The Application of Energy Consumption Prediction Model for Fresh Air System based on PSO-BP and BO-BP

Gonghao Li*

College of Information and Control Engineering, Qingdao University of Technology, Qingdao 266520, China.

*Corresponding Author Email: 18561161862@163.com

Abstract: With the fast progress of deep learning and neural network technology, neural networks are now being used in lots of different subjects. They have been especially popular in predicting energy use in fresh air systems, because of their ability to fit nonlinear data. However, the performance of neural networks largely depends on the selection of their hyperparameters, and how to effectively optimize these hyperparameters has become a key issue to improve the prediction accuracy of the model. This paper discusses the methods of optimizing the hyperparameters of neural networks and compares several optimization methods, and finally finds that the Bayesian optimization algorithm performs optimally in neural network optimization.

In the experiment, firstly, the relevant data of the new wind system were collected and data preprocessing was carried out, including missing value interpolation and outlier detection. The key factors affecting the energy consumption of the fresh air system were then used to train a prediction model. Then, PSO and BO were used to improve the neural network's hyperparameters. A Monte Carlo simulation showed the Bayesian algorithm identifies the global optimum more quickly, improving prediction accuracy.

Experimental findings show that factors like indoor-outdoor temperature, humidity and air quality influence the energy consumption of fresh air systems.The BP neural network model, optimised by the Bayesian optimisation algorithm, performs better when predicting energy consumption. Research results offer an effective energy consumption prediction method for fresh air systems and important experimental support for related energy prediction tasks.

Keywords: fresh air system, energy consumption prediction, BP neural network, Particle swarm optimization, Bayesian optimization.

1. Introduction

In the context of growing global concern over energy management and conservation, research into energy efficiency in buildings is now focused on predicting energy consumption in fresh air systems. As an integral part of modern buildings, fresh air systems involve indoor and outdoor air quality regulation, which directly affects the energy efficiency performance of buildings. Accurately predicting the energy consumption of fresh air systems can provide important support for building energy efficiency optimization and energy conservation measures. In recent years, the rapid development of deep learning and neural network technologies has provided new possibilities for energy consumption prediction. Through neural network models, especially BP neural networks, we can better deal with complex nonlinear relationships to achieve accurate energy consumption prediction. Therefore, exploring the application of deep learning neural networks in energy consumption prediction of fresh air system possesses important theoretical and practical values.

There have been a number of studies on energy consumption prediction, especially the exploration of energy consumption related features. For example, Huang Prejun,Lian Yong in the design and application of energy consumption prediction and energy saving management system for housing buildings showed [1] that there are many factors affecting the energy consumption of a building, including but not limited to indoor and outdoor temperature difference, air quality, humidity and so on. However, it's crucial to address the effective extraction of key features related to energy consumption. Meanwhile, many optimization strategies have been proposed by scholars for the optimization method of BP neural network model, such as Feng Xiaogang,Yan Fengying,Shen Dongkui,et al. Bayesian optimization was used in the study of minimum wall thickness prediction of pressure vessels based on Bayesian optimization BP neural network [2], and other optimization algorithms such as PSO, stochastic search and grid search. These methods perform well in some classical machine learning tasks, but there are fewer empirical studies in the field of energy consumption prediction for fresh air systems, and there are no clear conclusions stating which variables should be used as the key features for energy consumption prediction, or comparing the effectiveness of different optimization algorithms in improving the performance of neural networks.

Based on this, this paper uses GRA to identify key feature variables affecting the energy consumption of the fresh air system. These features are then used to train a Bayesian neural network (BNN), with PSO and BO optimisation algorithms compared. The Monte Carlo simulation results show that BA performs better in hyper-parameter tuning. This enhances the energy consumption model's accuracy and stability. The simulation supports the building's energy-saving control.

2. Fundamental Algorithm Analysis

2.1 Grey correlation analysis

Gray correlation analysis measures the degree of correlation between multiple factors and a reference sequence quantitatively,and is commonly used in contexts where the data sample size is small, information is incomplete, or uncertainty is high [3]. The core idea of the method is to examine the degree of similarity between each comparison series and the reference series: when the trend of a comparison series is more similar to that of the reference series, it means that the degree of correlation between the two is higher. By calculating and ranking the correlation degree of each comparison sequence, We can see which factors have a greater effect on the system's energy use.

First, the data after the cleaning has been completed is extracted and organized, and the energy consumption of the fresh air system is selected as the reference sequence, denoted as . The remaining factors to be evaluated (indoor-outdoor temperature difference, air quality, indoor-outdoor humidity difference) are denoted as X_1, X_2, X_3, respectively.In order to ensure the consistency and comparability of each scale, it is necessary to preprocess all the sequences with dimensionless, and Min-Max normalization is used for the processing[4-6].i.e, each sequence is processed in accordance with Eq.(1).

$$x_i'(k)=\frac{x_i(k)-\min_k x_i(k)}{\max_k x_i(k)-\min_k x_i(k)} \tag{1}$$

The values of each sequence are mapped to the interval [0,1]. After such normalization, the target sequence X_0 and the three comparison sequences X_1, X_2, X_3 all have the same range of values, which facilitates the subsequent calculation.

After completing the dimensionless quantization, the new air system energy consumption sequence is considered as the reference sequence; the remaining sequences of factors are considered as comparison sequences. At any particular point in time, each comparison series produces a particular correlation coefficient with the reference sequence. These correlation coefficients are not only numerical expressions, but also a kind of quantitative expression of the degree of correlation between the two at that node [7], such as equation (2).

$$\Delta_i(k)=\left|x_0(k)-x_i(k)\right| \tag{2}$$

Then collect all the values of $\Delta_i(k)$ and find the minimum and maximum values of them, which are recorded as

$$\Delta_{\min}=\min_{i,k}\Delta_i(k),\quad \Delta_{\max}=\max_{i,k}\Delta_i(k) \tag{3}$$

In order to distinguish the effect of the magnitude of the difference on the correlation coefficient, it is necessary to set a resolution coefficient ρ, which is generally taken as 0.5. In this case, the correlation coefficient of the comparative sequence with respect to the reference sequence at the sampling moment k can be written as

$$\xi_i(k)=\frac{\Delta_{\min}+\rho\Delta_{\max}}{\Delta_i(k)+\rho\Delta_{\max}} \tag{4}$$

where $\Delta_{\min}$ and $\Delta_{\max}$ as well as $\Delta_i(k)$ have been obtained in the previous step. Since each comparison sequence has a correlation coefficient at all moments, it is necessary to average its correlation coefficients over all moments to obtain the gray correlation

$$r_i=\frac{1}{n}\sum_{k=1}^{n}\xi_i(k) \tag{5}$$

When r_i is larger, it means that the ith comparison sequence tends to be more synchronized with the change of the reference sequence, which has a relatively greater impact on the energy consumption of the new air conditioning system. By doing a descending order ranking of the factors according to the calculated correlation, it is possible to identify which factors are the key influencing factors. In this study, the process is implemented in the MATLAB environment. The grey correlation between temperature, air quality and humidity differences and energy consumption is obtained, presented and interpreted in tabular form.

2.2 Monte Carlo Simulation

Monte Carlo simulation is a statistical method for estimating the expected value of quantities through multiple experiments. The Monte Carlo simulation method is popular in energy forecasting. It is useful in assessing the MSE and MAE of models. By dividing the training and test sets multiple times, Monte Carlo simulation can provide a comprehensive assessment of the model performance, which in turn helps to determine the stability and accuracy of the model.

The mean square error (MSE) is a measure of the difference between the predicted and true values. For each simulation, the MSE is calculated as in equation (6):

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_{\text{true},i} - y_{\text{pred},i})^2 \quad (6)$$

where $y_{\text{true},i}$ is used to denote the actual energy consumed by the ith example, $y_{\text{pred},i}$ The study predicts the energy expenditure of the ith sample, and N is the number of samples in the test set.The MSE measures the overall error of the model by calculating the square of the prediction error for each sample and averaging it. A lower MSE indicates that the model fits the data better.

Another metric of prediction accuracy is mean absolute error (MAE): the mean of the absolute value of the predicted and the actual values, as in equation (7)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}| y_{\text{true},i} - y_{\text{pred},i} | \quad (7)$$

In the process of energy consumption prediction, the data first needs to be prepared and divided into training and test sets several times. Each division of the training and test sets is randomized, and this randomness is able to simulate different data distributions that may be encountered in practice. After each division, the training set is used to train the model and the test set to evaluate its prediction efficacy. This approach avoids errors caused by data partitioning and ensures more stable and reliable evaluation outcomes.

