

Economic Load Forecasting based on IPSO-VMD-DBiLSTM Network

Jiaqi Yao

Daqing Treasury Department, Daqing 163311, China

ABSTRACT

Accurate short-term load forecasting is vital for the stable operation of power systems. Given the highly volatile and nonlinear nature of load sequences, we propose a novel Deep Bidirectional Long Short-Term Memory (DBiLSTM) network incorporating Variational Mode Decomposition (VMD) and an attention mechanism to address the aforementioned challenges. In this study, the model's hyperparameters are optimized using an Improved Particle Swarm Optimization (IPSO) technique. Notably, IPSO employs a nonlinearly decreasing inertia weight to overcome the drawbacks of premature convergence and local optima.

KEYWORDS

Economic Load Forecasting; Improved Particle Swarm Optimization (IPSO); Deep Bidirectional Long Short-Term Memory (DBiLSTM); Variational Mode Decomposition (VMD).

1. INTRODUCTION

Due to the ever-increasing national power supply demand, short-term load forecasting plays a significant role in ensuring power supply operations and safety management[1]. Accurate predictions serve as the primary guarantee for enhancing the efficiency of power generation equipment and economic operational adjustments[2]. Moreover, they play a critical role in system optimization, economic operation regulation, determining optimal trends, and electricity price trading [3].

In summary, this research presents a DBiLSTM network that integrates VMD and an attention mechanism, with model hyperparameters fine-tuned by IPSO, aimed at mitigating the difficulty in feature extraction from highly volatile load sequences and improving the prediction capability for nonlinear sequence patterns. This integrated approach thus enhances the accuracy and reliability of short-term load forecasting in power systems.

2. RELATED THEORIES AND METHODS

2.1. VMD

VMD is a non-recursive adaptive method for decomposing signal sequences, particularly designed to handle high volatility load data by breaking it down into components exhibiting both stationary and non-stationary properties[4]. Furthermore, VMD inherently mitigates modal aliasing and endpoint effects. The principle of modal decomposition is used to break down the load data using VMD because it is difficult to extract the characteristics of a loading series with high volatility. The load sequence f is decomposed into several finite-bandwidth modal functions $u_k(t)$, which are some times known

as IMFs, The result is an Amplitude Modulation and Frequency Modulation (AM-FM) signal with a finite bandwidth and a central frequency[5].

2.2. LSTM Neural Network

All LSTM improves the RNN and eliminates the gradient disappearance problem using multiple threshold gates, which is divided into three parts: the forget gate, input gate, and output gate[6]. The equations for the three calculations of the LSTM neural network are as follows:

$$\begin{aligned}
 f_t &= \text{sigmoid}(W_{fx}x_t + W_{fs}s_{t-1} + b_f) \\
 i_t &= \text{sigmoid}(W_{ix}x_t + W_{is}s_{t-1} + b_i) \\
 \tilde{c}_t &= \tanh(W_{cx}x_t + W_{cs}s_{t-1} + b_c) \\
 o_t &= \text{sigmoid}(W_{ox}x_t + W_{os}s_{t-1} + b_o) \\
 \hat{c}_t &= \tanh(c_t)
 \end{aligned} \tag{1}$$

2.3. BiLSTM Neural Network

The BiLSTM neural network is derived from the LSTM neural network, consisting of two LSTM networks in two directions[7]. The hidden layer of the BiLSTM network integrates forward and backward information, allowing the network units to simultaneously obtain the sequence's front and back information. Fig. 1 depicts the BiLSTM network[8].

