

AFDD-LSTSNET-SARIMA Based Short-Term Net Load Forecasting for Customers

Jingdong Wang

Jianchen Zhou (✉ 20102248@neepu.edu.cn)

Zhu Cheng

Guangqiang Qu

Research Article

Keywords: Short-term net load forecasting, Combinatorial model, Adaptive frequency domain decomposition method, ong- and short-term recurrent jump neural network model, seasonal differential integrated moving average autoregressive model

Posted Date: March 29th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-2739982/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

AFDD-LSTSNET-SARIMA Based Short-Term Net Load Forecasting for Customers

Jingdong Wang¹, Jianchen Zhou^{1*}, Zhu Cheng¹, Guangqiang Qu^{1†}

^{1*}School of Computer Science, Northeastern Electric Power University,
Changchun Street, Jilin, 132000, Jilin, China.

*Corresponding author(s). E-mail(s): 20102248@neepu.edu.cn;
Contributing authors: 707569380@qq.com; czzc898726506@qq.com;
[†]These authors contributed equally to this work.

Abstract

In order to improve the accuracy of short-term load forecasting for users, an adaptive frequency-domain decomposition (AFDD), long-short term skip neural network (LSTSNET) and the AFDD-LSTSNET-SARIMA short-term load forecasting model is proposed by combining the seasonal autoregressive integrated moving average model (SARIMA). Firstly, AFDD is used to find the optimal combination of all potential periodicity in the frequency domain by using genetic algorithm, thus decomposing the load into periodic and non-periodic components. Based on this, SARIMA and LSTSNET are used to predict the periodic and acyclic components respectively; finally, the prediction results of each component are recombined to obtain the final prediction results of short-term net load of users. The AFDD-LSTSNET-SARIMA model is validated on the measured data set of a community in New York State, U.S.A. The experimental results show that the proposed model overcomes the problem of high randomness and fluctuation of net load and effectively improves the accuracy of short-term net load prediction of customers.

Keywords: Short-term net load forecasting, Combinatorial model, Adaptive frequency domain decomposition method, ong- and short-term recurrent jump neural network model, seasonal differential integrated moving average autoregressive model

1 Introduction

"Carbon neutrality" is the vision of human society, and the efficient use of new energy is one of the effective ways to achieve the goal of carbon neutrality. At present, new energy is distributed in the form of "energy cells" on the customer side, and has the characteristics of plug-and-play[1]. Microgrid is the main form of new energy consumption, and its short-term load forecast is the basis for the stable operation of smart grid[2]. New microgrids are usually composed of a large number of customers with new energy generation equipment[3]. The ability to accurately predict the short-term net load of customers, i.e. the difference between electricity consumption and PV generation, will provide an important guarantee for the consumption of new energy in the whole region[4].

Researchers have conducted extensive studies on load forecasting. Prediction methods have transitioned from linear regression[5], differential integrated moving average autoregressive models[6], support vector machines[7], and BP neural networks[8] to recurrent neural networks[9] and convolution neural network (CNN)[10]; prediction models have evolved from single models such as long short term memory neural network (LSTM), gated recurrent unit neural network (GRU), to a single model such as gated recurrent unit neural network (GRU) . (LSTM)[11], gated recurrent unit neural network (GRU)[12], to RNN-CNN[13], GRU-CNN[14] and other combined models. However, the net load of customers is affected by many complex factors such as natural conditions, individual electricity consumption behavior, and electricity price, which makes it difficult for these methods to accurately predict the short-term net load of customers[15].

In recent years, wavelet transform[16, 17], empirical mode decomposition (EMD)[18, 19], frequency domain decomposition[20] and other methods adopt the idea of decomposing the load before predicting the components, which effectively solve the problems of strong randomness and large fluctuation of the load. However, there are still some shortcomings, for example, it is difficult to find the appropriate mother wavelet and decomposition layers for wavelet transform; EMD has problems of over-envelope, under-envelope, modal confusion and large amount of noise; frequency domain decomposition method has the disadvantage of difficult to choose the frequency domain.

In order to improve the accuracy of short-term net load prediction of users, this paper proposes a short-term load prediction model based on AFDD-LSTSNET-SARIMA. Firstly, the net load data is transferred from the time domain to the frequency domain by Fourier transform, and then the optimal combination of periods is found among all potential periodicity in the frequency domain by genetic algorithm, so that the net load data is decomposed into periodic and non-periodic components. In order to improve the prediction accuracy at the fluctuation of non-periodic components, an LSTSNET model consisting of convolutional layer, CRU unit and GRU jump unit, and fully connected layer is proposed. The periodic components have obvious regularity, and the SARIMA algorithm is chosen to predict the periodic components to ensure the prediction accuracy while saving computational resources. On this basis, the prediction results of each component are recombined to obtain the final prediction results of short-term net load of users. Compared with the single prediction model,

the combined AFDD-LSTSNET-SARIMA model can obtain higher prediction accuracy by combining linear and nonlinear prediction methods on the basis of solving the problem of difficult frequency domain selection of frequency domain decomposition method.