After each iteration, model prediction errors are calculated. Monte Carlo simulation is used to determine the predictive stability of the model and to assess its performance. It is important to note that a lower standard deviation is indicative of a higher predictive stability and a better performance of the model.

2.3 BP Neural Networks

2.3.1 BP Neural Network Fundamentals

The BP (backpropagation) neural net is a category of feed-forward artificial neural network [8],the core idea is to continuously adjust the network weights and bias through the two processes of forward propagation and backpropagation, so as to make the output gradually close to the target value [9]. Figure 1 shows the basic structure of a typical BP neural network.If the input of a layer is denoted as vector $\mathbf{x}$, the weight matrix as $\mathbf{W}$, and the bias vector as $\mathbf{b}$, the output of that layer can usually be expressed as $\phi(\mathbf{W}\mathbf{x}+\mathbf{b})$, where $\phi(\cdot)$ is a nonlinear activation function, such as ReLU, Sigmoid, or Tanh. After the network completes the forward propagation and obtains the predicted output, the loss function is computed by comparing with the real labels, and then the error is passed back to update the network parameters layer by layer by the backpropagation algorithm. The backpropagation calculates partial derivatives for the weights and biases of each layer by chaining the derivatives, and then iterates the parameters using gradient descent or its variant algorithms [11-13]. Specifically, if the loss function is denoted as L and the weights are denoted as $\mathbf{W}^{(l)}$ (the weight matrix of the lth layer), the update is usually $\mathbf{W}^{(l)} \leftarrow \mathbf{W}^{(l)} - \eta \nabla_{\mathbf{W}^{(l)}} L$, where η is the learning rate. By continuously repeating forward propagation and back propagation, BP neural networks are able to adaptively extract data features and approximate complex nonlinear mappings, which are widely used in tasks such as regression and classification.

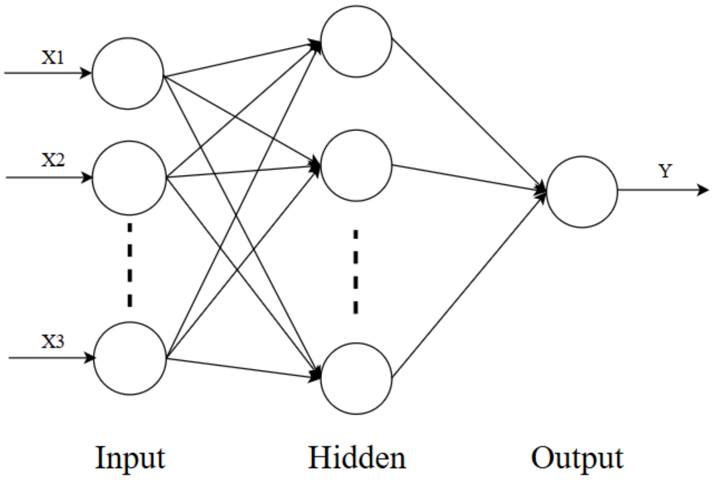

Figure 1. Basic structure of BP neural network

2.3.2 Determination of the number of hidden layers

A hidden-layer neural network approximates most nonlinear functions, if it contains enough neurons. Single layers work best for small and medium data sizes. Too many layers make training harder and cause overfitting. The model has a single hidden layer, which gives good expressiveness and balance between complexity and cost.

2.4 Particle Swarm Optimization for BP Neural Network Optimization

2.4.1 Fundamentals of the Particle Swarm Algorithm

The application of PSO has been demonstrated to enhance the performance of BP neural networks through the optimisation of their hyperparameters. Since the particle swarm algorithm has the characteristic of global search, based on this characteristic and the fast local search of BP neural network [14-17], the combination of the two can achieve better prediction results. The steps are outlined below:

(1) Process the input communication base station fresh air system operation data using formula (8) deviation standardization:

$$x = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{8}$$

(2) Particle initialization: a swarm of particles is first randomly generated. Each particle shows a possible neural network hyperparameter configuration.

It is hypothesised that the position of the particles is $x_i = [x_{i1}, x_{i2}, \ldots, x_{id}]$, where each x_{id} represents a hyperparameter.

The velocity of each particle is $v_i = [v_{i1}, v_{i2}, \ldots, v_{id}]$, which controls the movement of the particle in the solution space.

Particle fitness evaluation: train the BP neural network based on the position of each particle and calculate the error of the model on the validation set. This error is used as a function of particle fitness:

$$f(x_i) = \mathrm{MSE}(\text{validation set}) \tag{9}$$

Each particle has an individual best position $pbest_i$, i.e., the best adapted position in its history.

The position of the particle with the least adaptation among all particles is the global optimum position $gbest$.

(3) Update speed: particles update their velocity by current velocity, individual optimum position and global optimum position:

$$v_i(t+1) = \omega v_i(t) + c_1 r_1 (pbest_i - x_i(t)) + c_2 r_2 (gbest - x_i(t)) \tag{10}$$

where:

ω is the inertia weight, which controls the inertia of the particle along the current direction;

c_1 and c_2 are learning factors that control how the particle adjusts towards the individual optimum position and the global optimum position, respectively;

r_1 and r_2 are coincidental values in the interval [0,1] to increase the randomness.

Update position: it's clear the particle moves based on the new velocity.:

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{11}$$

(4) Adaptation evaluation and update

If the particle's current position is superior to the previous one, the particle's individual best position must be update:

$$pbest_i = x_i(t+1) \quad \text{if} \quad f(x_i(t+1)) < f(pbest_i)$$

If the current one is best, the global optimal position should change accordingly:

$$gbest = x_i(t+1) \quad \text{if} \quad f(x_i(t+1)) < f(gbest)$$

The iteration process continues until satisfaction or significant improvement.

2.4.2 Particle swarm optimization hidden layer neuron determination

Determining the quantity of neurons in the latent layer affects the accuracy of the fresh air system energy consumption prediction model. Using Equation (12), the number of neurons in the hidden layer was determined. The number of input layers, output layers, and hidden layers were previously set at 3, 1 and 1.

$$x = \frac{n}{\mathrm{c}(in + out)} \quad \mathrm{c} \in [2,10] \tag{12}$$

When the input layer is the same size as the target layer (and n is the number of samples), the constant c can be calculated using this formula. This formula allows values between 1 and 6 to be calculated, with the best choice being the one with the lowest mean square error (as shown in Figure 2).

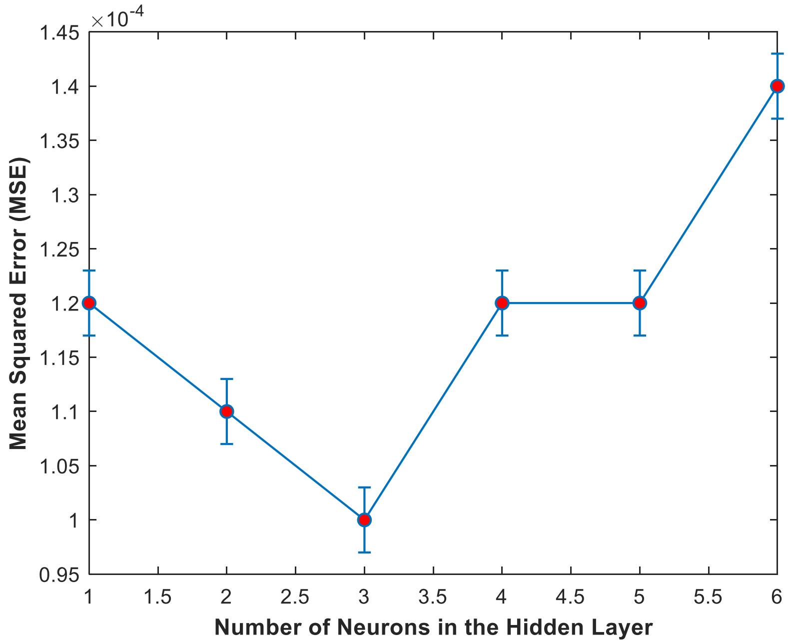

Figure 2. Plot of mean square error corresponding to each value

As shown, the mean square error is minimal when the hidden layer has 3 neurons. This is the optimum number of neurons. The model's parameters are in Table 1. See Figures 3 and 4 for the network structure and flowchart.

