Chapter 9
A Joint Data-Physics-Knowledge Driven Strategy for Electric Heating Load Forecasting and Scheduling

Jie Zhang, Hao Shen, Ling Qi, and Yongjie Chen

Abstract In the background of double carbon, electric heating technology is the development trend of heating method, which is conducive to energy saving and emission reduction. As the heating load is random, volatile, and not easy to regulate, a fused electric heating scheduling method is proposed. First, clustering is carried out according to load characteristics, and a power prediction algorithm is designed based on historical data, load demand, and heating trends with a fused data-physics-knowledge inference model. Then, a load scheduling model is designed at the dispatching end, with economy and comfort indicators as the objective functions. In the platform, a control terminal is installed at the user end to implement the dispatching strategy. The method can be used as a reference for the design of load scheduling strategies under the heating transition period.

Keywords Electric heating loads · Load forecasting · Heating demand · Scheduling indicator sets · Load scheduling

9.1 Introduction

China is in the stage of accelerated urbanization and the demand for heating is growing fast. At present, remote areas are mainly heated by burning coal and firewood, which has the problems of high heating costs and environmental pollution. Some areas use gas heating, which has the advantages of high energy utilization, high reliability of power supply, and low environmental pollution, but requires high architectural design and has safety risks (Nan et al. 2021). As a new type of heating technology, electric heating technology can reduce heating costs compared to traditional methods, and also has the advantages of cleanliness and automation, which helps to save energy and reduce emissions, and is a development trend (Zhongqi et al. 2021).

J. Zhang (✉) · H. Shen · L. Qi · Y. Chen
Nanjing University of Posts and Telecommunications, Nanjing 210023, China
e-mail: 843124494@qq.com

G. Huang (ed.), *Proceedings of 2022 7th International Conference on Environmental Engineering and Sustainable Development (CEESD 2022)*, Environmental Science and Engineering, https://doi.org/10.1007/978-3-031-28193-8_9

When the weather is cold, electric heating loads can cause shocks when they are connected to the grid on a large scale for a short period, so there is a need to study electric heating load scheduling technology. The research focuses on load forecasting and scheduling. The literature (Qing et al. 2019) proposes an RBF neural network prediction model based on ridge regression estimation, which can effectively eliminate the input multicollinearity problem and improve the prediction accuracy; the literature (Yi et al. 2019) proposes a digital building electricity model construction method based on physical and data fusion modeling, which can achieve refined energy efficiency analysis of electricity consumption; the literature (Yi et al. 2021) studies the progress of data and knowledge research progress of joint-driven methods and put forward the prospect of application in power systems. The above are relatively representative studies that have improved the accuracy of load prediction to a certain extent by means of neural networks, artificial intelligence, model fusion, etc.

In terms of load scheduling, existing research includes regulation and control of equipment, load, electricity price, scheduling index, etc. Literature (Shuai et al. 2017) optimizes the operation of heating equipment according to user comfort and realizes collaborative optimization of distributed load; literature (Yulong et al. 2020) establishes a distributed electric heating load model and realizes regulation and control of electric heating load by using the group control method; literature (Ning 2013) studies direct control of centralized load process; literature (Zhiqiang et al. 2019) proposes a modeling method for aggregated load characteristics of multiple types of users, which improved the accuracy of solving the thermal load characteristics; literature (Zhang et al. 2021) studies the evaluation method of the benefits of electric heating loads, and proposed evaluation indexes and evaluation models from various aspects such as comfort and policies.

Most of the above studies are direct load regulation, failing to adjust the strategy according to the dynamic demand and satisfaction of the load. The article, therefore, investigates load forecasting and scheduling strategies, firstly clustering multi-modal loads, then establishing forecasting algorithms that consider load forecasts, heating demand, and trends, followed by designing scheduling indicator sets and scheduling models to complete load schedules.

9.2 Electric Heating Load Forecasting Algorithm

9.2.1 Load Forecasting Model Architecture

The scheduling of electric heating loads relies on accurate heating power prediction techniques. Most of the existing prediction techniques use data models to train a large amount of historical power data and predict short-time power, but the prediction accuracy is not high. Therefore, the article uses the K-Means algorithm to cluster heating users, and then builds a fusion prediction model based on the data model (Weizhao et al. 2020).