2 Technical framework for combined predictive models

In this paper, after decomposing the net load data using the adaptive frequency domain decomposition method, the load components will show certain regularity. According to the different characteristics of the load components, different forecasting methods are used to solve the problems of large uncertainty and fluctuation of the net user load, so as to improve the forecasting accuracy. Figure 1 shows the flow diagram of the prediction model based on the adaptive frequency domain decomposition of the net user load combination.

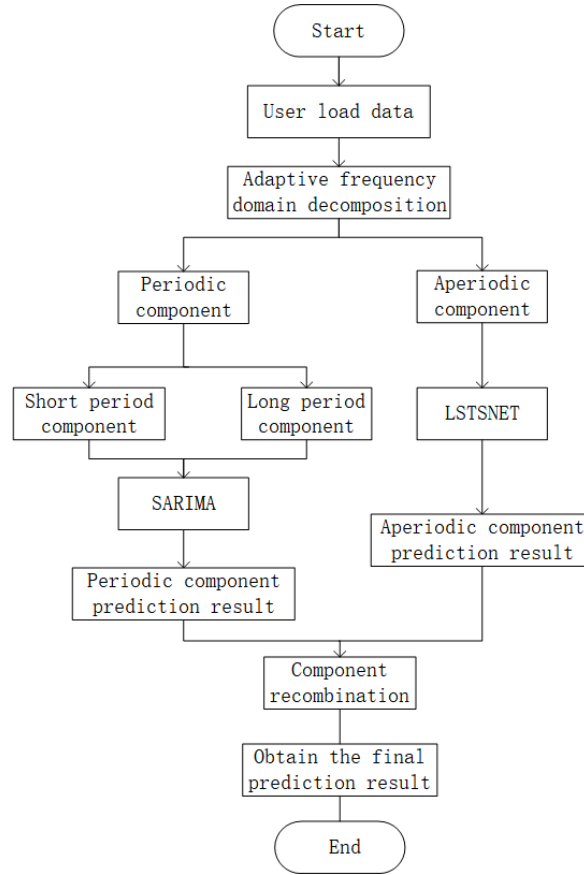


Fig. 1 Schematic diagram of the combined prediction process

The specific prediction process is as follows:

(1) The user load data is decomposed into periodic and non-periodic components using adaptive frequency domain decomposition.

(2) The periodic component has obvious regularity, and it can be accurately predicted only by mining the change pattern from the data itself. SARIMA predicts the data by mining the periodic pattern of the historical data itself, which is consistent with the characteristics of the periodic component itself. Therefore, the SARIMA model is used to predict the periodic components.

(3) To address the problems of non-periodic components with obvious volatility and instability, this paper proposes a long and short-term cyclic jump convolutional neural network, which uses convolutional layers to capture local feature relationships and adds a jump link in the cyclic process of GRU units to achieve the ability of mining the hidden and periodic relationships among load data while preserving the nonlinear fitting ability of GRU units, which improves the non-periodic components. The prediction accuracy is improved.

(4) The periodic and non-periodic components obtained from the above two algorithms are recombined to obtain the final net load prediction results.

3 Technical details of the combined predictive model

3.1 Adaptive frequency domain decomposition method based on genetic algorithm

The frequency domain decomposition method uses the Fourier transform to convert the load data from time domain to frequency domain, and then combines the components of different frequencies to obtain new load components and decompose them from the original load, so that the original irregular electric load is decomposed into components with different regular changes. Therefore, the frequency domain decomposition method is applicable to the time series data of power load with nonlinearity and non-smoothness. However, the frequency domain decomposition method is a data processing method that requires a priori experience, i.e., the frequency domain periodicity is determined before extraction. In the process of decomposition of net customer load, if only the daily and weekly components are extracted, the best regularity of the load cannot be found, i.e., the ideal non-periodic components are not obtained.

The adaptive frequency domain decomposition method based on genetic algorithm is an improved method of frequency domain decomposition method. After converting the load data to frequency domain, the genetic algorithm is used to extract all the periodic components in combination to find out the best regular components in the historical data, which improves the problem of high uncertainty and volatility of net load of users, and thus improves the prediction accuracy. Figure 2 shows the flow chart of the adaptive frequency domain decomposition method based on genetic algorithm.

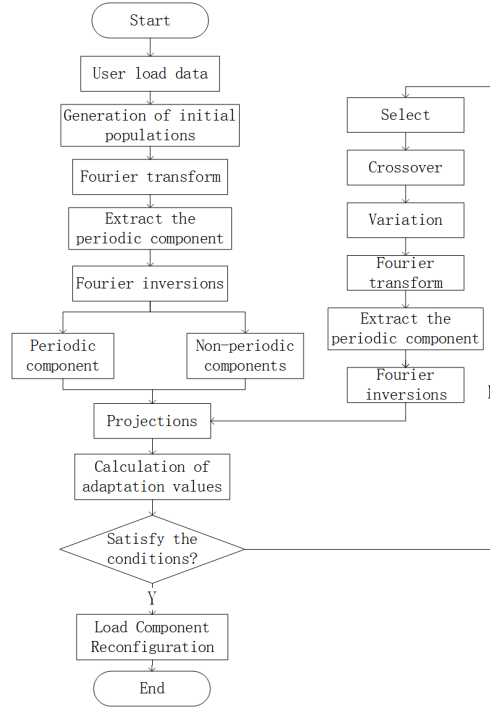


Fig. 2 Flow chart of adaptive frequency domain decomposition method based on genetic algorithm

The specific steps are:

(1) After Fourier decomposition, the user load data are orthogonal to each other, and the periodic components of the original load are extracted with different periods. The original load $P(t)$ is reconstructed as follows:

$$P(t) = a_0 + S(t) + L(t) + A(t) \quad (1)$$

where: $a_0 + S(t)$ constitutes the short-period component, $L(t)$ is the long-period component, and $A(t)$ is the acyclic component.