Table.1. Particle swarm optimization parameters of the BP Neural Network Model

Parameter Label	Parameter Setting
Input Layer Neurons	3
Hidden Layers	1
Hidden Layer Neuron Count	3
Target Layer Neurons	1
Training Iterations	1000
Learning Rate	0.2
Inertia Weight	0.8
Learning Factor	1.36

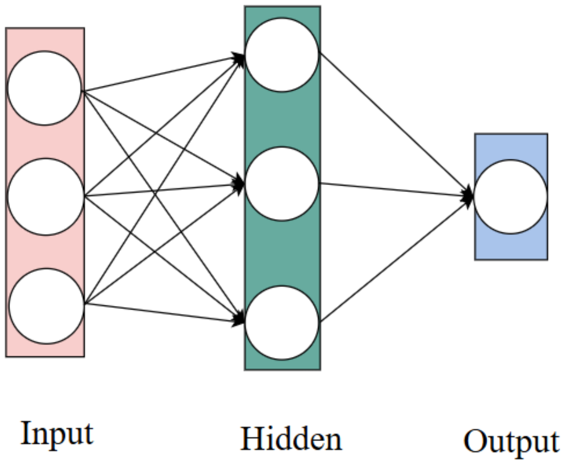

Figure 3. Particle swarm optimization BP neural network structure

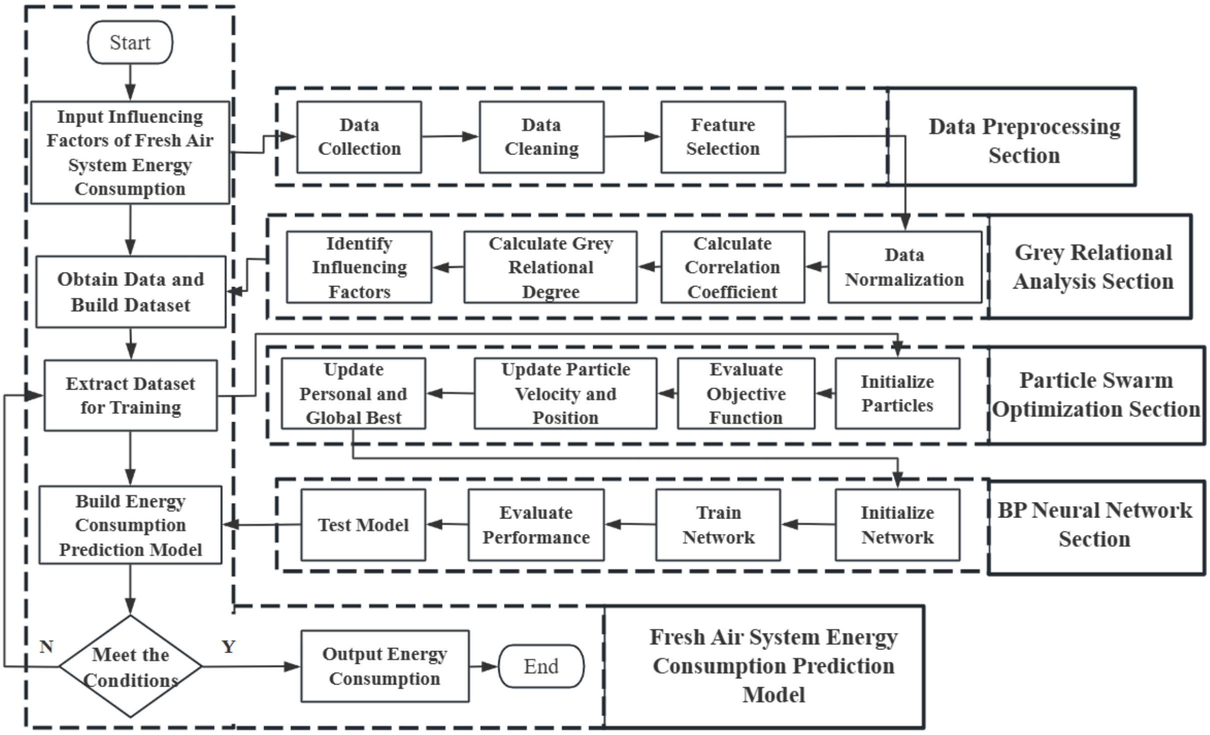

Figure 4. Particle Swarm Optimization BP Neural Network Model Flowchart

2.5 Bayesian optimization based BP neural network

2.5.1 Bayesian optimization fundamentals

Bayesian optimization is a global method that works well in many situations where the evaluation of the objective function is costly and cannot be derived explicitly or the gradient is not easily obtained [18]. Compared with the traditional BP neural network model, Bayesian optimization treats the combinations of hyperparameters that have not yet been evaluated as “points to be sampled” when tuning, and uses Gaussian processes to portray the overall trend and uncertainty of the objective function, Figure 5 illustrates the Gaussian regression process. As a small number of hyperparameter combinations are actually evaluated, the posterior distribution is updated to better approximate the objective function. Next, the next most valuable set of hyperparameters is evaluated on the posterior distribution through the acquisition function to balance “exploration” and “exploitation”. After a number of iterations in this way, Bayesian optimization can find hyperparameter settings that make the objective function perform better with relatively few attempts.

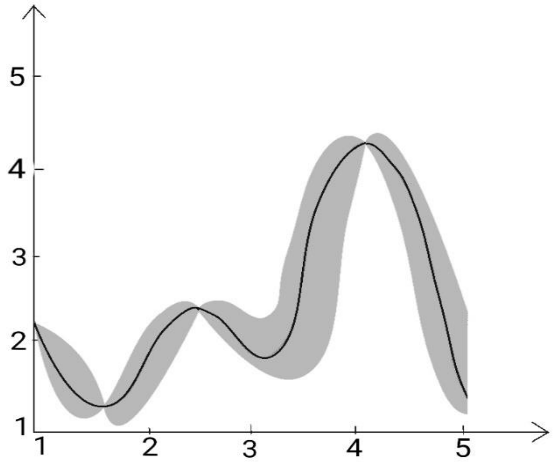

Figure 5. Gaussian process regression plot

2.5.2 Bayesian optimization of hidden layer neuron determination

The efficacy of Bayesian optimization is evaluated through the construction of a probabilistic model of the underlying goal function within the hyperparameter space during the process of parameter tuning. This is achieved by means of iterative selection of the potential optimal point for evaluation, and gradual improvement in understanding of the overall search space. In view of the sampling's limited scope and small data volume, the following observations can be made, in order to prevent the network structure from overcomplicating and bringing overfitting and training burden, the number of single hidden layer neurons can be reduced to 5 to 20 when combining such datasets with small data size and feature complexity such as air quality and temperature difference, etc. The MSE is the minimisation target and the objective function to be minimised. This allows the optimal hidden layer size to be found quickly and easily. It makes best use of existing data and improves prediction performance.

At each iteration, Bayesian optimization selects a new combination of hyperparameters and computes the objective function value. As the optimization proceeds, Bayesian optimization gradually improves the selection of hyperparameters and expects to find lower objective function values. Ideally, the objective function value should get smaller and smaller as the number of calculations increases, indicating that the hyperparameter combination found gradually improves the performance of the model. The graph of the optimization process in Fig. 6 shows that, the objective function value decreases significantly and tends to be stable, indicating that the Bayesian optimization effectively finds the appropriate hyperparameters, which makes the model's validation error decrease gradually and the optimization converges better.

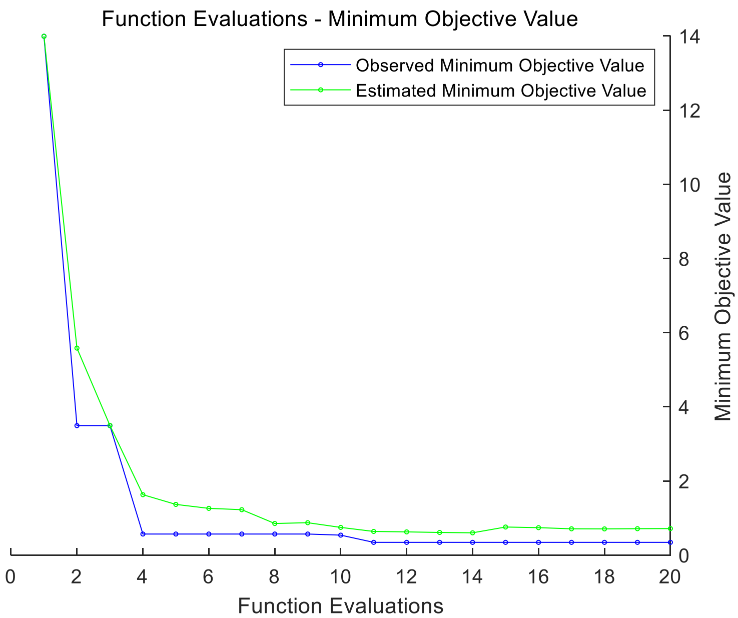

Figure 6. Optimization process graph

The BO method (0-25 range, MSE minimisation) is used to determine the number of neurons in the latent layer, and the number of neurons that makes the mean square error minimum is finally selected through continuous iterative loops, and Figure 7 illustrates the results of the BO experiments and Figure 8 illustrates the variation of the validation error with the number of iterations and its scatter plot.

