

Article

Hourly Heat Load Prediction for Residential Buildings based on Multiple Combination Models: A Comparative Study

Wenhan An ¹, Xiangyuan Zhu ¹, Kaimin Yang ¹, Moon Keun Kim ² and Jiyong Liu ^{1,*}

¹ School of Thermal Engineering, Shandong Jianzhu University, Jinan 250101, China; awh1633@163.com (W.A.), zhuxiangyuan20@sdjzu.edu.cn (X.Z.), yangkaimin@sdjzu.edu.cn (K.Y.)

² Department of Built Environment, Oslo Metropolitan University, Oslo, N-0130, Norway; moon.kim@oslomet.no

* Correspondence: JXL83@sdjzu.edu.cn

Abstract: The accurate prediction of residential heat load is crucial for effective heating system design, energy management and cost optimization. In order to further improve the prediction accuracy of the model, this study introduced principal component analysis (PCA), the minimum sum of squares of the combined prediction errors (minSSE), genetic algorithm (GA) and firefly algorithm (FA) into back propagation (BP) and ELMAN neural networks, and established three kinds of combined prediction models. The proposed methodologies are evaluated using real-world data collected from residential buildings over a period of one year. Obtained results of the PCA-BP-ELMAN, FA-ELMAN, GA-BP models are compared with the neural network before optimization. The experimental results show that the combined prediction models have higher prediction accuracy. The Mean Absolute Percentage Error (MAPE) evaluation indices of the three combined models are distributed between 5.95% and 7.05%. The FA-ELMAN model is the combination model with the highest prediction accuracy, and its MAPE is 5.95%, which is 2.25% lower than the MAPE of an individual neural network. This research contributes to the field by providing a comprehensive and effective framework for residential heat load prediction, which can be valuable for building energy management and optimization.

Keywords: heat load prediction; combined models; principal component analysis; genetic algorithm; firefly algorithm

1. Introduction

Currently, the proportion of energy consumption in the construction sector in our country's total energy consumption has reached 30%, and the energy consumption generated by heating accounts for 25% of the total energy consumption in buildings [1]. As of 2020, the national coverage of centralized heating has reached 12.266 billion square meters. However, the increase in carbon emissions and pollutant emissions caused by fuel combustion has resulted in serious environmental problems. In order to alleviate the energy crisis and improve environmental pollution, the government has implemented numerous measures in energy conservation and emission reduction. In our country, heating systems are mainly regulated through manual experience. However, due to the characteristics of centralized heating systems such as significant time delays, strong non-linearity, high energy consumption, and multiple influencing factors, the traditional inefficient heating regulation model has created issues such as significant heat waste and uneven distribution of heating. In practice, the energy consumption in heating processes still cannot meet the heating needs of users. Therefore, it is necessary to incorporate heat load forecasting into the heating system to guide heating operations. By utilizing a heat load prediction model, the heating system can adjust in advance based on the predicted load.

Citation: To be added by editorial staff during production.

Academic Editor: Firstname Last-name

Received: date

Revised: date

Accepted: date

Published: date



Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Building an accurate, efficient, and reliable heat load prediction model is crucial in ensuring the centralized heating system can meet heating demands effectively and on demand.

The commonly used heat load predicting models include data-driven model and mechanism model. By using building energy simulation software such as EnergyPlus and TRNSYS to evaluate building heat load, mechanism model requires a complex calculation process and usually takes a long time to get the result, so it is not suitable for real-time energy consumption prediction [2, 3]. However, the data-driven model can produce prediction results in a relatively short time, which is suitable for real-time energy consumption prediction and rapid decision making. Common data-driven models used in various fields include statistical analysis techniques, machine learning algorithms, and deep learning architectures. Statistical methods tend to overfit when handling diverse influencing factors, leading to poor predictive performance of the models [4-7]. With the wide application of artificial intelligence algorithm in the prediction field [8-11], more and more scholars have started researching load predicting methods founded on machine learning [12-14]. Some commonly used approaches in machine learning include artificial neural networks (ANN) [15], support vector machine [16], extreme learning machine [17], random forest (RF) [18] and regression tree [19]. Park et al. [20] separately used partial least squares, neural networks, and support vector regression (SVR) to establish load forecasting models for district heating, and the research results showed that SVR had the lowest average relative error. Meng et al. [21] built an ELMAN neural network prediction model, and on this basis proposed a DR control strategy for office buildings. Wang et al. [22] implemented 12 models for single-building heat load predicting, discussing the impact of prediction level and input uncertainty on load forecasting accuracy. Introducing these models for heat load forecasting has provided good predictive results, demonstrating that an individual model based on machine learning can derive feasible computational rules from a significant volume of historical operational data.

Several researchers have adopted hybrid methods that combine an optimization algorithm with a machine learning model [23]. In an improved model, various parameters can be optimized, such as weights, bias, and others [24]. The combination models can be divided into two classifications. The initial classification is the ensemble models that combine multiple individual models [25]. These models are trained on the same dataset using different individual models. The predictions from each individual model are then combined to generate an optimal output. The second category is the improved models that combine an individual model with optimization methods. For example, Fan et al. [18] developed an integrated model for predicting commercial building loads. They used a genetic algorithm (GA) to optimize the weights of the eight prediction models and combined the results to obtain the final prediction. The experimental results showed that the ensemble model outperformed typical individual models in terms of prediction accuracy. Qi et al. [26] introduced GA into back propagation (BP) neural network to establish a heat load prediction model. Meanwhile, the date type was quantified when selecting input variables, which ensured the prediction accuracy. Principal component analysis (PCA) was used to obtain reasonable model inputs and build a prediction model [27]. Zhao et al. [28] proposed a convolutional neural network (CNN) hybrid short-term heat load prediction model based on the adaptive T-distribution satin Bowerbird algorithm. Compared with other prediction models, The Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) of the mixed model decreased by 18.08% and 16.26% respectively. Song et al. [25] proposed a heat load prediction model based on convolution neural network-long short-term memory (CNN-LSTM), which obtains a more accurate dynamic heat load model. Al-shammari et al. [29] proposed a SVR regional heating system thermal load prediction model based on the firefly algorithm (FA).

In the field of heat load prediction, the innovation of this research lies in combining optimization algorithm with machine learning model to improve the accuracy and reliability of heat load prediction by comparing the predicting performance of different combination models. The heat load prediction models of PCA-BP-ELMAN, FA-ELMAN and

GA-BP were established respectively. We used the real heat load data and MATLAB platform for experimental simulation. By comparing and analyzing the performance of different combined prediction models, we explore the effectiveness and applicability of the combination of optimization algorithm and machine learning model in heat load prediction, and determine the best combined model to provide more accurate and reliable heat load prediction.

2. Methods

The research framework, depicted in Figure 1, encompasses key components such as data collection, data preprocessing, feature selection, implementation and prediction of classical individual machine learning models, as well as the implementation and prediction of combined models. To evaluate model performance, four evaluation indicators are employed to identify an individual model with superior predictive capabilities. Building upon this, PCA, minSSE algorithm, GA, and FA are incorporated into multiple efficient individual models for heat load prediction, resulting in the creation of three combined models. These models are assessed using the generated dataset and compared with both the pre-optimized prediction model and different combined models to validate the superior prediction accuracy of the combined models. Furthermore, the study investigates the optimization effects of various combination methods.

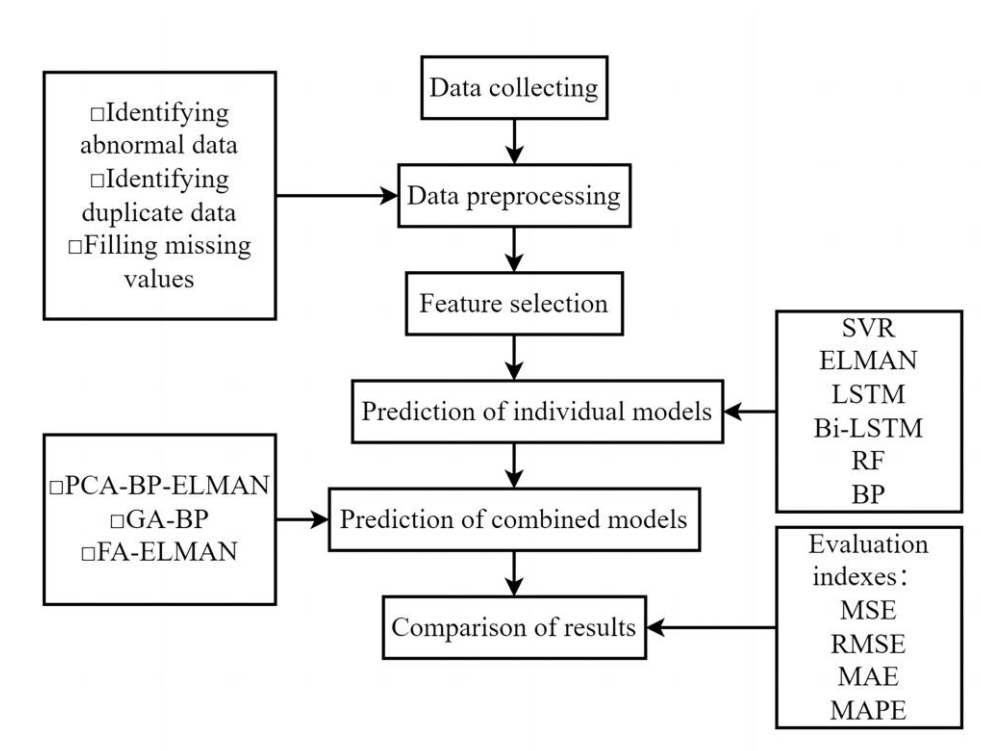


Figure 1. Research framework of this study.

