District Heating Load Prediction Algorithm Based on Feature Fusion LSTM Model

Guixiang Xue 1, Yu Pan 1, Tao Lin 1,*, Jiancai Song 2, Chengying Qi 3 and Zhipan Wang 4

1 School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China; xueguixiang_tju@163.com (G.X.); Panyu_top@outlook.com (Y.P.)
2 School of Information and Engineering, Tianjin University of Commerce, Tianjin 300134, China; cn_sjc@163.com
3 School of Energy and Environment Engineering, Hebei University of Technology, Tianjin 300401, China; qicy@163.com
4 School of International, Beijing University of Posts and Telecommunications, Beijing 100876, China; wangzhipan@bupt.edu.cn
* Correspondence: lintao@hebut.edu.cn

Received: 28 April 2019; Accepted: 30 May 2019; Published: 3 June 2019

Abstract: The smart district heating system (SDHS) is an important element of the construction of smart cities in Northern China; it plays a significant role in meeting heating requirements and green energy saving in winter. Various Internet of Things (IoT) sensors and wireless transmission technologies are applied to monitor data in real-time and to form a historical database. The accurate prediction of heating loads based on massive historical datasets is the necessary condition and key basis for formulating an optimal heating control strategy in the SDHS, which contributes to the reduction in the consumption of energy and the improvement in the energy dispatching efficiency and accuracy. In order to achieve the high prediction accuracy of SDHS and to improve the representation ability of multi-time-scale features, a novel short-term heating load prediction algorithm based on a feature fusion long short-term memory (LSTM) model (FFLSTM) is proposed. Three characteristics, namely proximity, periodicity, and trend, are found after analyzing the heating load data from the aspect of the hourly time dimension. In order to comprehensively utilize the data’s intrinsic characteristics, three LSTM models are employed to make separate predictions, and, then, the prediction results based on internal features and other external features at the corresponding moments are imported into the high-level LSTM model for fusion processing, which brings a more accurate prediction result of the heating load. Detailed comparisons between the proposed FFLSTM algorithm and the state-of-art algorithms are conducted in this paper. The experimental results show that the proposed FFLSTM algorithm outperforms others and can obtain a higher prediction accuracy. Furthermore, the impact of selecting different parameters of the FFLSTM model is also studied thoroughly.

Keywords: heat load prediction; deep learning; long short-term memory; feature fusion

1. Introduction

The world’s largest district heating system (DHS) is located in China. This system needs to consume a large quantity of fossil energy, such as coal and natural gas, during winter. This leads to a serious air pollution problem and extremely harmful effects on human health and has, therefore, attracted the widespread attention of the whole society. The analysis of [1] suggests that the Huai River Policy, which had the laudable goal of providing indoor heat, had disastrous consequences for human health. Specifically, it led to an increase in inhalable particles (particulate matter 10, PM10), with concentrations of 46%, and reductions in life expectancies of 3.1 years, caused by elevated rates of cardiorespiratory mortality, in the north.
With the rapid development of urbanization and the increasing demand for comfortable heating by residents, the energy consumption of the DHS in China will increase significantly in the coming decades, and this will inevitably put significant pressure on energy supplements and environmental protection. Energy consumption is closely related to healthy economic development and daily human life. Combining the latest information technologies such as the Internet, the Internet of Things (IoT), and automation and communication technologies to develop a smart district heating system (SDHS) and to achieve energy saving has become one of the important goals of the DHS for the future of China. Thus, the heating load prediction algorithm has become one of the necessary conditions and key technologies for an intelligent heating system contributing to energy planning, operation regulation, realization of system energy saving, and emission reduction [2].

In recent years, the prediction of the energy consumption load has attracted the attention of many scholars, and many studies have been devoted to this purpose. Since the dataset collected by various sensors in the DHS has significant time series characteristics, and the heat transfer process has an evident delay and nonlinear characteristics; different nonlinear time series prediction algorithms can be used to solve the problem of heating load prediction. The previous studies on heating load prediction were mainly focused on the classical machine learning algorithms and their improvements, such as artificial neural networks (ANN), a support vector machine (SVM), support vector regression (SVR), and extreme learning machines (ELM).

In the early research period, Dotzauer [3] analyzed several factors affecting the heating load. The ANN has a good performance of linear fit, which is suitable for heating load prediction. Karatasou et al. [4] proposed an energy load prediction algorithm based on ANN. In order to improve the prediction performance of ANN, Boithias et al. [5] proposed a genetic algorithm (GA) to optimize ANN training parameters and model inputs. However, traditional ANNs are prone to over-fitting, difficult to initialize, and are difficult to adjust the network parameters. Thus, researchers gradually turned their attention to SVRs. Cai et al. [6] proved the superiority of an SVM in predicting building energy consumption. Zhang et al. [7] applied an SVR to accomplish the nonlinear regression prediction of building energy consumption. The theoretical analysis for a prediction algorithm based on SVR was conducted by Fan et al. [8], who showed that the improvement strategies for SVR algorithms appeared in the short term. Hong [9] employed an immune algorithm (IA) to construct an SVR model to predict the annual electrical load in Taiwan. Ding et al. [10] proposed a short-term cooling load prediction algorithm for public buildings based on genetic algorithm–SVR (GA–SVR). Wang et al. [11] proposed a hybrid heating load prediction method based on particle swarm optimization and SVR. Al-Shammari et al. [12] proposed a hybrid prediction algorithm composed of an SVM and the firefly algorithm (FFA), which was named SVM–FFA and established seven heating load prediction models with different time scales. The FFA was employed to optimize SVM parameters to improve the prediction accuracy. Lee et al. [13] used a harmony search (HS) algorithm to determine the near-global optimal initial weights in the training of the model. The stratified sampling was used to sample the training data. A total of 25 HS–ANN hybrid models were tested with different combinations of DHS algorithm parameters. Dalipi et al. [14] proposed a data-driven supervised machine learning model to predict the heat load of buildings in a DHS. ELM was also employed for heating load prediction, as it is both simple and effective. Sajjadi et al. [15] proposed a short-term heat load prediction model of a DHS based on ELM. Roy et al. [16] applied a multivariate autoregressive (MVA) and an ELM hybrid model to predict a building’s thermal load.

However, traditional machine learning algorithms have certain limitations. Firstly, features need to be extracted manually, which makes it difficult to extract the features contained deep within the data. Then, the ability of nonlinear expression on the complex characteristic model is insufficient. Finally, the accuracy of the algorithm can reach a bottleneck and become difficult to improve.

After this, with the rapid development of deep learning (DL) technology in recent years, the related research based on DL algorithms has attracted the attention of many scholars. The recurrent neural networks (RNN) and long short-term memory (LSTM) are especially suitable for time series
analysis and prediction based on the ability of feature extraction and nonlinear representation, which can automatically extract the deep features from the data and improve prediction accuracy. Kato et al. [17] innovatively proposed a heating load prediction algorithm based on RNN and achieved good prediction accuracy.

We compared the input factors to the literature related to heating load prediction. The results are shown in Table 1.

| Literature | Algorithm | Influencing Factors |
|------------|-----------|---------------------|
| [3]        | ARMA      | outdoor temperature, heat load, behavior of the consumers |
| [11]       | PSO-SVR   | outdoor temperature, supply water temperature, supply water pressure, circular flow, heat load |
| [12]       | SVM–FFA   | time lagged heat load, outdoor temperature, primary return temperatures |
| [14]       | SVR, PLS, RT | forward temperature, return temperature, flow rate, heat load |
| [15]       | ELM       | outdoor temperature, primary supply temperature, primary return temperature, flow on primary side |
| [18]       | ANFIS     | outdoor temperature, primary supply temperature, primary return temperature, secondary supply temperature, secondary return temperature, flow on primary side |

In addition to the heating system, the time series prediction problems based on deep learning have also attracted the attention of many scholars in other fields—solar irradiance prediction [19], building energy consumption [20], cooling load [21], electrical load [22–27], and wind power [28,29], for example.