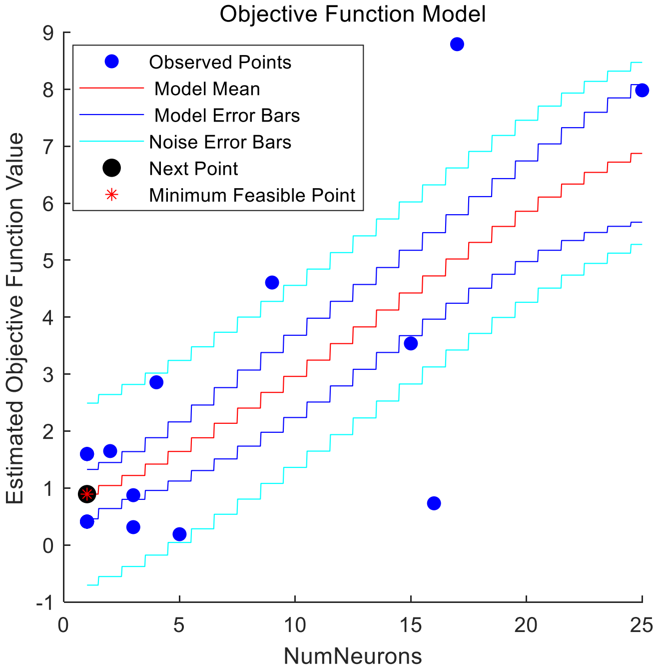

Figure 7. Bayesian optimization experiment results

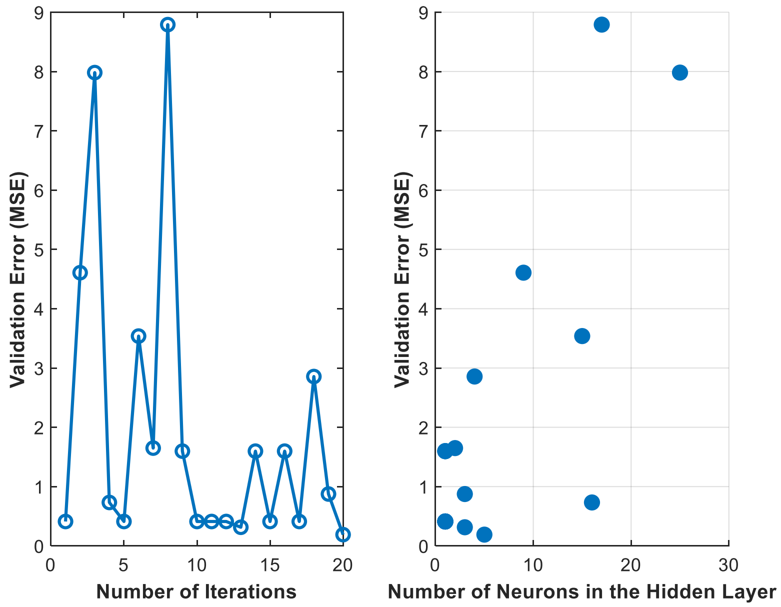

Figure 8. Validation error variation with number of iterations and its scatter plot

Figures 7 and 8 show that the MSE is at its lowest when there are 5 neurons in the latent layer (value: 0.143). Bayesian optimisation found that 5 is the best number of neurons. The best model was then trained with the training set. The model learnt the data features and patterns, enhancing its generalisation.

The basic parameters of the neural network model can finally be determined as in Table 2, The basic network diagram and structure of the model are shown in Figures 9 and 10, and the model implementation process as in Fig. 11.

ISSN: 1464-2859

Table.2. Bayesian optimization of BP neural network model parameters

Parameter	Value
Input Layer Neurons	3
Hidden Layers	1
Neurons in Hidden Layer	5
Output Layer Neurons	1
Training Iterations	1000