2.1. Optimization algorithm

2.1.1. Principal component analysis

The heating load is influenced by various factors, including meteorological, system, social, and building factors [30]. However, considering all these factors comprehensively when calculating it can lead to increased complexity and difficulty in problem-solving. PCA offers a solution by reducing multiple initial indicators that affect the problem to a few major components that are most effective. This simplifies the problem analysis [31, 32]. The main steps of this method include:

(1). The n samples with m influencing factors are subjected to data standardization, then:

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}, (j = 1, 2, \dots, m), \quad (1)$$

$$s_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}, (j = 1, 2, \dots, m), \quad (2)$$

$$\widetilde{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, (i = 1, 2, \dots, n; j = 1, 2, \dots, m), \quad (3)$$

where m is the number of influence factors; n is the number of influence factor samples; x_{ij} is the i th index of the j th influence factor, and \bar{x}_j is the sample mean of the j th influence factor. s_j is the sample standard deviation of the j th influence factor. \widetilde{x}_{ij} is a standardized indicator.

(2). The eigenvalues and eigenvectors of the correlation coefficient matrix $R = (r_{ij})_{m \times m}$ are calculated, where r_{ij} represents the correlation coefficient between the i th and j th influencing factors. A higher value indicates a stronger correlation between the two factors. To calculate the correlation coefficient, use the equation (4) and sort the eigenvalues in descending order, $\lambda_1, \lambda_2, \dots, \lambda_m$. From these feature vectors, m new index vectors can be formed, as shown in equation (5):

$$r_{ij} = \frac{\sum_{k=1}^n \widetilde{x}_{ki} \times \widetilde{x}_{kj}}{n-1}, (i, j = 1, 2, \dots, m), \quad (4)$$

where \widetilde{x}_{ki} is the k th standardized index of the i th influence factor; \widetilde{x}_{kj} is the k th standardized index of the j th influence factor.

$$\begin{cases} F_1 = u_{11}\widetilde{x}_1 + u_{21}\widetilde{x}_2 + \dots + u_{m1}\widetilde{x}_m \\ F_2 = u_{12}\widetilde{x}_1 + u_{22}\widetilde{x}_2 + \dots + u_{m2}\widetilde{x}_m \\ \vdots \\ F_m = u_{1m}\widetilde{x}_1 + u_{2m}\widetilde{x}_2 + \dots + u_{mm}\widetilde{x}_m \end{cases}, \quad (5)$$

where F_m is the m th principal component; $\widetilde{x}_1, \widetilde{x}_2, \dots, \widetilde{x}_m$ is the impact factor vector after standardization; u_j is the eigenvector corresponding to λ_j , $u_j = (u_{1j}, u_{2j}, \dots, u_{mj})$, $j = 1, 2, \dots, m$.

(3). The contribution rate is calculated and the principal component is extracted.

$$b_j = \frac{\lambda_j}{\sum_{k=1}^m \lambda_k}, (j = 1, 2, \dots, m), \quad (6)$$

$$\alpha_p = \frac{\sum_{k=1}^p \lambda_k}{\sum_{k=1}^m \lambda_k}, \quad (7)$$

where the information contribution rate of b_j as the main component F_j ; The cumulative contribution rate of α_p as the main component is taken as $\alpha_p > 90\%$.

2.1.2. The minSSE algorithm

The model is built by assigning distinct weight coefficients to the predictions generated by multiple individual models. Subsequently, a new predicted value is derived through the employment of the weighted average method. The main focus of the combined model lies in identifying the weight coefficient for each model, and there exist numerous calculation methods that can be utilized for this purpose. This study utilizes the minSSE algorithm, which aims to minimize the sum of squares of the combined prediction error. By applying this algorithm, the optimal weight coefficient can be determined.

Define the actual value at time t as x_t , $t = 1, 2, \dots, p$; Define the prediction of the i th model as x_{it} , $i = 1, 2, \dots, n$; $t = 1, 2, \dots, p$; Let $\{l_1, l_2, \dots, l_n\}$ denote the weight coefficients of the n models, and these coefficients must satisfy the following requirements.

$$\sum_{i=1}^n l_i = 1, i = 1, 2, \dots, n, \quad (8)$$

\hat{x}_t as the predicted value of the combined model at time t is calculated as follows:

$$\hat{x}_t = \sum_{i=1}^n l_i x_{it}, i = 1, 2, \dots, n, \quad (9)$$

The calculation process of minSSE algorithm is as follows:

(1). The prediction error of the combined prediction model at time t is calculated.

$$e_t = x_t - \hat{x}_t = \sum_{i=1}^n l_i e_{it}, \quad (10)$$

Where $e_{it} = x_t - x_{it}$ was defined as the prediction error of the i th individual model at time t .

(2). The combined prediction errors, denoted as e_t , are squared and then summed to calculate the value of J .

$$J = \sum_{t=1}^p e_t^2 = \sum_{t=1}^p \sum_{i=1}^n \sum_{j=1}^n l_i e_{it} l_j e_{jt},$$

$$= [l_1 l_2 \dots l_n] \begin{bmatrix} \sum_{t=1}^p e_{1t}^2 & \sum_{t=1}^p e_{1t} e_{2t} & K & \sum_{t=1}^p e_{1t} e_{nt} \\ \sum_{t=1}^p e_{2t} e_{1t} & \sum_{t=1}^p e_{2t}^2 & K & \sum_{t=1}^p e_{2t} e_{nt} \\ M & M & K & M \\ \sum_{t=1}^p e_{nt} e_{1t} & \sum_{t=1}^p e_{nt} e_{2t} & K & \sum_{t=1}^p e_{nt}^2 \end{bmatrix} \begin{bmatrix} l_1 \\ l_2 \\ M \\ l_n \end{bmatrix}, \quad (11)$$

$$\text{Let } L = \begin{bmatrix} l_1 \\ l_2 \\ M \\ l_n \end{bmatrix}, E = \begin{bmatrix} \sum_{t=1}^p e_{1t}^2 & \sum_{t=1}^p e_{1t} e_{2t} & K & \sum_{t=1}^p e_{1t} e_{nt} \\ \sum_{t=1}^p e_{2t} e_{1t} & \sum_{t=1}^p e_{2t}^2 & K & \sum_{t=1}^p e_{2t} e_{nt} \\ M & M & K & M \\ \sum_{t=1}^p e_{nt} e_{1t} & \sum_{t=1}^p e_{nt} e_{2t} & K & \sum_{t=1}^p e_{nt}^2 \end{bmatrix}, J = L^T E L$$

(3). Let $H = [1KK1]^T$; The following optimization model is constructed with the minimum sum of squares of prediction errors as the optimal criterion.

$$\min J = L^T E L, \quad (12)$$

$$\text{s.t.} \begin{cases} H^T L = 1 \\ L \geq 0 \end{cases}$$

(4). The Lagrange multiplier method is used to solve the model, and the optimal combinatorial weight vector L and the minimum sum of square of combinatorial prediction error J are obtained.

$$L = \frac{E^{-1}H}{H^T E^{-1}H}, \quad (13)$$

$$J = \frac{1}{H^T E^{-1}H}, \quad (14)$$

(5). Based on the weight coefficients assigned to each individual model, the predicted value of the combined model is calculated.

2.1.3. The genetic algorithm

The steps involved in the GA primarily consist of population initialization, fitness function, selection, crossover, and mutation [33]. The following sections outline each step in detail:

(1). Population initialization:

To facilitate computer recognition and storage, each individual in the population is encoded using a binary array. Each individual represents a real string consisting of four components: input layer and hidden layer link weight, hidden layer weight, hidden layer and output layer weight, output layer weight.

(2). Fitness function:

The fitness function is determined by the sum of squared errors between the predicted output value and the expected output value of the network. The calculation equation is as follows:

$$F = \frac{1}{\sum_{k=1}^n |y_k - x_k|}, \quad (15)$$

where n is the number of output nodes; y_k is the expected value of the k th node of the BP neural network. x_k is the predicted value of the k th node.

(3). Selection:

The roulette method, which employs a selection strategy based on the proportion of fitness values, is utilized for determining the selection probability (P) of each individual (i). This probability is calculated as follows:

$$p = \frac{f_i}{\sum_{j=1}^N f_j}, \quad (16)$$

where f_i is the fitness of individual X_i ; N is the number of population individuals.

(4). Cross:

The process involves randomly selecting an individual chromosome from a set of chromosomes and crossing it over with another chromosome, employing a specific method to generate a novel individual.

(5). Mutation:

The j th gene is selected for mutation in individual i th. At a particular locus, the gene is substituted with an allele, leading to the creation of a new individual.

2.1.4. The firefly algorithm

FA is a biologically-inspired optimization algorithm that utilizes group search. Within a specified range, fireflies emit light randomly, and their attractiveness is directly proportional to the intensity of their light emissions. During the process of movement and aggregation, each firefly's flight position changes based on the luminosity of the brightest individual in the current time zone. Ultimately, all fireflies converge towards the brightest individual.

The brightness function varying with distance in Gaussian form is defined as:

$$I = I_0 e^{-\gamma d^2}, \quad (17)$$

where I_0 is the maximum brightness of fireflies; d is the distance between two fireflies; γ is the absorption coefficient of light intensity.

The attractiveness function $\beta(d)$ of a firefly at a distance reference point d is defined as:

$$\beta(d) = \beta_0 e^{-\gamma d^2}, \quad (18)$$

where β_0 is the attraction of fireflies at the light source.

For two fireflies located in w_i and w_j , the distance between them is expressed as:

$$d_{ij} = \|w_i - w_j\| = \sqrt{\sum_{k=1}^n (w_{i,k} - w_{j,k})^2}, \quad (19)$$

where n is the dimension of the target problem; $w_{i,k}$ are the coordinates of the i th firefly on the k dimension; $w_{j,k}$ is the coordinates of the j th firefly on the k dimension.

When firefly i is attracted to another firefly j , which is brighter, the movement pattern is:

$$w_i = w_i + \beta_0 e^{-\gamma d_{ij}^2} (w_j - w_i) + \alpha \varepsilon_i, \quad (20)$$

where α is the step factor; ε_i is the vector of random numbers derived based on a Uniform distribution or Gaussian distribution.