Due to the fact that heating systems involve heat sources, heat exchange stations, primary pipe networks, secondary pipe networks, and thermal users, the nonlinearity and large delay characteristics of heat transfer make heating load prediction more complicated than for power systems. Fortunately, deep learning algorithms such as LSTM perform well in terms of time series feature extraction and nonlinear expression, which make them very suitable for heating load forecasting. However, the simple LSTM algorithm has a limited ability for historical features memory, and the large delay characteristic of heating systems makes the prediction possess a lower accuracy than that of the simple LSTM algorithm. Therefore, a feature fusion LSTM (FFLSTM) algorithm for the short-term prediction of a heating load is proposed in this paper, which combines the characteristic of proximity, periodicity, and trend, thus leading to a better performance compared with the state-of-the-art algorithms. The best mean absolute percentage (MAPE) that can be achieved is 0.009, which is a dramatic improvement compared with the state-of-the-art algorithms.

In general, the main contributions of our paper are demonstrated as follows.

Firstly, the data characteristics and correlation of the heating system supervisory control and data acquisition (SCADA) system is analyzed, and three features that have a significant influence on the heating load prediction are found: The proximity, periodicity, and trend, which provide the basis for constructing the accurate heat load prediction model.

Secondly, layered prediction algorithm architecture for the heating load is designed based on the nonlinear and large delay characteristics of the heating system. The bottom layer is the LSTM prediction model with three different time scales, and the top layer establishes the feature fusion model.

Finally, according to different delay characteristics, six kinds of prediction mathematical models with different parameters are designed, and the experimental tests are compared. The delay characteristics suitable for the heating system are obtained, which provides support for the subsequent realization of energy-saving optimization prediction control.
The remainder of this paper is organized as follows: Section 2 analyzes the correlation characteristics of the original data and applies a Gauss filter to reduce the noise interference; Section 3 demonstrates the architecture and mathematical model of the proposed FFLSTM methodology in detail; Section 4 represents the experimental analysis and evaluation of the prediction performance; Section 5 is the conclusion.

2. Feature Analysis and Selection of Heating Load in SDHS

The FFLSTM algorithm needs to fully explore the unique characteristics of a heat load. In this section, two datasets, which were obtained the two stations of Xingtai Heating Company hourly heat load from 2017 to 2018, are used as examples. We analyzed the heat load features and selected the inputs of the models. For a better visualization, the first 10 days of data are plotted Figure 1 to show the performance of these three characteristics.

![Figure 1. Time characteristics of the heat load.](image)

The three characteristics of the proximity, periodicity, and trend of the hourly data are considered in this paper. Figure 1 shows the periodic characteristics of the heating load data with two heat load fluctuations in 24 hours and a linear trend, shown as a red line. This is to ensure the thermal comfort of people after waking up in the morning, which involved increasing the secondary water supply temperature of the heat exchange station after 3:00 am in advance of them waking up. In the same manner, in order to ensure comfort heat levels after people get home from work every day, the secondary water supply temperature of the heat exchange station was increased after 15:00 pm. Therefore, the heat load of the heat exchange station has evident periodic characteristics. The heating load also has a significant correlation with the outdoor temperature, which leads to linear trend characteristics. From the perspective of the trend, in the early stage of heating, the heating load increases with the decrease in the room temperature; in the late cold stage, the heating load decreases with the increase of the outdoor temperature. However, due to the large hysteresis of the heating system, the daily energy consumption fluctuates differently.

According to thermodynamic heat conduction theory, the heat consumption of a heat exchange station is related to the secondary supply temperature, the secondary return temperature, and the flow rate. At the same time, due to the influence of thermal inertia and the meteorological factors of buildings, heat consumption is also related to historical heat consumption and outdoor temperature. The influencing factors are classified into the following two categories: Internal factors and external factors.

1. Internal factors:
   - Historical heat load (GJ)
   - Secondary supply temperature (supply temp for short) (°C)
• Secondary return temperature (return temp for short) (°C)
• Instantaneous flow rate (flow rate for short) (m³/h)

(2) External factors:
• Outdoor temperature (outdoor temp for short) (°C)

Figure 2 indicates a high correlation of the heat load with the characteristics of the historical heat load, outdoor temperature, second supply temperature, second return temperature, and instantaneous flow rate. Thus, these characteristics might be useful predictors of the dependent variable.

![Figure 2](image)

**Figure 2.** Correlation of heat load versus each of the influencing factors. (a) Historical heat load, (b) outdoor temperature, (c) second supply temperature, (d) second return temperature, and (e) instantaneous flow rate.
3. Methodology and Analysis

3.1. The Architecture of the Proposed FFLSTM

The intelligent district heating system is a complex nonlinear dynamic system, and the energy-saving regulation operation of the heat exchange station involves the characteristics and transmission delay of the primary network and the secondary network, the thermal inertia of the building, meteorological factors, and the personalized requirements of the terminal thermal users. Because the distance between each heat exchange station and the heat source is different and the fact that the buildings possess different energy saving characteristics, the historical dataset corresponding to the heat exchange station can essentially represent the comprehensive enclosure structural characteristics of buildings in the corresponding heating area. The prediction algorithm should automatically extract the corresponding delay characteristics and provide an accurate heating load prediction. This paper proposes a feature fusion integrated learning algorithm based on LSTM which can extract and process the characteristics of proximity, periodicity, and the trend of the heat exchange station load dataset and, then, obtain more accurate heating load prediction results after fusion processing so as to improve the accuracy of the energy saving regulation of heating. The overall architecture of the proposed FFLSTM model is presented in Figure 3.

![Figure 3](image)

*Figure 3.* The overall architecture of the prediction based on the fusion algorithm of time feature extraction.

As shown in Figure 3, the FFLSTM is a hierarchical structure with five layers. The input layer includes the time series of the secondary heating network and meteorology. The feature selection layer is used to select the characteristics of proximity, periodicity, and trend of the historical heat data. The feature processing layer uses LSTM to deal with three time-series with different time scale features. The three prediction models are called the N-LSTM, P-LSTM, and T-LSTM, and they deal with the three different data characteristics of proximity, periodicity, and trend, respectively. The feature fusion layer is used to process the prediction results of the N-LSTM, P-LSTM, and T-LSTM. At this stage, the external feature outdoor temperature is considered. Then, the prediction result of the FFLSTM is obtained in the output layer.
3.2. The Mathematical Model of FFLSTM

The FFLSTM is a feature-based ensemble learning algorithm whose mathematical principles can be expressed as follows:

\[
\begin{align*}
Q_h & = f(T_h, Q_h^N, Q_h^P, Q_h^T) \\
Q_h^N & = f_1(x_{h-1}, x_{h-2}, \ldots, x_{h-p}) \\
Q_h^P & = f_2(x_{h-24x1}, x_{h-24x2}, \ldots, x_{h-24q}) \\
Q_h^T & = f_3(x_{h-24x7x1}, x_{h-24x7x2}, \ldots, x_{h-24x7r})
\end{align*}
\]

where

- \(Q_h\) denotes predicted heat consumption at the hour \(h\);
- \(Q_h^N\) denotes the predicted heat load from proximity data at hour \(h\);
- \(Q_h^P\) denotes the predicted heat load from the periodic data at hour \(h\);
- \(Q_h^T\) denotes the predicted heat load from the trend data at hour \(h\);
- \(f\) indicates a predictive model of the fusion data;
- \(f_1\) indicates a predictive model of the proximity data;
- \(f_2\) indicates a predictive model of periodic data;
- \(f_3\) indicates a predictive model of trend data;
- \(x_{h-i}\) represents the proximity hourly input vector;
- \(x_{h-24j}\) represents the periodic daily input vector;
- \(x_{h-24x7k}\) represents the trend weekly input vector;
- \(p = (1,2,\ldots,24)\) represents the hours factor of the model;
- \(q = (1,2,\ldots,7)\) represents the days factor of the model; and
- \(r = (1,2,3,4)\) represents the weeks factor of the model.