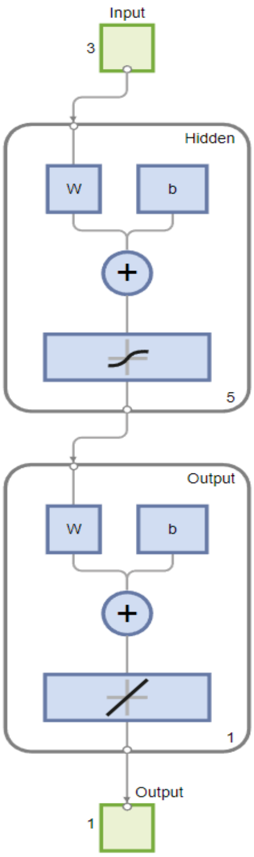

Figure 9. Bayesian optimization neural network diagram

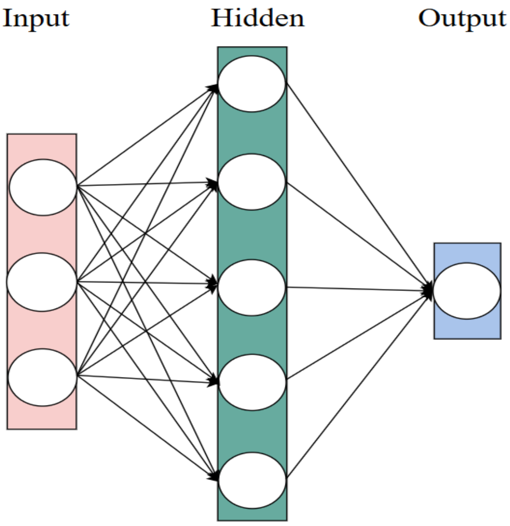

Figure 10. Bayesian optimization of neural network structures

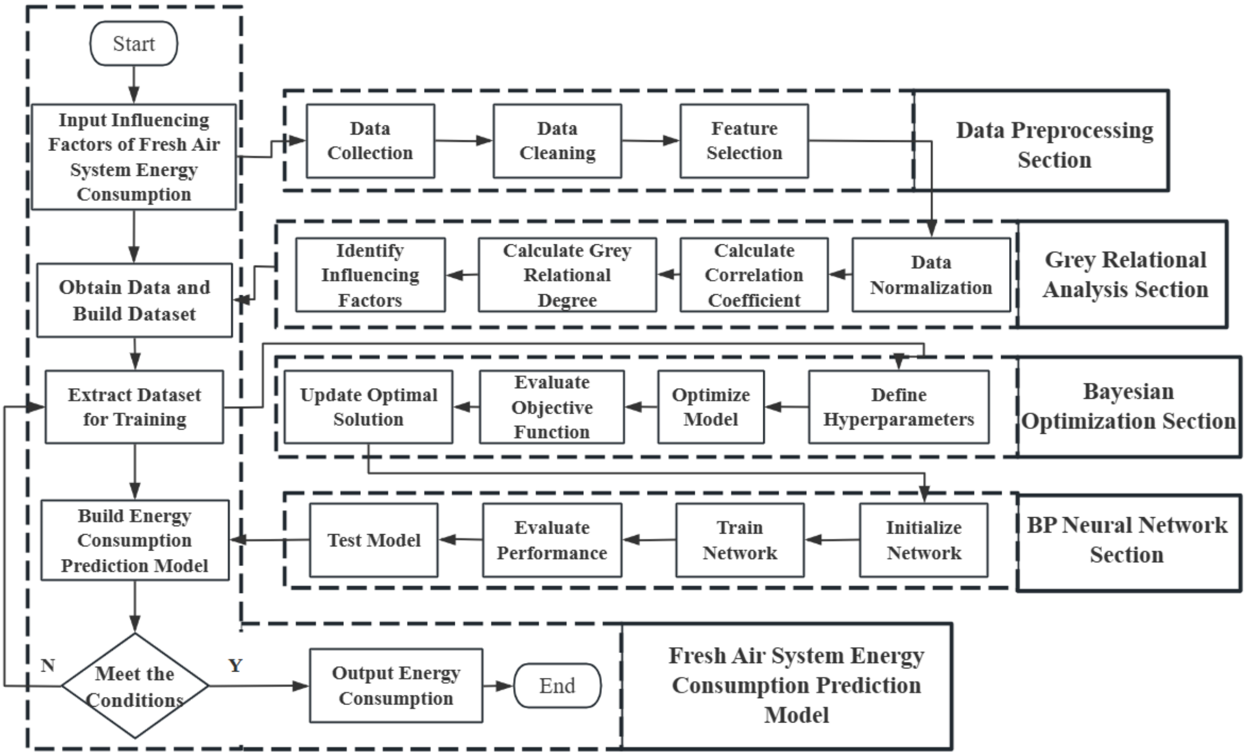

Figure 11. PSO-Optimized BP Neural Network Model Flowchart

3. Results

3.1 Data Pre-processing Outcomes

First, we collected five data points from a shopping mall in Jinan with a new air conditioning system installed between July 1 and 10, between 10 a.m. and 2 p.m. each day. Due to the generally high temperatures at noon, the data showed a clear up and down trend. We obtained the energy consumption data of the fresh air system through the energy management system of this mall. By collecting and analyzing these data, we were able to accurately grasp the changes in energy consumption of the fresh air system at different times, thus providing data support for subsequent energy optimization and energy saving analysis. In order to better analyze these data, we chose one of the day's data to show in detail, as shown in Table 3. In the MATLAB environment, line graphs showed the indoor-outdoor temperature and humidity differences of the shopping mall (see Figures 12 and 13). The results showed missing values and outliers, which could affect the analysis results. We used KNN and Z-Score interpolation to ensure the accuracy of the data. These methods enabled us to prepare and analyse the data.

Table.3. Selected raw data tables

Time	Indoor Temperature (°C)	Indoor Humidity (%)	Outdoor Temperature (°C)	Outdoor Humidity (%)	Air Quality (AQI)	Energy Consumption (kWh)
10:00	25.5	45	29.5	58	60	10.2
11:00	24.4	46	30.7	57	65	11.1
12:00	24.7	45	31.8	56	72	12.3
13:00	25.1	44	32.5	55	70	13.5
14:00	25.4	44	33.2	54	68	14.2

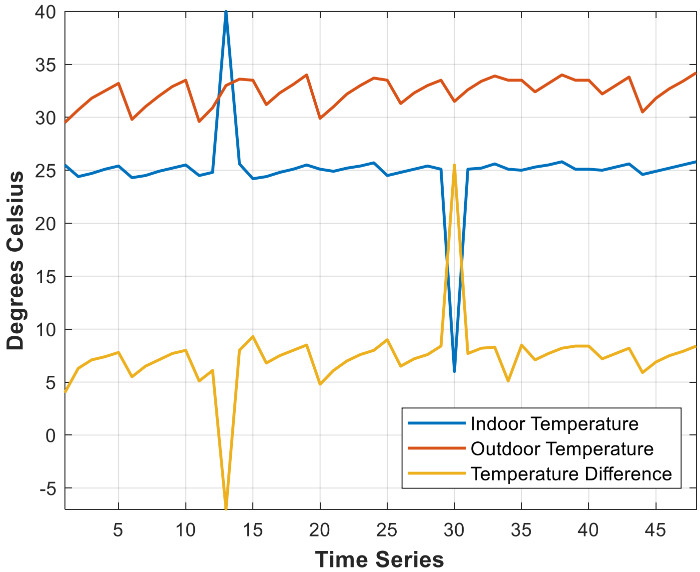

Figure 12. Line chart of raw indoor and outdoor temperatures and their differences

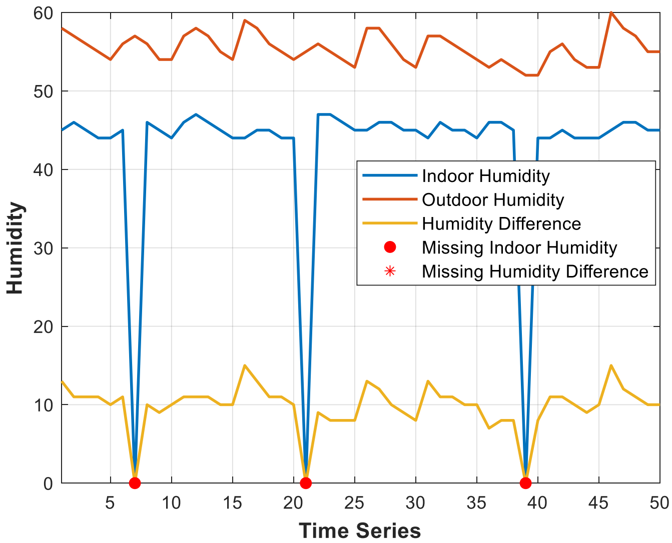

Figure 13. Line chart of original indoor and outdoor humidity and its difference

KNN (K-Nearest Neighbors) interpolation is a distance-based algorithm for filling in missing values. The basic idea is that for the sample points containing missing values, the distances of other sample points are used to infer the possible values of the missing values.

Firstly, the similarity calculation is performed, for the samples containing missing values, the distance between the sample to be interpolated and all other samples is calculated, the commonly used similarity measure is the Euclidean distance [19], as in equation (13):

$$d(x,y)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \tag{13}$$

where x and y are the feature vectors of the two sample points, and x_i and y_i denote the values of the ith feature, respectively. Based on the calculated distances, the K closest neighboring samples are selected.K is a hyperparameter, and cross-validation is usually required to select the optimal K value. For numerical data like temperature values, the mean of the K neighbors is used to fill in the missing values:

$$\hat{x}_j=\frac{1}{K}\sum_{i=1}^{K}x_{ij} \tag{14}$$

where $\hat{x}_j$ is the predicted value of the jth missing value and x_{ij} is the jth eigenvalue of the i-th neighbor.

The Z-Score method is an outlier detection method based on the distribution of the data and is applicable when the data conforms to a normal distribution.The Z-Score represents the distance between the standard deviation of a data point and the mean of the distribution in which it is located. By calculating the Z-Score value of each data point, it is possible to determine whether the point is an outlier [20]. Calculate each data point's deviation from the sample mean and compare with the standard deviation. Data points significantly deviant can be identified. If the magnitude of the deviation exceeds a predefined threshold, the data point can be designated as an outlier.

For a characteristic column $X = \{x_1, x_2, ..., x_n\}$, compute its mean μ and standard deviation σ.

Mean value formula:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{15}$$

Standard deviation formula:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2} \tag{16}$$

Z-Score indicates how far the data point x_i is from the mean μ in standard deviation, and for each data point x_i, its Z-Score value is calculated:

$$Z_i = \frac{x_i - \mu}{\sigma} \tag{17}$$

For the indoor temperature anomalies, we used the Z-Score method to detect them, and the results are shown in Fig. 14. After calculation, the Z-Score values of the two data points exceeded 4 to 6 standard deviations, respectively, indicating that these data points significantly deviated from the normal range. In order to improve data quality and model accuracy, we decided to remove these two outliers from the dataset, and the processed data folds are shown in Figure 15.