2.2. Construction of models

2.2.1. Construction of individual models

(1). SVR is a machine learning technique that is based on statistical learning theory and the principle of minimizing structural risk. Its primary objective is to prevent overfitting and ensure that the model can generalize and approximate accurately. SVR is particularly effective for identifying non-linear relationships between the dependent variable and multiple influencing factors, especially when working with limited sample sizes [34–36]. The essence is to project the training samples onto a high-dimensional plane and identify an appropriate hyperplane that can effectively separate the samples.

(2). Elman neural network is a recurrent neural network composed of four layers. The input layer is responsible for signal transmission, while the output layer performs linear weighting. The hidden layer units are capable of employing either linear or non-linear transfer functions. The context layer is utilized to store the previous moment's output values of the hidden layer units and subsequently feed them back to the network's input [37].

(3). LSTM is an improvement of recurrent neural networks (RNNs) that addresses the issue of long-term dependencies. It not only maintains the connectivity among hidden layer nodes in RNNs but also introduces the concept of filtering past states [38, 39]. The results of long time series training with good continuity are better than those of ordinary RNNs models [40, 41]. The key of LSTM is the cell state, which carries pertinent information outward. The cell state undergoes updates through the forget gate, input gate, and output gate.

(4). Based on LSTM, bidirectional long short-term memory (Bi-LSTM) transforms the single-direction LSTM layer into both Backward and Forward layers. The Forward layer processes the sequential information at the current time step, while the Backward layer reads the same sequence in reverse, incorporating the inverse order information. The final output combines the respective outputs from the Backward and Forward layers at each time step. In LSTM, the hidden output signals can be transmitted not only to neighboring grids but also to the next layer's grids. During the training process, the two state neurons do not interact, allowing for the expansion of the network into a typical feed-forward neural network. Building upon this, the adjustment of network weights is achieved through both forward and backward propagation methods [42, 43].

(5). RF is an improved algorithm of ensemble learning. It involves randomly sampling a training set of m samples, with replacement, from the original dataset. The size of the sample matches that of the original dataset. Then, it constructs m decision trees $\{h(x, \theta_n), n = 1, 2, \dots, m\}$ (where θ_n is an independently and identically distributed random vector). The final output $\hat{y}(x)$ of RF regression is obtained by averaging the outputs of the m decision trees.

(6). BP neural network structure mainly consists of three layers: the input layer, the output layer, and the hidden layer. These layers are fully interconnected, and there is no coupling between neurons within the same layer. The network's functionality is achieved through forward signal propagation and backward error propagation [44, 45].

2.2.2. Construction of combined models

In order to improve the accuracy of artificial neural network in building energy prediction, the following three problems need to be solved: First, it is easy to fall into the local optimal problem in the training stage; Secondly, the problem of input selection; Finally, the question of hyperparameter setting [46]. To solve the above problems, different optimization algorithms are introduced into building energy consumption prediction. The following is the construction process of the combined models.

(1). PCA-BP-ELMAN

The parameters of the input layer were optimized using PCA, and the optimal weight coefficient of the model was obtained through the minSSE algorithm, which utilized the predicted and actual values of the BP and ELMAN neural network models. The PCA-BP-ELMAN model was then constructed based on these weight coefficients for prediction. The specific calculation process is depicted in Figure 2.

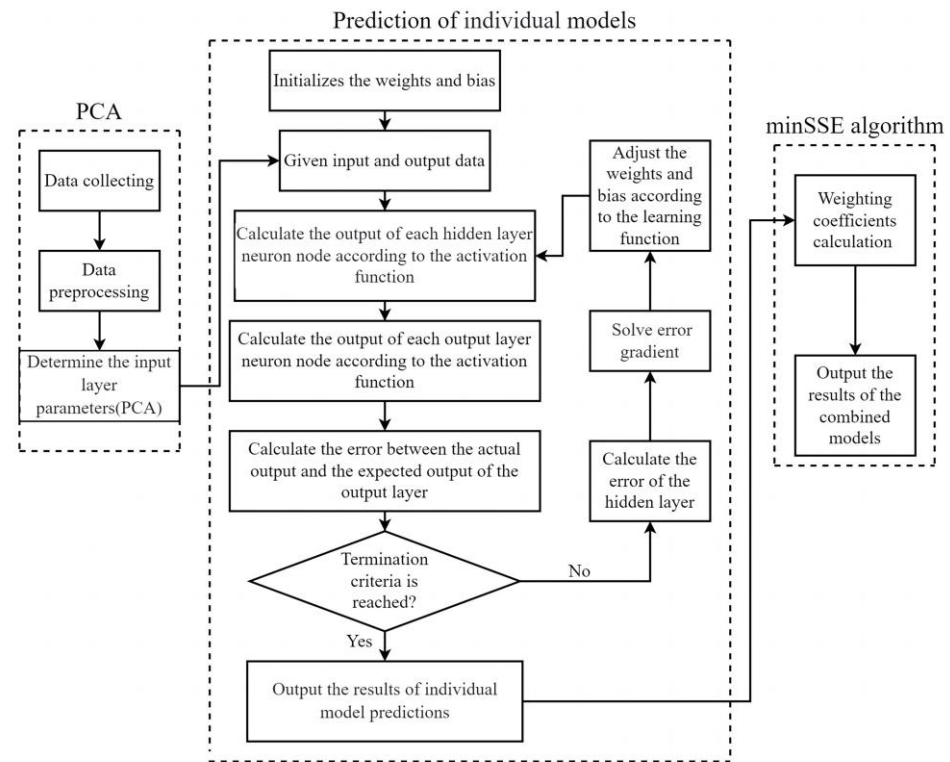


Figure 2. A flowchart of the PCA-BP-ELMAN model.

(2). FA-ELMAN

The optimal firefly individual is selected by changing the position and brightness of fireflies, followed by the optimal replacement of initial weights and bias [47]. Then the nonlinear relationship between each input parameter and the hourly heat load is established by ELMAN neural network, so as to enable accurate prediction. The ELMAN neural network is optimized by using FA, as shown in Figure 3.

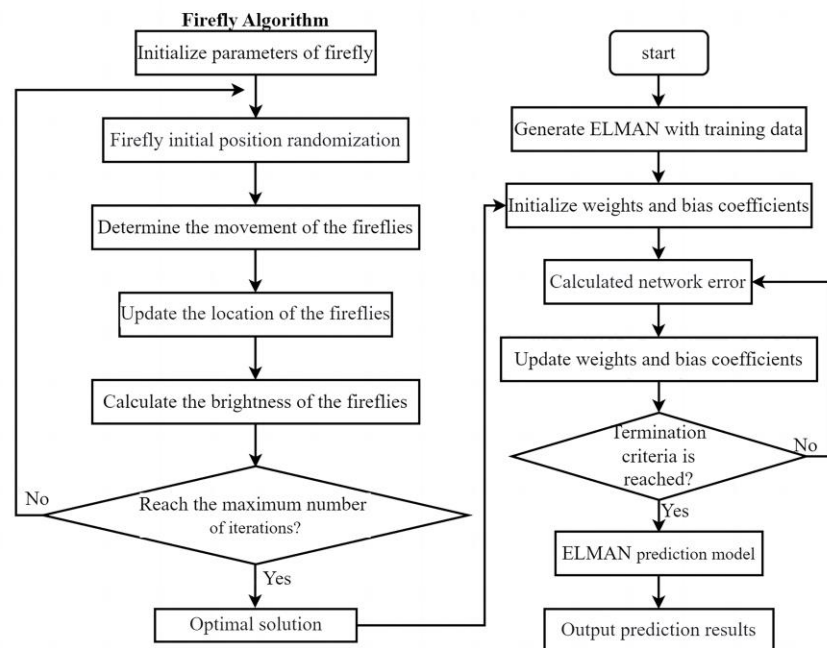


Figure 3. A flowchart of the FA-ELMAN model.

(3). GA-BP

GA is a methodology that simulates natural biological evolution and proves to be effective in solving complex problems, including nonlinear and global optimization [48]. In this study, GA is used to optimize the weights and bias of the BP neural network. Subsequently, the optimized weights and bias are employed to train the network and obtain the optimal solution. The specific calculation process is depicted in Figure 4.

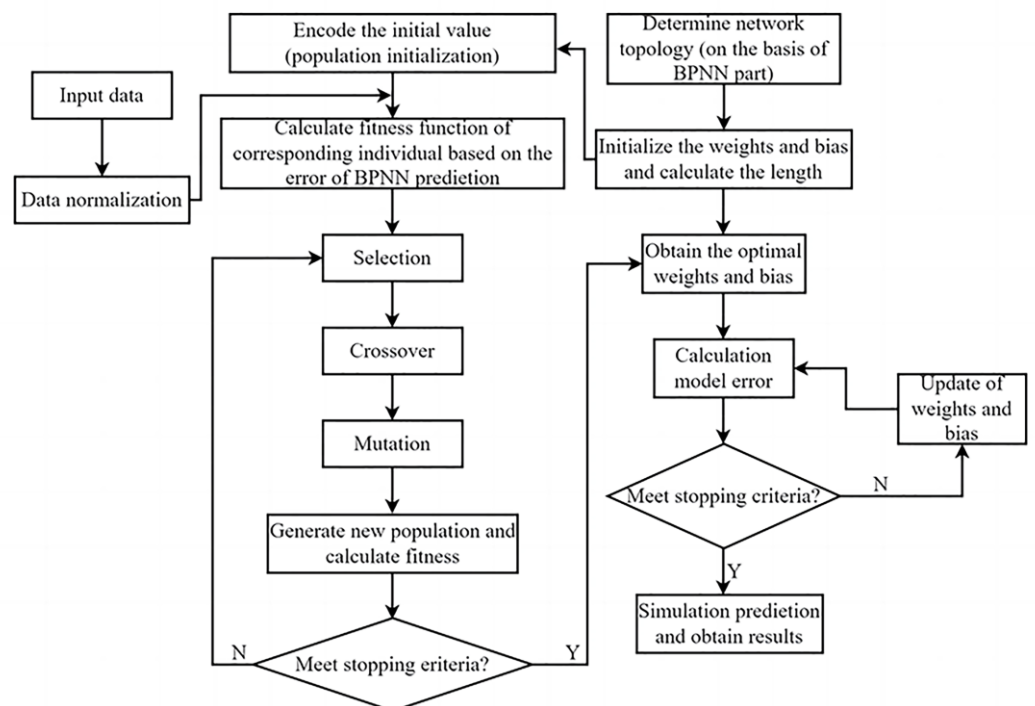


Figure 4. A flowchart of the GA-BP model.