The three different time-scale feature vectors \(x_{h-i}\), \(x_{h-24j}\), and \(x_{h-24x7k}\) can be defined as shown in Equation (2), respectively:

\[
\begin{align*}
x_{h-i} & = (T_{h-i}^o, Q_{h-i}^o, T_{h-i}^S, T_{h-i}^R, F_{h-i})(i = 1, 2, \cdots, p) \\
x_{h-24j} & = (T_{h-24j}^o, Q_{h-24j}^o, T_{h-24j}^S, T_{h-24j}^R, F_{h-24j})(j = 1, 2, \cdots, q) \\
x_{h-24x7k} & = (T_{h-24x7k}^o, Q_{h-24x7k}^o, T_{h-24x7k}^S, T_{h-24x7k}^R, F_{h-24x7k})(k = 1, 2, \cdots, r)
\end{align*}
\]

where

- \(x_{h-i}\) represents the input vector composed of the outdoor temperature, heat consumption, secondary supply temperature, secondary return temperature, and instantaneous flow rate in the (h-i)th hour. The hours factor \(p = (1,2,\ldots,24)\) represents the interval from the predicted time of the proximity input data, ranging from 1 to 24 h.
- \(x_{h-24j}\) represents the input vectors such as the outdoor temperature, heat consumption, secondary supply temperature, secondary return temperature, and instantaneous flow in the (h-24j)th hour. The days factor \(q = (1,2,\ldots,7)\) represents the interval from the predicted time of the periodic input data, ranging from 1 to 7 days.
- \(x_{h-24x7k}\) represents the input vector of the outdoor temperature, heat consumption, secondary water supply temperature, secondary water return temperature, and instantaneous flow in the (h-24x7k)th hour. The weeks factor \(r = (1,2,3,4)\) represents the interval from the predicted time of the trend input data, ranging from 1 to 4 weeks.

According to the analysis of section II, the factors of outdoor temperature \((T^o)\), heat load \((Q)\), secondary supply temperature \((T^S)\), secondary return temperature \((T^R)\), and instantaneous flow \((F)\) are considered to build the predictive model for the purpose of achieving a more accurate prediction performance.
The SCADA system developed by our team collects these parameters once every 20 s and stores them in the database every 10 min. Thus, there are 24 groups of hourly datasets, which contain six measurement values for each parameter (hours group \( h = 1, 2, 3, \ldots, 24 \), internal index \( i = 1, 2, \ldots, 6 \)). The original monitoring historical data are represented as the following:

- \( T_{oh} \) denotes the \( i \)th outdoor temperature measurement value at the \( h \)th hour;
- \( T_{sh} \) denotes the \( i \)th secondary supply temperature measurement value at the \( h \)th hour;
- \( T_{rh} \) denotes the \( i \)th secondary return temperature measurement value at the \( h \)th hour;
- \( F_{h} \) denotes the \( i \)th instantaneous flow rate measurement value at the \( h \)th hour; and
- \( Q_{h} \) denotes the \( i \)th cumulate heat load measurement value at the \( h \)th hour.

Because the heating system is a complex nonlinear system, and the heat transfer has a large delay, there will be a significant deviation in the real-time data collected by the SCADA system to predict the heating load. In general, hourly average measurements are required as the inputs to the prediction data. We preprocessed the original measurement data to form hourly statistical data.

\[
\begin{align*}
T_{oh} & = \frac{1}{6} \sum_{i=1}^{6} T_{oh,i} \\
T_{sh} & = \frac{1}{6} \sum_{i=1}^{6} T_{sh,i} \\
T_{rh} & = \frac{1}{6} \sum_{i=1}^{6} T_{rh,i} \\
F_{h} & = \frac{1}{6} \sum_{i=1}^{6} F_{h,i} \\
Q_{h} & = Q_{h}^{c}(h+1) - Q_{h}^{c}(h) 
\end{align*}
\]

where,
- \( T_{oh} \) denotes the average outdoor temperature value at the \( h \)th hour;
- \( T_{sh} \) denotes the average secondary supply temperature value at the \( h \)th hour;
- \( T_{rh} \) denotes the average secondary return temperature value at the \( h \)th hour;
- \( F_{h} \) denotes the average instantaneous flow rate value at the \( h \)th hour; and
- \( Q_{h} \) denotes the average heat load value at the \( h \)th hour.

### 3.3. Evaluation Criteria

In this paper, the root-mean-square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) were selected as the evaluation indices of the heat load prediction model, which are defined as:

\[
\begin{align*}
RMSE & = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - O_i)^2} \\
MAE & = \frac{1}{n} \sum_{i=1}^{n} |P_i - O_i| \\
MAPE & = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{P_i - O_i}{P_i} \right|
\end{align*}
\]

where \( O_i \) and \( P_i \) represent the true value and predicted value of the heat load, respectively, and \( n \) represents the total number of the test samples.

### 4. Experiments and Discussion

#### 4.1. System Background and Data Description

We have established a smart district heating system for the Xingtai Heating Company. With a heating area of about 16 million square meters with a total of 382 heat exchange stations, the real-time data acquisition and remote automatic control of the unattended heat exchange station were
realized, and all data analysis and control strategies were completed in the monitoring data center. The datasets were collected from the heat exchange stations in Xingtai during the 2017–2018 heating season. The various sensors monitoring data were collected every 20 s and stored into the rational Microsoft SQL Server database every 10 min. There are 382 heat exchange stations and 120 days of each heating season, which ran from 15 October 2017 to 15 March 2018. Thus, the dataset size was 24 hour/day × 120 day × 6 records/hour = 17,280 records for each substation, and the total records were about 17,280 × 382 = 6.6 million for the whole enterprise in one heating season, which can meet the requirement of the training and testing process of proposed FFLSTM algorithms.

The heat load predictions of each heat exchange station were independent of each other. A cross-validation method was used to verify the performance of the proposed algorithm. In our experiment, each dataset was split into a training dataset (70%), validation dataset (10%), and a testing dataset (20%).

4.2. Time Delay Factors Selection and Different Time-Scale Models

To determine the influence of the time delay characteristics on the heat load, correlation analysis was conducted on data of different time scales using Equation (5):

$$\text{corr} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

where $x_i$ and $y_i$ represent two different variables, $\bar{x}$ represents the mean of $x_i$, $\bar{y}$ represents the mean of $y_i$, and corr represents the correlation between variables $x_i$ and $y_i$.

The results of the correlation analysis are shown in Table 2. Here, $p$, $q$, and $r$ represent proximity data, periodic data, and trend data intervals with the predicted time, respectively; the intervals of these three time characteristics are 1–5 h, 1–4 days, and 1–2 weeks, respectively.