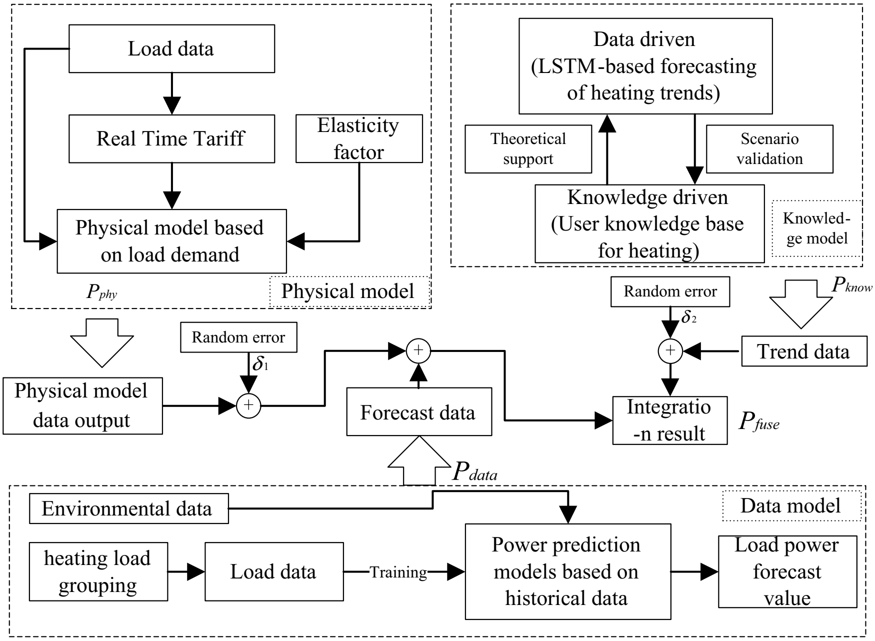

Fig. 9.1 Power forecasting based on physical data fusion

The process of the power forecasting algorithm is shown in Fig. 9.1. Firstly, a load forecasting data model based on generalized regression neural network (GRNN) and historical data of heating power is built to output power forecasting data, and is denoted as P_{data}; then a physical model based on time-of-use tariff and load demand elasticity coefficient is built to output heating demand data; then a load heating trend projection model based on knowledge inference algorithm is built to output heating trend data; finally, the data model is modified by fusing demand data and trend inference data to output short-time power forecasting data.

9.2.2 Physical Modelling in Heating Demand Forecasting

Under the influence of weather, electricity price, and other factors, there is a sudden increase and decrease in user demand for heating, and the prediction accuracy of the data model will decrease.

First, the time-sharing price calculation method on the user side is designed based on the load proportion factor. Select typical days that can reflect the user's heating condition in a certain period, and then set up the calculation formula of electricity price according to the purchase cost C_1, transmission and distribution loss C_2, transmission and distribution price C_3, and government fund C_4, as shown in Eq. (9.1) (Yujie 2019):

$$C_{i,t} = \frac{24 \times P_{i,t}}{\sum P_{i,t}} \times (C_1 + C_2 + C_3 + C_4) \tag{9.1}$$

where $C_{i,t}$ denotes the real-time tariff of load i at moment t; $P_{i,t}$ denotes the real-time power of load i at moment t.

The price elasticity coefficient E is introduced to characterize the relationship between the load's heating demand and the real-time electricity price. According to the effect of the fluctuation of the electricity price at time t on the electricity demand at time t and time h, the price elasticity coefficient can be divided into the self-elasticity coefficient and the other elasticity coefficient, which are denoted by $E(t, t)$ and $E(t, h)$ respectively:

$$E(t, t) = \frac{Ci, t}{Pi, t = 0} \times \frac{\partial Pi, t}{\partial Ci, t}, E(t, h) = \frac{Ci, h}{Pi, h = 0} \times \frac{\partial Pi, h}{\partial Ci, h} \tag{9.2}$$

where $P_{i,t-0}$ represents the power at the initial moment ($t = 0$) of the i-th load.