2.3. Data processing and analysis

2.3.1. Data collecting and preprocessing

Weifang City is located in the western part of the Shandong Peninsula and serves as a central city within the cluster. Geographically, it spans from 118°10' to 120°01' east longitude and 35°41' to 37°26' north latitude. The city covers a total land area of 1,616,723.99 hectares. Weifang experiences specific climatic conditions characterized by high winter temperatures, relatively abundant precipitation, and reduced sunlight. Overall, the climate is favorable. During the spring season, temperatures rise, precipitation increases, and sunlight is abundant, often leading to frequent meteorological disasters. In summer, the temperature becomes pleasant, with higher precipitation but less sunlight. Nevertheless, the climate remains suitable. In autumn, temperatures remain high, precipitation decreases slightly, and sunlight diminishes further. Weifang was selected as the research area for forecasting the heat supply load.

The data utilized in this study were obtained from two sources - the China Meteorological Data Network and the Weifang thermal data visualization monitoring platform. The dataset used for analysis was collected from a specific community. The type of enclosure structure is an ordinary external wall, the thickness of the wall is 240 mm, and the heat transfer coefficient is 2.03 W/m²/°C. The research focused on the complete heating season, with the query time range set from November 15, 2021, to March 15, 2022. To ensure data accuracy, a query interval of 3 minutes was used. Historical operating data and heat loads were sampled at 1-hour intervals.

Since there are instances of abnormal and missing data, it is essential to apply proper data preprocessing techniques. For missing data, interpolation methods are used to fill in the gaps. Regarding abnormal data, the 3 σ criterion is initially used to identify outliers, and then interpolation methods are employed to determine appropriate replacements. When long-term repetitions or omissions are detected, these occurrences are regarded as anomalies in the data collection process and are subsequently eliminated.

2.3.2. Feature selection

The heating load is affected by many factors such as meteorological factors, system factors, social factors and building factors. Temperature, relative humidity, wind speed, solar radiation and other meteorological factors will affect the heating load. With the increase of temperature, the value of heating load gradually decreases, and with the decrease of temperature, the value of heating load gradually increases. When the temperature is unchanged, with the reduction of relative humidity, the evaporation of sweat is enhanced, the human body will feel cold, and the heating load will increase. As the wind speed increases, the heating load will also increase. Solar radiation also has a certain influence on the heating load, when the solar radiation is small, the heating load is larger, but its influence on the heat load is small. The influence of some characteristics of the heating network itself on the heat load is mainly reflected in the supply water temperature, return water temperature, circulation flow and so on. Social factors such as residents' heating habits, local policies, number of inhabitants and development of the population will affect the change of heating load, and building factors such as its structure, function, geographical location, number of flats, the size of the heated volume and the quality of thermal insulation will also affect the heating load [30]. However, the influence of social factors on the heat load makes its change relatively slow, so it cannot be taken into account in the short-term heat load prediction, and once the building is formed, the influence of its own factors on the heat load can also be ignored.

To sum up, the preliminary analysis of the factors affecting heating load and the related literature review [49-51], Eight characteristics, including supply water temperature at time t-1, return water temperature at time t-1, heat load at time t-1, circulation flow at time t-1, solar radiation at time t-1, outdoor temperature at time t, wind speed at time t,

and relative humidity at time t , were determined as input variables of the prediction model, and heat load at time t , were determined as output variables.

2.4. Evaluation indices

The four selected accuracy evaluation indices for assessing the precision of the prediction model are as follows: Mean Square Error (MSE), RMSE, Mean Absolute Error (MAE), and MAPE [52].

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2, \quad (21)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2}, \quad (22)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|, \quad (23)$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{x_i - y_i}{y_i} \right|, \quad (24)$$

where n —the number of samples; x_i —the i^{th} simulated value, GJ/h; y_i —the i^{th} actual measured value, GJ/h; \bar{x} —the mean of the simulated values, GJ/h; \bar{y} —the mean value of the measured value, GJ/h.

For the same dataset, smaller values of MSE, RMSE, MAE, and MAPE indicate reduced discrepancies between the predicted and actual values, implying a higher level of precision in the model.

3. Results

3.1. Comparisons of individual models

Using MATLAB as the calculation platform, the heat and meteorological data of a community in Weifang from November 15, 2021, to March 15, 2022 were used as training data to train the model. On this basis, the optimal input parameters of each model were selected as follows: (1). SVR: The kernel function is RBF, the loss function $p=0.4$, the penalty factor $C=1000$, and the kernel parameter $\gamma=0.01$; (2). ELMAN: the number of neurons in the input layer is 8, the number of neurons in the output layer is 1, the number of neurons in the hidden layer is 8, the maximum number of iterations is 1000, the learning rate is 0.01, and the minimum error of the training target is 10^{-5} ; (3). LSTM and Bi-LSTM: The optimization method was adam algorithm, the maximum number of iterations was 300, the initial learning rate was 0.1, and the learning rate dropped to 0.1×0.1 after 50 training sessions. The optimized neural network was trained using trainNet function. The model data is in a period of 24 hours, and the input data formats are batch_size:24, time_step:1, iput_size:8; (4) RF: The number of decision trees is 500, the minimum number of leaves of each decision tree is 3, and the regRF_train function is called to train the model; (5). BP: Set the maximum number of iterations to 1000, the learning rate to 0.01 and the BP neural network structure to 8-14-1.

To assess the accuracy of these models, MSE, RMSE, MAE, and MAPE were used as evaluation indices, and their corresponding error indicators are presented in Table 1. The results in Table 1 indicate that the MAPE values for the RF, SVR, LSTM, and Bi-LSTM models are all greater than 15%. Moreover, as shown in Figure 5, these four models exhibit a significant deviation between the predicted and actual values, consistently overestimating the heat load. Consequently, these models demonstrate poor prediction performance and are not recommended for heat load prediction. Conversely, the BP and ELMAN neural network prediction models display lower evaluation indices with MAPE values less than 10%, MSE values below $0.1 \text{ GJ}^2/\text{h}^2$, RMSE values less than 0.3 GJ/h , and MAE values less than 0.2 GJ/h . Additionally, the predicted values from these two models exhibit minimal deviation from the actual values. Therefore, these models exhibit good prediction

performance and are suitable for heat load prediction. Furthermore, in comparison to the other four models with inadequate prediction performance, the BP and ELMAN neural network models demonstrate a maximum reduction in MAPE of 20.5% and 19.75%, respectively.

Table 1. Prediction results of the individual models.

Model	MSE (GJ ² /h ²)	RMSE (GJ/h)	MAE (GJ/h)	MAPE (%)
BP	0.052	0.229	0.169	7.45
ELMAN	0.061	0.246	0.184	8.20
RF	0.162	0.403	0.351	16.25
SVR	0.435	0.660	0.601	27.95
LSTM	0.202	0.450	0.412	19.03
Bi-LSTM	0.187	0.432	0.392	18.02

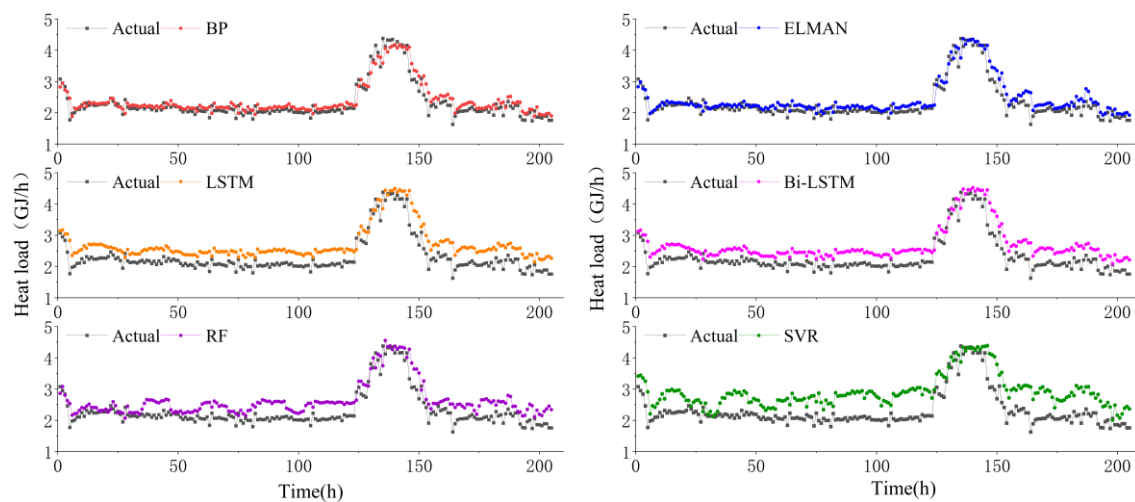


Figure 5. Comparison of the performance between proposed and individual models.

3.2. Results of combined models

(1). PCA-BP-ELMAN

To effectively reduce redundancy and correlation in the initial characteristic variable data for heat load prediction, dimensionality reduction processing was performed using SPSS software. The results, summarized in Table 2, showed that the first four principal components accounted for a cumulative contribution rate of 90.610%. This indicates that these principal components effectively captured and represented the majority of information contained in the initial eight feature variables. As a result, principal components Y₁ to Y₄ were selected as the input data for further analysis.