### Table 2. The correlation value of heat load versus other factors.

| Time intervals | Historical Heat load | Outdoor Temp | Supply Temp | Return Temp | Flow Rate |
|----------------|---------------------|--------------|-------------|-------------|-----------|
| $p=1$ h        | 0.913               | −0.339       | 0.767       | 0.710       | 0.783     |
| $p=2$ h        | 0.729               | −0.301       | 0.577       | 0.538       | 0.549     |
| $p=3$ h        | 0.553               | −0.275       | 0.379       | 0.360       | 0.320     |
| $p=4$ h        | 0.408               | 0.271        | 0.206       | 0.208       | 0.124     |
| $p=5$ h        | 0.291               | 0.294        | 0.090       | 0.123       | 0.027     |
| $q=1$ d        | 0.765               | −0.224       | 0.768       | 0.717       | 0.693     |
| $q=2$ d        | 0.626               | −0.219       | 0.579       | 0.396       | 0.524     |
| $q=3$ d        | 0.151               | −0.157       | 0.045       | −0.086      | 0.140     |
| $q=4$ d        | −0.128              | −0.117       | −0.189      | −0.283      | 0.069     |
| $r=1$ w        | −0.529              | 0.257        | −0.545      | −0.578      | 0.390     |
| $r=2$ w        | 0.395               | −0.126       | 0.726       | 0.691       | 0.377     |

* Note: h represents the hour, d represents the day and w represents the week in Table 2.

Due to the different thermal inertia of the heating system and the different path-length from the heat source, the heat transfer has the characteristics of being non-linear and having a large delay. In addition, since the energy consumption of the heating system is significantly affected by meteorological factors, the time series data collected from the sensor cannot directly and accurately describe such complex factors.

To better extract the nonlinear time delay characteristics of the physical heating pipe network, we chose the time intervals ($p=1, 2, 3; q=1, 2$, and $r=1$) based on the analysis the correlation of the characteristics of the original data shown in Table 2.

In this paper, a total of six different models are defined; this is shown in Table 3.
Table 3. Six models with different inputs.

| Model   | Proximity Factor (p) | Periodic Factor (q) | Trend Factor (r) |
|---------|----------------------|---------------------|------------------|
| model 1 | 1 h                  | 1 d                 | 1 w              |
| model 2 | 2 h                  | 1 d                 | 1 w              |
| model 3 | 3 h                  | 1 d                 | 1 w              |
| model 4 | 1 h                  | 2 d                 | 1 w              |
| model 5 | 2 h                  | 2 d                 | 1 w              |
| model 6 | 3 h                  | 2 d                 | 1 w              |

4.3. Parameter Selection and Performance Evaluation

In order to achieve the optimal prediction performance of the heat load model, the parameters of the LSTM network need to be selected in advance. In this paper, the selection of parameters of an LSTM model was based on the proximity data, and the initial parameters are shown in Table 4.

Table 4. Initial parameters of the FFLSTM.

| Parameters      | Value | Description             |
|-----------------|-------|-------------------------|
| hide_unit       | 20    | the number of hidden cells |
| learn_rate      | 0.006 | the value of the learning rate |
| time_step       | 5     | the value of the time step |
| batch_size      | 20    | the value of the batch size |
| iter_count      | 200   | the number of iterations |

To determine the optimal value of these parameters for the experiments, every candidate is tested according to the sequence of learn_rate, batch_size, time_step, hide unit, and iter_count. The best performance of each parameter is taken as the optimal selection based on the previous experimental results.

As is shown in Table 5, when learn_rate is used as a single regulation variable, the best result is obtained, 0.002; learn_rate = 0.002 was used in the experiment and in the subsequent experiments. When batch_size was adjusted, the results obtained were not as good as those under learn_rate = 0.002 and batch_size = 20. Therefore, when time_step was used as a single variable, the parameter batch_size was still selected as 20.

Table 5. The performance of heat load prediction for different selected parameters.

| Parameters Set | Turning Parameter | Value | Train | Test |
|----------------|-------------------|-------|-------|------|
|                |                   |       | RMSE  | MAE  | MAPE |
| hide_unit = 20 | learn_rate        | 0.0006| 0.027 | 0.020 | 0.007 | 0.049 | 0.039 | 0.014 |
| time_step = 5  |                   | 0.0001| 0.021 | 0.021 | 0.014 | 0.005 | 0.048 | 0.039 | 0.014 |
| batch_size = 20|                   | 0.003 | 0.027 | 0.019 | 0.006 | 0.048 | 0.040 | 0.014 |
| iter_count = 200|                  | 0.001 | 0.001 | 0.001 | 0.001 | 0.005 | 0.044 | 0.035 | 0.013 |

Note: The text in red bold font is the optimal prediction result of the given algorithm for each model.
The final selected experimental parameters are shown in Table 6. In order to ensure consistency, each LSTM model used the same parameters, which are shown in Table 6, in the entire fusion algorithm.

| Parameters      | Value | Description                      |
|-----------------|-------|----------------------------------|
| hide_unit       | 5     | the number of hidden cells       |
| learn_rate      | 0.002 | the value of the learning rate   |
| time_step       | 12    | the value of the time step       |
| batch_size      | 20    | the value of the batch size      |
| iter_count      | 200   | the number of iterations with early stopping |

4.4. Compared with the Base LSTM Models

The integrated learning approach uses a series of independent models for prediction and then combines them to integrate a model. It can generate better prediction results. A single model sometimes only obtains local optimal results or over-fitting; therefore, the practice of combining multiple models can improve accuracy. The basic models in the FFLSTM feature processing layer are independent of each other. To achieve a better integration effect for determining the FFLSTM, the prediction results of the FFLSTM are compared with those of the three basic models, which are shown in Table 7.

| Algorithm | Model   | Station A | Station B |
|-----------|---------|-----------|-----------|
|           |         | RMSE      | MAE       | MAPE      | RMSE      | MAE       | MAPE      |
| FFLSTM    | model 1 | 0.043     | 0.035     | 0.012     | 0.624     | 0.491     | 0.064     |
| N-LSTM    |         | 0.050     | 0.039     | 0.014     | 0.707     | 0.575     | 0.074     |
| P-LSTM    |         | 0.099     | 0.087     | 0.031     | 1.205     | 0.953     | 0.128     |
| T-LSTM    |         | 0.129     | 0.113     | 0.040     | 1.061     | 0.883     | 0.118     |
| FFLSTM    | model 2 | 0.044     | 0.040     | 0.014     | 0.518     | 0.414     | 0.055     |
| N-LSTM    |         | 0.045     | 0.034     | 0.012     | 0.694     | 0.593     | 0.079     |
| P-LSTM    |         | 0.104     | 0.089     | 0.031     | 1.468     | 1.086     | 0.149     |
| T-LSTM    |         | 0.131     | 0.107     | 0.038     | 1.038     | 0.882     | 0.117     |
| FFLSTM    | model 3 | 0.039     | 0.028     | 0.010     | 0.589     | 0.453     | 0.060     |
| N-LSTM    |         | 0.039     | 0.029     | 0.010     | 0.920     | 0.699     | 0.093     |
| P-LSTM    |         | 0.106     | 0.095     | 0.033     | 1.325     | 0.955     | 0.128     |
| T-LSTM    |         | 0.136     | 0.105     | 0.037     | 1.223     | 1.010     | 0.131     |
| FFLSTM    | model 4 | 0.048     | 0.035     | 0.012     | 0.512     | 0.422     | 0.056     |
| N-LSTM    |         | 0.048     | 0.035     | 0.013     | 0.834     | 0.597     | 0.081     |
| P-LSTM    |         | 0.072     | 0.063     | 0.022     | 1.448     | 1.129     | 0.153     |
| T-LSTM    |         | 0.097     | 0.081     | 0.029     | 1.511     | 1.130     | 0.157     |
| FFLSTM    | model 5 | 0.065     | 0.050     | 0.018     | 0.428     | 0.335     | 0.044     |
| N-LSTM    |         | 0.066     | 0.055     | 0.020     | 0.763     | 0.614     | 0.084     |
| P-LSTM    |         | 0.088     | 0.076     | 0.027     | 0.964     | 0.760     | 0.102     |
| T-LSTM    |         | 0.133     | 0.101     | 0.035     | 0.999     | 0.798     | 0.109     |
| FFLSTM    | model 6 | 0.042     | 0.032     | 0.012     | 0.485     | 0.381     | 0.049     |
| N-LSTM    |         | 0.045     | 0.035     | 0.013     | 0.699     | 0.579     | 0.078     |
| P-LSTM    |         | 0.090     | 0.074     | 0.026     | 1.071     | 0.878     | 0.115     |
| T-LSTM    |         | 0.097     | 0.078     | 0.028     | 1.232     | 1.014     | 0.136     |

