Research Article

Energy Loss Calculation of Low Voltage Distribution Area Based on Variational Mode Decomposition and Least Squares Support Vector Machine

Keyan Liu ¹, Dongli Jia ¹, Fengzhan Zhao ², Qicheng Zhang ²,
Shuai Hao ² and Shuai Zhang ²

¹China Electric Power Research Institute, Beijing 100192, China

²College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

Correspondence should be addressed to Fengzhan Zhao; zhaofz@cau.edu.cn

Received 11 May 2021; Revised 8 July 2021; Accepted 14 July 2021; Published 27 July 2021

Academic Editor: Nianyin Zeng

Copyright © 2021 Keyan Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the accuracy of theoretical energy loss calculation in low voltage distribution area (LV-area), this paper proposes a new prediction method based on variational mode decomposition (VMD) and particle swarm optimization (PSO) least squares support vector machine (LSSVM). Firstly, the main influencing factors of energy loss calculation in LV-area are determined by the grey correlation method, which reflects the data-driven characteristic of the method and ensures the objectivity of the prediction results and the generalization of the calculation model. Secondly, the trend component and fluctuation component are obtained by VMD of daily energy loss series in the LV-area. The variable set of main influencing factors of energy losses is used as the input variable of LSSVM, and the VMD result of the energy loss sequence is used as the output. The theoretical energy loss training and calculation model of LV-area is established. Compared with the traditional calculation model, this model has more accurate calculation accuracy by taking into account the frequency characteristics of energy losses in different frequency bands. PSO is used to optimize the parameters of LSSVM for the purpose of improving the accuracy of LSSVM. Finally, an example of 252 LV-area in a city in northern China is given to verify the validity of the proposed method. The results indicate that the proposed method generates more accurate results.

1. Introduction

In recent years, the development of intelligent distribution network has carried out a simulation of construction and energy saving simultaneously. At the same time, the accuracy and timeliness of energy loss calculation are required to be higher. As an important part of the distribution network directly connected with load, the analysis and calculation of the power loss in the LV-area has always been one of the important research topics [1]. The traditional energy loss calculation method needs a lot of real-time operation data and very detailed system basic data; however, these data are not easily available. It is difficult to get the accurate calculation results by using traditional strategy [2–5]. Artificial intelligence algorithm, which has been successfully applied in load forecasting and other fields [6, 7], provides a new way

to solve this problem, which could generate accurate prediction of energy loss using based on limited data.

In the past decades, researchers have proposed lots of methodologies to improve energy loss prediction accuracy. For example, Bai and Liu [8] proposed an evaluation method based on BP neural network improved by particle swarm optimization. Zhong and Chen [9] selected eleven energy loss features from LV-area sample data, constructed a relatively perfect electrical index evaluation system of the LV-area, took them as the input variable of the training model, and proposed a line loss calculation method based on deep learning. But the actual distribution network is extremely complex; it is extremely difficult to acquire such a large amount of distribution operation data from courts users. Qin and Liu [10] applied extreme learning machines in line loss prediction, and the author also selected the particle

swarm optimization method to optimize hidden layer parameters of ELM. Some methods using the artificial intelligence algorithm achieve accurate distribution network energy loss calculation by using limited data. Data-driven feature selection methods such as the grey correlation method provide a solution to this problem. Yao and Zhu [11] applied gradient boosting decision tree model in line loss predicting in LV-area. And the author demonstrated that the method proposed performs better than others' traditional subjective feature selection methods. Actually, the data-driven feature extraction method makes the calculation model more accurate, can eliminate the coupling relationship between feature quantities, and reduce the subjective influence, which could make the prediction model more universal and generalizable [12, 13]. Actually, the energy loss data series of LV-area have significant frequency characteristics. But in the past studies, the frequency characteristics of energy loss series are often ignored. Sufficiently considering the characteristics of energy loss sequence in different frequency bands, more accurate calculation results can be obtained. Fan et al. [6] applied the empirical mode decomposition (EMD) method in electric load signal decomposition. The author divided the data into eight groups, including a trend term and seven high-frequency items, which are simultaneously employed to predict modelling. The result proved that the method aimed at the fact that different frequency characteristics could get a more accurate predictive value. VMD is evolved from EMD and it can decompose the source signal into physically significant subsignals of different frequency bands, which is more suitable for power signals analysis.

