Short-term wind power prediction based on hybrid variational mode decomposition and least squares support vector machine optimized by improved salp swarm algorithm model

Zhongde Su*, Huacai Lu

Key Laboratory of Electric Drive and Control of Anhui Higher Education Institutes, Anhui Polytechnic University, Wuhu, 241000, China

*Corresponding author’s e-mail: 2200220131@stu.ahpu.edu.cn

Abstract. To improve the accuracy of wind power prediction, a short-term wind power prediction model based on variational mode decomposition (VMD) and improved salp swarm algorithm (ISSA) optimized least squares support vector machine (LSSVM) is proposed. In the model, the variational modal decomposition is used to decompose the wind power sequence into multiple eigenmode components with limited bandwidth. The improved salp swarm algorithm is employed to tune the regularization parameter and kernel parameter in LSSVM. The proposed wind power prediction strategy using mean one-hour historical wind power data collected from a wind farm located in zhejiang, China. Compared with other prediction models illustrate the better prediction performance of VMD-ISSA-LSSVM.

1. Introduction

Wind energy is rich in resources and is considered as a substitute for fossil fuels. As an important renewable energy, wind power has the advantages of wide distribution, renewable, pollution-free and so on [1]. At present, the methods of wind power forecasting mainly include statistical methods, physical methods and combined forecasting. Since wind power forecasting is affected by many factors, each forecasting model has its own advantages and disadvantages. It is often difficult for a single forecasting model to achieve the best forecasting effect, and the combined forecasting model combines the advantages of different models [2]. Least Squares Support Vector Machine (LSSVM) has the advantages of fast training speed and good generalization performance, which are widely used in wind power forecasting. Wu et al. [3] proposed Variational mode decomposition (VMD), least squares support vector machine (LSSVM) optimized by bat algorithm (BA) to improve the efficiency and accuracy of wind speed prediction. Zhang et al. [4] using chaos firefly algorithm to optimize LSSVM layer parameters and established wind speed prediction model based on LSTM-LSSVM-CFA. Compared with the basic firefly algorithm, the chaos firefly algorithm has better global convergence ability and enhance the prediction accuracy. Zhang et al. [5] proposed CEMDAN to decompose the wind power data, then used LASSO method to eliminate the noise signal and refit, and finally established the DE-optimized DNN neural network prediction model to predict the wind power.

In this paper, a novel hybrid model VMD-ISSA-LSSVM is proposed for short-term wind power prediction. The main works of this paper are summarized as follows: (1) VMD decomposition method is used to decompose wind power data to realize adaptive signal decomposition and noise reduction; (2) ISSA was used to optimize the LSSVM, the regularization parameters and kernel parameters of the LSSVM model are calculated; (3) To examine the performance of the proposed prediction model,
compared with DBN model, LSSVM model, PSO-LSSVM model, SSA-LSSVM model and ISSA-LSSVM model, the hybrid model proposed has better wind power Prediction effect.

2. Variational mode decomposition
VMD is a new adaptive signal processing method used to decompose non-stationary signals into multiple modes. Therefore, when the sum of all modes is equal to the original signal, each mode has a finite bandwidth corresponding to the center frequency, and the sum of bandwidths of all modes is the smallest. To solve the constrained variational problem, the objective function is as follows:

$$\min_{\{w_k\}, \{u_k\}} \sum_k \left[ \sum \left( \hat{\partial}_x \left[ \left( \hat{\partial}_t + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right)^2 \right], \quad \text{st.} \quad \sum_k u_k = f$$  \hspace{1cm} (1)

Where $u_k$ represents the band-limited modal functions, $\{u_k\} = \{u_1, \ldots, u_k\}$, $w_k$ represents center frequencies, $\{w_k\} = \{w_1, \ldots, w_k\}$.