(2) Divide all extractable periods into long-period components and short-period components. $\text{mod}(m, n)mn$ The short period $S(t)$ is α and contains the DC component a_0 . $a_0 + S(t)$ The set of frequencies of :

$$\Omega_{short} = \{w_0\} \cup \left\{ w_i \mid \text{mod}\left(\alpha, \frac{2\pi}{w_i}\right) = 0 \right\} \quad (2)$$

In this paper, we set the long period $L(t)$ period to b . $L(t)$ The frequency set of :

$$\Omega_{long} = \left\{ w_i \mid \text{mod}\left(b, \frac{2\pi}{w_i}\right) = 0 \cap \text{mod}\left(a, \frac{2\pi}{w_i}\right) \neq 0 \right\} \quad (3)$$

(3) All potential long and short periods are encoded using 0 or 1, and then the initial population is generated, and the corresponding period combination scheme of the

population is transformed to the time domain by Fourier inverse transform to obtain the short period component, long period component, and non-periodic component.

(4) The SARIMA and LSTSNET prediction models are used to predict the periodic and non-periodic components, respectively, and the obtained results are used to find the optimal periodic components using the fitness function, i.e., to find a and b in Eqs. (4) and (5). the fitness functions are shown below:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y_i^*)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

Where: y_i^* is the predicted value at the i time point; y_i is the actual value at the i time point; $\bar{y}$ is the mean of the actual values; n is the sample size.

3.2 Periodic component forecasting based on SARIMA model

In the field of load forecasting, SARIMA is the most commonly used seasonal time series forecasting technique. SARIMA is obtained by adding seasonality analysis based on ARIMA model. SARIMA model is generally expressed as SARIMA (p, d, q) (P, D, Q) (s), which is in the form of:

$$\varphi_p(B)\phi_P(B^s)(1 - B^S)^D y_t = \theta(B)\Theta_Q(B^s)\varepsilon_t \quad (5)$$

Where, Y_t is the predicted value, ε_t is the residual, P is the autoregressive model order, P is the seasonal autoregressive model order, q is the moving average model order, Q is the seasonal moving average model order, d is the number of differences, D is the number of seasonal differences, $\varphi_p(B)$ is the autoregressive model, $\phi_P(B^s)$ is the seasonal autoregressive model, $\theta_q(B)$ is the moving average model, and $\Theta_Q(B^s)$ is the seasonal moving average model.

After testing the data for smoothness using the unit root test, SARIMA converts the non-smooth series into a smooth series by using the difference operation and the seasonal difference operation. The autocorrelation and partial autocorrelation plots are used to find the possible parameters of p, q, P, Q in the SARIMA model. When the SARIMA algorithm predicts the periodic components, the model parameters are determined by the above process, and the optimal model is found by using the deficit pool information criterion and Bayesian information criterion, and the degree of fit of SARIMA to the data is checked by the coefficient of determination. When the fit degree meets the requirement, the prediction result of the periodic component is output.

3.3 Non-periodic component prediction based on LSTSNET model

The non-periodic components obtained by the adaptive frequency domain decomposition method reflect the stochastic fluctuations of PV power generation and customer load, which have strong stochasticity. In response to the above situation, we propose a deep learning prediction framework based on GRU unit improvement, i.e., long and short term recurrent jump neural network model, whose model is shown in Fig. 3.