The least squares support vector machine is an improved machine prediction algorithm based on the support vector machine (SVM) [14], which has the advantages of minimizing the structural risk and adapting to small samples, and can greatly reduce the solving complexity. LSSVM has been widely used in many fields, such as in load prediction and other fields [15–18], but it is seldom used in the energy loss prediction field. Prediction model parameters have a great influence on the accuracy of the prediction model; thus, it is of great significance to optimize the parameters of the prediction model. Researchers have put forward various methodologies in order to optimize model parameters. Most of the methods are based on metaheuristics modelling [19–21]. Many researchers applied various methodologies in parameters optimization for SVM, and the prediction effect of the SVM model got good results. Deng et al. [22] proposed a method to allocate novel gate resource in airport, which is based on the quantum evolutionary algorithm improved by PSO. The result showed that the algorithm optimized by PSO has a good optimal scheduling effect. PSO has the advantages of simple operation and fast convergence and has been widely applied in many fields.

Based on this, the paper proposes the energy loss calculation method based on variational mode decomposition and particle swarm optimization least squares support vector machine for LV-area. Firstly, the factors affecting the energy losses of LV-area are listed, and the main influencing factors

of energy loss calculation of LV-area are determined by the grey correlation method. Then, the daily LV-area energy loss series of LV-area is decomposed into trend components and fluctuation components with different frequency characteristics by VMD, the main influencing factors of energy losses obtained by grey correlation method are used as input variables of LSSVM, and the VMD score of energy loss series is calculated. The component obtained from the solution is taken as the output, and the theoretical energy loss calculation model of LV-area is established. In order to improve the accuracy of LSSVM, PSO is used to optimize the parameters of LSSVM. Finally, the effectiveness of the proposed method is verified by taking 252 LV-areas in a city in northern China.

The structure of this paper is as shown in Figure 1. Sections 2–4 introduce the theoretical principles of the grey correlation method, VMD, and PSO optimization of LSSVM, respectively. Section 5 introduces the establishment of the energy loss prediction model based on the above methods. In Section 6, a numerical example is used to analyse the validity of the model.

2. Grey Relational Principal Component Analysis of Influencing Factors of Energy Losses

The structure of the distribution network is often complex, and there are many factors that affect the LV-area energy loss. They mainly include line length, insulation rate of distribution network, cable rate, line section, active energy supply, reactive energy supply, energy factor, distribution transformer load rate, working voltage of distribution transformer, and the number of users (equal to the number of meters), some of which play a major role in the change of energy loss rate. In this paper, the grey correlation method is used to find out the first five main factors affecting energy losses.

The energy loss rate in the LV-area is taken as the reference sequence, and the factors affecting the energy losses in the LV-area and the station area are taken as the comparison sequence. The formula is expressed as follows:

$$\begin{cases} x_0(k) = \{x_0(k_1), x_0(k_2), x_0(k_3), \dots, x_0(k_n)\}, \\ x_i(k) = \{x_i(k_1), x_i(k_2), x_i(k_3), \dots, x_i(k_n)\}, \end{cases} \quad (1)$$

where x_0 denotes the real energy loss rate sequence; x_i denotes the i -th affecting factors; and k represents the data at the k -th measuring moment.

Before using the grey correlation method to analyse the correlation degree of the data, due to the different measurement units of each influencing factor, the data cannot be directly calculated, and the data needs to be dedimensioned. In this paper, the initial value operator is used to normalize the influencing factors and LV-area energy loss series.

The correlation coefficients of influencing factor series and daily energy loss rate series at different times can be expressed as follows [13]:

FIGURE 1: Structure of this paper.

$$\zeta_i(k) = \frac{\min_i(\min_k |x_0(k) - x_i(k)|) + \rho \max_i(\max_k |x_0(k) - x_i(k)|)}{\Delta_{0i}(k) + \rho \Delta(\max)}, \quad (2)$$

where ρ is the resolution coefficient, generally between 0 and 1, usually 0.5.

The correlation coefficient represents the influence degree of the influencing factors of the daily LV-area energy loss, and the average value of the correlation coefficient is taken as the correlation degree [13]:

$$\gamma_i = \frac{1}{N} \sum_{k=1}^N \zeta_i(k). \quad (3)$$

Here, γ_i is the grey correlation coefficient of the influencing factor series to the daily energy loss rate series. The closer the value of γ_i is to 1, the greater the influence of this factor on the energy loss in the LV-area. According to the calculated grey correlation degree, the top 50% influencing factors of the correlation degree are selected as the input data of theoretical energy loss calculation. In this paper, five of the above 10 main influencing factors of energy losses are selected; they are active load supply, reactive power supply, transformer voltage, transformer power factor, and transformer load rate, separately.

3. Variational Mode Decomposition

VMD algorithm can find the centre frequency and bandwidth of each modal component adaptively, and VMD has good frequency separation ability. In this paper, by means of VMD, a group of energy loss rate subsequences representing different frequency components are obtained [23–25].