Based on real-time electricity prices and load demand elasticity coefficients, the heating demand data of the load at moment t under the influence of electricity price fluctuations is calculated using Eq. (9.3), and is denoted as $Pphy$, which reflects the impact of real-time electricity prices on customer demand and can correct the electric heating load power forecast data output by the data model.

$$P_{phy} = P_{i,t} \times \left(1 + E(t, t) \times \frac{\left[C_{i,t} - C_{i,t=0}\right]}{C_{i,t=0}}\right) + P_{i,t} \times \sum_{\substack{h=1 \\ h \neq t}}^{24} E(t, h) \times \frac{\left[C_{i,h} - C_{i,h=0}\right]}{C_{i,h=0}} \tag{9.3}$$

9.2.3 Knowledge Modelling in Heating Demand Forecasting

Due to the influence of electricity price, temperature and other factors, the instantaneous trend change of heating power of different electric heating loads is not consistent, the knowledge inference algorithm can reason about the users' heating trend according to the change of temperature, electricity price, etc. Therefore, the article designs a heating trend inference model based on the knowledge inference

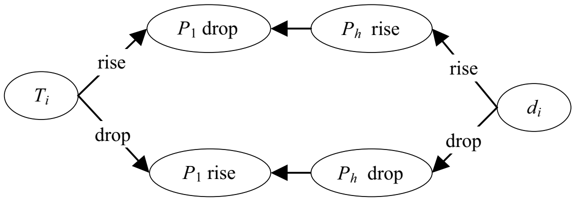

Fig. 9.2 Triadic relationship diagram

algorithm, including the knowledge inference model with and without rules (design intelligent algorithm), to obtain the users' heating trend inference data.

Knowledge Base on Changes in Heating Trends. The article constructs an incremental knowledge base for describing the links between historical data on electric heating loads, heating trends, and external environmental factors to form a knowledge system, including a rule base, a fact database, and a model algorithm base (Chunlei et al. 2010).

The rule base consists of two parts, premises and conclusions, and is represented using the concept of a triad, specifically a knowledge graph containing n entities, m relationships, and facts stored as a triad $D = \{(h, r, t)|h \in E, r \in R, t \in E\}$, each triple consists of a head entity $h \in E$, a tail entity $t \in E$ and a relationship $r \in R$ between them, where E denotes the set of entities and R the set of relationships. The article defines the following triad to construct the incremental knowledge base based on the relationship between real-time temperature T_i, load power d_i, real-time electricity price $P(h)$ and load trend $P_1(t)$, as shown in Fig. 9.2.

The factual database uses the historical data collected to store information on the process of scheduling policy and user changes, with the following formula.

$$F = \begin{bmatrix} T_1 & \dots & T_n \\ C_1 & \dots & C_n \\ P_1 & \dots & P_n \end{bmatrix} \tag{9.4}$$

where T denotes the external temperature, C denotes the change in dynamic tariff in the dispatch strategy and P denotes the change in the customer's heating power.

The library of model algorithms includes inference algorithms, i.e. the ability to use existing knowledge from the rule base and fact database to reason about load heating trends when the electric heating load is influenced by environmental and self-inflicted factors, as well as the ability to optimize the internal parameters of the model using heuristic knowledge and experience in the inference process.

Knowledge-based Reasoning on Heating Trend. In the constructed knowledge inference model, both regular and irregular cases are included. If the knowledge base already contains relevant knowledge, the information available in the knowledge base is used to reason about the trend in heating power.

Regular inference includes the following types: (1) Consistent matching: The data to be inferred and the head entity of the knowledge are matched exactly, the content

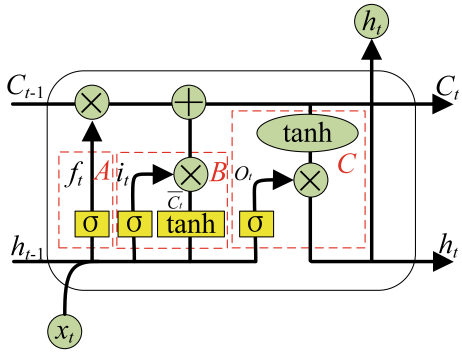

Fig. 9.3 Cell structure of LSTM

includes quarter, month, hour, temperature, and average temperature; (2) Domain value range matching: When the input quarter, month, hour, temperature, average temperature data is matched within a certain domain value range, within the set range is to be converted into the input value defined by the rule, for the more volatile data to be inferred, there are different matching objects and domain value ranges.