Note: The text in red bold font is the optimal prediction result of the given algorithm for each model.
In Table 7, the N-LSTM, P-LSTM, and T-LSTM represent LSTM proximity prediction, periodic prediction, and LSTM trend prediction, respectively. It can be seen that, in station A and station B, the N-LSTM shows better prediction results compared to those of P-LSTM and T-LSTM. This also indicates that the heating load of the heat exchange station is more correlated with the recent historical data of influential factors, such as the heat load and outdoor temperature. The FFLSTM obtained the highest prediction accuracy after the correction of periodicity and trend.

4.5. Compared with Other Algorithms

To evaluate the heat load performance of the proposed FFLSTM algorithm, we compared it with the state-of-the-art algorithms which are widely used in the forecasting field, such as traditional LSTM, back propagation (BP), SVR, a regression tree (RT), random forest regression (RFR), gradient boosting regression (GBR), and extra trees regression (ETR). The radial kernel function (RBF) was used in SVR for this paper; it can map a sample to a higher dimensional space and is suitable for training and testing with moderate data samples. The sigmoid activation function was used for BP. The size of the hidden layer and the number of iterations in the gradient descent loop of BP were 3 and 20,000, respectively.

In order to prove the superiority of the proposed FFLSTM, the datasets of the two heat exchange stations were used to validate the model. The RMSE, MAE, and MAPE of station A and station B are shown in Figures 4–6, respectively. In order to compare the prediction performance of different heat exchange stations in the same dimension for convenience, the normalized MAPE was selected as the main index parameter, and MAE and RMSE were selected as the auxiliary index parameters. Comprehensive experiments have shown that the FFLSTM has the best predictive power on both datasets. RMSE, MAE, and MAPE are smaller than the other comparison algorithms.

**Figure 4.** Evaluation of the results of the root-mean-square error (RMSE) for station A and B, respectively.
Figure 5. Evaluation of the results of the mean absolute error (MAE) for station A and B, respectively.

Table 8 lists the RMSE, MAE, and MAPE values of the proposed algorithm and the state-of-the-art comparison algorithms. Table 8 shows that the performance of the FFLSTM was optimal in the models. The RMSE, MAE, and MAPE of heat exchange station A reached 0.03 GJ, 0.02 GJ, and 0.009 GJ, respectively. The three evaluation indexes are significantly improved compared with other algorithms. The prediction accuracy of station A was above 98%, and station B was above 92%. Because the energy consumption prediction of heat exchange station B uses the network parameters trained by historical data of heat exchange station A, the prediction accuracy of heat exchange station B was slightly lower but also reached over 92%. This shows that the FFLSTM algorithm also has a good generalization ability. We chose the average of the six mode prediction results as the final prediction value, as shown in Table 9.
Figure 5. Evaluation of the results of the mean absolute error (MAE) for station A and B, respectively.

Figure 6. Evaluation of the results of the mean absolute percentage error (MAPE) for station A and B, respectively.

Table 8 lists the RMSE, MAE, and MAPE values of the proposed algorithm and the state-of-the-art comparison algorithms. Table 8 shows that the performance of the FFLSTM was optimal in the models. The RMSE, MAE, and MAPE of heat exchange station A reached 0.03 GJ,
Table 8. The performance comparison of different algorithm for six models.