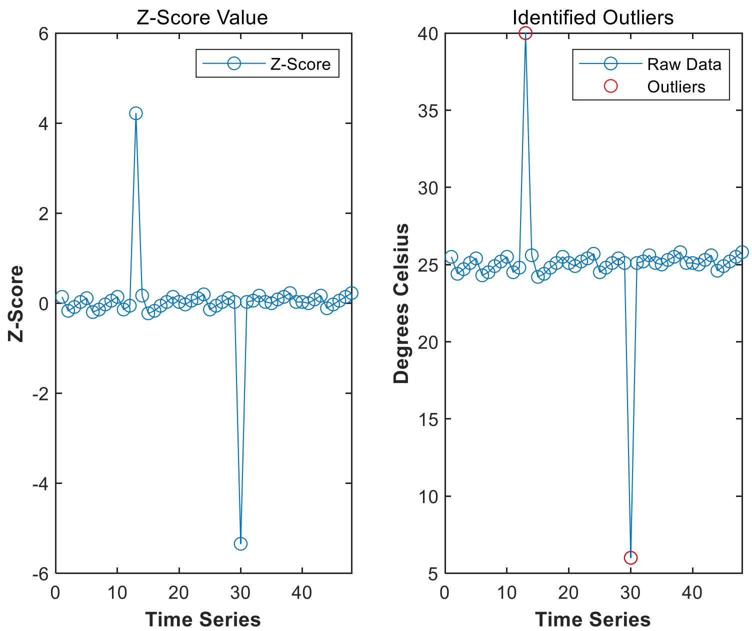

Figure 14. Z-Score method result chart

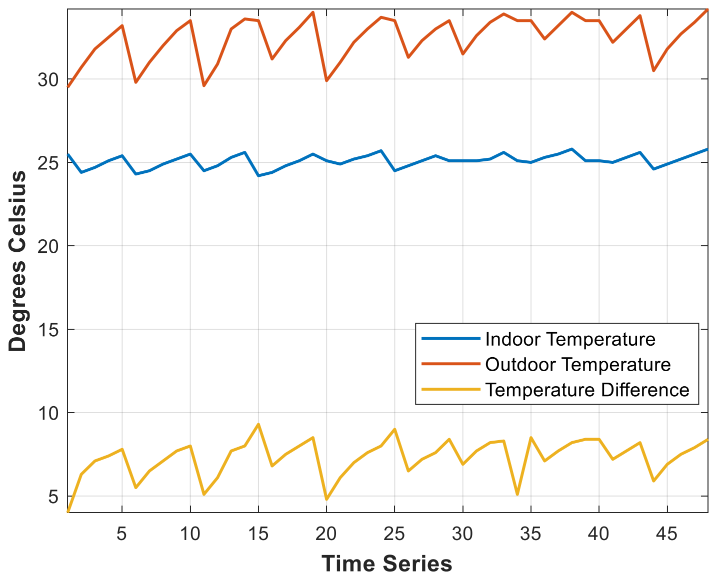

Figure 15. Line graph of data after removal of outliers

The KNN interpolation method can effectively fill the missing values in humidity data by considering the similarity of neighboring data points to infer the missing values. After experiments, the interpolation effect is best when the K value is taken as 3. The results are shown in Fig. 16, and the corresponding interpolated humidity values are 47, 45, and 45 when the time series are 7, 21, and 39, respectively. it can balance the computational efficiency and the filling accuracy, and make sure that the interpolated humidity data are more reliable, and the folds of the processed data are shown in Fig. 17.

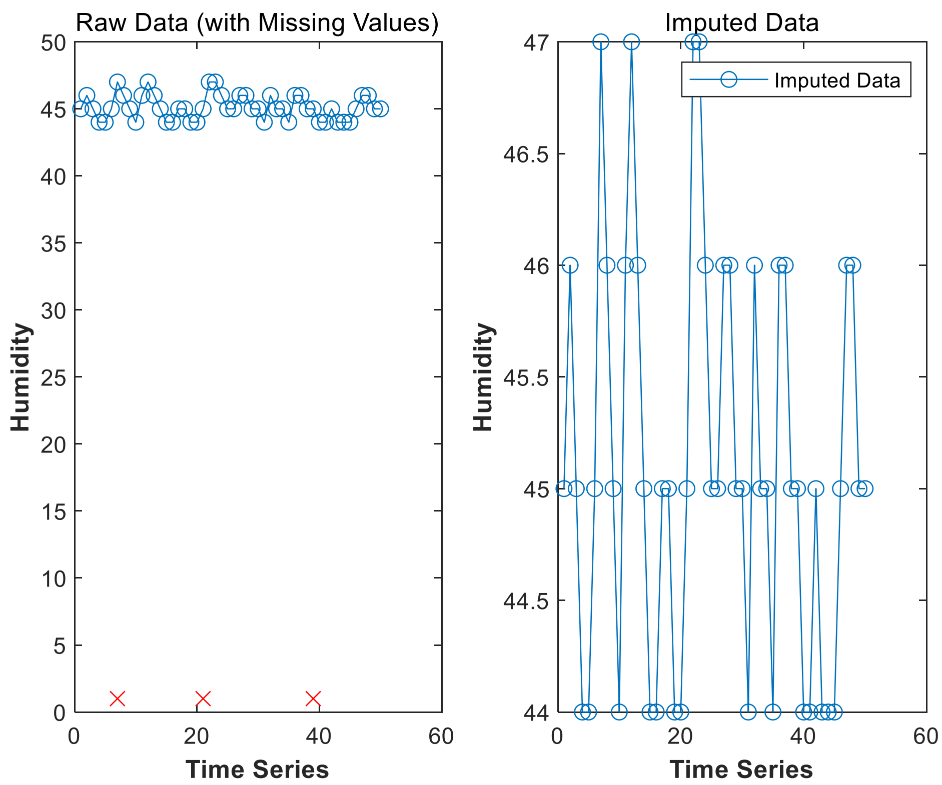

Figure 16. KNN interpolation method result plot

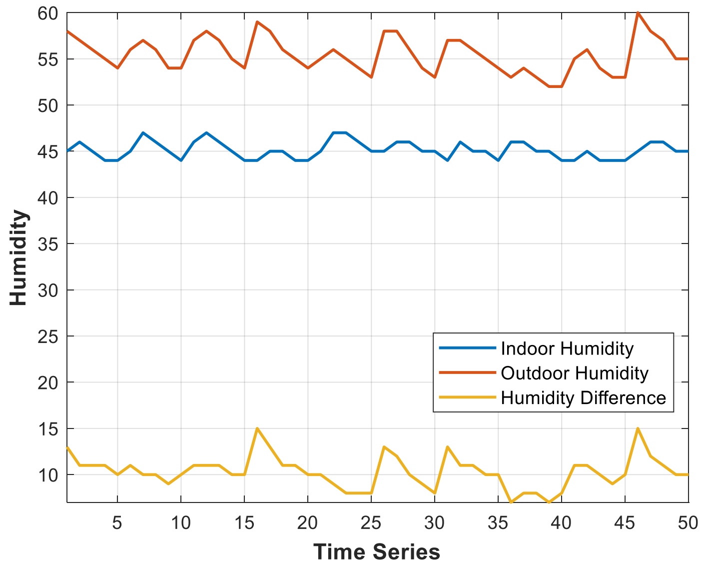

Figure 17. Line graph of data after filling in missing values

3.2 Gray correlation analysis results

Combined with the 48 sets of data collected from 1 to 10 July in this study (excluding data outliers), and in accordance with the above process, the grey correlation between each influence factor in the operation data of the system for fresh air and the energy consumption of the system for fresh air at the moment of T+1 was calculated as shown in Table 4, and the results of the histogram in the MATLAB environment are shown in Figure 18. Among them, the indoor-outdoor temperature difference at the moment t, the indoor-outdoor humidity difference at the moment t, the air quality at the moment t, and the energy consumption of the fresh air system at the instant t are represented by T(t), H(t), A(t), and P(t), accordingly [21-22].

Table.4. Gray correlation of each influence factor with P(t)

Variable	Grey Relational Degree
T(t)	0.842
H(t)	0.612
A(t)	0.793

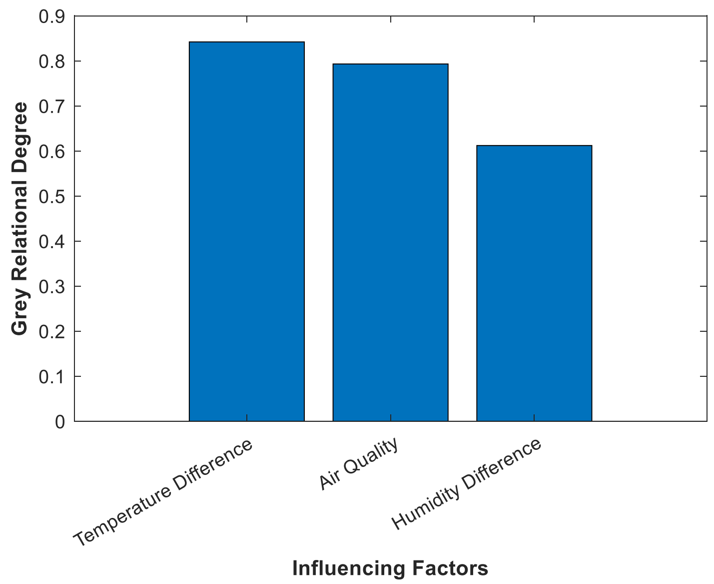

Figure 18. Histogram of gray correlation of each influencing factor with P(t)

Table 4 shows a 0.8 correlation between air quality and temperature difference and energy usage. This suggests that air quality and temperature affect energy consumption. The correlation between humidity difference and energy usage is weaker at 0.6. Combined with the actual situation of commercial buildings, the larger temperature difference often means higher cooling or ventilation energy consumption, plus air quality factors in some extreme weather will also affect the fresh air demand.