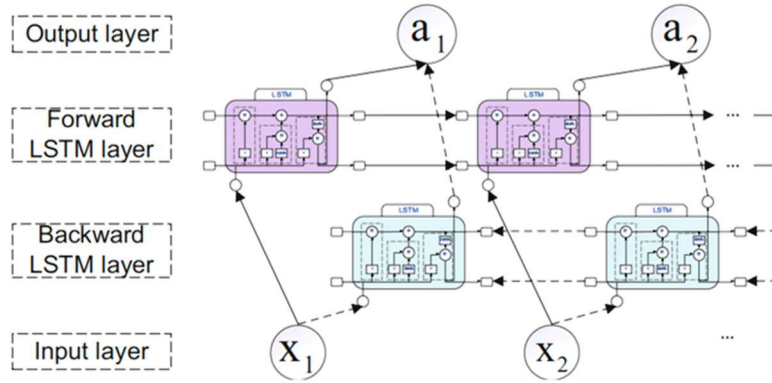


Fig. 1 Basic structure of BiLSTM

2.4. PSO

The PSO process uses a velocity and position search model. Each particle represents a candidate solution in the solution space, where the degree of solution is defined by the fitness function after the optimization objective[9]. By continuously updating the velocity and position of the particles, calculating the fitness value of the objective function, and comparing between the optimization variables of multiple-dimensional particles, the global optimum is achieved and better model parameters are obtained[10]. Equations (2) give the velocity and position update formulas for the PSO algorithm, respectively.

$$\begin{aligned}
 v_{i,j}^{t+1} &= \omega v_{i,j}^t + c_1 r_1 (pbest_{i,j}^t - x_{i,j}^t) + c_2 r_2 (gbest_j^t - x_{i,j}^t) \\
 x_{i,j}^{t+1} &= x_{i,j}^t + v_{i,j}^{t+1}
 \end{aligned} \tag{2}$$

In equation (2), ω represents the inertia weight; c_1 and c_2 are learning factors; r_1 and r_2 are two independent random numbers distributed within the interval $[0,1]$; $v_{t,i,j}$, $x_{i,j}$, $pbest_{i,j}$, and $gbest_j$ respectively denote the velocity component, position component, personal best value, and global best value of particle i at iteration t .

3. IPSO-DBiLSTM NETWORK WITH VMD

The flowchart of the deep network power load forecasting using the IPSO-DBiLSTM network with VMD and the attention mechanism is illustrated in Fig.2.