Table 2. Contribution rate of each component.

Component	Characteristic value	Rate of contribution (%)	Accumulating contribution rate (%)
1	3.655	45.684	45.684
2	1.872	23.406	69.090
3	0.937	11.710	80.800
4	0.785	9.810	90.610
5	0.425	5.312	95.922

6	0.168	2.102	98.024
7	0.118	1.479	99.503
8	0.040	0.497	100.000

Table 3 presents the linear relationship between each principal component and its respective variable. In particular, principal component Y_1 , which accounts for a contribution rate of 45.684%, is largely associated with return water temperature, circulation flow rate, historical heat load, and outdoor temperature. Principal component Y_2 is primarily linked to relative humidity and solar radiation. Additionally, principal components Y_3 and Y_4 are mainly influenced by wind speed and water supply temperature, respectively.

Table 3. Principal component coefficient.

Characteristic variable	Y_1	Y_2	Y_3	Y_4
$T_s(t-1)$.334	.088	.576	-.564
$T_R(t-1)$.482	.174	-.022	-.086
$F(t-1)$.410	.164	-.455	.371
$S(t-1)$	-.167	.603	-.198	-.251
$T(t)$	-.451	.171	-.115	-.172
$W(t)$	-.091	.360	.631	.660
$H(t)$.084	-.620	.082	.093
$Q(t-1)$	-.491	.169	-.061	.021

Each model in this study was trained using data that underwent PCA dimensionality reduction. The initial parameters for both the BP and ELMAN neural networks were set as follows: the transfer function was chosen as tansing, the training was conducted for 1000 iterations, the target error was defined as 10^{-5} , the learning rate was set to 0.001, and the normalization range was $[-1, 1]$. The minSSE algorithm was employed to obtain the optimal weight coefficients for each model, based on the predicted and actual values. Table 4 presents these weight coefficients, which range between 0 and 1, and their sum totals to 1.

Table 4. Weighting coefficients of individual models.

Model	PCA-BP	PCA-ELAMN
PCA-BP-ELMAN	0.52	0.48

Each individual model is used to predict the heat load, resulting in forecasted outcomes. These prediction results are then combined and weighted to calculate the final prediction results, which represent the prediction of the PCA-BP-ELMAN model. The respective evaluation indicators can be observed in Table 5, and a comparison between the predicted and actual results is presented in Figure 6.

Table 5. Prediction results of the PCA-BP-ELMAN model.

Model	MSE (GJ ² /h ²)	RMSE (GJ/h)	MAE (GJ/h)	MAPE (%)
PCA-BP-ELMAN	0.062	0.248	0.151	6.29

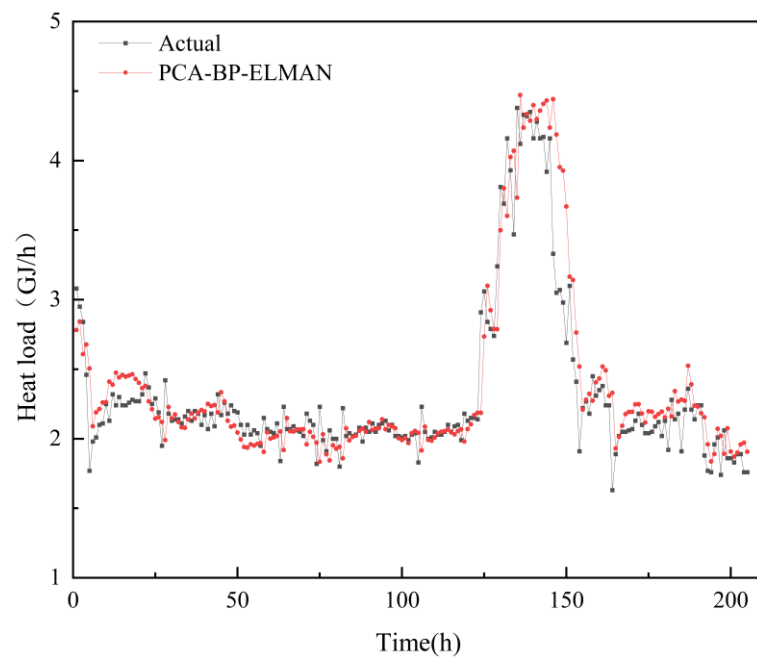


Figure 6. Comparison between the actual data of heat load and results of a one-hour- ahead heat load prediction using the PCA-BP-ELMAN method.

(2). FA-ELMAN

The initial parameters for the firefly algorithm are set as follows: a population size of 50, a maximum iteration limit of 30, a critical error threshold of 0.01, a maximum attraction value (β_0) of 2, a light intensity absorption coefficient (γ) of 1, and a step factor (α) of 0.02. Similarly, the initial parameter settings for the ELMAN neural network are the same as described in the PCA-BP-ELMAN section. The respective evaluation indicators can be observed in Table 6, and a comparison between the predicted and actual results is presented in Figure 7.

Table 6. Prediction results of the FA-ELMAN model.

Model	MSE (GJ ² /h ²)	RMSE (GJ/h)	MAE (GJ/h)	MAPE (%)
FA-ELMAN	0.050	0.214	0.142	5.95

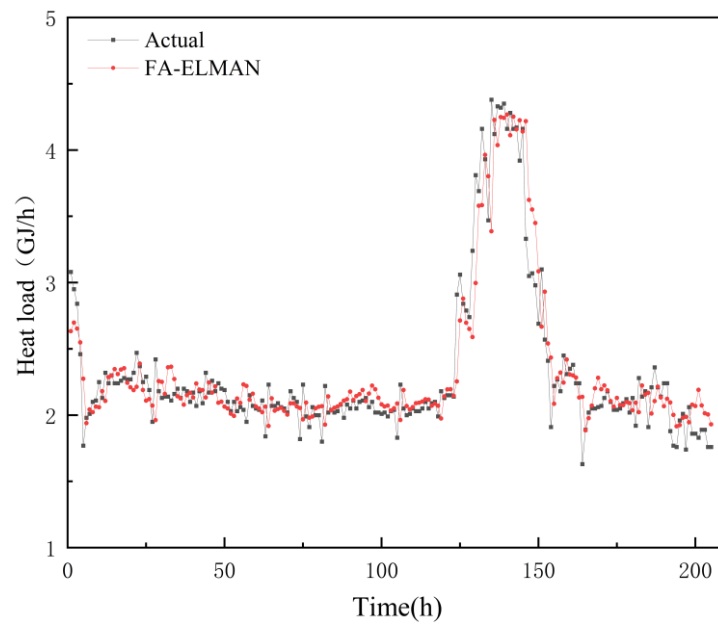


Figure 7. Comparison between the actual data of heat load and results of a one-hour- ahead heat load prediction using the FA-ELMAN method.

(3). GA-BP

The initial parameters for the genetic algorithm are set as follows: the number of iterations is 30, the population size is 10, the crossover probability is 0.8, and the mutation probability is 0.2. The initial parameter setting for the BP neural network section is the same as described in the PCA-BP-ELMAN method. The respective evaluation indicators can be observed in Table 7, and a comparison between the predicted and actual results is presented in Figure 8.

Table 7. Prediction results of the GA-BP model.

Model	MSE (GJ ² /h ²)	RMSE (GJ/h)	MAE (GJ/h)	MAPE (%)
GA-BP	0.049	0.222	0.161	7.05

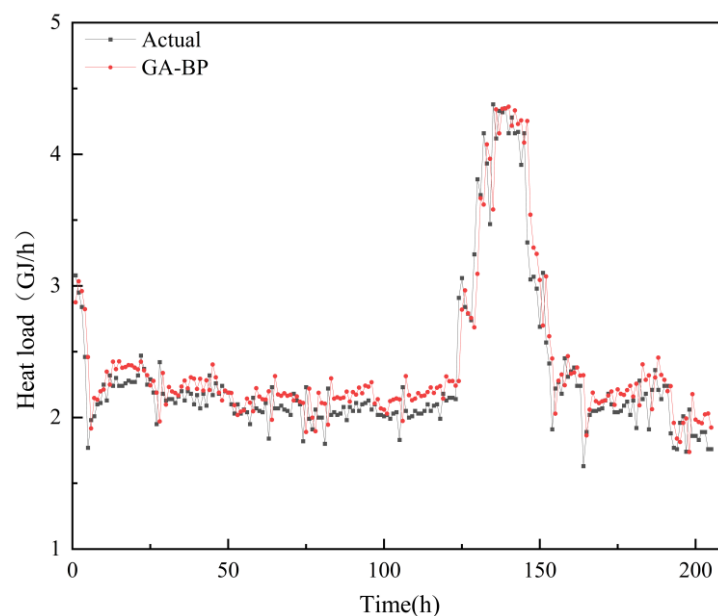


Figure 8. Comparison between the actual data of heat load and results of a one-hour-ahead heat load prediction using the GA-BP method.

3.3. Comparisons of individual and combined models

In order to visually and effectively analyze and compare the optimization effects of different combined prediction models, namely the BP and ELMAN neural network models, column charts and dot plots are utilized to assess the MSE, RMSE, MAE, and MAPE. These charts serve as clear and intuitive representations of the performance of the various models.

Figure 9 presents a comparison of evaluation errors between the BP model, ELMAN neural network model, and three combined prediction models. It is evident from the figure that the ELMAN neural network model exhibits the highest error index among the various prediction models, while the FA-ELMAN model shows the lowest error index. Additionally, the MSE, RMSE, MAE, and MAPE values of the three combined models consistently prove to be smaller than those of the traditional BP and ELMAN neural network models. This confirms the effectiveness of each employed optimization algorithm. Moreover, it is noteworthy that the model optimized by the FA algorithm demonstrates the most substantial reduction in each index, followed by the PCA-minSSE algorithm. Conversely, the model optimized by the GA algorithm displays the smallest reduction in each index. These findings suggest that the predictive performance of the three combined prediction models follows the order of FA-ELMAN, PCA-BP-ELMAN, and GA-BP, from high to low.