3.3 Analysis of model training and prediction experiment results

In the previous section, we have identified the optimal structure of the model. Based on this structure, the model is retrained in this section, aiming to obtain the best performance. From the training results as shown in Fig. 19, the regression coefficient on the training set reaches 0.97342, indicating a good fit between the model prediction results and the real data, thus verifying the reasonableness of the model construction.

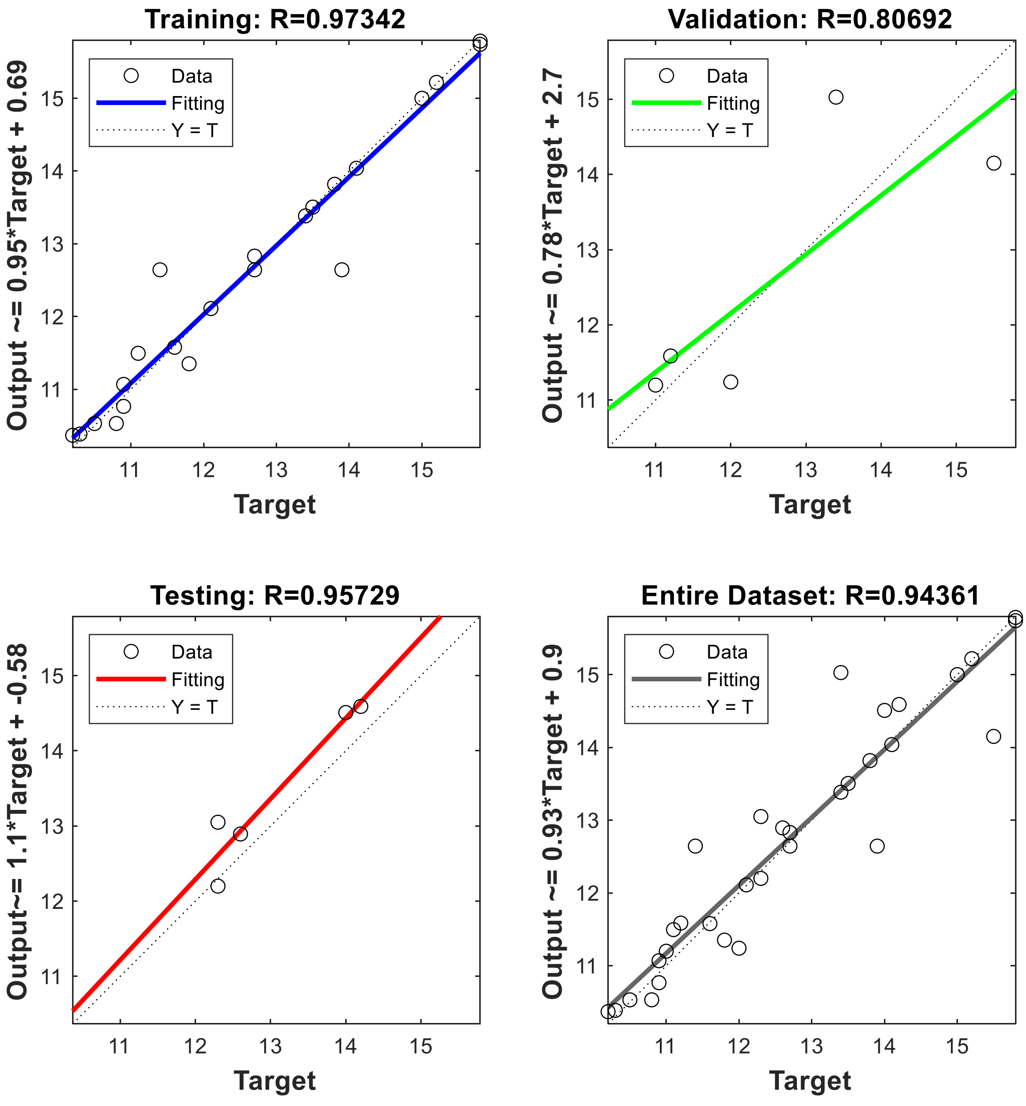

Figure 19. Model regression curve

To assess the accuracy of the model, energy consumption data from the mall's fresh air system was collected from 11 to 20 July. The validation set data was then fed into the model and the error between them was analysed. Figures 20 and 21 show the comparison between the two models.

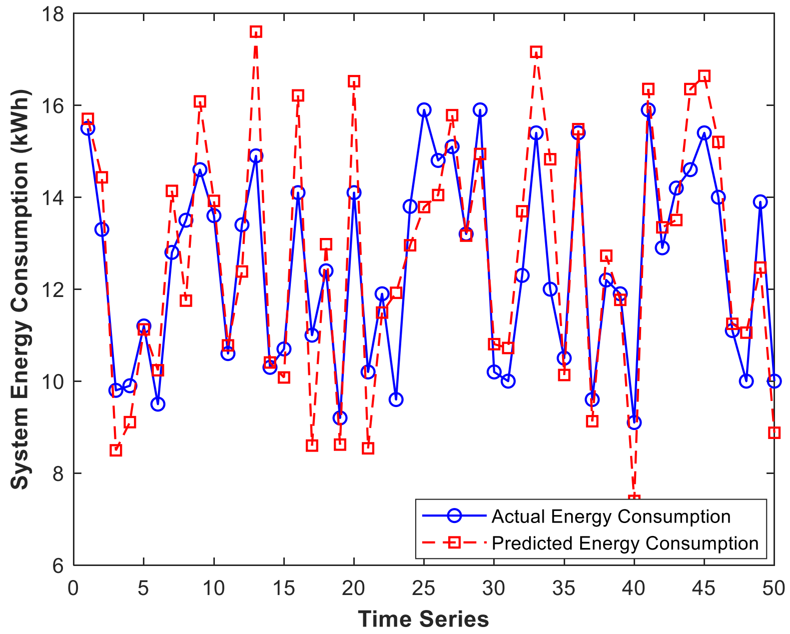

Figure 20. Comparison of prediction results based on particle swarm optimization

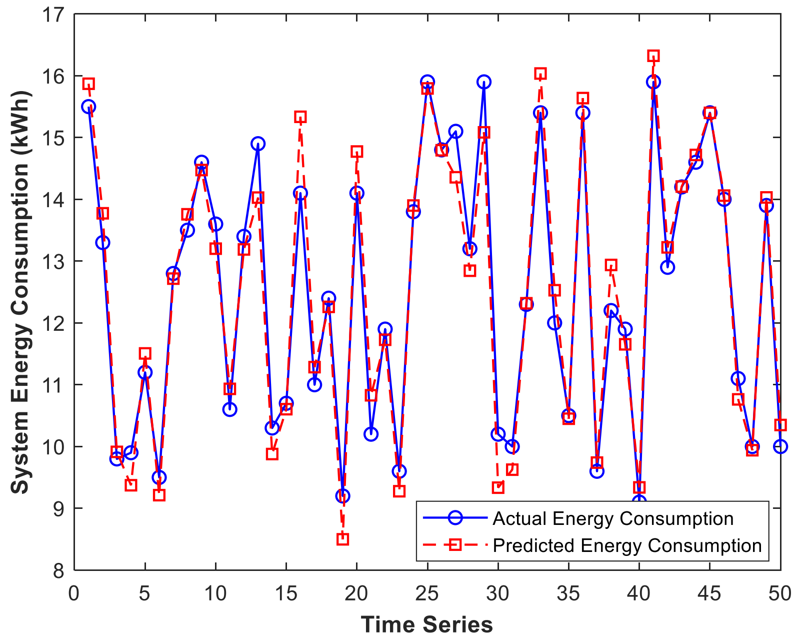

Figure 21. Comparison of prediction results based on Bayesian optimization

Fig. 20 and Fig. 21 show the comparison between the prediction results with particle swarm optimization BP neural network and Bayesian optimization BP neural network and the actual data, respectively, from which it can be observed that the particle swarm optimization BP neural network model has a large difference with the actual energy consumption, while the Bayesian optimization BP neural network has a small difference with the actual energy consumption.