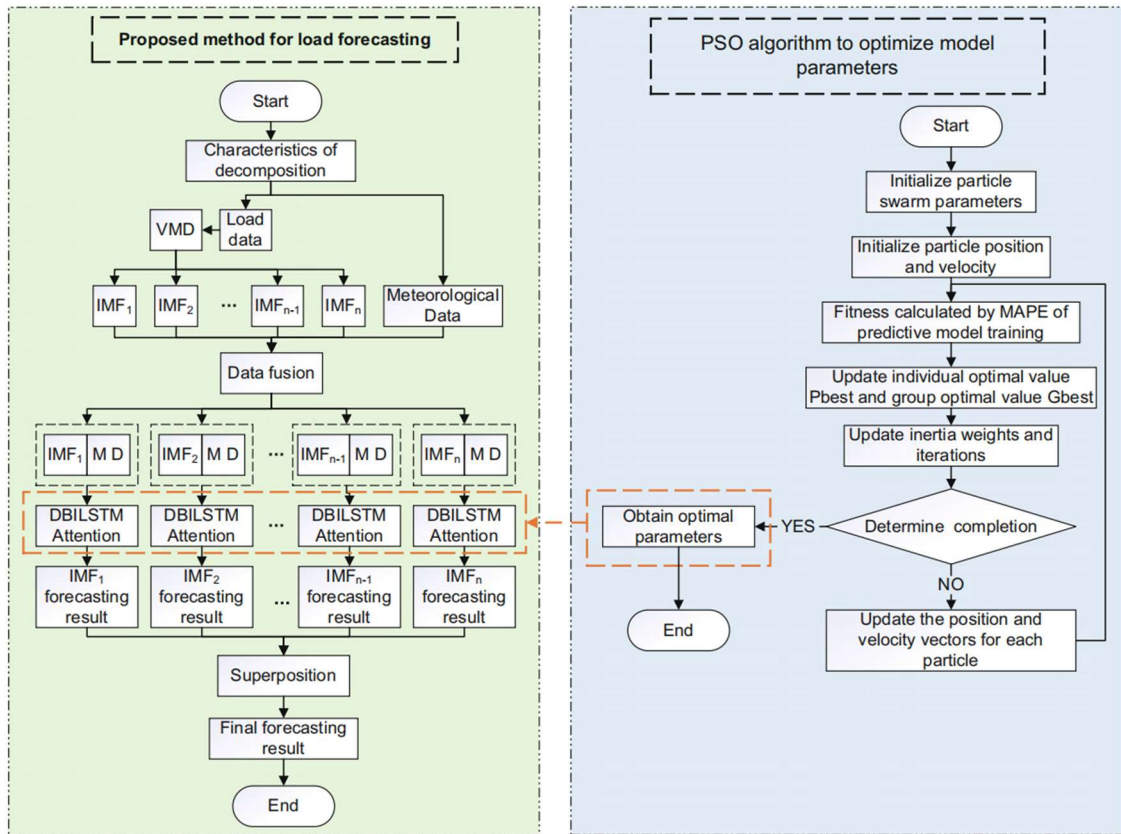


Fig. 2 Flowchart of forecasting model

The experiment is based on the short-term load data sets of the regional 1 and regional 2 power systems from the 9th Power Attribute Modeling Competition. The data collection for each region was selected to include load and meteorological data from January 1, 2012 to January 1, 2013. The load data was collected at 24 sampling points every day and measured in megawatts. The meteorological statistics include daily maximum temperature, minimum temperature, average temperature, relative humidity, and rainfall. The training set contains the experimental data of the first 355 days, while the test set contains the last 10 days.

In order to evaluate the performance of the power load prediction model proposed in this work, we evaluated LSTM, BiLSTM, BiLSTM-attention, DBiLSTM-attention, and TCN techniques. By comparing the evaluation metrics of multiple models, the accuracy of the experiment was verified. The following are comparisons of models:

The outcomes of several models' predictions are shown in Fig. 3 for the test set. The approach provided in this research shows excellent outcomes in various indicators when compared to the other

models. The evaluation index data in Table 1, On the evaluation index, we can see that the MAPE of this model is 1.07%, which is 62.5%, 63.2%, 57.9%, 50.2%, and 59.5% better than the other models, respectively. The R2 of this experiment reaches 99.57%, which is almost close to 1. All evaluation indicators show that the model proposed in this paper has better prediction results.

This experiment's accuracy is demonstrated by the comparison of the several experiments in this section.