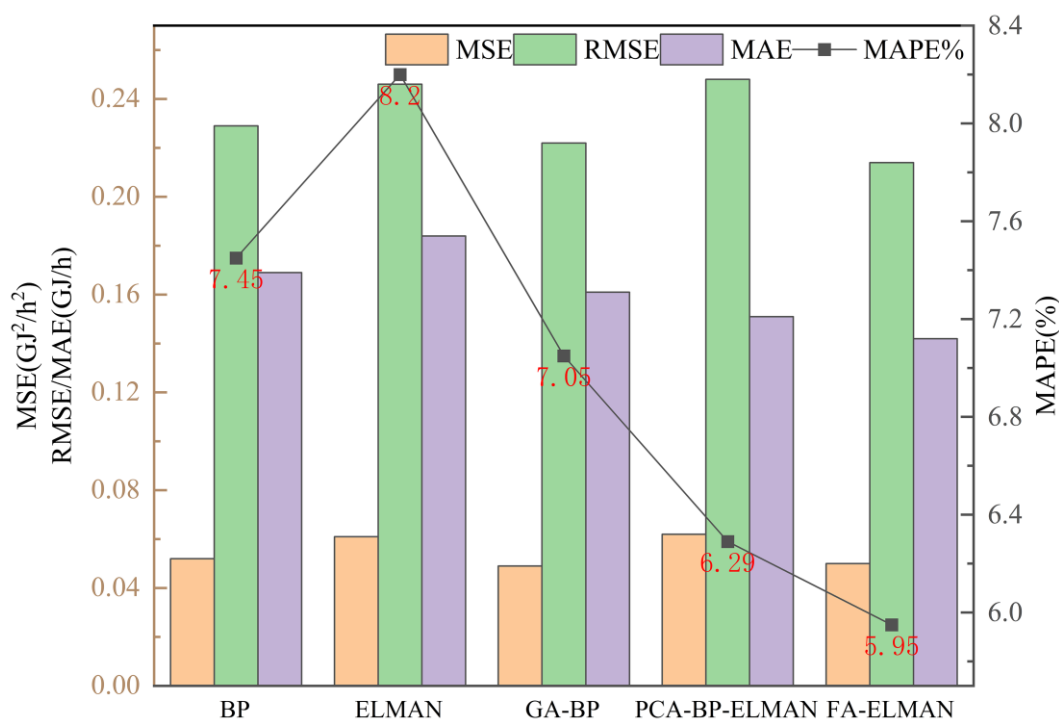


Figure 9. Comparison performance of BP, ELMAN, combined models.

The error index of the FA-ELMAN combined prediction model (MSE=0.046 GJ²/h², RMSE=0.214 GJ/h, MAE=0.142 GJ/h, MAPE=5.95%) was compared to that of the traditional ELMAN neural network model (MSE=0.060 GJ²/h², RMSE=0.246 GJ/h, MAE=0.184 GJ/h, MAPE=8.20%). The results showed that the FA-ELMAN combined prediction model achieved significant improvements. Specifically, the MSE decreased by 0.014 GJ²/h², the RMSE decreased by 0.032 GJ/h, the MAE decreased by 0.042 GJ/h, and the MAPE decreased by 2.25%.

4. Discussion

Accurate and reasonable heat load prediction models serve as a robust foundation for ensuring on-demand heating in centralized heating systems. In this study, we leverage winter heat load data from a residential building in Weifang City and integrate it with hourly meteorological data. This study primarily proposes three combined prediction models and examines the effectiveness of optimizing the combination of load prediction models, which improves the accuracy of heat load prediction. Nevertheless, further research can be undertaken in the following areas:

(1). Due to data availability limitations, we can only use limited historical data for model training and validation. Therefore, further research may consider using longer time ranges of data as well as more frequent sampling to improve the generalization ability of the model.

(2). The influence of indoor temperature and wind direction on heat load is disregarded due to limited data availability. Additionally, factors such as day-night variations, seasonal changes, and other potential influences are ignored to simplify the prediction model, owing to the complexity of its design. To address these limitations, future studies should expand the range of collected data by enhancing monitoring conditions. Furthermore, the use of sophisticated methods can help effectively handle day-night and seasonal changes, thereby enabling a more comprehensive exploration of their impact.

(3). In order to develop a predictive model that can be broadly applied to common building types, we focus our research on a specific type of building structure, namely residential buildings with standard building materials and insulation. However, it is worth mentioning that the principles and methods used in our study can potentially be applied to other types of buildings as well. We acknowledge that the building structure and its thermal properties play a crucial role in heat load estimation. In future research, we recognize the importance of investigating how different types of building structures, materials, and insulation systems impact heat load prediction. This will further enhance the applicability and accuracy of our predictive models across diverse building types.

(4). The predictive models developed during this study demonstrated excellent prediction performance in simulation experiments. However, they have not yet undergone practical testing by heating supply companies. If future research allows for continuous adjustments of the models based on real-time conditions during practical experiments, it could truly enable on-demand heating.

(5). Due to space constraints, we have not been able to cover the findings of all countries in detail in the introductory chapter. However, in our future research, we will comprehensively analyze the research status and results of heat load forecasting in various countries and regions from an international perspective.

(6). In order to enhance the rationality of experimental design, PCA, as a feature selection method, was introduced into FA-ELAMN and GA-BP models. The respective evaluation indicators can be observed in Table 8, and a comparison between the predicted and actual results is presented in Figure 10. According to Table 8, after PCA method was introduced into the models, the value of MAPE decreased by 0.31% and 0.35% respectively. Therefore, by introducing PCA into GA-BP and FA-ELMAN models, the prediction accuracy of the new combined models has been improved slightly.

Table 8. Prediction results of the PCA-GA-BP and PCA-FA-ELMAN models.

Model	MSE (GJ ² /h ²)	RMSE (GJ/h)	MAE (GJ/h)	MAPE (%)
PCA-GA-BP	0.061	0.246	0.159	6.74
PCA-FA-ELMAN	0.043	0.208	0.133	5.60

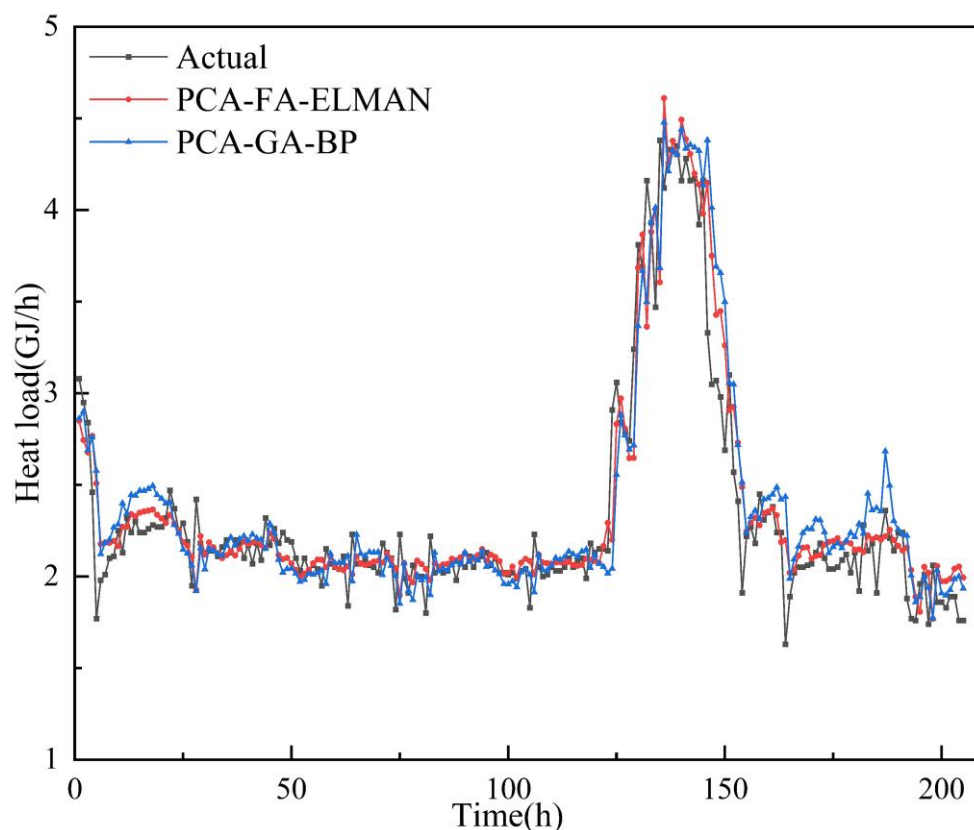


Figure 10. Comparison between the actual data of heat load and results of a one-hour-ahead heat load prediction using the PCA-GA-BP and PCA-FA-ELMAN methods.

5. Conclusions

PCA-BP-ELMAN, FA-ELMAN, and GA-BP are established to apply heat load prediction. This study aimed to compare the prediction accuracy of these three combined models. The main conclusions are as follows:

(1). To assess the prediction performance of six commonly used individual models, we selected eight initial characteristics: water supply temperature, return water temperature, circulation flow, solar radiation, heat load, outdoor temperature, wind speed, and relative humidity at time $t-1$. The results indicated that the BP and ELMAN neural network prediction models had a MAPE of less than 10% and a MSE of less than $0.1 \text{ GJ}^2/\text{h}^2$. Furthermore, the RMSE and MAE were both less than 0.3 GJ/h and 0.2 GJ/h respectively, suggesting that these models demonstrate high prediction accuracy.

(2). The dimensionality of the characteristic variables that impact heating load is reduced from 8 to 4 using PCA, and the input parameters of the neural network are optimized accordingly. In predicting heat load, the most significant principal components are return water temperature, circulation flow, historical heat load, and outdoor temperature. Following these are relative humidity, solar radiation, water supply temperature, and wind speed.

