* After step (1) and step (2) are completed, the candidate features are gradually substituted into the IR feature selection model. Calculate the inconsistency rate of datasets $G_1, G_2, \cdots, G_g$ which belongs to the subset mode that

consists of the subset Γ and each remaining feature. The feature G_i corresponding to the minimum inconsistency rate is selected as the optimal feature. $\Gamma = \{\Gamma, G_i\}$ is the updated optimal feature set.

(4) *Determine the Optimal Feature Subset and Parameters of the Optimal Regression Model.* The current feature subset is substituted into the ILSSVM model to calculate the forecasting accuracy $r(j)$ in the learning process of the current cycle training sample, and the fitness function value Fitness (j) of each cycle can be obtained. The optimal feature subset can be determined through the comparison of fitness function values between each generation. Determine whether each iteration reaches the stopping condition of the algorithm. If it is not satisfied, the new feature subset is reinitialized and the new cycle is entered until the optimal subset of the global optimal feature is obtained. It is important to note that the parameters of the ILSSVM model also need to be optimized and the initial value of γ and σ will be allocated randomly. The fitness function based on double factors of forecasting accuracy and feature selection quantity is constructed as follows:

$$\text{fitness}(j) = -\left(a + r(j) + b \times \frac{1}{\text{Numfeature}(j)}\right), \quad (33)$$

where Numfeature (x_i) is the optimal number of features selected for each iteration and a and b are constants between $[0, 1]$. The optimal feature quantity and the fitness function value are proportional to each other in each iteration, and the cost

forecasting accuracy of substation projects is inversely proportional to the fitness function value.

(5) *Stop Optimization and Start the Cost Forecasting.* Circulation ends at the maximum number of iteration. Here, the optimal feature subset and the best value of γ and σ can be substituted into ILSSVM model for cost forecasting of substation projects.

4. Case Study

4.1. Collection and Processing of Data. The relevant cost data of substation projects from 2015 to 2017 in different regions is collected in this paper, including 88 voltage level substation projects. The cost levels and the influencing factors of the first 66 substation projects are used as the training set, and the last 22 data sets are used as the test set. The collected original data of the substation projects cost are shown in Table 1.

The explanation of the processing of input indicators is shown as follows. The data of the construction properties of substation projects are mainly divided into three categories, in which the assignment of new substation is 1, the main transformer is 2, and t the interval project is 3. The data of the substation types are mainly divided into three categories, in which the assignment of the indoor type is 1, the half indoor type is 2, and the outdoor type is 3. The topography and landform are mainly divided into the following eight situations, in which the assignment of hillock is 1, slope is 2, plain is 3, flat ground is 4, paddy is 5, dry land is 6, mountain is 7, and muddy land is 8. The level of economic development in the construction area is based on the data of the local GNP and the technical level of the designers is based on the proportion of employees with bachelor degree or above in this project. The difference between the actual progress of the construction and the scheduled progress plan is selected to represent the construction progress level. And, the data are normalized according to the following formula (34):

$$Y = \{y_i\} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \quad i = 1, 2, 3, \ldots, n, \quad (34)$$

where x_i is the actual value, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the sample data, and y_i is the normalized value.

4.2. Evaluation Indicators of the Forecasting Results. Evaluation indicators of the cost forecasting results of substation projects adopted in this paper are shown as follows.

(1) Relative error (RE):

$$\text{RE} = \frac{x_i - \widehat{x}_i}{x_i} \times 100\%. \quad (35)$$

(2) Root-mean-square error (RMSE):

TABLE 1: The collected original data of the substation project cost (unit: yuan/kV·A).

Serial number	Cost
1	299.3
2	285.6
3	305.5
4	310.3
5	405.6
6	358.3
7	256.9
8	305.8
9	356.9
10	1058.6
11	501.2
12	208.6
13	356.2
14	401.5
15	378.6
16	369.5
17	301.6
18	325.8
19	337.1
20	368.9
21	370.2
22	450.1
23	980.6
24	286.8
25	279.5
26	308.6
27	312.8
28	315.9
29	364.2
30	372.5
31	383.9
32	295.6
33	270.2
34	260.8
35	239.3
36	381.7
37	406.9
38	315.6
39	285.5
40	333.7
41	336.3
42	309.0
43	341.6
44	244.2
45	346.8
46	349.5
47	352.1
48	394.7
49	405.6
50	428.2
51	449.3
52	470.4
53	491.5
54	212.6
55	353.7
56	254.8
57	375.9
58	397.0
59	418.1