When consistent matching and threshold matching are not possible, a knowledge inference model based on a long and short-term memory neural network (LSTM) is used (Zhongwei et al. 2019), with the model inputs being temporal data, temperature data, and load power and the outputs being load heating trend prediction data. The LSTM model uses a cellular structure as shown in Fig. 9.3 and includes three controls, the forgetting gate, the input gate, and the output gate, corresponding to parts A, B and C in the figure. In the forgetting gate, the load information is selectively allowed to pass forward through a neural layer using a sigmoid function and a point-by-point multiplication operation. The input gate is used to determine the load information to be added inside the cell state. The output gate is box C in Fig. 9.3, where a sigmoid function is used to determine the fraction of cell states that need to be output, and then the tanh function is used to process the cell states and output the load trend results.

9.2.4 Data-Physics-Knowledge Fusion Power Prediction Algorithm

Combining the data, physical and knowledge models in Fig.n9.1, we can obtain the accurately predicted power of the electric heating load. The output of the power prediction model based on historical data is denoted as $P_{data}(t)$, which represents the power prediction data at time t. The physical model outputs the heating demand power data of users under the influence of real-time electricity price, which is denoted as $P_{phy}(t)$, and during the experimental process, there may be random errors caused by environmental conditions and unstable measuring instruments, so random errors δ_1 are added; the knowledge inference model outputs the heating trend data of users under the influence of electricity price, time and temperature, which is denoted as $P_{know}(t)$, and random errors δ_2 are added, and Eq. (9.5) is used for fusion calculation.

$$P_{fuse}(t) = P_{data}(t) \times m/p + P'_{phy}(t) \times n/p + P'_{know}(t) \times k/p \qquad (9.5)$$

In formula (9.9), $P'_{phy}(t)$ represents the sum of $P_{phy}(t)$ and δ_1; $P'_{know}(t)$ denotes the sum of $P_{know}(t)$ and δ_2; m、n、k are the ratios of $P_{data}(t)$, $P'_{phy}(t)$ and $P'_{know}(t)$, and p is the sum of the three. The scale coefficients predicted by the model are fused and used as the fused predicted power output.

9.3 The Scheduling Strategy for Electric Heating Loads

9.3.1 Electric Heating Load Dispatch Model Architecture

The designed load scheduling model includes two parts, the scheduling layer, and the user layer, as shown in Fig. 9.4.

The scheduling layer deploys a load scheduling model, which consists of three parts: Load prediction, calculation of dispatch indicator set, and optimal power allocation. First, the fusion algorithm is used to predict the heating power of the electric heating load and calculate the value of the dispatch indicator based on the