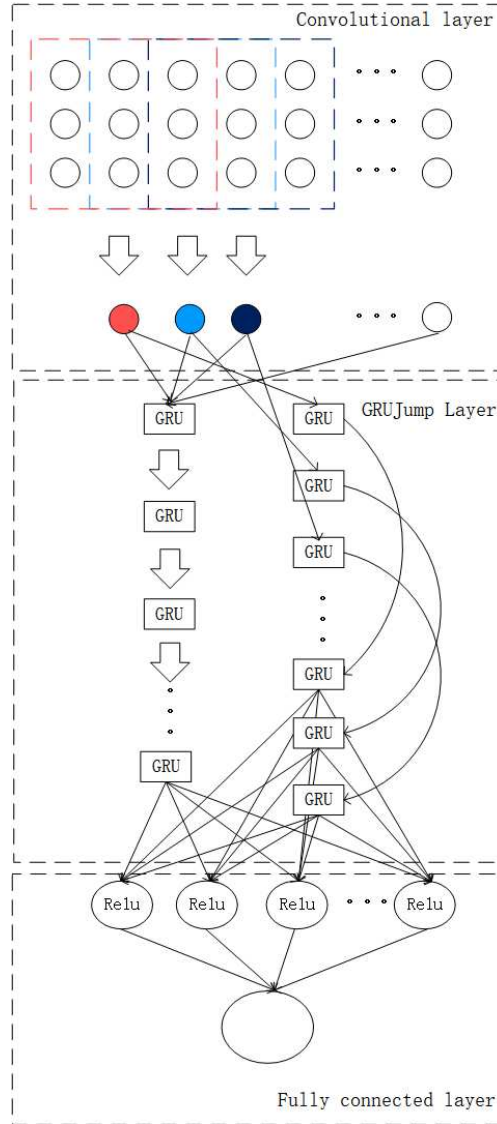


Fig. 3 LSTSNET model framework diagram

The first layer of the LSTSNET algorithm is a convolutional layer without pooling, which captures the local dependencies between load data and features using convolution. The convolutional layer consists of multiple filters with a width of w (determined by the number of features) and a height of h . The second layer of the LSTSNET algorithm is a GRU unit and a GRU jumping unit. GRU is designed to mine the relationship between data in time series, but when the length of time series features is too large, GRU is often unable to capture the long-term hidden relationships between data as well. By adding a time-hopping loop structure, GRU can improve the ability of

mining data with a large time span. The mathematical expression of the loop-hopping process is shown below:

$$\begin{aligned}
r_t &= \sigma(x_t w_{xr} + h_{t-p} w_{hr} + b_r) \\
u_t &= \sigma(x_t w_{xu} + h_{t-p} w_{hu} + b_u) \\
c_t &= RELU(x_t W_{xc} + r_t \odot (h_t - pW_h v) + b_c) \\
h_t &= (1 - u_t) \odot h_t - p + u_t \odot c_t.
\end{aligned} \tag{6}$$

where $\odot$ is the product of elements, σ is the sigmoid function, RELU is the RELU function, $w_{xr}, w_{xu}, w_{hr}, w_{hu}, w_{xc}, w_{hc}$ is the weight of the input variables in the corresponding gate, b_r, b_u, b_c is the bias term of the corresponding gate, x_t is the input variable, $h_t - p$ is the state memory variable of the previous p moment, h_t is the state memory variable of the current moment, r_t is the reset gate state, u_t is the update gate state, c_t is the current waiting candidate set state.

The third layer of the LSTSNET algorithm is the fully-connected layer, and the inputs of the fully-connected layer are the predicted values of GRU units and jump GRU units. The input expressions of the fully connected layer are shown below:

$$h_t = W^R h_t^R + \sum_{i=0}^{p-1} W_i^S h_{t-i}^S + b \tag{7}$$

Where: h_t^R is the output value of GRU unit, $h_t - i^S$ is the output value of long jump cycle unit.

4 Example analysis

4.1 Experimental data presentation and analysis

The data in this study are obtained from 35 days of measured net load data and weather data of 26 households in a smart community in Brookdale, New York, USA, from May 1, 2019 to June 4, 2019. All 26 customers are equipped with distributed power (photovoltaic generation equipment) and self-generated. The smart meter collection period is 15 minutes, that is, 96 points a day for one household, a total of 3360 points for 35 days, of which 30 days is the training set and 5 days is the test set, and the raw data are shown in Figure 4

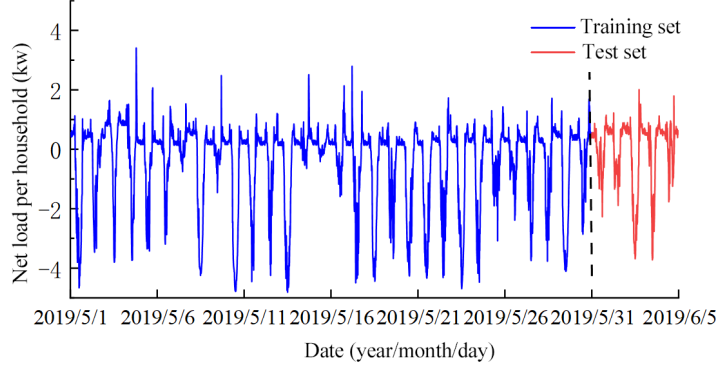


Fig. 4 Dataset of 1 household net load

From Fig. 4 , it can be seen that neither the single household net load data nor the total net load data of 26 households are cyclical, and the daily peak-to-valley differences fluctuate significantly and in different ranges.