In order to simplify the above constrained optimization problem, Lagrangian multipliers and penalty factors are used to construct non-constrained problems. The expression is as follows:

$$L(\{u_k\}, \{w_k\}, \lambda) := \alpha \sum_k \left[ \sum \left( \hat{\partial}_x \left[ \left( \hat{\partial}_t + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right)^2 \right] + \left[ f(t) - \sum_k u_k(t) \right]^2 + \left( \lambda(t), f(t) - \sum_k u_k(t) \right)$$  \hspace{1cm} (2)

To solve the optimization problem, the Alternate Direction Method of Multiplier (ADMM) is proposed to update $u_k^{n+1}$, $w_k^{n+1}$ and $\lambda^{n+1}$ to obtain the saddle point of the augmented Lagrangian expression. According to ADMM, the update of $u_k$ is shown in equation (3), and the update of $w_k$ is shown in equation (4).

$$\hat{u}_k^{n+1} = \left( \hat{f}(w) - \sum_{i \neq k} u_i(w) + \hat{\lambda}(w)/2 \right) \left( 1 + 2 \alpha (w - w_k)^2 \right)$$  \hspace{1cm} (3)

$$w_k^{n+1} = \int_0^\infty w \hat{u}_k(w) \hat{w} dw / \int_0^\infty \hat{u}_k(w) \hat{w} dw$$  \hspace{1cm} (4)

Where $\hat{u}_k(w)$, $\hat{u}(w)$, $\hat{\lambda}(w)$ and $\hat{\lambda}(w)$ is the Fourier transform of $u_k^{n+1}(w)$, $u(w)$, $u_i(w)$ and $\lambda(w)$.

3. Least squares support vector machine
Least Squares Support Vector Machine is an improved algorithm based on Support Vector Machine. The principle is to map data to a high-dimensional space through a non-linear function and perform linear regression. The regression function is shown in equation (5).

$$f(x) = w^T \phi(x) + b$$  \hspace{1cm} (5)

Where $w$ is weight vector, and $b$ is bias term.

According to the principle of risk minimization, the constraint problem of LSSVM can be expressed as:

$$\min J(w, \varepsilon) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^N \varepsilon_i^2, \quad \text{s.t.} \quad y_i \left[ w^T \phi(x_i) + b \right] = 1 - \varepsilon_i, \quad i = 1, \ldots, N$$  \hspace{1cm} (6)

Where $\gamma$ is a regularization constant and $\varepsilon_i$ is the slack variable.

It is difficult to solve the high dimension of the above optimization problem directly. The Lagrange function is established as follows:

$$L(w, b, \varepsilon, \alpha) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^N \varepsilon_i^2 - \sum_{i=1}^N \alpha_i \left[ w^T \phi(x_i) + b + \varepsilon_i - y_i \right]$$  \hspace{1cm} (7)

Where $\alpha_i$ is the Lagrangian multiplier.

According to the optimal conditions of KKT:
Then we can get the following equation:

$$w = \sum_{i=1}^{N} \alpha_i y_i \phi(x_i); \quad \sum_{i=1}^{N} \alpha_i = 0; \quad \alpha_i = \gamma e_i; \quad y_i [w^T \phi(x_i) + b] - 1 + e_i$$

(9)

By Eliminating $w$ and $e$, the equation (9) can be transformed into a linear equation system as follows:

$$\begin{bmatrix} 0 & I^T \\ I & ZZ^T + \frac{I}{\gamma} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}$$

(10)

Where $Z = [\phi(x_1)^T y_1, L, \phi(x_N)^T y_N]$, $I = [1, L, 1]$, $y = [y_1, L, y_N]^T$, $\alpha = [\alpha_1, L, \alpha_N]$.

According to the Mercer condition, the kernel function $k(x_i, x_j)$ is equivalent to the scalar product calculation in the above formula to linearize the nonlinear problem.

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$$

(11)

The LSSVM model expression is:

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

(12)

The performance of LSSVM is affected by regularization parameters, kernel function types and parameters. Radial basis function (RBF) has better local development ability, so this paper uses RBF as the kernel function in LSSVM.