Further Monte Carlo simulations of the above two optimization algorithms prediction models are conducted for several experimental evaluations to compare the effectiveness of the two algorithms, PSO and BO, in the prediction of energy consumption of fresh air systems. In each experiment, we optimize the hyperparameters of the BP neural network using different optimization algorithms by randomly dividing the training data and test data, and calculate the MSE and MAE, and the obtained experimental results are shown in Fig. 22.

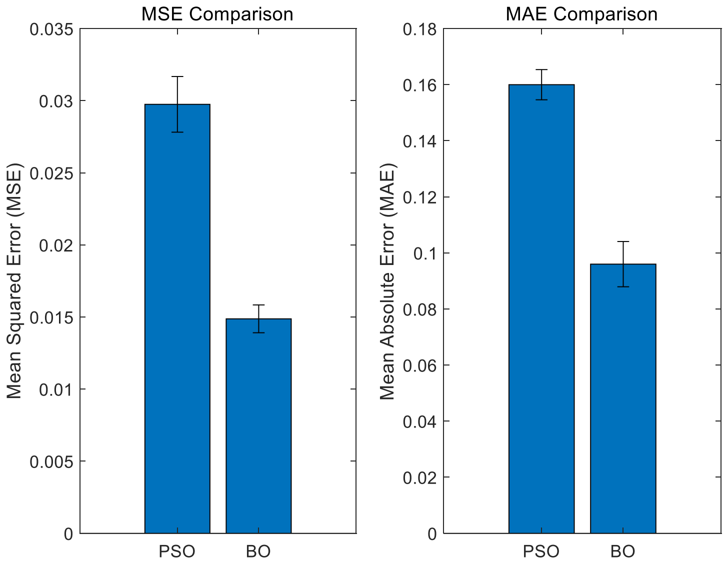

Figure 22. Monte Carlo simulation experiment results graph

BO outperforms PSO in optimising hyperparameters, as shown by lower MSE (1.5%) and MAE (6%) in a Monte Carlo simulation. PSO has higher MSE and MAE, suggesting susceptibility and reduced stability when anomalous data is present. BO can adjust neural network hyperparameters more effectively through a global search and optimization strategy to enhance prediction accuracy and model generalisation. PSO doesn't perform as well here, but can still optimize.

4. Conclusions

In this paper, we propose an energy consumption prediction model based on BP neural network for the energy consumption prediction problem of fresh air system, and explore the effects of different hyperparameter optimization algorithms on the model performance. We extracted key features that are closely related to the energy consumption of the fresh air system through GRA, including the indoor-outdoor temperature difference, air quality, humidity difference and other important factors. These features were used as input variables of the model, and the BP neural network was optimized after hyper-parameter tuning of two optimization algorithms, PSO and BO, respectively. Monte Carlo simulation experimental results show that the BP neural network based on BO exhibits better performance in the energy consumption prediction task compared to PSO, and is able to fit the data more accurately and improve the prediction accuracy and stability.

Through the research in this paper, we demonstrate the advantages of BO algorithm in the application of energy usage prediction of fresh air system, which provides an effective method for hyperparameter optimization. The successful optimization of the model not only improves the prediction accuracy, but also provides reliable data support for the intelligent energy-saving control of the new air system. In practice, based on these optimization results, building energy efficiency management and energy saving strategies can be formulated more accurately, providing an important basis for achieving building energy saving goals.

The study's findings offer room for improvement. The feature extraction technique could be refined with methods like automatic feature learning from deep learning. The energy consumption data has significant time-series characteristics that are not currently considered. Research should explore integrating time series analysis with advanced deep learning models, like LSTM and Transformer. These can better capture temporal data dependencies, enhancing predictions and generalisability.

As more data is collected, the model will adapt to changes in energy use over time and under different conditions through online learning. Researchers can use more data types (e.g., building data) to create more detailed, accurate, and intelligent analyses of energy predictions, contributing to the reduction of energy consumption in buildings and the protection of the environment.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

Data Sharing Agreement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

[1] HUANG Prejun,LIAN Yong. Design and Application of Energy Consumption Prediction and Energy Saving Management System for Housing Buildings[J]. Building Materials Technology and Application,2024,(06):76-80.

[2] Feng Xiaogang,Yan Fengying,Shen Dongkui,et al. Research on minimum wall thickness prediction of pressure vessel based on Bayesian optimization BP neural network[J]. Science and Technology Innovation,2024,(22):69-72.

[3] LI Gang,JIANG Boyang,ZHOU Gang,et al. Research on spray characteristics of two-fluid nozzle based on gray correlation analysis[J/OL]. Metal Mining,1-10[2025-03-12].

[4] GU Rongzhi, WU Hongyan, WU Haiying. Risk assessment and zoning of thunderstorm and gale disasters in Jiangsu based on gray correlation degree and other methods[J/OL]. Storm Disaster,1-10[2025-03-17].

[5] MUNAJIDIN MUNUR, QIN Kaixuan, GAO Fei, et al. An empirical study on the evaluation of the implementation effect of the special fund policy for the convention and exhibition industry--an analysis based on gray correlation[J/OL]. Practice and Understanding of Mathematics,1-16[2025-03-17].

[6] Ma Shiya. Gray correlation analysis and countermeasures of carbon emission factors in the construction industry[J]. Brick and tile,2025,(03):102-105.

[7] LIANG Wei-Ye,TIAN Century,YANG Xiaoyan,et al. Research on the Integrated Development of Port, Industry and City Based on Gray Correlation Analysis--Taking Beibuwan Port as an Example[J]. China Business Journal,2025,34(03):142-146.

[8] MENG Tong, LEI Ming, SONG Wenguang, et al. Improved KPCA combined with multi-objective dragonfly algorithm to optimize BP neural network for combine harvester fault diagnosis[J/OL]. Mechanical and Electrical Engineering,1-10[2025-03-17].

[9] ZHANG Chuanyuan, CHEN Yatian, GAO Zhenwei, et al. Adaptive forecasting method of consumer-side electricity load based on BP neural network[J]. Information Technology,2025,(02):187-192.

[10] XIA Zhenglan,WANG Zhaoyuan. Research on carbon trading price prediction based on LASSO-BP neural network[J]. Henan Science,2025,43(03):427-436.

[11] Sheng Jiahui,Cai Yong. Research on predictive analysis of gas sensor model based on BP neural network[J]. Industrial Control Computer,2025,38(02):19-20+24.

[12] WU Yong. Improved seagull algorithm to optimize BP neural network for GPS elevation fitting[J]. Geospatial Information,2025,23(02):108-111.

[13] ZHU Kaijia,LI Huateng,WANG Haoran. Research on the method of BP neural network speed prediction model based on comprehensive improvement[J]. Inner Mongolia Highway and Transportation,2025,(01):52-57.

[14] JIA Xingli, XIAO Zhan, GUO Junheng, et al. Intelligent optimization method of highway route corridor belt based on particle swarm optimization[J/OL]. Journal of Chang'an University(Natural Science Edition),1-10[2025-03-12].

[15] Wang Xiaoman,Lv Jianxiong,Li Xianjun,et al. Prediction of moisture content during drying of glue-impregnated bamboo bundles based on PSO-BP neural network model[J/OL]. Forestry Science,1-13[2025-03-12].

[16] YIN Yuntong, MA Jian, BAI Zhentao, et al. Intelligent prediction of anchorage force for soil sites based on improved PSO-BP neural network[J/OL]. Journal of Mechanics,1-17[2025-03-12].

[17] Yan Xishan,Han Weiguang. An improved particle swarm optimization algorithm for service migration in multi-access edge computing[J/OL]. Small Microcomputer Systems,1-8[2025-03-17].

[18] ZHU Congcong,GUO Sheng,CHANG Haitao,et al. Bayesian regularization optimization of BP neural network for SOH estimation[J]. Battery,2025,55(01):25-31.

[19] Song Yalin. Research on completely random missing data interpolation based on generative adversarial network [D]. Shanxi University of Finance and Economics,2024.

[20] Huang Le. Analysis of financial status of N new energy enterprises based on Z-score model[J]. Old Brand Marketing,2025,(03):82-84.

[21] YI Bingjie, TAN Pingan, WU Yanming, et al. Load prediction model of fresh air system for communication base station[J]. Electromechanical Engineering Technology,2022,51(07):81-85.

[22] Yi Bingjie. Research on temperature field analysis and energy saving optimization of fresh air system in communication base station[D]. Xiangtan University,2022.