3.1. Realization of Variational Mode Decomposition. The VMD algorithm consists of three steps: constructing a constrained variational model, Lagrange transformation, and alternating optimization:

Step 1: a constrained variational model is constructed. The original sequence is decomposed into K modal subsequences after VMD, and u_k is the modal component with a specific centre frequency. Under the condition that the sum of all modal components is equal to the original sequence, the constrained variational expression is obtained by minimizing the sum of the estimated bandwidth of each model [26]:

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad (4)$$

s.t. $\sum_{k=1}^K u_k = f(t),$

where K is the number of modes to be decomposed ($K = 4$ in this paper); $\{u_k\}$ and $\{\omega_k\}$ are the k -th ($k = 1, 2, \dots, K$) mode sequence and the center frequency of each mode sequence after VMD; $\delta(t)$ is Dirac distribution; and $f(t)$ is the original daily energy loss rate sequence.

Step 2: *Lagrange Transformation.* The above-mentioned constrained variational problems are solved by constructing Lagrange multiplication operators to transform them into unconstrained variational problems. The variational constraint model represented by the Lagrange function is as follows [26]:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \|f(t) - \sum u_k(t)\|_2^2 + \langle \lambda(t), f(t) - \sum u_k(t) \rangle, \quad (5)$$

where λ is the Lagrange multiplication operator and α is the quadratic penalty factor to reduce the interference of Gaussian noise.

Step 3: *Alternating Optimization.* The variational constraint model (5) in Step 2 can be solved according to the following two iterative equations [26]:

$$\hat{u}_k^{n+1}(\omega) = \left(\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right) \frac{1}{1 + 2\alpha(\omega - \omega_k)^2}, \quad (6)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega}, \quad (7)$$

where $\hat{u}_k^{n+1}(\omega)$ is the Wiener filter of the modal sequence, $u_k(t)$ is the real part of the modal sequence, which is calculated by the inverse Fourier transform, $\omega_k^{n+1}(\omega)$ is the frequency centre of the corresponding modal sequence, and n is the number of iterations. After iteration, when the accuracy meets the convergence criterion, the feasible solution of the constrained variational model is obtained.

3.2. Variational Mode Decomposition of Daily Energy Loss Series. The VMD process of the daily energy loss series is summarized as follows:

Step 1: input the daily LA-area energy loss series $f(n)$, $n = 1, 2, \dots, 24$, to be decomposed.

Step 2: initialize $\{u_k^1\}$, $\{\omega_k^1\}$, $\{\hat{\lambda}_k^1\}$, and set n to zero.

Step 3: update $\hat{u}_k^{n+1}(\omega)$ and ω_k^{n+1} according to equations (6) and (7).

Step 4: update $\hat{\lambda}$ according to the following equation:

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau \left(\hat{f}(\omega) - \sum_k^K \hat{u}_k^{n+1} \right). \quad (8)$$

Step 5: set $n = n + 1$, repeat Steps 3 and Steps 4 for iteration until the following formula is satisfied [26]:

$$\sum_{k=1}^K \frac{\|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < \varepsilon. \quad (9)$$

Based on equation (9), the energy loss subsequences are obtained. The modal function is u_k , and the central frequency is ω_k , $k = 1, 2, \dots, K$. Based on the analysis of the VMD principle, when the value of K increases to a certain amount, the curve has an obvious downward bending phenomenon, and the critical value of K is the optimal number of VMD. If the number of decomposition amount is too large, the component will show intermittent phenomenon, resulting in the average instantaneous frequency of the high-frequency signal decrease, which causes the downward bending phenomenon. Confirmed by test, the best decomposition amount K for energy loss series is 4.

4. Particle Swarm Optimization for Least Squares Support Vector Machine

In this paper, the kernel function density and penalty coefficient of the least squares support vector machine (LSSVM) are determined by iterative optimization of particle swarm optimization (PSO). This method retains the advantages of the least squares support vector machine, such as small sample size demand, fast calculation speed, and high accuracy and also takes advantage of the fast search speed of particle swarm optimization.

4.1. Least Squares Support Vector Machine and Its Regression Process. The least squares support vector machine is an improved machine prediction algorithm based on SVM. The

loss function is the least squares support vector machine loss function, which transforms the inequality constraints of the SVM calculation method into equality constraints, which facilitates the solution of Lagrange multiplier α and transforms the original quadratic programming problem into the problem of solving linear equations. LSSVM also has the advantages of minimizing the structural risk of SVM, adapting to small samples, and greatly reducing the complexity of the solution.

The regression process of LSSVM is as follows:

Step 1: for a given sample set $\{(\mathbf{x}_i, y_i), i = 1, 2, \dots, N\}$, $\mathbf{x}_i \in \mathbf{R}^m$ are the training samples of LSSVM, and $y_i \in \mathbf{R}$ are the corresponding calculation samples. For the nonlinear separable training sample, it can be mapped to a higher dimensional space through the function $\varphi(\mathbf{x})$, and the optimal linear regression function is expressed as follows:

$$f(x) = \omega^T \varphi(x) + b, \quad (10)$$

where ω is the vector of weight coefficient in feature space and $b \in \mathbf{R}$ is deviation.