4.2 User net load forecasting feature set

The net user load consists of both PV power generation and load, so in addition to factors affecting load such as temperature, humidity and previous hour load, factors affecting PV power generation such as PV panel temperature and sunlight radiation should also be added. In summary, the input parameters of the model constructed in this paper are shown in Table 1:

Table 1 Model input parameters

Parameter	Meaning
x_1	Dew point temperature
x_2	Humidity
x_3	Photovoltaic panel temperature
x_4	Amount of sunlight radiation
x_5	First 15 minutes load
x_6	First 30 minutes load
x_7	First 45 minutes of load
x_8	First 60 minutes load

The above five eigenvectors are arranged into 3360 96×8 two-dimensional matrices on the principle of time series, as shown in equations (8) and (9).

$$X_{N,96 \times 8} = \begin{pmatrix} X_{1,1} & \cdots & X_{1,1} \\ \vdots & \ddots & \vdots \\ X_{1,1} & \cdots & X_{1,1} \end{pmatrix} \quad (8)$$

$$f = \begin{pmatrix} X_{1,96 \times 8} \\ \vdots \\ X_{3360,96 \times 8} \end{pmatrix} \quad (9)$$

4.3 SARIMA and LSTNET hyperparameter settings

The details of the implementation of several major algorithms in this paper are as follows:

The LSTNET model is implemented using the framework keras and tensorflow, and the hyperparameters of LSTNET mainly include, convolutional kernel size, training number, learning rate and so on. All hyperparameters are set as shown in Table 2.

Table 2 Model input parameters

Parameter Name	Parameter Value	Parameter Name	Parameter Value
Epoch	50	GRU-unit	10
Batch size	128	Skip-GRU-unit	10
Learn rate	0.01	dropout	0.5
Kernel size	3*5	skip	24

The SARIMA model is implemented using the framework sklearn. The hyperparameters of SARIMA include p = regression order, d = difference order and q = moving average order, P = seasonal autoregressive order, D = seasonal difference order and Q = seasonal moving average order. The hyperparameters of SARIMA include regression order, difference order and moving average order, seasonal autoregressive order, seasonal difference order and seasonal moving average order.

In this paper, the mean absolute percentage error E_{MAPE} , root mean square error E_{RMSE} , is used as the performance evaluation index of the prediction model, and the mathematical expression is as follows:

$$E_{MAPE} = \frac{\sum_{i=1}^n \left(\frac{|y_i - y_i^*|}{y_i} \right)}{n} \quad (10)$$

$$E_{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - y_i^*)^2}{n}} \quad (11)$$

4.4 Comparison and analysis of prediction results

In this paper, the load components obtained from the adaptive frequency domain decomposition of the data in Figs. 4 above are used to train the corresponding prediction models, and the previous hourly load data are processed in a dynamic way during the training process, i.e., after the prediction points are obtained, the prediction points are added to the feature series as new features to continue predicting the

next value. Fig. 5 to Fig.10 show the results of adaptive frequency domain decomposition for single-user and 26-user net load, in which the optimal decomposition frequency domain for 26-user total net load data is, $a = 480$ $b = 1680$, and the optimal decomposition frequency domain for 1-user load data is, $a = 480$ $b = 672$.