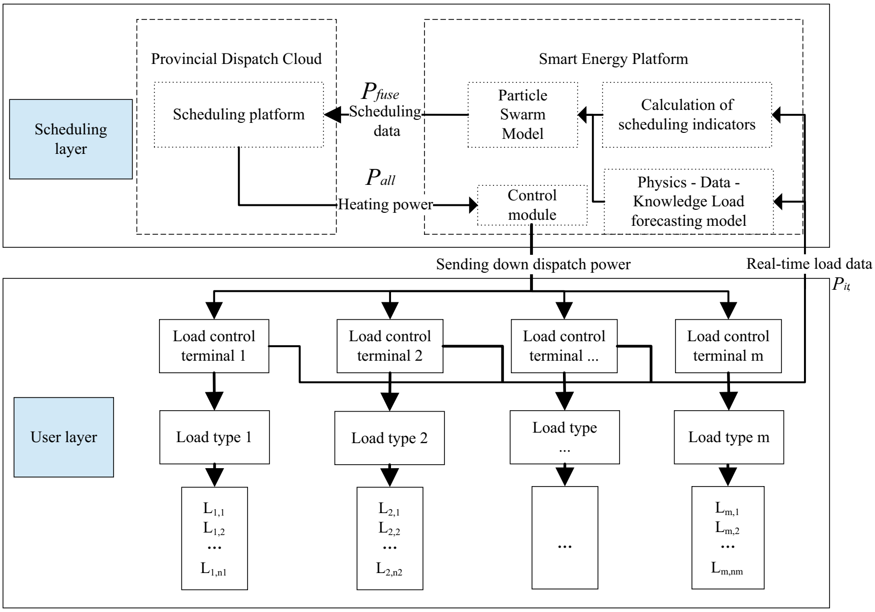

Fig. 9.4 Implementation architecture of two-layer scheduling strategy

real-time heating power fed back from the load control. Then, using the particle swarm optimization algorithm, the power prediction value is used as the initial power dispatch value, and the data of the dispatch indicator is set as the constraint. The final model outputs the power allocation result that maximizes the dispatching benefit and uploads it to the provincial dispatching cloud; the provincial dispatching cloud then issues the heating power $Pall$ according to the uploaded dispatching demand.

9.3.2 Scheduling Indicator Set

Economic Indicator. The economy indicator reflects the economic consumption of the customer in terms of heating before and after the use of the scheduling strategy. The article uses the consumption with the scheduling strategy in place and without the heating equipment participating in the scheduling to calculate it, as shown in Eq. (9.6). The smaller the value of this value, the lower the cost of heating at the current moment after the scheduling.

$$E^{cost} = \frac{[C_{i,t}P(i,t) + C_{i,t}L_{fit}]\Delta t}{[C_{i,t}P_{fit} + C_{i,t}L_{fit}]\Delta t} \tag{9.6}$$

where E^{cost} is the electricity economy indicator; $C_{i,t}$ is the real-time electricity price for load i at moments t; $P(i,t)$ is the load consumed by the heating equipment at moment t; $Pfit$ and is the operating power of heating and other equipment under non-participating dispatch, respectively; Δt is 60 min.

Comfort Indicator. Electricity comfort is the average percentage error between the dispatched power and the power consumed by the equipment in normal operation after the dispatch strategy has been implemented. Electricity comfort can be expressed as Eq. (9.7).

$$E^{fit} = \frac{P(i,t) - P_{fit}}{P_{fit}} \tag{9.7}$$

where $Pfit$ is the power consumption of the heating equipment running all the time, $P(i,t)$ is the dispatching power, E^{fit} is less than 0, and the smaller it is the less the dispatching power is sent down to meet the heating demand and the lower the comfort level of the user.

Scheduling Strategy Implementation. The article designs a particle swarm algorithm-based load optimization scheduling model. In the implementation of the load scheduling strategy, the load control terminal at the user layer sends the real-time heating power of the user to the scheduling layer, and the load scheduling model at the scheduling layer predicts the short-time heating power using a data-physics-knowledge inference model. The particle swarm algorithm in the scheduling model then takes the short-time heating power as the initial scheduling value and uses the set of scheduling indicators as the model constraint to output a power scheduling value that meets the objective function.

Taking into account the economic and comfort requirements of users with heating, the objective function used in the scheduling model is as follows, where F is the output of the objective function, and the larger the value, the better the scheduling effect.

$$F = 1/E^{cost} + E^{fit} \tag{9.8}$$

When implementing the scheduling strategy, the constraints of the particle swarm algorithm include the following three: (1) Economical index requirements, the economical index reflects the consumption ratio after optimization and before optimization, the index range needs to be 0.5–1.5 to guarantee the economic benefits of the scheduling layer and the user layer in an integrated manner. (2) Fairness indicator requirements. The fairness indicator is used to avoid the situation where the heating is turned off for too long when the heating demand is high, and the value should be in the range of 0–1.

$$\begin{aligned} &0.5 \le E^{fit} \le 1.5 \\ &C = Ci, t + \Delta Ci, t \end{aligned} \tag{9.9}$$

where, $\Delta Ci, t$ is the amount of tariff adjustment, when the comfort level is less than 1, reduce the heating tariff, otherwise increase the tariff appropriately, so that the comfort index is maintained at around 1.

9.4 Algorithm Simulation

To verify the dispatching strategy proposed in the article, a customer load in a region of Northern Europe was selected for testing, where the energy meter is used as a load control terminal to collect data such as voltage, current, power and ambient temperature, and can issue control commands to the circuit breaker for the throwing and cutting control of controllable loads.