Serial number	Cost
60	335.3
61	326.2
62	317.1
63	308.0
64	298.9
65	289.9
66	280.8
67	271.7
68	262.6
69	253.5
70	244.5
71	235.4
72	326.3
73	217.2
74	208.1
75	199.1
76	390.0
77	280.9
78	285.1
79	289.4
80	293.7
81	297.9
82	402.2
83	306.4
84	310.7
85	274.9
86	319.2
87	283.4
88	369.5

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\frac{x_i - \widehat{x}_i}{x_i} \right)^2}. \tag{36}$$

(3) Mean absolute percentage error (MAPE):

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{(x_i - \widehat{x}_i)}{x_i} \right| \cdot 100\%. \tag{37}$$

(4) Average absolute error (AAE):

$$\text{AAE} = \frac{1}{n} \frac{\left(\sum_{i=1}^{n} |x_i - \widehat{x}_i| \right)}{\left((1/n) \sum_{i=1}^{n} x_i \right)}. \tag{38}$$

In equations (35)–(38), x is the actual value of the substation projects cost, $\widehat{x}$ is the forecasting value of the substation projects cost, and n is the number of data groups. The smaller the value of above indicators, the higher the accuracy of cost forecasting.

4.3. Feature Selection. The main content of this section is the selection of the optimal feature subset based on the DIR model, which helps determine the input indicators of the forecasting model. In this paper, Matlab R2014b is used to carry on the programming computation. The test platform

environment is based on Intel Core i5-6300U, 4G memory, and Windows 10 Pro edition system.

The iterative process diagram of the extraction of training sample features based on the WPA-DIR-ILSSVM model is shown as Figure 3. The accuracy curve shown in the figure describes the forecasting accuracy of ILSSVM for training samples in different iterations. The fitness curve describes the fitness function value calculated during each iteration. The selected number is the optimal number of features calculated by the DIR model in the process of convergence. And, the reduced number of features refers to the number of features eliminated by the WPA algorithm during the convergence process.

From the figure above, it can be seen that the WPA algorithm converges when the number of iterations is 39 and the optimal fitness function value is −0.91 at this time. The forecasting accuracy of the training sample is 98.9% at the 39th iteration, which shows that the fitting ability of ILSSVM is enhanced through the learning and training of the algorithm and the forecasting accuracy of the training sample is the highest. Meanwhile, the number of selected features also tends to be stable when WPA reaches the 39th time. The algorithm eliminates 26 redundant features from the 33 candidate features, and the final selected features are construction properties, substation voltage level, main transformer capacity, substation type, number of transformers, main transformer single unit price, and floor space.

4.4. Cost Forecasting of Substation Projects and Result Analysis. The input vectors are substituted into the ILSSVM model for training and testing after obtaining all the optimal features of the sample data. In this paper, the self-written program is used to run and calculate in Matlab software. It is worth that the wavelet kernel function is chosen as the kernel function of the ILSSVM regression model in this paper. And, the important parameters of the model are obtained through optimization based on WPA algorithm to ensure the accuracy of ILSSVM. The parameter settings of WPA have been given in Section 3, and parameters of ILSSVM calculated by running the program are $\gamma = 43.0126$ and $\sigma = 19.0382$.