| Algorithm | Model | Station A |          |          | Station B |          |          |
|-----------|-------|-----------|----------|----------|-----------|----------|----------|
|           |       | RMSE      | MAE      | MAPE     | RMSE      | MAE      | MAPE     |
| FFLSTM    |       | 0.043     | 0.035    | 0.012    | 0.707     | 0.575    | 0.074    |
| LSTM      |       | 0.070     | 0.056    | 0.020    | 1.109     | 0.850    | 0.112    |
| BP        |       | 0.077     | 0.065    | 0.023    | 0.750     | 0.549    | 0.073    |
| SVR(RBF)  | model 1 | 0.124     | 0.100    | 0.036    | 1.121     | 0.930    | 0.122    |
| RT        |       | 0.133     | 0.106    | 0.038    | 1.216     | 0.994    | 0.131    |
| RFR       |       | 0.133     | 0.107    | 0.038    | 1.172     | 0.957    | 0.126    |
| GBR       |       | 0.127     | 0.103    | 0.037    | 1.201     | 0.980    | 0.128    |
| ETR       |       | 0.139     | 0.111    | 0.040    | 1.155     | 0.947    | 0.124    |
| FFLSTM    |       | 0.044     | 0.040    | 0.014    | 0.518     | 0.414    | 0.055    |
| LSTM      |       | 0.071     | 0.056    | 0.020    | 0.766     | 0.590    | 0.077    |
| BP        |       | 0.060     | 0.047    | 0.017    | 1.191     | 0.966    | 0.137    |
| SVR(RBF)  | model 2 | 0.123     | 0.100    | 0.036    | 1.109     | 0.921    | 0.122    |
| RT        |       | 0.140     | 0.110    | 0.039    | 1.282     | 1.042    | 0.140    |
| RFR       |       | 0.137     | 0.111    | 0.040    | 1.210     | 0.983    | 0.128    |
| GBR       |       | 0.133     | 0.107    | 0.039    | 1.176     | 0.960    | 0.126    |
| ETR       |       | 0.133     | 0.108    | 0.039    | 1.157     | 0.948    | 0.124    |
| FFLSTM    |       | 0.039     | 0.028    | 0.010    | 0.589     | 0.453    | 0.060    |
| LSTM      |       | 0.050     | 0.042    | 0.015    | 1.279     | 1.041    | 0.141    |
| BP        |       | 0.058     | 0.046    | 0.016    | 1.191     | 0.966    | 0.137    |
| SVR(RBF)  | model 3 | 0.122     | 0.122    | 0.036    | 1.096     | 0.912    | 0.121    |
| RT        |       | 0.144     | 0.114    | 0.041    | 1.393     | 1.143    | 0.150    |
| RFR       |       | 0.131     | 0.105    | 0.038    | 1.170     | 0.964    | 0.127    |
| GBR       |       | 0.132     | 0.107    | 0.038    | 1.173     | 0.960    | 0.127    |
| ETR       |       | 0.132     | 0.107    | 0.038    | 1.189     | 0.972    | 0.127    |
| FFLSTM    |       | 0.033     | 0.027    | 0.010    | 0.556     | 0.445    | 0.058    |
| LSTM      |       | 0.090     | 0.068    | 0.025    | 1.059     | 0.832    | 0.108    |
| BP        |       | 0.066     | 0.050    | 0.018    | 1.191     | 0.966    | 0.137    |
| SVR(RBF)  | model 4 | 0.123     | 0.100    | 0.036    | 1.110     | 0.923    | 0.120    |
| RT        |       | 0.133     | 0.106    | 0.038    | 1.234     | 1.006    | 0.132    |
| RFR       |       | 0.122     | 0.098    | 0.035    | 1.210     | 0.982    | 0.128    |
| GBR       |       | 0.142     | 0.116    | 0.041    | 1.174     | 0.963    | 0.125    |
| ETR       |       | 0.130     | 0.106    | 0.038    | 1.167     | 0.962    | 0.126    |
| FFLSTM    |       | 0.055     | 0.037    | 0.013    | 0.572     | 0.414    | 0.056    |
| LSTM      |       | 0.105     | 0.075    | 0.027    | 1.169     | 0.969    | 0.129    |
| BP        |       | 0.068     | 0.051    | 0.018    | 1.191     | 0.966    | 0.137    |
| SVR(RBF)  | model 5 | 0.122     | 0.099    | 0.036    | 1.093     | 0.913    | 0.119    |
| RT        |       | 0.136     | 0.107    | 0.039    | 1.283     | 1.051    | 0.139    |
| RFR       |       | 0.131     | 0.105    | 0.038    | 1.169     | 0.962    | 0.126    |
| GBR       |       | 0.141     | 0.115    | 0.041    | 1.149     | 0.944    | 0.123    |
| ETR       |       | 0.129     | 0.104    | 0.038    | 1.106     | 0.918    | 0.119    |
| FFLSTM    |       | 0.030     | 0.026    | 0.009    | 0.535     | 0.404    | 0.055    |
| LSTM      |       | 0.068     | 0.054    | 0.019    | 1.540     | 1.193    | 0.166    |
| BP        |       | 0.066     | 0.052    | 0.019    | 1.191     | 0.966    | 0.137    |
| SVR(RBF)  | model 6 | 0.122     | 0.098    | 0.035    | 1.086     | 0.909    | 0.119    |
| RT        |       | 0.134     | 0.106    | 0.038    | 1.467     | 1.162    | 0.154    |
| RFR       |       | 0.128     | 0.103    | 0.037    | 1.173     | 0.961    | 0.127    |
| GBR       |       | 0.140     | 0.114    | 0.041    | 1.150     | 0.944    | 0.124    |
| ETR       |       | 0.125     | 0.101    | 0.036    | 1.123     | 0.933    | 0.122    |

Note: The text in red bold font is the optimal prediction result of given algorithm for each model.
Table 9. The accuracy for the final prediction model.

| Algorithm   | Station A | Station B |
|-------------|-----------|-----------|
|             | Accuracy (%) | Accuracy (%) |         |
| FFLSTM      | 98.87     | 94.03     |
| LSTM        | 97.9      | 87.78     |
| BP          | 98.15     | 87.37     |
| SVR(RBF)    | 96.42     | 87.95     |
| RT          | 96.12     | 85.9      |
| RFR         | 96.87     | 87.3      |
| GBR         | 96.05     | 87.45     |
| ETR         | 96.18     | 87.63     |

Note: The text in red bold font is the optimal prediction result of the given algorithm for each model.

4.6. System Verification and Energy Saving Analysis

In order to further verify the prediction accuracy of the FFLSTM algorithm in the actual SDHS and its performance in formulating energy scheduling plans and improving energy efficiency, we embedded the FFLSTM algorithm into the smart district heating system of the Xingtai Heating Company for preliminary engineering verification. The operation process of the FFLSTM prediction algorithm in SDHS is shown in Figure 7.

![Figure 7](image_url)

Figure 7. The operation process of the FFLSTM prediction algorithm in a smart district heating system (SDHS).

The SCADA system collects and stores the state data of the system in real time. Every night, the system regularly retrains the FFLSTM deep neural network according to the historical data collected by SCADA. With the entry of the latest proximity data, the weights of the network model will be updated and more accurate. According to the input characteristic parameters of DHS and the outdoor temperature, the FFLSTM algorithm accurately predicts the hourly heating load for the next 24 h, which is shown in Figure 8, and then the heating dispatching department formulates the next energy-saving control strategy and completes the energy production scheduling in advance.
In order to meet the demand adjustment of heating load caused by the rapid change of outdoor temperature, the feed forward control logic in the system responds two hours ahead of time based on the results of the heating load obtained by the FFLSTM prediction algorithm. In this way, the SDHS response speed to the outdoor temperature is accelerated, the energy-saving control effect is improved, and the indoor temperature of the user side is more stable.

Through comprehensive monitoring of the data quality, the energy consumption data of heat exchange stations are evaluated, and the energy saving and consumption reduction results are comprehensively and accurately reflected. By comparing the energy utilization before and after the SDHS, we found that FFLSTM has a high accuracy for predicting the heating load for the next 24 h, and the heating load has a significant negative correlation with the outdoor temperature—that is, the change of the heating load and outdoor temperature show a reverse trend.

Figure 8 shows that the FFLSTM has a high accuracy for predicting the heating load for the next 24 h, and the heating load has a significant negative correlation with the outdoor temperature—that is, the change of the heating load and outdoor temperature show a reverse trend.

In order to meet the demand adjustment of heating load caused by the rapid change of outdoor temperature, the feed forward control logic in the system responds two hours ahead of time based on the results of the heating load obtained by the FFLSTM prediction algorithm. In this way, the SDHS response speed to the outdoor temperature is accelerated, the energy-saving control effect is improved, and the indoor temperature of the user side is more stable.

Through comprehensive monitoring of the data quality, the energy consumption data of heat exchange stations are evaluated, and the energy saving and consumption reduction results are comprehensively and accurately reflected. By comparing the energy utilization before and after the SDHS, we found that FFLSTM has a high accuracy for predicting the heating load for the next 24 h, and the heating load has a significant negative correlation with the outdoor temperature—that is, the change of the heating load and outdoor temperature show a reverse trend.

In order to meet the demand adjustment of heating load caused by the rapid change of outdoor temperature, the feed forward control logic in the system responds two hours ahead of time based on the results of the heating load obtained by the FFLSTM prediction algorithm. In this way, the SDHS response speed to the outdoor temperature is accelerated, the energy-saving control effect is improved, and the indoor temperature of the user side is more stable.

Through comprehensive monitoring of the data quality, the energy consumption data of heat exchange stations are evaluated, and the energy saving and consumption reduction results are comprehensively and accurately reflected. By comparing the energy utilization before and after the SDHS, we found that FFLSTM has a high accuracy for predicting the heating load for the next 24 h, and the heating load has a significant negative correlation with the outdoor temperature—that is, the change of the heating load and outdoor temperature show a reverse trend.

In order to meet the demand adjustment of heating load caused by the rapid change of outdoor temperature, the feed forward control logic in the system responds two hours ahead of time based on the results of the heating load obtained by the FFLSTM prediction algorithm. In this way, the SDHS response speed to the outdoor temperature is accelerated, the energy-saving control effect is improved, and the indoor temperature of the user side is more stable.