Step 2: according to the structural risk minimization criterion, the planning problem J of equation (10) is transformed into the following equation:

$$\min_{\omega, b, e} J(\omega, e) = \frac{1}{2} \omega^T \omega + \frac{\xi}{2} \sum_{i=1}^n e_i^2, \quad (11)$$

where $e_i \in \mathbf{R}$ is the error variable, ξ is the penalty coefficient, and $\xi > 0$.

Step 3: to solve the above optimization problem, the Lagrange function is constructed as follows:

$$L(\omega, b, e_i, \alpha_i) = J(\omega, e) - \sum_{i=1}^N \alpha_i [\omega^T \varphi(x_i) + b + e_i - y_i], \quad (12)$$

where α_i is the Lagrange multiplication operator.

Step 4: equation (12) is solved by Corner-Kuhn-Tucker condition to obtain α and b , and the regression function is obtained by calculation:

$$f(x) = \sum_{i=1}^N \alpha_i K(x_i, x_j) + b. \quad (13)$$

In formula (13), $K(\mathbf{x}_i, \mathbf{x}_j)$ is the kernel function. The commonly used kernel functions are linear kernel function like formula (14) and RBF kernel function like formula (15):

$$K(x_i, x_j) = x_i \cdot x_j^T, \quad (14)$$

$$K(x_i, x_j) = \exp \left\{ \frac{-\|x_i - x_j\|^2}{2\sigma^2} \right\}, \quad (15)$$

where σ is the width of the kernel function.

The penalty coefficient σ and kernel density σ have a great influence on the accuracy of the LSSVM prediction model. The generalization ability of the model increases with the decrease of ξ , and the sample error increases. The smaller σ , the higher the complexity of the model, while greater σ leads to low learning efficiency. A reasonable selection of parameters ξ and σ can improve the accuracy of LSSVM. In this paper, the particle swarm optimization algorithm is used to optimize these two parameters.

4.2. Particle Swarm Optimization. The principle of particle swarm optimization algorithm is based on the study of the foraging behavior of birds. The parameters ξ and σ that need to be optimized are regarded as a search individual in the whole region, that is, particles. In the iterative process, the particle can dynamically adjust its velocity and position according to the particle optimal position p_{best} and the population optimal position g_{best} :

$$v_i^{k+1} = \omega v_i^k + c_1 \mu (p_i^k - x_i^k) + c_2 \eta (p_g^k - x_i^k), \quad (16)$$

$$x_i^{k+1} = x_i^k + p v_i^k, \quad (17)$$

$$\omega = \omega_{\max} - \frac{k}{\text{iter max}} (\omega_{\max} - \omega_{\min}). \quad (18)$$

Equations (16) and (17) are the updated formulas of particle velocity and particle position, respectively. Equation (18) is the inertia weight updating formula, where ω is the inertia weight, and the larger the value, the stronger the global optimization ability of particles, the weaker the local optimization ability, and vice versa. c_1 and c_2 are acceleration constants, c_1 reflects the individual's own cognitive ability, and c_2 reflects the individual's cognitive ability of the whole group. μ and η are random numbers between them $[0, 1]$ and p is the constraint factor.

4.3. Least Squares Support Vector Machine Optimized by Particle Swarm Optimization. The parameters of LSSVM model are optimized by using particle swarm optimization algorithm, and parameter (ξ, σ) is optimized by continuous iteration to minimize the objective function, as follows:

$$\min f(\xi, \sigma) = \sum_{i=1}^M (X_i - X'_i)^2, \quad (19)$$

where X_i and X'_i are the corresponding outputs of the i -th sample data and the calculated outputs of LSSVM based on the inputs.

The process of PSO optimizing the ξ and σ parameters of LSSVM is as follows:

Step 1: initialize PSO parameters, including the number of particles L , individual extreme acceleration coefficient c_1 , global acceleration coefficient c_2 , and the maximum number of iterations iter max .

Step 2: all particle position vectors (ξ, σ) are initialized randomly.

Step 3: select formula (19) as the fitness function, train LSSVM through the current particle position vector, calculate fitness value, and update the optimal value and global optimal value of each particle.

Step 4: the optimal position of particle swarm optimization is calculated and the new position of each particle is updated.

Step 5: the particle position is checked to judge whether it meets the end of iteration condition. If the fitness value meets the convergence criterion or the number of iterations reaches the maximum set value, the iteration process is ended. The global optimal solution is obtained from the feasible solutions obtained in each iteration process, and the optimal values of ξ and σ are output. Otherwise, go back to Step 3 and continue the iteration.