The unoptimized ILSSVM, LSSVM, SVM, and BP neural networks (BPNNs) are also selected to conduct the cost forecasting of substation projects, which helps prove the forecasting performance of the cost forecasting model proposed in this paper. The topological structure of the BPNN model is 7-5-1, and the transfer functions of the hidden layer and the output layer adopt the tansig function and purelin function, respectively. The maximum number of training is 200, the minimum error of the training target is 0.0001, the training rate is 0.1, and the initial weight and threshold are obtained by its own training. In the SVM model, the penalty parameter c is 10.276, the kernel function parameter σ is 0.0013 and the loss function parameter ε is 2.4375. In the LSSVM model, c is 10.108, σ is 0.0026, and ε is 0.0018. And, in the ILLSVM model, c is 16.263, σ is 0.0012, and ε is 0.0015.
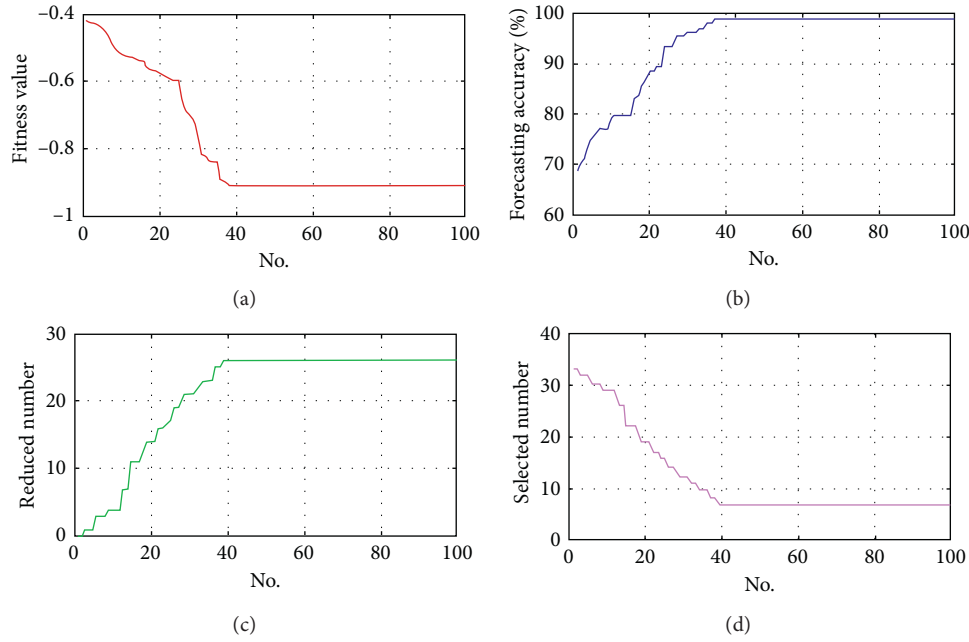
(a)

(b)

(c)

(d)

FIGURE 3: The curve of convergence for feature selection.

The forecasting results of BPNN, SVM, LSSVM, ILSSVM, and the model constructed in this paper for cost forecasting of the test set are shown in Table 2.

The forecasting results in Table 2 are plotted as shown in Figure 4 to facilitate the analysis more intuitively. The RE of each forecasting model is shown in Figure 5. It can be seen from Table 2 and Figures 4 and 5 that the forecasting error of all WPA-DIR-ILSSVM models is within [−3%, 3%], and the minimum value of the absolute relative errors is 0.06% and the maximum value is 1.30%. Among them, the number of errors outside [−1%, 1%] is 5, which are substation projects with serial numbers 67, 70, 72, 79, and 83, with relative errors of −1.14% and 1.13%, 1.30%, 1.28%, and 1.08%, respectively. The relative errors of the 11 sample points in the ILSSVM forecasting model is controlled within the range of [−3%, 3%], and the error range of 3 sample points is [−2%, 2%], with the serial number of 70, 77, and 83, respectively. For the substation projects above, the relative errors are −1.94%, 1.68%, and −0.89%, respectively, and the minimum value of the absolute relative error is 0.89% and the maximum is 3.66%. The relative errors of the 4 sample points in the LSSVM forecasting model is controlled within the range of [−3%, 3%], which are substation projects with serial numbers 68, 76, 72, 80, and 86, with relative errors of −2.78%, 2.19%, −2.74%, and 1.36%, respectively. However, all of them are outside the range of [−1%, 1%], the minimum value of the absolute relative error is 1.36%, and the maximum is 6.27%. The minimum value of the absolute relative error in the SVM forecasting model is 1.44% and the maximum is 7.10%, in which the errors of most sample points are between [−6%, −4%] and [4%, 6%]. The minimum value of the absolute relative error in the BPNN forecasting model is 2.57% and the maximum is 7.71%, in which the errors of most sample points are between [−7%, −5%] and