Through comprehensive monitoring of the data quality, the energy consumption data of heat exchange stations are evaluated, and the energy saving and consumption reduction results are comprehensively and accurately reflected. By comparing the energy utilization before and after the SDHS, we found that FFLSTM has a high accuracy for predicting the heating load for the next 24 h, and the heating load has a significant negative correlation with the outdoor temperature—that is, the change of the heating load and outdoor temperature show a reverse trend.

In order to meet the demand adjustment of heating load caused by the rapid change of outdoor temperature, the feed forward control logic in the system responds two hours ahead of time based on the results of the heating load obtained by the FFLSTM prediction algorithm. In this way, the SDHS response speed to the outdoor temperature is accelerated, the energy-saving control effect is improved, and the indoor temperature of the user side is more stable.

Through comprehensive monitoring of the data quality, the energy consumption data of heat exchange stations are evaluated, and the energy saving and consumption reduction results are comprehensively and accurately reflected. By comparing the energy utilization before and after the SDHS, we found that FFLSTM has a high accuracy for predicting the heating load for the next 24 h, and the heating load has a significant negative correlation with the outdoor temperature—that is, the change of the heating load and outdoor temperature show a reverse trend.

In order to meet the demand adjustment of heating load caused by the rapid change of outdoor temperature, the feed forward control logic in the system responds two hours ahead of time based on the results of the heating load obtained by the FFLSTM prediction algorithm. In this way, the SDHS response speed to the outdoor temperature is accelerated, the energy-saving control effect is improved, and the indoor temperature of the user side is more stable.

Through comprehensive monitoring of the data quality, the energy consumption data of heat exchange stations are evaluated, and the energy saving and consumption reduction results are comprehensively and accurately reflected. By comparing the energy utilization before and after the SDHS, we found that FFLSTM has a high accuracy for predicting the heating load for the next 24 h, and the heating load has a significant negative correlation with the outdoor temperature—that is, the change of the heating load and outdoor temperature show a reverse trend.

In order to meet the demand adjustment of heating load caused by the rapid change of outdoor temperature, the feed forward control logic in the system responds two hours ahead of time based on the results of the heating load obtained by the FFLSTM prediction algorithm. In this way, the SDHS response speed to the outdoor temperature is accelerated, the energy-saving control effect is improved, and the indoor temperature of the user side is more stable.

Through comprehensive monitoring of the data quality, the energy consumption data of heat exchange stations are evaluated, and the energy saving and consumption reduction results are comprehensively and accurately reflected. By comparing the energy utilization before and after the SDHS, we found that FFLSTM has a high accuracy for predicting the heating load for the next 24 h, and the heating load has a significant negative correlation with the outdoor temperature—that is, the change of the heating load and outdoor temperature show a reverse trend.

In order to meet the demand adjustment of heating load caused by the rapid change of outdoor temperature, the feed forward control logic in the system responds two hours ahead of time based on the results of the heating load obtained by the FFLSTM prediction algorithm. In this way, the SDHS response speed to the outdoor temperature is accelerated, the energy-saving control effect is improved, and the indoor temperature of the user side is more stable.

Through comprehensive monitoring of the data quality, the energy consumption data of heat exchange stations are evaluated, and the energy saving and consumption reduction results are comprehensively and accurately reflected. By comparing the energy utilization before and after the SDHS, we found that FFLSTM has a high accuracy for predicting the heating load for the next 24 h, and the heating load has a significant negative correlation with the outdoor temperature—that is, the change of the heating load and outdoor temperature show a reverse trend.

In order to meet the demand adjustment of heating load caused by the rapid change of outdoor temperature, the feed forward control logic in the system responds two hours ahead of time based on the results of the heating load obtained by the FFLSTM prediction algorithm. In this way, the SDHS response speed to the outdoor temperature is accelerated, the energy-saving control effect is improved, and the indoor temperature of the user side is more stable.

Through comprehensive monitoring of the data quality, the energy consumption data of heat exchange stations are evaluated, and the energy saving and consumption reduction results are comprehensively and accurately reflected. By comparing the energy utilization before and after the SDHS, we found that FFLSTM has a high accuracy for predicting the heating load for the next 24 h, and the heating load has a significant negative correlation with the outdoor temperature—that is, the change of the heating load and outdoor temperature show a reverse trend.

In order to meet the demand adjustment of heating load caused by the rapid change of outdoor temperature, the feed forward control logic in the system responds two hours ahead of time based on the results of the heating load obtained by the FFLSTM prediction algorithm. In this way, the SDHS response speed to the outdoor temperature is accelerated, the energy-saving control effect is improved, and the indoor temperature of the user side is more stable.
after the system-embedded FFLSTM prediction algorithm, compared with the previous year, the heat consumption of station A and station B in 2017–2018 heating season decreased by 9.7% and 8.2%, respectively, which is shown in Table 10. The corresponding comprehensive economic cost decreased significantly.

### Table 10. The comparison of the energy utilization for station A and station B.

| Year | Heat Consumption (GJ) |
|------|-----------------------|
|      | Station A | Station B |
| 2017 | 10,762 | 31,601 |
| 2018 | 9,711  | 29,010 |
| Energy saving rate | 9.7% | 8.2% |

5. Conclusions

The development of intelligent district heating systems will be a trend in the future. Accurate forecasting of the heating load based on artificial intelligence technology is the basis and guarantee of energy-saving control strategies. Such strategies can improve the energy efficiency and formulate an energy dispatching plan. The FFLSTM model for heat load prediction in a DHS was proposed in this paper. The biggest advantage of the FFLSTM algorithm is that it is a unified representation method that integrates three different time-scale characteristics—proximity, periodicity, and trend—when multiple impact factors of the heating load of the district heating system are considered. Various parameters, such as the learning rate, batch size, time step, hide unit, and iteration count, were thoroughly investigated. Detailed experimental comparisons between the proposed FFLSTM algorithm and state-of-the-art algorithms, such as LSTM, BP, SVR, RT, RFR, GBR, and ETR, were performed to verify the superiority of the proposed algorithm. It can be concluded that FFLSTM has a better prediction performance and higher accuracy compared to other prediction algorithms. Furthermore, a total of six prediction models with different time-scale features were proposed and compared in detail. Finally, we embedded the algorithm into the real SDHS and carried out preliminary engineering verification. The results show that the algorithm has a better prediction accuracy of the heating load. Combined with the feed-forward control technology, an energy-saving control strategy was constructed, which improves the energy utilization of the system. Compared with last year, the comprehensive heat consumption index was reduced by more than 8%.

In future work, we will continue to explore the optimization of an energy-saving control strategy based on artificial intelligence. In the parameter selection, the factor of indoor temperature will be added to improve the reliability and accuracy of the prediction.

**Author Contributions:** This paper is the fruit of all authors’ collaboration. Problem formulation and Conceptualization, G.X., T.L., Y.P., and J.S.; methodology, G.X., Y.P., and J.S.; software, Y.P. and J.S.; validation, G.X., T.L., and J.S.; data curation, G.X., T.L., and J.S.; writing—original draft preparation, G.X., T.L., Y.P., and J.S.; writing—review and editing, J.S., C.Q., and Z.W.; supervision, G.X., T.L.; project administration, G.X., T.L., and J.S.

**Funding:** This work was supported by the National Natural Science Foundation of China (Grant No.61702157) and the Postdoctoral Fund of Hebei Province of China (No. B2017005002).