5. Energy Loss Calculation Method Based on Grey Correlation Method and VMD-PSO-LSSVM Model

A theoretical energy loss calculation method of LV-area based on grey correlation method and VMD-PSO-LSSVM model is designed in this paper. Firstly, the energy loss rate sequence is decomposed by VMD, and four groups of subsequences reflecting different frequency components are obtained. Then, the four subsequences are combined with the energy loss influencing factor set (obtained by the grey correlation method) as the input of four different LSSVM prediction calculation models. Finally, the results of each LSSVM model are added to get the final result. In conclusion, in order to get more accurate calculation results, this method designs energy loss prediction calculation models for different frequency components of the energy loss rate series.

The calculation process of this method is shown in Figure 2, and the detailed steps are as follows:

Step 1: the main influencing factors of energy losses are determined by the grey correlation method, including active energy supply, reactive energy supply, transformer voltage, transformer energy factor, and transformer load rate.

Step 2: the training samples are the daily energy loss rate of LV-area 24 hours a day for 30 days, and they are input into the VMD model as input data.

Step 3: the daily energy loss rate data series are decomposed by VMD with $K=4$, and four groups of daily energy loss rate series with different frequency components are obtained, which are expressed as four groups of 30 rows and 24 columns (" 30×24 ") daily energy loss rate matrix.

Step 4: the five main influencing factors determined by the grey correlation method are combined with four groups of matrices obtained from the VMD of the daily

FIGURE 2: Flow chart of energy loss calculation method based on the grey correlation method VMD-PSO-LSSVM model.

energy loss rate of the training samples, which are respectively input into four different PSO optimized LSSVM models for energy loss calculation.

Step 5: four groups of energy loss calculation results of different frequency scales are accumulated to get the theoretical energy loss calculation results. The time interval of daily energy loss rate and influencing factor data is 1 hour.

In order to evaluate the accuracy of energy loss calculation, the average absolute percentage error (MAPE) and the maximum relative error ($E_{\max}$) are used to evaluate the accuracy of energy loss calculation. The index calculation formula is as follows [6]:

$$\text{MAPE} = \frac{1}{M} \sum_{i=1}^M \left| \frac{X_i - X'_i}{X_i} \right|, \quad (20)$$

$$E_{\max} = \max \left| \frac{X_i - X'_i}{X_i} \right|,$$

where X_i and X'_i are the theoretical energy loss actual value and theoretical energy loss calculated value of the i -th test sample, respectively.

6. Numerical Example and Analysis

6.1. An Overview of the Numerical Example.

- (1) Data source: the algorithm is verified by using the energy loss data of 252 LV-area in a city in northern China. These data have been processed with bad data. Here, the energy loss estimation method of one of the distribution transformers is introduced in detail.
- (2) Input data: the influencing factors of energy loss from May 1, 2020, to May 31, 2020 (obtained by the grey correlation method), including the active energy supply, reactive energy supply, transformer voltage, transformer energy factor, transformer load rate, and daily energy loss rate at the low voltage side of the distribution transformer. A total of 31 groups of experimental data are selected. The first 30 days of energy loss rate data and its influencing factors are selected as the training set; the energy loss rate data on the 31st day is taken as the test set, and some data are shown in Table 1.

Taking the daily energy loss rate series on May 31, 2020, as an example, the daily energy loss series is decomposed into four 1×24 subsequences by the

TABLE 1: Daily energy loss rate and influencing factors from May 1, 2020, to May 31, 2020 (partial data).

No.	Date	Time	Active energy supply	Reactive energy supply	Transformer voltage (pu)	Transformer energy factor	Transformer load rate	Energy loss rate (%)
1	2020-5-1	1:00	179.07	156.94	0.921	0.790	0.462	3.69
2	2020-5-1	2:00	171.99	168.41	0.921	0.758	0.455	3.95
3	2020-5-1	3:00	153.39	125.77	0.927	0.807	0.425	2.96
4	2020-5-1	4:00	127.76	129.32	0.929	0.749	0.406	3.01
5	2020-5-1	5:00	114.92	148.20	0.929	0.670	0.401	3.61
6	2020-5-1	6:00	130.53	142.57	0.928	0.725	0.407	3.35
7	2020-5-1	7:00	152.78	124.12	0.927	0.810	0.424	2.93
8	2020-5-1	8:00	169.49	130.33	0.924	0.823	0.449	3.11
9	2020-5-1	9:00	167.51	113.85	0.926	0.852	0.446	2.81
10	2020-5-1	10:00	165.19	121.75	0.925	0.834	0.442	2.93
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
742	2020-5-31	22:00	356.68	170.24	0.900	0.915	0.599	5.10
743	2020-5-31	23:00	308.80	189.90	0.904	0.872	0.585	4.98
744	2020-5-31	24:00	257.68	129.00	0.915	0.907	0.556	3.69