[5%, 7%]. Meanwhile, the fluctuation range of errors is relatively large in the BPNN model. According to the results of the absolute relative error, the WPA-DIR-ILSSVM model has the highest forecasting accuracy, next are the ILSSVM, LSSVM, and SVM models, and the worst is the BNPP model. Therefore, the forecasting accuracy and stability of WPA-DIR-ILSSVM model are the best, which reflects that the WPA algorithm can effectively enhance the training and learning process so as to avoid falling into a local optimum and improve the global searching ability of ILSSVM. Simultaneously, satisfactory forecasting results can be obtained by using the input vectors based on the DIR model. Additionally, ILSSVM presents more satisfactory performance than LSSVM, SVM, and BPNN. This result indicates that the more accurate forecasting results can be achieved by means of the improvement of LSSVM.

The RMSE, MAPE, and AAE of BPNN, SVM, LSSVM, ILSSVM, and WPA-DIR-ILSSVM are shown in Figure 6. From Figure 6, we can conclude that the RMSE, MAPE, and AAE of the proposed model are 0.8025%, 0.7159%, and 0.7157%, which are all the smallest among the above five models. In addition, the RMSE, MAPE, and AAE of the ILSSVM model are 2.6858%, 2.7961%, and 2.7956%, respectively. The RMSE, MAPE, and AAE of the LSSVM model are 4.5163%, 4.3614%, and 4.2778%, respectively. The RMSE, MAPE, and AAE of the SVM model are 4.5558%, 4.3895%, and 4.3203%. The RMSE, MAPE, and AAE of the BPNN model are 5.5044%, 5.2589%, and 5.1402%, respectively. The overall forecasting error of model the degree of error dispersion can be reflected through these indicators. Therefore, the overall forecasting effect of ILSSVM is better than that of LSSVM, SVM, and BPNN, and the overall forecasting effect of LSSVM is better than that of SVM and BPNN, which indicates that the overall forecasting

TABLE 2: The forecasting values and actual values of the test set (unit: yuan/kV·A).

| Serial number	Actual value	BPNN	SVM	LSSVM	ILSSVM	WPA-DIR-ILSSVM
67	271.7	286.4	287.1	284.6	280.5	268.6
68	262.6	251.3	273.1	255.3	269.6	264.2
69	253.5	270.0	239.6	264.0	259.6	255.9
70	244.5	225.6	254.5	259.5	239.7	247.2
71	235.4	251.4	245.9	246.5	243.1	236.7
72	326.3	335.6	335.2	341.1	335.7	330.5
73	217.2	208.3	201.8	203.6	211.6	215.6
74	208.1	219.0	216.5	218.0	213.9	209.5
75	199.1	210.2	195.2	208.6	204.9	200.6
76	390.0	376.9	395.6	398.5	381.6	387.6
77	280.9	296.1	294.8	269.6	285.6	280.5
78	285.1	302.8	269.6	296.9	295.0	288.0
79	289.4	271.9	302.1	303.2	280.2	293.1
80	293.7	310.5	281.0	285.6	302.8	291.3
81	297.9	313.7	312.3	311.1	307.6	298.1
82	402.2	412.5	389.5	385.6	413.6	405.9
83	306.4	325.3	321.6	322.6	303.7	309.7
84	310.7	298.7	327.0	292.7	320.2	309.1
85	274.9	290.3	287.2	288.6	283.2	276.7
86	319.2	338.1	302.9	323.5	307.5	318.6
87	283.4	300.1	295.7	270.6	293.2	283.6
88	369.5	350.2	385.9	387.1	380.9	372.0
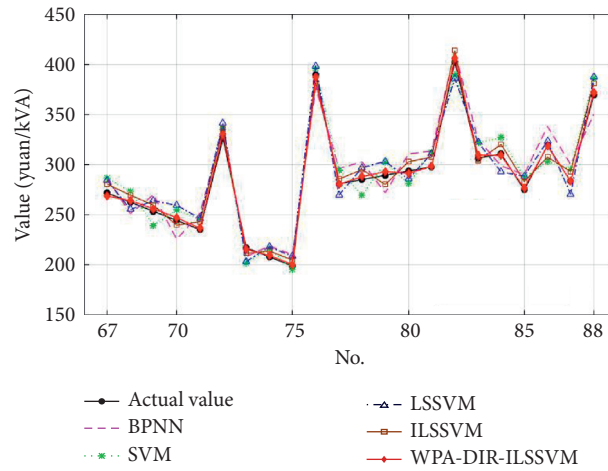
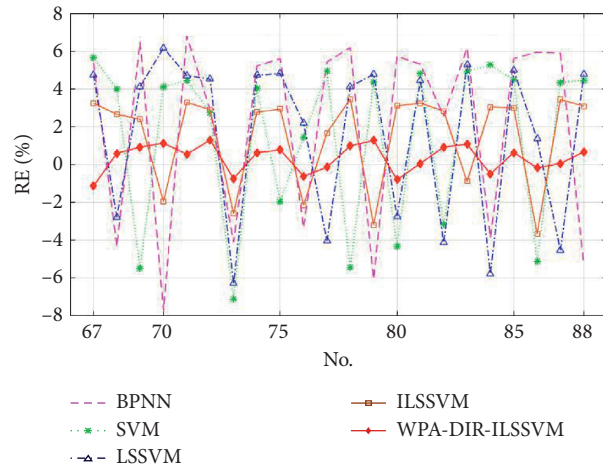FIGURE 4: Comparison of forecasting results of the test sets.