4. Salp swarm algorithm

Salp Swarm Algorithm is a swarm intelligence optimization algorithm proposed by Mirjalili et al. It is mainly inspired by the grouping and predation behavior of Salp Swarm in the sea[6]. Mathematical modeling is used to model the salps chain, which is divided into two types according to the position in the chain structure formed by salps, namely leader and follower. The optimal solution of the problem is taken as the food source of salps group. The specific process is as follows:

$$x_j^l = \begin{cases} F_j + c_1 ((u_b_j - l_b_j)c_2 + l_b_j), & c_3 \geq 0.5 \\ F_j - c_1 ((u_b_j - l_b_j)c_2 + l_b_j), & c_3 < 0.5 \end{cases}$$

(13)

Where $X_j^l$ is the position of the first salp in $j$-th dimension; $u_b$ is the upper bounds of the $j$-th dimension, $l_b$ is the lower bounds of the $j$-th dimension, $F_j$ is the position of the food in $j$-th dimension; $c_1, c_2$ and $c_3$ are random numbers. Parameter $c_1$ is calculated as follows:

$$c_1 = 2e^{\frac{t}{T}}$$

(14)

Where $t$ is the current iteration and $T$ is the total number iterations.

To update the follower position, use Newton's laws of motion:

$$X_j^f = \frac{1}{2} a t^2 + v_0 t$$

(15)

Where $i \geq 2$, $X_j^f$ is the position of $i$-th followers in $j$-th dimension, $a = \frac{v_{final}}{v_0}$, $v = x - \frac{x_0}{t}$ and $v_0$ is the initial speed. Because the optimization time in the algorithm is iteration, and the difference between iterations is 1, considering $v_0 = 0$, the equation can be expressed as:
\[ x'_j = \frac{1}{2} (x'_j + x'^{-1}_j) \]  

\[ (16) \]

5. Improved salp swarm algorithm

5.1. Tent map

The initial population of salp swarm algorithm is generated by random strategy, which has limited global search ability and is easy to fall into local optimal. Chaos was introduced to improve the optimization accuracy of salp swarm algorithm. The purpose of chaotic mapping is to generate chaotic sequences in the interval \([0, 1]\) through mapping relations. The randomness and ergodicity of chaos can be used to initialize the population to effectively improve the diversity of the population, so that the algorithm can jump out of local optimal and improve the global search ability.

Tent mapping has better ergodic uniformity than Logistic mapping, and can search for the global optimal solution, and has a faster iteration speed. The mathematical expression of Tent mapping is as follows:

\[ x_{k+1} = \begin{cases} 
\mu x_k, & 0 \leq x_k \leq 0.5 \\
\mu(1 - x_k), & 0.5 < x_k \leq 0.5 
\end{cases} \]  

\[ (17) \]

Where \(\mu \in (0, 2]\), \(k=1, 2, 3 \ldots N\), \(\mu\) is the control parameter reflecting the chaos degree of the system, The larger \(\mu\) is the better the chaos of the system, and \(\mu=2\) in the paper.

5.2. Inertia weight strategy

In the basic salp swarm algorithm, the position update of the follower is related to the position of the individual and the position of the previous individual, and is not affected by any parameters. If the solution of the previous individual is poor, the algorithm will fall into local optimum. The inertia weight factor is introduced into the position update of followers, and \(w\) decreases linearly with the increase of iterations. When \(w\) is large, it is conducive to global search, while when \(w\) is small, it is conducive to local search. The new follower position update formula is as follows:

\[ x'_j = \frac{1}{2} (x'_j + w \cdot x'^{-1}_j) \]  

\[ (18) \]

Where \(w\) is the inertia weight coefficient, and the calculation method is as follows:

\[ w = w_{\text{max}} \cdot \left( w_{\text{max}} - w_{\text{min}} \right) \cdot \frac{(T-t)}{T} \]  

\[ (19) \]

Where \(w_{\text{max}}\) is 0.92 and \(w_{\text{min}}\) is 0.01.

5.3. Procedure of improved salp swarm algorithm

The procedure of the ISSA is described as follows:

Step 1: In the first step of Improved Salp Swarm Algorithm, Initialize the parameters of ISSA.