(3). The application of PCA-minSSE, FA, and GA algorithms in heat load prediction has been shown to improve the accuracy of the prediction model. The models with the highest prediction performance, ranked from high to low, are FA-ELMAN, PCA-BP-ELMAN, and GA-BP. When compared to individual neural network models with superior prediction performance, the combined model exhibits a decrease in MAPE of 2.25%.

Author Contributions: Conceptualization, W.A., X.Z. and J.L.; methodology, J.L., M.K. and K.Y.; software, W.A.; validation, W.A. and J.L.; formal analysis, W.A.; investigation, W.A. K.Y., M.K. and J.L.; resources, J.L.; data curation, W.A. and J.L.; writing—original draft preparation, W.A., X.Z.,

M.K. and J.L.; writing—review and editing, W.A., X.Z., K.Y., M.K. and J.L.; visualization, W.A.; supervision, J.L., M.K. and K.Y.; project administration, J.L.; funding acquisition, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Natural Science Foundation of Shandong Province (ZR2021ME199, ZR2020ME211 and ZR2021ME237) and the Support Plan for Outstanding Youth Innovation Team in Colleges and Universities of Shandong Province (2019KJG005).

Data Availability Statement: Not applicable.

Acknowledgments: This work was also supported by the Plan of Introduction and Cultivation for Young Innovative Talents in Colleges and Universities of Shandong Province.

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

Nomenclature

b_j	information contribution rate of principal component F_j
d_{ij}	the distance between any two fireflies i and j
e_t	the prediction error of the combined prediction model at time t
e_{it}	the prediction error of the i^{th} individual model at time t .
f_i	the fitness of individual X_i
F	the fitness of individual
F_m	the m^{th} principal component
$F(t-1)$	circulation flow at time $t-1$, m^3/h
H	let $H=[1KK1]^T$
$H(t)$	relative humidity at time t , %
I_0	the maximum brightness of fireflies
J	the sum of squares of combined prediction errors
l_i	the weight coefficient
m	the number of influence factors
n	the number of influence factor samples
N	the number of population individuals
P	the selection probability
$Q(t-1)$	heat load at time $t-1$, GJ/h
r_{ij}	the correlation coefficient
R	the correlation coefficient matrix
s_j	the sample standard deviation of the j^{th} influence factor
$S(t-1)$	solar radiation at $t-1$ time, W/m^2
$T(t)$	outside temperature at time t , $^{\circ}\text{C}$
$T_R(t-1)$	return water temperature at time $t-1$, $^{\circ}\text{C}$
$T_S(t-1)$	water supply temperature at time $t-1$, $^{\circ}\text{C}$
u_{ij}	the eigenvector
w_i	the pattern of movement for the reference firefly i
$w_{i,k}$	the coordinates of the i^{th} firefly on the k dimension
$w_{j,k}$	the coordinates of the j^{th} firefly on the k dimension
W_t	wind speed at time t , m/s
\bar{x}	the mean of the simulated values
x_i	the i^{th} simulated value
\bar{x}_j	the sample mean of the j^{th} influence factor
x_t	the actual value of the heat load at time t
\hat{x}_t	the combined predicted value of heat load at time t
x_{ij}	the i^{th} index of the j^{th} influence factor
x_{it}	the heat load prediction of the i^{th} model at time t
\tilde{x}_{ij}	a standardized indicator
\tilde{x}_{ki}	the k^{th} standardized index of the i^{th} influence factor
\tilde{x}_{kj}	the k^{th} standardized index of the j^{th} influence factor
\bar{y}	the mean value of the measured value
y_i	the i^{th} actual measured value

Abbreviation

Bi-LSTM	bidirectional long short-term memory
BP	back propagation
CNN	convolutional neural network
FA	firefly algorithm
GA	genetic algorithm
LSTM	long short-term memory
minSSE	the minimum sum of squares of the combined prediction errors
MAE	Mean Absolute Error, GJ/h
MAPE	Mean Absolute Percentage Error, %
MSE	Mean Square Error, GJ ² /h ²
PCA	principal component analysis
RF	Random Forest
RMSE	Root Mean Square Error, GJ/h
RNNs	recurrent neural networks
SVR	support vector regression
Greek Symbols	
α	the step factor
α_p	the cumulative contribution rate
β_0	the attraction of fireflies at the light source
$\beta(d)$	the attractiveness function
γ	the absorption coefficient of light intensity
ε_i	the vector of random numbers
λ_j	the eigenvalues
Subscripts	
i	represents the i^{th}
j	represents the j^{th}
k	represents the k^{th}
t	time

References

- Zhang, Y.; Xia, J.; Fang, H.; Jiang, Y.; Liang, Z. Field tests on the operational energy consumption of Chinese district heating systems and evaluation of typical associated problems. *Energy and Buildings*. **2020**, *224*, 110269. <https://doi.org/10.1016/j.enbuild.2020.110269>.
- Eguía, P.; Granada, E.; Alonso, J.M.; Arce, E.; Saavedra, A. Weather datasets generated using kriging techniques to calibrate building thermal simulations with TRNSYS. *Journal of Building Engineering*. **2016**, *7*, 78–91. <https://doi.org/10.1016/j.jobee.2016.05.007>.
- Yu, S.; Cui, Y.; Xu, X.; Feng, G. Impact of Civil Envelope on Energy Consumption based on EnergyPlus. *Procedia Engineering*. **2015**, *121*, 1528–1534. <https://doi.org/10.1016/j.proeng.2015.09.130>.
- Amasyali, K.; El-Gohary, N.M. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*. **2018**, *81*, 1192–1205. <https://doi.org/10.1016/j.rser.2017.04.095>.
- Mat Daut, M.A.; Hassan, M.Y.; Abdullah, H.; Rahman, H.A.; Abdullah, M.P.; Hussin, F. Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review. *Renewable and Sustainable Energy Reviews*. **2017**, *70*, 1108–1118. <https://doi.org/10.1016/j.rser.2016.12.015>.
- Park, B.R.; Choi, E.J.; Hong, J.; Lee, J.H.; Moon, J.W. Development of an energy cost prediction model for a VRF heating system. *Applied Thermal Engineering*. **2018**, *140*, 476–486. <https://doi.org/10.1016/j.applthermaleng.2018.05.068>.
- Fan, C.; Ding, Y. Cooling load prediction and optimal operation of HVAC systems using a multiple nonlinear regression model. *Energy and Buildings*. **2019**, *197*, 7–17. <https://doi.org/10.1016/j.enbuild.2019.05.043>.
- Su, M.; Liu, J.; Zhou, S.; Miao, J.; Kim, M.K. Dynamic prediction of the pre-dehumidification of a radiant floor cooling and displacement ventilation system based on computational fluid dynamics and a back-propagation neural network: A case study of an office room. *Indoor and Built Environment*. **2022**, *31*, 2386–2410. <https://doi.org/10.1177/1420326X221107110>.

9. Ren, L.; An, F.; Su, M.; Liu, J. Exposure Assessment of Traffic-Related Air Pollution Based on CFD and BP Neural Network and Artificial Intelligence Prediction of Optimal Route in an Urban Area. *Buildings*. **2022**, *12*, 1227. <https://doi.org/10.3390/buildings12081227>.
10. Kim, M.K.; Cremers, B.; Liu, J.; Zhang, J.; Wang, J. Prediction and correlation analysis of ventilation performance in a residential building using artificial neural network models based on data-driven analysis. *Sustainable Cities and Society*. **2022**, *83*, 103981. <https://doi.org/10.1016/j.scs.2022.103981>.
11. Su, M.; Liu, J.; Kim, M.K.; Wu, X. Predicting moisture condensation risk on the radiant cooling floor of an office using integration of a genetic algorithm-back-propagation neural network with sensitivity analysis. *Energy and Built Environment*. **2024**, *5*, 110-129. <https://doi.org/10.1016/j.enbenv.2022.08.004>.
12. Liu, J.; Yin, Y. Power Load Forecasting Considering Climate Factors Based on IPSO-Elman Method in China. *Energies*. **2022**, *15*, 1236. <https://doi.org/10.3390/en15031236>.
13. Moayed, H.; Mosavi, A. Suggesting a Stochastic Fractal Search Paradigm in Combination with Artificial Neural Network for Early Prediction of Cooling Load in Residential Buildings. *Energies*. **2021**, *14*, 1649. <https://doi.org/10.3390/en14061649>.
14. Yousaf, A.; Asif, R.M.; Shakir, M.; Rehman, A.U.; S. Adrees, M. An Improved Residential Electricity Load Forecasting Using a Machine-Learning-Based Feature Selection Approach and a Proposed Integration Strategy. *Sustainability*. **2021**, *13*, 6199. <https://doi.org/10.3390/su13116199>.
15. Wei, Y.; Xia, L.; Pan, S.; Wu, J.; Zhang, X.; Han, M.; Zhang, W.; Xie, J.; Li, Q. Prediction of occupancy level and energy consumption in office building using blind system identification and neural networks. *Applied Energy*. **2019**, *240*, 276-294. <https://doi.org/10.1016/j.apenergy.2019.02.056>.
16. Zhao, H.X.; Magoules, F. Parallel Support Vector Machines Applied to the Prediction of Multiple Buildings Energy Consumption. *Journal of Algorithms Computational Technology*. **2010**, *4*, 231-250. <https://doi.org/10.1260/1748-3018.4.2.231>.
17. Guo, Y.; Wang, J.; Chen, H.; Li, G.; Liu, J.; Xu, C.; Huang, R.; Huang, Y. Machine learning-based thermal response time ahead energy demand prediction for building heating systems. *Applied Energy*. **2018**, *221*, 16-27. <https://doi.org/10.1016/j.apenergy.2018.03.125>.
18. Fan, C.; Xiao, F.; Wang, S. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. *Applied Energy*. **2014**, *127*, 1-10. <https://doi.org/10.1016/j.apenergy.2014.04.016>.
19. Chou, J-S.; Bui, D-K. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy and Buildings*. **2014**, *82*, 437-446. <https://doi.org/10.1016/j.enbuild.2014.07.036>.
20. Park, T.C.; Kim, U.S.; Kim, L-H.; Jo, B.W.; Yeo, Y.K. Heat consumption forecasting using partial least squares, artificial neural network and support vector regression techniques in district heating systems. *Korean Journal of Chemical Engineering*. **2010**, *27*, 1063-1071. <https://doi.org/10.1007/s11814-010-0220-9>.
21. Meng, Q.; Xi, Y.; Ren, X.; Li, H.; Jiang, L.; Yang, L. Thermal Energy Storage Air-conditioning Demand Response Control Using Elman Neural Network Prediction Model. *Sustainable Cities and Society*. **2022**, *76*, 103480. <https://doi.org/10.1016/j.scs.2021.103480>.
22. Wang, Z.; Hong, T.; Piette, M.A. Building thermal load prediction through shallow machine learning and deep learning. *Applied Energy*. **2020**, *263*, 114683. <https://doi.org/10.1016/j.apenergy.2020.114683>.
23. Bourdeau, M.; Zhai, X.; Nefzaoui, E.; Guo, X.; Chatellier, P. Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustainable Cities and Society*. **2019**, *48*, 101533. <https://doi.org/10.1016/j.scs.2019.101533>.
24. Yu, W.; Li, B.; Jia, H.; Zhang, M.; Wang, D. Application of multi-objective genetic algorithm to optimize energy efficiency and thermal comfort in building design. *Energy and Buildings*. **2015**, *88*, 135-143. <https://doi.org/10.1016/j.enbuild.2014.11.063>.
25. Song, J.; Zhang, L.; Xue, G.; Ma, Y.; Gao, S.; Jiang, Q. Predicting hourly heating load in a district heating system based on a hybrid CNN-LSTM model. *Energy and Buildings*. **2021**, *243*, 110998. <https://doi.org/10.1016/j.enbuild.2021.110998>.