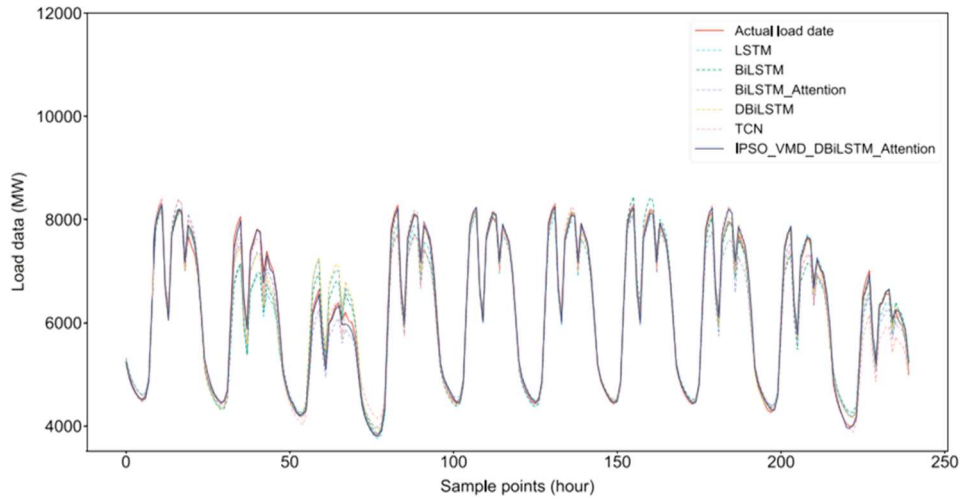


Fig. 3 load power of different forecasting methods

Table 1. evaluation indicators of different forecasting models

Model	MAPE (%)	RMSE(MW)	MAE(MW)	R ₂	time(s)
LSTM	2.84	277.3381	182.8632	0.9596	11.6719
BiLSTM	2.91	270.7034	186.6629	0.9616	18.3278
BiLSTM-Attention	2.53	210.3687	160.3975	0.9768	18.4167
DBiLSTM	2.15	213.6752	133.5019	0.976	28.851
TCN	2.64	242.4138	166.5851	0.9692	23.7522
IPSO-VMD-DBiLSTM-Attention	1.08	90.6845	67.1109	0.9957	500.4397

The single-model approach's DBiLSTM has a relatively long run time, while the hybrid prediction model based on VMD decomposition has a higher computational complexity than the single model. Clearly, the method proposed in this article has the highest computational cost in terms of run time. However, in practical applications, the accuracy of power load forecasting will be prioritized, and the need for computational cost will be moderate.

4. CONCLUSION

This study proposes a DBiLSTM network based on VMD and attention mechanism, and optimizes the hyperparameters of the model using IPSO. This method improves the modality extraction of strong volatile load sequences and the prediction of nonlinear sequences. First, VMD is used to decompose strong volatile load sequences into stationary and non-stationary sequences with different characteristics. Secondly, the decomposition pattern is passed to DBiLSTM for training, and then

selective attention to hidden layer state features is passed to the attention mechanism. Finally, the model hyperparameters are optimized using IPSO algorithm, effectively improving the prediction results of load sequences with strong volatility and nonlinearity.

REFERENCES

- [1] BOZORG M, BRACALE A, CARAMIA P, et al. Bayesian bootstrap quantile regression for probabilistic photovoltaic power forecasting[J]. *Protection and Control of Modern Power Systems*, 2020, 5(3):218-229.
- [2] SHA B, YNA C, MN A, et al. A Comparative Study of Predictive Approaches for Load Forecasting in Smart Buildings[J]. *Procedia Computer Science*, 2019, 160 :173-180.
- [3] HE F F, ZHOU J Z, MO L, et al. Day-ahead short-term load probability density forecasting method with a decomposition based quantile regression forest[J]. *Applied Energy*, 2020, 262(9):234-246.
- [4] JUAN F, RENDON S. Structural combination of seasonal exponential smoothing forecasts applied to load forecasting[J]. *European Journal of Operational Research*, 2019, 275(3):916-924.
- [5] Fallah SN, Deo RC, Shojafar M, Conti M, Shamshirband S (2018) Computational intelligence approaches for energy load forecasting in smart energy management grids: State of the art, future challenges, and research directions. *Energies* 11(3):596.
- [6] Huang N, Lu G, Xu D (2016) A permutation importance based feature selection method for short-term electricity load forecasting using random forest. *Energies* 9(10):767.
- [7] Zheng Z, Chen H, Luo X (2018) Subsampled support vector regression ensemble for short term electric load forecasting. *Energy* 164(10):160–170. <https://doi.org/10.1016/j.energy.2018.08.169>.
- [8] Wang K, Qi X, Liu H (2019) A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network. *Appl Energ* 251:113315.
- [9] Sehovac L, Grolinger K (2020) Deep learning for load forecasting: Sequence to sequence recurrent neural networks with attention. *IEEE Access* 8:36411–36426.
- [10] Wang S, Wang X, Wang S, Wang D (2019) Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting. *Int J Electr Power Energy Syst* 109:470–479.