Step 2: Evaluate the fitness of each salp position.

Step 3: select leader and follower.

Step 4: Update \(c_1\) according to equation (14) and generate random number \(c_2\) and \(c_3\).

Step 5: if \(i=1\), update the position of the leader salp by equation (13). If \(i \geq 1\), update the position of the follower salp by equation (18).

Step 6: Evaluate the individual fitness of the updated population, and select the individual with the lowest fitness value compared with the current food location. If there is an individual fitness value less than the current food location, the individual fitness value will be used as the new food location.

Step 7: Continue to iterate steps 4 to 6 until the maximum number of iterations is reached, and finally output the optimal food position.
6. Experimental analysis

6.1. Selection of samples
In this study, the wind power dataset was collected from a wind farm in Zhejiang, China. As shown in Figure 1, the original wind power has strong fluctuation and poor stationarity. In order to obtain better prediction effect, wind power data should be decomposed by VMD. VMD decomposition needs to determine the bandwidth parameter and the number of modes. The bandwidth parameter is used in the article, and the value of the mode number is determined by the center frequency method. As shown in Table 1, take \(k=1, 2, 3, 4, 5\) to perform VMD decomposition on wind power. It can be seen that when \(k=5\), the center frequency values of IMF4 and IMF5 are similar, mode mixing occurs, and the wind power signal is over-decomposed, so take the number of modes \(k=4\).

Figure 1. Time series of original wind power.

The four components after VMD decomposition are shown in Figure 2. It can be seen from the figure that sequence 1 reflects the change trend of original wind power data, while sequence 2, 3 and 4 belong to the fluctuation component. It can be found from Figure 2 that the fluctuation range of sequence 1 to sequence 4 after VMD decomposition gradually becomes smaller, and the smaller the fluctuation range, the smaller the impact on the prediction result.

6.2. Selection of prediction error index
In this paper, the mean absolute error MAE, mean absolute percentage error MAPE and root mean square error RMSE are used to evaluate the predict results.

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |Y_{di} - Y_{Fi}|, \quad RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (Y_{di} - Y_{Fi})^2}, \quad MAPE = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{Y_{di} - Y_{Fi}}{Y_{di}}\right) \times 100\% \quad (20)
\]

Where \(N\) is the number of samples, \(Y_{di}\) is the wind power real value, \(Y_{Fi}\) is the predicted value.

6.3. Prediction model based on VMD-ISSA-LSSVM
LSSVM prediction model produces different results under different regularization parameters and kernel function parameters. In order to improve the accuracy of LSSVM model prediction, VMD decomposition and improved SSA optimization parameters of LSSVM were used to establish the prediction model. The flow chart is shown in Figure 3.
6.4. Results and discussion
The ranges of search space used in this paper are set in the ranges [0.01 150], population size \( N = 30 \), and maximum iteration times \( T = 100 \). In order to verify the superiority of the VMD-ISSA-LSSVM model proposed in the article, compare with DBN, LSSVM, PSO-LSSVM, SSA-LSSVM, ISSA-LSSVM models. Set the regularization parameter \( \gamma = 10 \) and the kernel parameter \( \sigma = 90 \) of LSSVM, and the optimization results of LSSVM parameters of other models are shown in Table 2.

| Modes | IMF1 | IMF2 | IMF3 | IMF4 | IMF5 |
|-------|------|------|------|------|------|
| K=2   | 0.05 | 164.62 |      |      |      |
| K=3   | 0.05 | 164.63 | 336.93 |      |      |
| K=4   | 0.003 | 42.12 | 166.33 | 374.78 |      |
| K=5   | 0.003 | 42.12 | 166.33 | 335.59 | 378.80 |

| Modes  | \( \gamma \) | \( \sigma \) |
|--------|---------------|---------------|
| PSO-LSSVM | 0.1           | 0.558         |
| SSA-LSSVM | 80.889        | 68.639        |
| ISSA-LSSVM | 27.487        | 62.296        |
| VMD-ISSA-LSSVM | 146.625 | 3.059 |

Table 1. Centre frequency corresponding to different K.