26. Qi, L.; Feng, Z. Improved BP Neural Network of Heat Load Forecasting Based on Temperature and Date Type. *Journal of System Simulation*. **2018**, *30*, 1464-1472. <https://doi.org/10.16182/j.issn1004731x.joss.201804031>.
27. Ding, Y.; Zhang, Q.; Yuan, T.; Yang, K. Model input selection for building heating load prediction: A case study for an office building in Tianjin. *Energy and Buildings*. **2018**, *159*, 254-270. <https://doi.org/10.1016/j.enbuild.2017.11.002>.
28. Zhao, A.; Mi, L.; Xue, X.; Xi, J.; Jiao, Y. Heating load prediction of residential district using hybrid model based on CNN. *Energy and Buildings*. **2022**, *266*, 112122. <https://doi.org/10.1016/j.enbuild.2022.112122>.
29. Al-Shammari, E.T.; Keivani, A.; Shamshirband, S.; Mostafaeipour, A.; Yee, P.L.; Petković, D.; Ch, S. Prediction of heat load in district heating systems by Support Vector Machine with Firefly searching algorithm. *Energy*. **2016**, *95*, 266-273. <https://doi.org/10.1016/j.energy.2015.11.079>.
30. Kapalo, P.; Adamski, M. The Analysis of Heat Consumption in the Selected City. *Lecture Notes in Civil Engineering*. **2021**, *100 LNCE*, 158-165. https://doi.org/10.1007/978-3-030-57340-9_20.
31. Yang, M. Numerical study of the heat transfer to carbon dioxide in horizontal helically coiled tubes under supercritical pressure. *Applied Thermal Engineering*. **2016**, *109*, 685-696. <https://doi.org/10.1016/j.applthermaleng.2016.08.121>.
32. Lam, J.C.; Wan, K.K.W.; Wong, S.L.; Lam, T.N.T. Principal component analysis and long-term building energy simulation correlation. *Energy Conversion and Management*. **2010**, *51*, 135-139. <https://doi.org/10.1016/j.enconman.2009.09.004>.
33. Tang, D.; Zhang, Z.; Hua, L.; Pan, J.; Xiao, Y. Prediction of cold start emissions for hybrid electric vehicles based on genetic algorithms and neural networks. *Journal of Cleaner Production*. **2023**, *420*, 138403. <https://doi.org/10.1016/j.jclepro.2023.138403>.
34. Protić, M.; Shamshirband, S.; Petković, D.; Abbasi, A.; Mat Kiah, M.L.; Unar, J.A.; Živković, L.; Raos, M. Forecasting of consumers heat load in district heating systems using the support vector machine with a discrete wavelet transform algorithm. *Energy*. **2015**, *87*, 343-351. <https://doi.org/10.1016/j.energy.2015.04.109>.
35. Paudel, S.; Elmitri, M.; Couturier, S.; Nguyen, P.H.; Kamphuis, R.; Lacarriere, B.; Le Corre, O. A relevant data selection method for energy consumption prediction of low energy building based on support vector machine. *Energy and Buildings*. **2017**, *138*, 240-256. <https://doi.org/10.1016/j.enbuild.2016.11.009>.
36. Li, Q.; Meng, Q.; Cai, J.; Yoshino, H.; Mochida, A. Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks. *Energy Conversion and Management*. **2009**, *50*, 90-96. <https://doi.org/10.1016/j.enconman.2008.08.033>.
37. Cai, C.; Qian, Q.; Fu, Y. Application of BAS-Elman Neural Network in Prediction of Blasting Vibration Velocity. *Procedia Computer Science*. **2020**, *166*, 491-495. <https://doi.org/10.1016/j.procs.2020.02.059>.
38. Wang, Y.; Zhan, C.; Li, G.; Zhang, D.; Han, X. Physics-guided LSTM model for heat load prediction of buildings. *Energy and Buildings*. **2023**, *294*, 113169. <https://doi.org/10.1016/j.enbuild.2023.113169>.
39. Xue, G.; Qi, C.; Li, H.; Kong, X.; Song, J. Heating load prediction based on attention long short term memory: A case study of Xingtai. *Energy*. **2020**, *203*, 117846. <https://doi.org/10.1016/j.energy.2020.117846>.
40. Li, J.; Zhu, D.; Li, C. Comparative analysis of BPNN, SVR, LSTM, Random Forest, and LSTM-SVR for conditional simulation of non-Gaussian measured fluctuating wind pressures. *Mechanical Systems and Signal Processing*. **2022**, *178*, 109285. <https://doi.org/10.1016/j.ymsp.2022.109285>.
41. Lin, K.; Zhao, Y.; Tian, L.; Zhao, C.; Zhang, M.; Zhou, T. Estimation of municipal solid waste amount based on one-dimension convolutional neural network and long short-term memory with attention mechanism model: A case study of Shanghai. *Science of The Total Environment*. **2021**, *791*, 148088. <https://doi.org/10.1016/j.scitotenv.2021.148088>.
42. Cui, M. District heating load prediction algorithm based on bidirectional long short-term memory network model. *Energy*. **2022**, *254*, 124283. <https://doi.org/10.1016/j.energy.2022.124283>.

43. Yan, Q.; Lu, Z.; Liu, H.; He, X.; Zhang, X.; Guo, J. An improved feature-time Transformer encoder-Bi-LSTM for short-term forecasting of user-level integrated energy loads. *Energy and Buildings*. **2023**, *297*, 113396. <https://doi.org/10.1016/j.enbuild.2023.113396>.
44. Yang, Y.; Zheng, X.; Sun, Z. Coal Resource Security Assessment in China: A Study Using Entropy-Weight-Based TOPSIS and BP Neural Network. *Sustainability*. **2020**, *12*, 2294. <https://doi.org/10.3390/su12062294>.
45. Xie, L. The Heat load Prediction Model based on BP Neural Network-markov Model. *Procedia Computer Science*. **2017**, *107*, 296-300. <https://doi.org/10.1016/j.procs.2017.03.108>.
46. Lu, C.; Li, S.; Lu, Z. Building energy prediction using artificial neural networks: A literature survey. *Energy and Buildings*. **2022**, *262*, 111718. <https://doi.org/10.1016/j.enbuild.2021.111718>.
47. Chaki, S.; Biswas, T.K. An ANN-entropy-FA model for prediction and optimization of biodiesel-based engine performance. *Applied Soft Computing*. **2023**, *133*, 109929. <https://doi.org/10.1016/j.asoc.2022.109929>.
48. Ding, C.; Xia, Y.; Yuan, Z.; Yang, H.; Fu, J.; Chen, Z. Performance prediction for a fuel cell air compressor based on the combination of backpropagation neural network optimized by genetic algorithm (GA-BP) and support vector machine (SVM) algorithms. *Thermal Science and Engineering Progress*. **2023**, *44*, 102070. <https://doi.org/10.1016/j.tsep.2023.102070>.
49. Ling, J.; Dai, N.; Xing, J.; Tong, H. An improved input variable selection method of the data-driven model for building heating load prediction. *Journal of Building Engineering*. **2021**, *44*, 103255. <https://doi.org/10.1016/j.jobe.2021.103255>.
50. Wei, Z.; Zhang, T.; Yue, B.; Ding, Y.; Xiao, R.; Wang, R.; Zhai, X. Prediction of residential district heating load based on machine learning: A case study. *Energy*. **2021**, *231*, 120950. <https://doi.org/10.1016/j.energy.2021.120950>.
51. Fu, X.; Huang, S.; Li, R.; Guo, Q. Thermal Load Prediction Considering Solar Radiation and Weather. *Energy Procedia*. **2016**, *103*, 3-8. <https://doi.org/10.1016/j.egypro.2016.11.240>.
52. Sun, J.; Gong, M.; Zhao, Y.; Han, C.; Jing, L.; Yang, P. A hybrid deep reinforcement learning ensemble optimization model for heat load energy-saving prediction. *Journal of Building Engineering*. **2022**, *58*, 105031. <https://doi.org/10.1016/j.jobe.2022.105031>.