VMD method. The original series of daily energy loss rate and four subsequences of daily energy loss rate decomposed by VMD are shown in Figure 3. As can be seen from Figure 3, four subsequences of daily energy loss rate represent different frequency components obtained from VMD. Subsequence 1 is the trend component, which is similar to the change trend of the actual value of daily energy loss rate; subsequences 2~4 are the fluctuation components, in which subsequence 2 reflects the shape of the curve, and subsequences 3 and 4 reflect the random fluctuation details of the curve.

Because subsequence 1 and subsequence 2 fluctuate more smoothly than other subsequences, LSSVM1 and LSSVM2 use linear kernel function (formula (14)), while LSSVM3 and LSSVM4 use RBF kernel function (formula (15)) because subsequence 3 and subsequence 4 fluctuate more strongly.

- (3) Calculation model: in order to analyse the effectiveness of the model, this paper compares the algorithm based on VMD-PSO-LSSVM (the algorithm of this paper), PSO-LSSVM without variational mode decomposition (PSO-LSSVM), and LSSVM without any optimization algorithm (LSSVM) comparative study.
- (4) The parameters of the calculation model are as follows: in the VMD-PSO-LSSVM algorithm, the modal function K is 4, the number of particles L is 30, the individual extreme acceleration coefficient is 1.5; the global acceleration coefficient is 1.7; the maximum iteration times are 100; the initial parameters of the four LSSVM algorithms are 30, 4; and the PSO-LSSVM and LSSVM parameters are set the same as the VMD-PSO-LSSVM parameters.
- (5) Calculation objective: the daily energy loss rate series on May 31, 2020, is calculated from the series data, the influencing factors of daily energy loss rate from

May 1 to May 30, 2020, and the influencing factors of energy loss on May 31, 2020.

6.2. Comparative Analysis of Calculation Results. The daily energy loss rate on May 31, 2020, is obtained by using the above four calculation models, as shown in Figure 4. The prediction results are shown in Table 2. It can be seen from Figure 4 that the results calculated by VMD-PSO-LSSVM in this paper are closer to the actual values than the other two calculation methods, indicating that the calculation results obtained by VMD-PSO-LSSVM in this paper are more accurate.

6.3. Error Comparison and Efficiency Analysis. The average calculation error and average forecast time of the above three calculation models on May 31, 2020, are shown in Table 3.

Compared with the calculation results of the above three models, the calculation error indexes MAPE and Emax of VMD-PSO-LSSVM are the smallest, which indicates that the theoretical energy loss calculation accuracy of this method is the highest; the calculation error index of VMD-PSO-LSSVM is smaller than that of PSO-LSSVM, which verifies the necessity of combining VMD algorithm. Comparing the calculation time of the three models, VMD-PSO-LSSVM takes the most time, but its accuracy is significantly improved compared with the other two models.

Statistical hypothesis testing is also used to verify the prediction accuracy advantage of the model. In this paper, the T-double-tail test method is used to test whether there is a significant difference in the prediction difference of the three methods, so as to judge whether VMD and PSO have a significant improvement in the prediction accuracy of the model. The results of hypothesis testing are shown in Table 4.

It can be seen from Table 4 that the prediction accuracy of the VMD-PSO-LSSVM model is significantly different from that of the other two models. At the same time, the

FIGURE 3: Daily energy loss rate series on May 31 and 4 subsequences after VMD.

FIGURE 4: Daily energy loss rate series and calculation series of each model on May 31, 2020.

TABLE 2: Daily energy loss rate calculation result of each model on May 31, 2020.

Time	Daily energy loss rate of LV-area			
	Origin data	PSO-LSSVM	VMD-PSO-LSSVM	LSSVM
1:00	3.07	3.18	3.10	2.93
2:00	3.93	4.08	3.85	3.74
3:00	2.67	2.77	2.64	2.80
4:00	2.72	2.62	2.76	2.85
5:00	2.87	2.76	2.89	2.73
6:00	2.71	2.60	2.77	2.58
7:00	2.88	2.78	2.90	3.03
8:00	3.33	3.46	3.34	3.49
9:00	3.71	3.57	3.70	3.88
10:00	3.58	3.71	3.59	3.40
11:00	4.09	4.25	4.06	3.91

TABLE 2: Continued.