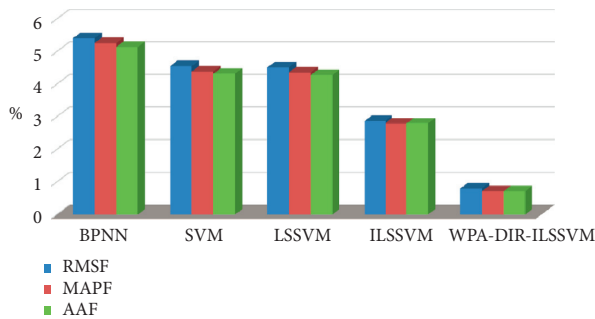FIGURE 5: RE of the forecasting results of the test set.

Figure 6: RMSE, MAPE, and AAE of the forecasting results.

performance of LSSVM is significantly improved after the weighted improvement. The forecasting accuracy of WPA-DIR-ILSSVM is better than that of ILSSVM, which proves that the parameters γ and σ of ILSSVM selected by the WPA algorithm have a good optimization effect. And, the DIR model ensures the integrity of the input information while reducing redundant data, in which the ideal forecasting results are achieved. In addition, K-fold is utilized in this paper to improve generalization performance, which advances the accuracy of prediction.

5. Conclusions

This paper presents a hybrid cost forecasting model that combines DIR with ILSSVM optimized by WPA. First, in order to forecast the substation projects cost, the DIR combined with the WPA is employed to select the input feature. Furthermore, the WPA is also adopted to optimize the parameters of the ILSSVM. Finally, after obtaining the optimized input subset and the best value of γ and σ, the proposed model is utilized for the cost forecasting of substation projects. Several conclusions based on the studies can be obtained as follows: (a) by the utilization of DIR, the influence of unrelated noises can be reduced and the forecasting performance can be effectively improved. (b) The optimization algorithm WPA adds the model with strong global searching capability and the ILLSVM model optimized by WPA shows good performance. (c) Based on the error valuation criteria, compared with LSSVM, a better forecasting result can be achieved based on ILSSVM, which shows that the method of improving LSSVM by replacing the traditional radial basis kernel function with wavelet kernel function is effective. (d) Through the example verification of substation projects in different regions, different voltage grades, and different scales, an ideal forecasting effect is obtained, which shows that the model proposed in this paper is more adaptable and stable. Hence, the proposed cost forecasting method of WPA-DIR-ILSSVM is effective and feasible, and it may be an effective alternative for the cost forecasting in the electric-power industry. However, more sample data are needed for verification. At the same time, adopting more intelligent models to forecast the substation projects cost is also our next work.

References

[1] A. S. Sidhu, M. G. Pollitt, and K. L. Anaya, "A social cost benefit analysis of grid-scale electrical energy storage projects: a case study," *Applied Energy*, vol. 212, no. 15, pp. 881–894, 2018.

[2] Z. Zhang, Z. Yu, W. Li et al., "Research on the whole life cycle cost management and maintenance decision of secondary equipment based on physical ID," in *Proceedings of the 2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)*, IEEE, Changsha, China, June 2020.