**Conflicts of Interest:** The authors declare no conflict of interest.
Nomenclature

ANFIS adaptive neuro-fuzzy inferences system  
ANN artificial neural networks  
BP back propagation  
DHS district heating system  
DL deep learning  
ELM extreme learning machines  
ETR extra trees regression  
FFA firefly algorithm  
FFLSTM feature fusion long short-term memory  
GA genetic algorithm  
GA–SVR genetic algorithm–support vector regression  
GBR gradient boosting regression  
IA immune algorithm  
IoT Internet of Things  
LSTM long short-term memory  
MAE mean absolute error  
MAPE mean absolute percentage error  
MVA multivariate autoregressive  
N-LSTM proximity LSTM  
PLS Partial Least Square  
P-LSTM periodic LSTM  
PM10 particulate matter 10  
PSO particle swarm optimization  
RBF radial kernel function  
RFR random forest regression  
RMSE root-mean-square error  
RNN recurrent neural networks  
RT regression tree  
SCADA supervisory control and data acquisition  
SDHS smart district heating system  
SVM support vector machine

References

1. Ebenstein, A.; Fan, M.Y.; Greenstone, M.; He, G.J.; Zhou, M.G. New evidence on the impact of sustained exposure to air pollution on life expectancy from China’s Huai River Policy. Proc. Natl. Acad. Sci. USA 2017, 114, 10384–10389. [CrossRef] [PubMed]
2. Chou, J.-S.; Bui, D.-K. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. Energy Build. 2014, 82, 437–446. [CrossRef]
3. Dotzauer, E. Simple model for prediction of loads in district-heating systems. Appl. Energy 2002, 73, 277–284. [CrossRef]
4. Karatasou, S.; Santamouris, M.; Geros, V. Modeling and predicting building’s energy use with artificial neural networks: Methods and results. Energy Build. 2006, 38, 949–958. [CrossRef]
5. Boithias, F.; El Mankibi, M.; Michel, P. Genetic algorithms based optimization of artificial neural network architecture for buildings’ indoor discomfort and energy consumption prediction. Build. Simul. 2012, 5, 95–106. [CrossRef]
6. Cai, H.; Shen, S.; Lin, Q.; Li, X.; Xiao, H. Predicting the energy consumption of residential buildings for regional electricity supply-side and demand-side management. IEEE Access 2019, 7, 30386–30397. [CrossRef]
7. Zhang, F.; Deb, C.; Lee, S.E.; Yang, J.; Shah, K.W. Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. Energy Build. 2016, 126, 94–103. [CrossRef]
8. Fan, G.-F.; Peng, L.-L.; Hong, W.-C.; Sun, F. Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression. Neurocomputing 2016, 173, 958–970. [CrossRef]
9. Hong, W.-C. Electric load forecasting by support vector model. Appl. Math. Model. 2009, 33, 2444–2454. [CrossRef]
10. Ding, Y.; Zhang, Q.; Yuan, T. Research on short-term and ultra-short-term cooling load prediction models for office buildings. Energy Build. 2017, 154, 254–267. [CrossRef]
11. Wang, M.; Tian, Q. Dynamic Heat Supply Prediction Using Support Vector Regression Optimized by Particle Swarm Optimization Algorithm. Math. Probl. Eng. 2016, 2016, 3968324. [CrossRef]
12. Al-Shammari, E.T.; Keivani, A.; Shamshirband, S.; Mostafeipour, A.; Yee, P.L.; Petkovic, D.; Ch, S. Prediction of heat load in district heating systems by Support Vector Machine with Firefly searching algorithm. Energy 2016, 95, 266–273. [CrossRef]
13. Lee, A.; Geem, Z.; Suh, K.-D. Determination of Optimal Initial Weights of an Artificial Neural Network by Using the Harmony Search Algorithm: Application to Breakwater Armor Stones. Appl. Sci. 2016, 6, 164. [CrossRef]
14. Dalipi, F.; Yildirim Yayilgan, S.; Gebremedhin, A. Data-Driven Machine-Learning Model in District Heating System for Heat Load Prediction: A Comparison Study. Appl. Comput. Intell. Soft Comput. 2016, 2016, 3403150. [CrossRef]
15. Sajjadi, S.; Shamshirband, S.; Alizamir, M.; Yee, P.L.; Mansor, Z.; Manaf, A.A.; Altameem, T.A.; Mostafeipour, A. Extreme learning machine for prediction of heat load in district heating systems. Energy Build. 2016, 122, 222–227. [CrossRef]
16. Roy, S.S.; Roy, R.; Balas, V.E. Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of MARS and ELM. Renew. Sustain. Energy Rev. 2018, 82, 4256–4268.
17. Kato, K.; Sakawa, M.; Ishimaru, K.; Ushiro, S.; Shibano, T. Heat Load Prediction through Recurrent Neural Network in District Heating and Cooling Systems. In Proceedings of the 2008 IEEE International Conference on Systems, Man And Cybernetics, Singapore, 12–15 October 2008; p. 1400.
18. Shamshirband, S.; Petković, D.; Enayatifar, R.; Hanan Abdullah, A.; Marković, D.; Lee, M.; Ahmad, R. Heat load prediction in district heating systems with adaptive neuro-fuzzy method. Renew. Sustain. Energy Rev. 2015, 48, 760–767. [CrossRef]
19. Qing, X.; Niu, Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy 2018, 148, 461–468. [CrossRef]
20. Li, C.; Ding, Z.; Zhao, D.; Yi, J.; Zhang, G. Building Energy Consumption Prediction: An Extreme Deep Learning Approach. Energies 2017, 10, 1025. [CrossRef]
21. Fu, G. Deep belief network based ensemble approach for cooling load forecasting of air-conditioning system. Energy 2018, 148, 269–282. [CrossRef]
22. Liu, F.; Cai, M.; Wang, L.; Lu, Y. An ensemble model based on adaptive noise reducer and over-fitting prevention LSTM for multivariate time series forecasting. IEEE Access 2019, 7, 26102–26115. [CrossRef]
23. Han, L.Y.; Peng, Y.X.; Li, Y.H.; Yong, B.B.; Zhou, Q.G.; Shu, L. Enhanced Deep Networks for short-Term and Medium-Term Load Forecasting. IEEE Access 2019, 7, 4045–4055. [CrossRef]
24. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. Energies 2018, 11, 1626. [CrossRef]
25. Tian, C.; Ma, J.; Zhang, C.; Zhan, P. A Deep Neural Network Model for Short-Term Load Forecast Based on Long Short-Term Memory Network and Convolutional Neural Network. Energies 2018, 11, 3493. [CrossRef]
26. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M. Single and Multi-Sequence Deep Learning Models for Short and Medium Term Electric Load Forecasting. Energies 2019, 12, 149. [CrossRef]
27. Liang, Y.; Ke, S.; Zhang, J.; Yi, X.; Zheng, Y. GeoMAN: Multi-Level Attention Networks for Geo-Sensory Time Series Prediction; IJCAI: Stockholm, Sweden, 19 July 2018; pp. 3428–3434.
28. Lopez, E.; Valle, C.; Allende, H.; Gil, E.; Madsen, H. Wind Power Forecasting Based on Echo State Networks and Long Short-Term Memory. *Energies* 2018, 11, 526. [CrossRef]

29. Shi, X.; Lei, X.; Huang, Q.; Huang, S.; Ren, K.; Hu, Y. Hourly Day-Ahead Wind Power Prediction Using the Hybrid Model of Variational Model Decomposition and Long Short-Term Memory. *Energies* 2018, 11, 3227. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).