Time	Daily energy loss rate of LV-area			
	Origin data	PSO-LSSVM	VMD-PSO-LSSVM	LSSVM
12:00	3.27	3.39	3.27	3.11
13:00	3.67	3.53	3.70	3.49
14:00	3.38	3.26	3.40	3.21
15:00	3.53	3.66	3.55	3.36
16:00	3.51	3.37	3.52	3.34
17:00	3.45	3.58	3.44	3.61
18:00	3.87	3.73	3.91	3.69
19:00	4.62	4.45	4.68	4.83
20:00	5.06	4.86	4.87	4.81
21:00	4.52	4.69	4.55	4.74
22:00	5.10	4.90	5.21	5.35
23:00	4.98	4.79	5.06	4.74
24:00	3.69	3.56	3.70	3.52

TABLE 3: The average error and calculation time of three times for each combination model.

Model	Calculation error		Calculation Time (s)
	MAPE (%)	$E_{\max}$	
VMD-PSO-LSSVM	1.03	0.0216	3.27
LSSVM	4.79	0.0496	1.63
PSO-LSSVM	3.77	0.0393	2.53

TABLE 4: The results of the T-double-tail test of three different prediction models.

Compared models	T-double-tail test		
	P value	P compared with $\alpha = 0.05$	P compared with $\alpha = 0.01$
VMD-PSO-LSSVM and PSO-LSSVM	9×10^{-12}	<	<
VMD-PSO-LSSVM and LSSVM	1×10^{-15}	<	<
PSO-LSSVM and LSSVM	2×10^{-4}	<	<

α denotes significance level. There is a significant difference between the two samples if the P value is lower than α .

LSSVM model optimized by PSO also has significant accuracy improvement compared with the model without parameter optimization. The results demonstrate that the VMD and PSO methods greatly improve the prediction accuracy of the prediction model.

Using the same method to estimate the energy losses of other 251 LV-areas, the average calculation error MAPE is 1.203%. The relative advantages and disadvantages of the three methods are similar to those in Table 3, which shows that the method proposed in this paper is effective.

7. Conclusion

In this paper, an energy loss calculation method of LV-area based on variational mode decomposition and PSO optimized LSSVM is proposed.

Firstly, the grey correlation method is used to analyse a series of factors that may affect the calculation of energy losses and, finally, five main factors are determined to simplify the calculation. This data-driven feature selection method makes the model more objective and enhances the universality of the calculation model. It is verified by the author that this method not only can be used for energy loss

prediction in LV-area but also has good applicability for energy loss prediction in 10 kV distribution network.

Secondly, different subsequences of daily energy loss rate decomposed by VMD combined with influencing factors are taken as the input of four different LSSVM calculation models, and then the calculation results of each model are added to get the final result. Considering the different frequency characteristics of the energy loss sequence, it is decomposed into components of different frequency bands and predicted by model matching, respectively. Energy loss calculation models designed for different frequency components of energy loss rate series make the calculation more refined. Compared with the traditional algorithm, the prediction model proposed in this paper has better adaptability and calculation accuracy.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Keyan Liu conceptualized the study, designed the methodology, conducted formal analysis, wrote the original draft of the study. Dongli Jia curated the data; administered the project; was responsible for resources; wrote, reviewed, and edited the article. Fengzhan Zhao supervised the project. Qicheng Zhang arranged software. Shuai Hao investigated the study. Shuai Zhang validated the findings.

Acknowledgments

This research was funded by the Science and Technology Project of State Grid Corporation of China (SGTJDK00DWJS1800014): "Lean Line Loss Management Technology and Application of Distribution Network Based on Big Data."