[3] L. Liu and G. Jiang, "LCC prediction model of substation in high latitude and cold area based on SSA-LS-SVM," *Smart Power*, vol. 48, no. 6, pp. 54–60, 2020.

[4] W. Li, "A model of building project cost estimation based on multiple structure integral linear regression," *Archit Technologies*, vol. 46, pp. 846–849, 2015.

[5] B.-C. Kim, "Integrating risk assessment and actual performance for probabilistic project cost forecasting: a second moment bayesian model," *IEEE Transactions on Engineering Management*, vol. 62, no. 2, pp. 158–170, 2015.

[6] L. Gao and Y. Niu, "Intelligent estimation of high-speed railway cost based om fuzzy C means clustering-fuzzy inference system," *Journal of Shijiazhuang Tiedao University (Social Science Edition)*, vol. 14, no. 2, pp. 36–43, 2020.

[7] S. Dong, Y. Zhang, Z. He, N. Deng, X. Yu, and S. Yao, "Investigation of support vector machine and back propagation artificial neural network for performance prediction of the organic rankine cycle system," *Energy*, vol. 144, pp. 851–864, 2018.

[8] D. Keles, J. Scelle, F. Paraschiv, and W. Fichtner, "Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks," *Applied Energy*, vol. 162, pp. 218–230, 2016.

[9] Z. Dong, Y. Q. Wang, Y. Wang, X. Zhao, C. Y. Rong, and J. Nie, "Cost forcast of power grid transformation project based on DEA and genetic BP neural network," *Hebei Electric Power*, vol. 39, no. 1, pp. 26–29, 2020.

[10] Y. P. Ling, P. F. Yan, C. Z. Han, and C. G. Yang, "BP neural network based cost prediction model for transmission projects," *Electric Power*, vol. 45, pp. 95–99, 2012.

[11] D. X. Niu, H. C. Wang, H. Y. Chen, and Y. Liang, "The general regression neural network based on the fruit fly optimization algorithm and the data inconsistency rate for transmission line icing prediction," *Energies*, vol. 10, p. 2066, 2017.

[12] R. C. Deo, X. Wen, and F. Qi, "A wavelet-coupled support vector machine model for forecasting global incident solar

radiation using limited meteorological dataset," *Applied Energy*, vol. 168, pp. 568–593, 2016.

[13] N. N. Wang, F. Wang, Y. T. Yin, H. Li, and Y. Hou, "Research on cost predicting of power transformation projects based on SVM," *Construction Economics & Finance*, vol. 37, pp. 48–52, 2016.

[14] Z. Y. Song, D. X. Niu, X. L. Xiao, and L. Zhu, "Substation engineering cost forecasting method based on modified firefly algorithm and support vector machine," *Electric Power*, vol. 50, pp. 168–173, 2017.

[15] X. Yuan, Q. Tan, X. Lei, Y. Yuan, and X. Wu, "Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine," *Energy*, vol. 129, pp. 122–137, 2017.

[16] Z. Tian, "Short-term wind speed prediction based on LMD and improved FA optimized combined kernel function LSSVM," *Engineering Applications of Artificial Intelligence*, vol. 91, Article ID 103573, 2020.

[17] W. T. Huang, P. Zhou, and J. X. Cheng, "An estimation method of engineering cost based on adaboost and variable selection with LSSVM," *Journal of Chongqing Jiaotong University (Natural Sciences)*, vol. 35, pp. 54–57, 2016.

[18] M. Liu, "Prediction model of project cost of based on chaotic theory and least square support vector machine," *Journal of Inner Mongolia Normal University (Natural Science)*, vol. 44, pp. 333–338, 2015.

[19] Z. Tian, "Kernel principal component analysis-based least squares support vector machine optimized by improved grey wolf optimization algorithm and application in dynamic liquid level forecasting of beam pump," *Transactions of the Institute of Measurement and Control*, vol. 42, no. 6, pp. 1135–1150, 2020.

[20] G. R. Ruiz, C. F. Bandera, G. A. Temes, and S. O. Gutierrez, "Genetic algorithm for building envelope calibration," *Applied Energy*, vol. 168, pp. 691–705, 2016.