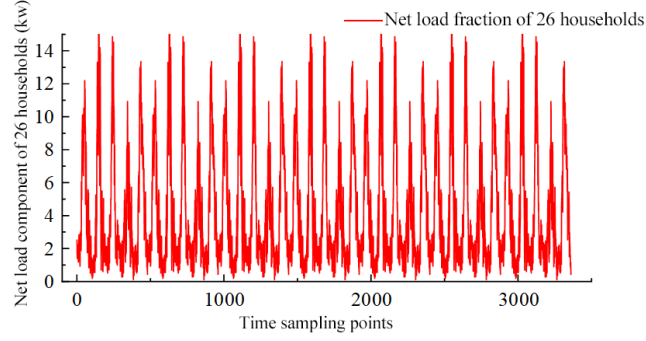


Fig. 5 Short period component

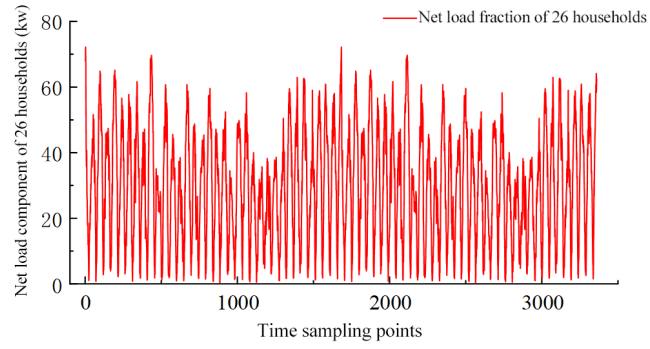


Fig. 6 Long period component

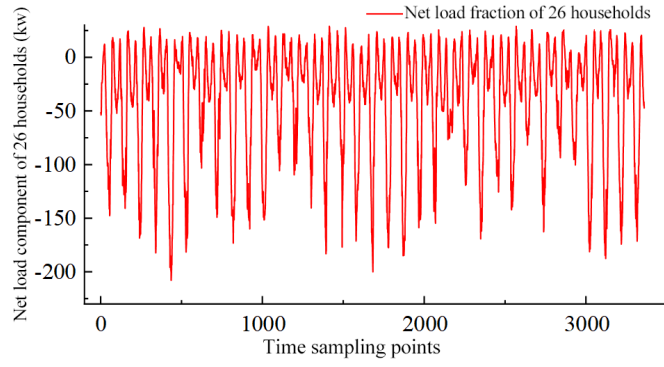


Fig. 7 Non-periodic component

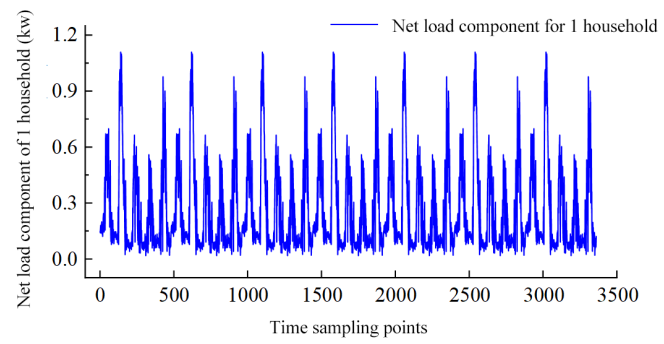


Fig. 8 Short period component

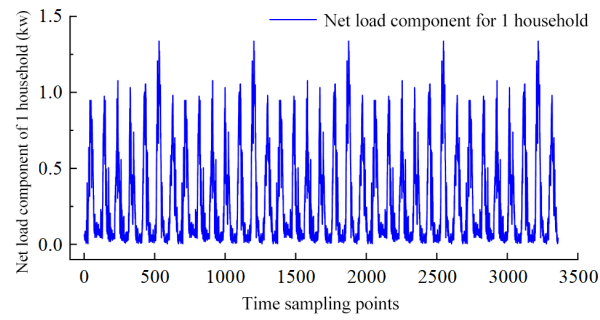


Fig. 9 Long period component

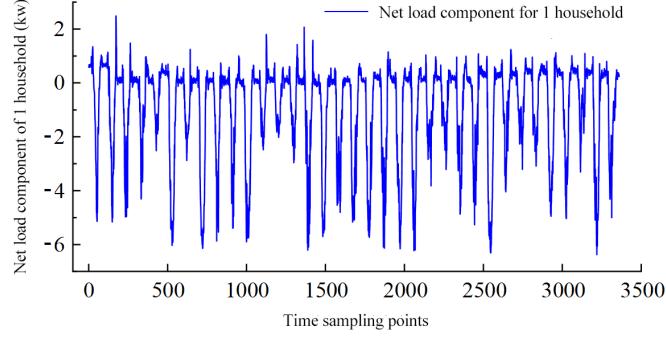


Fig. 10 Non-periodic component

From Fig. 5 to Fig. 10, it can be seen that after the adaptive frequency domain decomposition, the non-periodic load components peaks and troughs of both 26-home net load and single-home net load appear with certain patterns, which reduces the prediction difficulty. Short-term net load forecasting is carried out for single households of 1 day and 5 consecutive days in the dataset using single forecasting models such as LSTM, GRU, LSTSNET and combined models such as frequency domain decomposition method, CNN-LSTM, AFDD-LSTSNET-SARIMA. Among them, the periodic components obtained by frequency domain decomposition method and adaptive frequency domain decomposition method are predicted by SARIMA algorithm, and the non-periodic components are predicted by LSTSNET. The mean absolute percentage error E_{MAPE} and root mean square error E_{RMSE} of the above models on the test set were calculated and the results are shown in Table 3.

Table 3 Evaluation indicators of different models for single households

Predictive Models	2019/5/31		2019/5/31-2019/6/5	
	RMSE	MAPE	RMSE	MAPE
LSTM	0.1987	0.4411	0.2576	0.4493
GRU	0.1913	0.4316	0.2628	0.4376
CNN-GRU	0.1747	0.3707	0.2229	0.3667
LSTSNET	0.1111	0.2473	0.1346	0.2318
Frequency domain decomposition method	0.1052	0.2590	0.1282	0.2383
AFDD-LSTSNET-SARIMA	0.0855	0.1818	0.1195	0.1270

As shown in Table 3, from the analysis of E_{MAPE} indicators, the prediction effect of single algorithms such as LSTM and GRU is not as good as the combined models such as CNN-LSTM, LSTSNET, frequency domain decomposition and adaptive frequency domain decomposition method. The reason is that the combined models incorporate the convolutional structure of the convolutional neural network, and the feature vectors are more closely related to each other after passing through the convolutional layer, which has a better effect in predicting the peaks and valleys of waves.

The LSTSNET is more effective in capturing the relationship between long interval data by adding jump GRU units during training. The frequency domain decomposition method and adaptive frequency domain decomposition method are more effective than other methods because the load data are decomposed in the frequency domain, which changes the original pattern of the load, making the peaks and valleys regular in terms of values and trends. Fig. 11 shows the prediction curves comparing the prediction model and the model proposed in this paper for 1 day, and Fig. 11 shows the prediction curves for 5 days.