9.4.1 Load Forecasting Experiment

The article called the K-means algorithm package of the SK-learn library to classify the loads into four categories. Then for each class of load, a prediction algorithm was used for power prediction. As shown in Fig. 9.5, the load raw data curve form "-" and the data model prediction results are shown in the figure in the curve form "."; the physical model corrects the data and the correction results are in the form of the curve "-."; the data is then corrected with the knowledge-based inference of the load

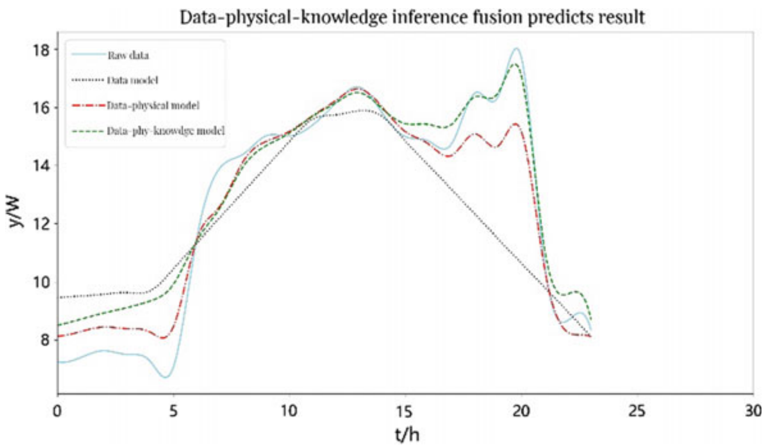

Fig. 9.5 Load forecast results of the next day

Table 9.1 Comparison of prediction standard deviation

Methods	Prediction standard deviation
Traditional forecasting method	4.948
Data-Physics Fusion	0.859
Data-Physics-Knowledge Inference	0.867

heating trend data in the form of the curve "–". The predicted standard deviations are shown in Table 9.1.

Combined with the graphical analysis, the fusion forecasting algorithm is more accurate. In addition, the prediction accuracy of the modified knowledge inference model is not significantly different from the physical model correction, but the knowledge inference model is more accurate for the inference of the heating trend of the load, and better reflects the change of the load trend between 17:00 and 21:00.

9.4.2 Scheduling Strategy Model Solver

The particle swarm algorithm is used to solve the scheduling model, and the solution results are shown in Fig. 9.6, where the red and blue curves are the heating power data before and after optimization respectively. It can be seen that after using the scheduling strategy, the load power fluctuation situation is reduced and the power mutation is less, which is conducive to the stable operation of the grid.

In addition, according to the set of scheduling indicators designed in the article, the implementation of the strategy was analyzed. Taking the 7th h as an example, the

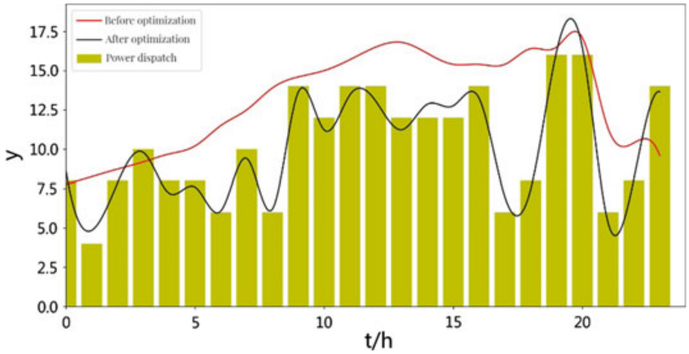

Fig. 9.6 Scheduling strategy model solving

user's power before scheduling was 4.67 kWh with a consumption of 5.51 RMB, and after scheduling was 3.035 kWh with consumption of 1.902 RMB, the value of the user's economic indicator was 0.345, indicating that the user's heating consumption expenditure situation had decreased; the comfort indicator was 0.65, indicating that the scheduling strategy was effective in ensuring The final calculated *W* value is 3.549, which is greater than 1, indicating that the scheduling strategy is effective. Compared with traditional methods, the dispatching strategy proposed in the article can reduce consumption to a greater extent and meet the requirements of economy and comfort while ensuring user satisfaction.