References

- [1] Y. W. Peng, "Theoretical line loss calculation method of distribution network based on improved kernel vector machine," *Proceedings of the Chinese Society for Electrical Engineering*, vol. 31, no. 34, pp. 120–126, 2011, in Chinese.
- [2] B. J. Chen and K. L. Xiang, "Theoretical line loss calculation of distribution network based on the integrated electricity and line loss management system," in *Proceedings of the 2018 China International Conference on Electricity Distribution (CICED)*, pp. 2531–2535, IEEE Press, Tianjin, China, September 2018.
- [3] S. Y. Shi and X. L. Shen, "Research on calculation of low voltage distribution network theoretical line loss based on matpower," in *Proceedings of the 2011 International Conference on Advanced Power System Automation and Protection*, pp. 22–25, IEEE Press, Beijing, China, October 2011.
- [4] Y. N. Yang and Y. Xue, "A line loss calculation method based on big data and load curve," *Journal of Physics: Conference Series*, vol. 1865, no. 4, 2021.
- [5] Z. X. Wang, Y. N. Li, and H. X. Liu, "Analysis and calculation of line loss data based on hybrid clustering," in *Proceedings of the 2019 IEEE/ACIS 18th International Conference on Computer and Information Science (ICIS)*, pp. 310–314, IEEE Press, Beijing, China, June 2019.
- [6] G.-F. Fan, S. Qing, H. Wang, W.-C. Hong, and H.-J. Li, "Support vector regression model based on empirical mode decomposition and auto regression for electric load forecasting," *Energies*, vol. 6, no. 4, pp. 1887–1901, 2013.
- [7] Y. H. Chen and W. C. Hong, "Electric load forecasting based on a least squares support vector machine with fuzzy time series and global harmony search algorithm," *Energies*, vol. 9, no. 2, pp. 1–13, 2016.
- [8] J. H. Bai and L. P. Liu, "A method of evaluating 10 kV distribution network line losses based on intelligent algorithm," in *Proceedings of the 2018 International Conference on Power System Technology (POWERCON)*, pp. 3324–3329, IEEE Press, Guangzhou, China, November 2018.
- [9] X. Q. Zhong and J. Chen, "A method of line loss analysis in station area based on deep learning," *Power System Technology*, vol. 44, no. 2, pp. 769–774, 2020, in Chinese.
- [10] H. Q. Qin and J. F. Liu, "Line loss prediction based on particle swarm optimization combined with extreme learning machine," *Journal of Physics: Conference Series*, vol. 1802, no. 3, Article ID 032087, 2021.
- [11] M. T. Yao and Y. Zhu, "Research on predicting line loss rate in low voltage distribution network based on gradient boosting decision tree," *Energies*, vol. 12, no. 13, 2019.
- [12] H. Zhao, H. Liu, Y. Jin, X. Dang, and W. Deng, "Feature extraction for data-driven remaining useful life prediction of rolling bearings," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, Article ID 3511910, 2021.
- [13] Y. T. Zhang and Z. Z. Wang, "A 10 kV distribution network line loss prediction method based on grey correlation analysis and improved artificial neural network," *Power System Technology*, vol. 43, no. 4, pp. 1404–1410, 2019, in Chinese.
- [14] T. A. Liu and Y. J. Wang, "Research on least squares support vector machine combinatorial optimization algorithm," in *Proceedings of the 2009 International Forum on Computer Science-Technology and Applications*, pp. 452–454, IEEE Press, Chongqing, China, December 2009.
- [15] Z. C. Cao and L. Liu, "Short-term load forecasting based on variational modal decomposition and optimization model," in *Proceedings of the 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, pp. 121–126, IEEE Press, Vancouver, Canada, August 2019.
- [16] P. L. Li and H. C. Liu, "Power load forecasting based on VMD multifrequency combinations," in *Proceedings of the 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, pp. 1214–1218, IEEE Press, Chongqing, China, October 2019.
- [17] Y. W. Cao and M. Zeng, "A short-term electric load forecasting method based on improved VMD-LSSVR considering comprehensive energy correlation," in *Proceedings of the 2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*, pp. 1382–1388, IEEE Press, Wuhan, China, October 2020.
- [18] H. S. Wu and X. L. Chang, "Power load forecasting with least squares support vector machines and chaos theory," in *Proceedings of the 2006 6th World Congress on Intelligent Control and Automation*, pp. 4369–4373, IEEE Press, Dalian, China, June 2006.
- [19] M.-W. Li, Y.-T. Wang, J. Geng, and W.-C. Hong, "Chaos cloud quantum bat hybrid optimization algorithm," *Non-linear Dynamics*, vol. 103, no. 1, pp. 1167–1193, 2021.
- [20] W. Deng, J. Xu, X.-Z. Gao, and H. Zhao, "An enhanced MSIQDE algorithm with novel multiple strategies for global optimization problems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–10, 2020.
- [21] W. Deng and S. F. Shang, "Quantum differential evolution with cooperative coevolution framework and hybrid mutation strategy for large scale optimization," *Knowledge-Based Systems*, vol. 224, no. 19, 2021.
- [22] W. Deng, J. Xu, H. Zhao, and Y. Song, "A novel gate resource allocation method using improved PSO-based QEA," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–9, 2020.
- [23] A. A. Abdoos, "A new intelligent method based on combination of VMD and ELM for short term wind power forecasting," *Neurocomputing*, vol. 203, pp. 111–120, 2016.
- [24] F. Z. Zhao and S. Hao, "Short-term load forecasting for distribution transformer based on VMD-BA-LSSVM algorithm," *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, vol. 35, no. 14, pp. 190–197, 2019, in Chinese.
- [25] F. Z. Zhao and Q. C. Zhang, "Multi-time scale coordinated optimal energy dispatch of grid-connected microgrid using VMD-MPC," *Transactions of the Chinese Society of*

Agriculture Engineering (Transactions of the CSAE), vol. 37, no. 7, pp. 190–198, 2021, in Chinese.

- [26] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.