[21] G. Li, L. Boukhatem, and J. Wu, "Adaptive quality-of-service-based routing for vehicular ad hoc networks with ant colony optimization," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3249–3264, 2017.

[22] D. Niu, Y. Liang, and W.-C. Hong, "Wind speed forecasting based on EMD and GRNN optimized by FOA," *Energies*, vol. 10, no. 12, p. 2001, 2017.

[23] M. Azaza and F. Wallin, "Multi objective particle swarm optimization of hybrid micro-grid system: a case study in Sweden," *Energy*, vol. 123, pp. 108–118, 2017.

[24] Y.-C. Li and P.-D. Xu, "Improved wolf pack algorithm for optimum design of truss structures," *Civil Engineering Journal*, vol. 6, no. 8, pp. 1411–1427, 2020.

[25] X. Zhang, C. Mei, D. Chen, and J. Li, "Feature selection in mixed data: a method using a novel fuzzy rough set-based information entropy," *Pattern Recognition*, vol. 56, pp. 1–15, 2016.

[26] J. Liu, J. Ma, Y. Zhang et al., "Discriminative feature representation to improve projection data inconsistency for low dose CT imaging," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2499–2509, 2018.

[27] T. N. Ma, D. X. Niu, Y. L. Huang, and Z. D. Du, "Short-term load forecasting for distributed energy system based on spark platform and multi-variable L2-boosting regression model," *Power System Technology*, vol. 40, pp. 1642–1649, 2016.

[28] J.-P. Liu and C.-L. Li, "The short-term power load forecasting based on sperm whale algorithm and wavelet least square support vector machine with DWT-IR for feature selection," *Sustainability*, vol. 9, no. 7, p. 1188, 2017.

[29] J. J. Xue, Y. Wang, H. Li, and J. Y. Xiao, "A smart wolf pack algorithm and its convergence analysis," *Control and Decision*, vol. 31, pp. 2131–2139, 2016.

[30] Y. Lv, F. Hong, T. Yang, F. Fang, and J. Liu, "A dynamic model for the bed temperature prediction of circulating fluidized bed boilers based on least squares support vector machine with real operational data," *Energy*, vol. 124, pp. 284–294, 2017.

[31] A. Barati-Harooni, A. Najafi-Marghmaleki, M. Arabloo, and A. H. Mohammadi, "An accurate CSA-LSSVM model for estimation of densities of ionic liquids," *Journal of Molecular Liquids*, vol. 224, pp. 954–964, 2016.

[32] M. Malvoni, M. G. De Giorgi, and P. M. Congedo, "Photovoltaic forecast based on hybrid PCA-LSSVM using dimensionality reduced data," *Neurocomputing*, vol. 211, pp. 72–83, 2016.

[33] R. Dong, J. Xu, and B. Lin, "ROI-based study on impact factors of distributed PV projects by LSSVM-PSO," *Energy*, vol. 124, pp. 336–349, 2017.

[34] Q. Wu and C. Peng, "Wind power grid connected capacity prediction using LSSVM optimized by the bat algorithm," *Energies*, vol. 8, no. 12, pp. 14346–14360, 2016.

[35] X. Song, J. J. Zhao, J. Song et al., "Local demagnetization fault recognition of permanent magnet synchronous linear motor based on S-transform and PSO-LSSVM," *IEEE Transactions on Power Electronics*, vol. 35, no. 8, pp. 7816–7825, 2020.

[36] Z. Zhao, "Backtracking search optimization algorithm-based least square support vector machine and its applications," *Engineering Applications of Artificial Intelligence*, vol. 94, Article ID 103801, 2020.

[37] X. Zhang, "Analysis on the influencing factors of project cost in substation construction stage," *Low Carbon World*, vol. 33, pp. 64-65, 2017.

[38] S. Montoya-Bueno, J. Muñoz-Hernandez, and L. Baringo, "A benders' decomposition approach for renewable generation investment in distribution systems," *Energies*, vol. 13, no. 5, p. 1225, 2020.

[39] X. Yi, X. Chuanyu, W. Wu et al., "Life cycle cost prediction of substation based on advanced PSO and least squares support vector machine," *IOP Conference Series Earth and Environmental Ence*, vol. 461, Article ID 012083, 2020.