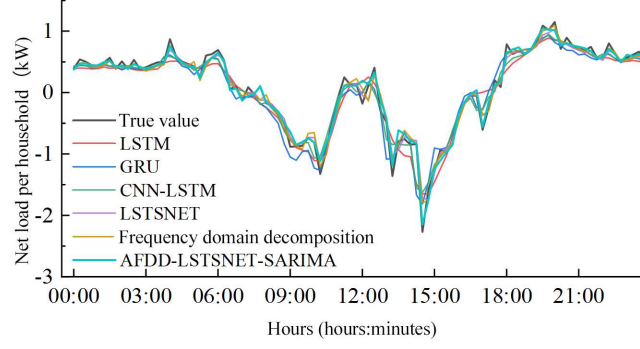


Fig. 11 Single household 1-day multiple model prediction results

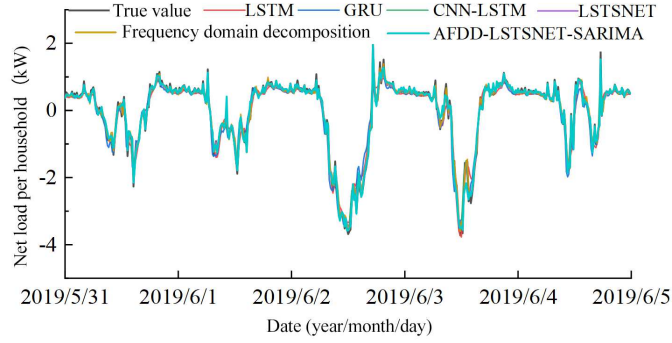


Fig. 12 Single Household 5-Day Multiple Model Forecast Results

AFDD-LSTSNET-SA RIMA model proposed in this paper predicts better than the other five models at the peaks and troughs, and also restores the trend of net user load. Fig. 12 proves that the AFDD-LSTSNET-SA RIMA model can also predict the sudden load changes better in the long-term prediction process, but there are also prediction errors, which may be caused by the sharp increase in load at the corresponding moment in the training set, and therefore follow the previous pattern in

the smooth section. In order to further verify the effectiveness of the proposed model and the prediction accuracy of the model under the fluctuation of large values of load, the total net load of 26 households in the district was predicted by using the above six models, and the prediction evaluation indexes are shown in Table 4.

Table 4 Evaluation indexes of different models for 26 households

Predictive Models	2019/5/31		2019/5/31-2019/6/5	
	RMSE	MAPE	RMSE	MAPE
LSTM	4.595	0.2583	7.448	0.2840
GRU	3.6491	0.2419	6.5043	0.2722
CNN-GRU	3.6774	0.2387	5.6546	0.3300
LSTSNET	3.278	0.2202	5.4350	0.2300
Frequency domain decomposition method	2.8642	0.3272	4.1800	0.1745
AFDD-LSTSNET-SARIMA	2.5658	0.1663	3.3742	0.1521

single-family, all the prediction models except the AFDD-LSTSNET-SARIMA model have decreased in E_{MAPE} , but the AFDD-LSTSNET-SARIMA model still has the lowest values of E_{MAPE} and E_{RMSE} , which once again proves the superiority of the proposed method in the paper.

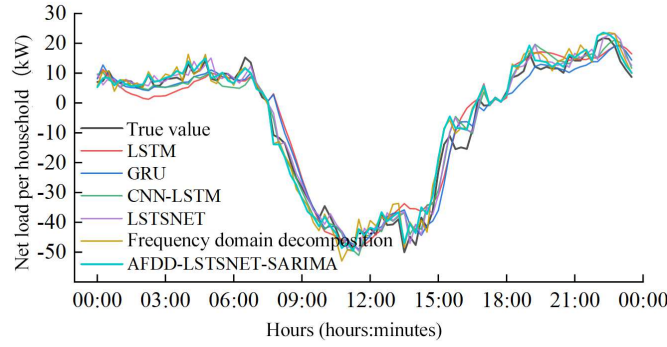


Fig. 13 Forecasting results of different models for 26 households 1 day

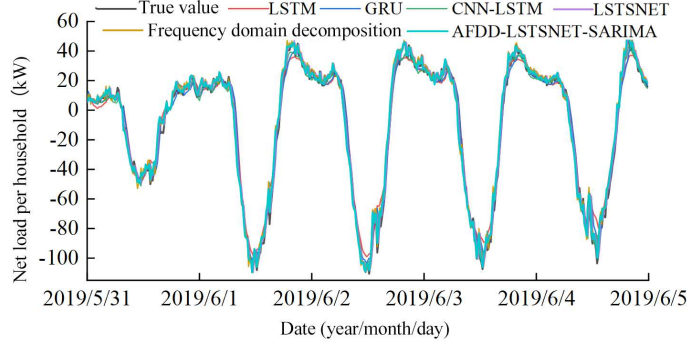


Fig. 14 Forecast results of multiple models for 26 households for 5 days

It can be seen from Fig. 10 and Fig. 11 that the load fluctuates continuously at the peaks and troughs, and the single prediction model is less effective in facing such situations, while the combined model performs better. The AFDD-LSTS-NET-SARIMA proposed in this paper is more stable, without large error points, and almost restores the load variation trend.