9.5 Conclusion

To address the existing scheduling challenges of electric heating loads, the paper innovatively proposes a joint data-physics-knowledge inference model-driven electric heating load scheduling strategy, which improves the accuracy of load power prediction by jointly modifying the power prediction results of traditional load prediction algorithms through physical models and knowledge inference models. In addition, the paper proposes an optimal scheduling strategy based on the particle swarm algorithm by fusing power prediction data and index set data and implements the optimal scheduling of electric heating loads in combination with load control terminals. The results show that the algorithm can output more accurate load power prediction information, and the two-tier scheduling strategy can meet the requirements of the scheduling indexes and obtain better scheduling results.

The load prediction algorithm and the scheduling strategy designed in the paper can provide more accurate power prediction data and scheduling data for grid dispatching companies and load aggregators and provide loads with more comfortable and less expensive heating methods.

Funding This research was funded by Jiangsu Postgraduate Practice Innovation Program (SJCX21_0304): Research on double-level scheduling of electric heating loads and information security protection strategies.

References

Chunlei H, Feng W, Jianjun W (2010) Intelligent online hydropower dispatching based on knowledge reasoning[J]. Automat Electric Power Syst **34**(21): 50–54, 115

Nan Z, Bowen L, Huan L, Gang L, Rucong W, Quan H, Philbert M, Shan L, Yuguang Z, Riaz A, Mohammed IZA, Crispin P, Renjie D (2021) The potential co-benefits for health, economy and climate by substi-tuting raw coal with waste cooking oil as a winter heating fuel in rural households of northern China[J]. Environ Res **194**

Ning L (2013) Design considerations of a centralized load controller using thermostatically controlled appli-ances for continuous regulation reserves[J]. IEEE Trans Smart Grid 4(2):914–921

Qing X, Chao Z, Shuangshuang Z, Jian L, Dan G, Yongchun Z (2019) Research on short-term power load forecasting method based on machine learning [J]. Electric Measure Instrument 56(23):70–75

Shuai F, Kunqi J, Bingqing G, Limin J, Zhihua W, Guangyu H (2017) Collaborative optimal operation strategy for decentralized electric heating loads[J]. Automat Electric Power Systems 41(19):20–29

Weizhao X, Huan T, Min C (2020) Research on evaluation model of three-phase electric energy meter based on K-means++ algorithm[J]. Electric Measure Instrument **57**(17): 142–146

Yi T, Xiao H, Chaohai Z (2019) Power consumption modeling method of digital buildings based on physical-statistical model [J]. Distribut Utilizat **36**(10): 16–21, 29

Yi T, Rui G, Jianfeng D, Chenyi Z, Chaoming Z, Jie D (2021) Subsequent commutation failure prediction of HVDC by integrating physical-driven and model-driven methods[J]. Electric Power Construct **42**(05): 69–80

Yujie Z (2019) Reaearch on electric heating load adjust ability evaluation and cluster control strategy[D]. Northeast Electric Power University

Yulong Y, Tong W, Leiyang Z, Jinsong L, Yue H, Ruitong L (2020) Distributed electric heating load group modeling and standby optimization[J]. Electric Measure Instrument 57(02):81–87

Zhang Z, An J, Zhou X, Sun P, Rong C (2021) Wei Ling. Research on comprehensive benefit evaluation of electric heating based on AHP and TOPSIS method[J/OL]. Electrical Measure Instrument **1–7**: 12–27

Zhiqiang W, Shan W, Xinyue Z, Wenxia L, Qi G, Qifang C (2019) Load characteristics modeling of regional electric heating system considering difference of users response behaviors[J]. Automat Electric Power Systems 43(07):67–73

Zhongqi Z, Xiqiang C, Kaining S, Mao F, Jinsong Y (2021) Economic analysis of time shifting electric heating load under time-of-use pricing policy[J]. Shan-dong Electric Power **48**(03): 6–10

Zhongwei Z, Lei C, Xiliang C, Dalei K, Tianting S (2019) Survey of knowledge reasoning based on neural network[J]. Comput Eng Appl **55**(12): 8–19, 36