Figure 4 shows the prediction curves of the six models. It can be seen from the figure that the prediction curves of DBN model and LSSVM model without optimization are significantly different from the real curve. The prediction model of LSSVM optimized by SSA and ISSA is better. Compared with other models, the prediction curve of VMD-ISSA-LSSVM is closer to the real curve.

Figure 3. Flow chart based on VMD-ISSA-LSSVM. Figure 4. Prediction results of different models.

It can be seen from Table 3 that the indicators based on VMD-ISSA-LSSVM prediction have better improvement and more accurate prediction accuracy compared with the other five prediction models.
Table 3. Prediction indexes of different models.

| Models          | RMSE  | MAE   | MAPE/% |
|-----------------|-------|-------|--------|
| DBN             | 7.4948| 5.7155| 2.4319 |
| LSSVM           | 3.4672| 2.8245| 1.2059 |
| PSO-LSSVM       | 3.6845| 2.6447| 1.1455 |
| SSA-LSSVM       | 2.8887| 2.5208| 1.09   |
| ISSA-LSSVM      | 2.8372| 2.4867| 1.0693 |
| VMD-ISSA-LSSVM  | 1.811 | 1.2514| 0.4853|

7. Conclusion

In this paper, a hybrid model is proposed to improve the efficiency and accuracy of wind power prediction. First, VMD technology was used to decompose the original wind power data, and then ISSA was used to optimize the LSSVM parameters. Finally, a hybrid VMD and ISSA-LSSVM model with good predictive effect is established. Comparing with other prediction models, some conclusions can be obtained as follows.

1) VMD decomposition can decompose the wind power sequence with volatility and non-smoothness into several relatively stable sub-sequences. Simulation results illustrate the VMD decomposition contribute to enhance the prediction performance.

2) Compared with the DBN prediction model and the LSSVM prediction model, the model optimized by PSO and SSA has better prediction performance.

3) In this paper, Tent chaos mapping and inertial weight strategy are introduced to improve the standard SSA, which improves the local search ability and global search ability. Simulation results show that the prediction performance of ISSA-LSSVM model is better than that of SSA-LSSVM model.

4) By comparing the simulation test results of DBN model, LSSVM model, PSO-LSSVM model, SSA-LSSVM model, ISSA-LSSVM model and VMD-ISSA-LSSVM model, it can be seen that the prediction model based on VMD-ISSA-LSSVM proposed in this paper has better prediction performance.

References

[1] L. Li, X. Zhao, M. Tseng and R.R. Tan, (2020) Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm, J. Clean. Prod., vol. 242, pp. 118447.

[2] Y. Cui, C. Huang and Y. Cui, (2020) A novel compound wind speed forecasting model based on the back propagation neural network optimized by bat algorithm, Environ. Sci. Pollut. R., vol. 27, no. 7, pp. 7353-7365.

[3] Q. Wu and H. Lin, (2019) Short-Term Wind Speed Forecasting Based on Hybrid Variational Mode Decomposition and Least Squares Support Vector Machine Optimized by Bat Algorithm Model, Sustainability-Basel, vol. 11, no. 3, pp. 652.

[4] L. Zhang, B. Wang, B. Fang, H. Ma, Z. Yang and Y. Xu, (2018) Two-Stage Short-Term Wind Speed Prediction Based on LSTM-LSSVM-CFA, IEEE, pp. 1-6.

[5] Q. Zhang, Z. Tang, S. Cao and G. Wang, (2019) Wind farm wind power prediction method based on CEEMDAN and DE optimized DNN neural network, IEEE, pp. 1626-1630.

[6] S. Mirjalili, A.H. Gandomi, S.Z. Mirjalili, S. Saremi, H. Faris and S.M. Mirjalili, (2017) Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems, Adv. Eng. Softw., vol. 114, pp. 163-191.