5 Conclusion

In this paper, we propose a combined prediction model based on AFDD-LSTSNET-SARIMA. The adaptive frequency domain decomposition method finds two optimal frequency domains among all decomposable frequency domains, i.e., the prediction accuracy of the separated non-periodic variation components is the highest. The LSTSNET model uses convolutional layers to explore the relationship between adjacent data, while adding jumping GRU units to strengthen the connection between data with large span, which improves the prediction accuracy of non-periodic components. The experiments show that the integrated prediction model proposed in this paper is better than the frequency domain decomposition method with fixed frequency domain and other common models.

6 Declarations

Project Support: Architecture Design and Verification of Multi-Value Chain Collaborative Data Space Management Engine and Management System of Ministry of Science and Technology (Project No. : 2020YFB1707804)

References

- [1] Iván Patrao, Emilio Figueres, Gabriel Garcerá, and Raúl González-Medina. Microgrid architectures for low voltage distributed generation. *Renewable and Sustainable Energy Reviews*, 43:415–424, 2015.
- [2] Ding Yong Mao Meiqin and Zhang Luchen Wang Yangyang. Microgrid-an “organic cell” for future energy interconnection system. *CITIES*, 9:2, 2017.

- [3] Liang Caihao and Duan Xianzhong. Distributed generation and its impact on power system. *Automation of Electric Power Systems*, 25(12):53–56, 2001.
- [4] Aqdas Naz, Nadeem Javaid, Muhammad Babar Rasheed, Abdul Haseeb, Musaed Alhussein, and Khursheed Aurangzeb. Game theoretical energy management with storage capacity optimization and photo-voltaic cell generated power forecasting in micro grid. *Sustainability*, 11(10):2763, 2019.
- [5] Grzegorz Dudek. Pattern-based local linear regression models for short-term load forecasting. *Electric power systems research*, 130:139–147, 2016.
- [6] Li Wei and Zhang Zhen-gang. Based on time sequence of arima model in the application of short-term electricity load forecasting. In *2009 International Conference on Research Challenges in Computer Science*, pages 11–14. IEEE, 2009.
- [7] Sajjad Fattaheian-Dehkordi, Alireza Fereidunian, Hamid Gholami-Dehkordi, and Hamid Lesani. Hour-ahead demand forecasting in smart grid using support vector regression (svr). *International transactions on electrical energy systems*, 24(12):1650–1663, 2014.
- [8] Ling Juan Li and Wen Huang. A short-term power load forecasting method based on bp neural network. In *Applied Mechanics and Materials*, volume 494, pages 1647–1650. Trans Tech Publ, 2014.
- [9] Mohammad Navid Fekri, Harsh Patel, Katarina Grolinger, and Vinay Sharma. Deep learning for load forecasting with smart meter data: Online adaptive recurrent neural network. *Applied Energy*, 282:116177, 2021.
- [10] Hossein Javedani Sadaei, Petrônio Cândido de Lima e Silva, Frederico Gadelha Guimaraes, and Muhammad Hisyam Lee. Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series. *Energy*, 175:365–377, 2019.
- [11] Runhai Jiao, Tianming Zhang, Yizhi Jiang, and Hui He. Short-term non-residential load forecasting based on multiple sequences lstm recurrent neural network. *IEEE Access*, 6:59438–59448, 2018.
- [12] Yusen Wang, Wenlong Liao, and Yuqing Chang. Gated recurrent unit network-based short-term photovoltaic forecasting. *Energies*, 11(8):2163, 2018.
- [13] Junhong Kim, Jihoon Moon, Eenjun Hwang, and Pilsung Kang. Recurrent inception convolution neural network for multi short-term load forecasting. *Energy and buildings*, 194:328–341, 2019.
- [14] Lizhen Wu, Chun Kong, Xiaohong Hao, and Wei Chen. A short-term load forecasting method based on gru-cnn hybrid neural network model. *Mathematical problems in engineering*, 2020:1–10, 2020.
- [15] Ke Wang, Jianguo Yao, Liangzhong Yao, Shenchun Yang, and Taiyou Yong. Survey of research on flexible loads scheduling technologies. *Automation of electric power systems*, 38(20):127–135, 2014.
- [16] Eric Ofori-Ntow Jnr, Yao Yeveyo Ziggah, and Susana Relvas. Hybrid ensemble intelligent model based on wavelet transform, swarm intelligence and artificial neural network for electricity demand forecasting. *Sustainable Cities and Society*, 66:102679, 2021.
- [17] Yang Zhang, Caibo Deng, Ran Zhao, et al. A novel integrated price and load forecasting method in smart grid environment based on multi-level structure.

- Engineering Applications of Artificial Intelligence*, 95:103852, 2020.
- [18] Luca Ghelardoni, Alessandro Ghio, and Davide Anguita. Energy load forecasting using empirical mode decomposition and support vector regression. *IEEE Transactions on Smart Grid*, 4(1):549–556, 2013.
 - [19] Ye Ren, PN Suganthan, and Narasimalu Srikanth. A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods. *IEEE Transactions on Sustainable Energy*, 6(1):236–244, 2014.
 - [20] Qian Zhang, Yuan Ma, Guoli Li, Jinhui Ma, and Jinjin Ding. Short-term load forecasting based on frequency domain decomposition and deep learning. *Mathematical Problems in Engineering*, 2020:1–9, 2020.