

PAPER • OPEN ACCESS

## Hybrid Modeling of Central Air-Conditioning Cold Source System Energy Consumption with K-means Cluster Algorithm

To cite this article: Zhihao Zhou 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **295** 052035

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

# Hybrid Modeling of Central Air-Conditioning Cold Source System Energy Consumption with K-means Cluster Algorithm

Zhihao Zhou<sup>1\*</sup>

<sup>1</sup>School of Energy and Environment, Southeast University, Nanjing 210096, China.

\*Corresponding author: 15365188194@163.com

**Abstract.** Gray box model and artificial neural network (ANN) model of central air-conditioning cold source system have been established with simulation data to predict system operation energy consumption and compare the prediction accuracy difference of models under variable training data. A hybrid model is proposed to combine gray box model and ANN model based on K-means cluster algorithm. The results show that ANN model has higher prediction accuracy than gray box model when the membership of model input variables to trained input data is greater than 0.4. Otherwise, the gray box model has higher energy consumption prediction accuracy. Compared with gray box model and ANN model, prediction accuracy of hybrid model increases 27.7% and 33.85% on average under different training data.

## 1. Introduction

With environmental comfort requirement increasing, central air-conditioning system widely apply in large buildings, consuming vast energy. The cold source system includes the main energy-consuming equipment in the central air-conditioning system, and its operating energy consumption accounts for about 60% of the central air-conditioning operating energy consumption[1]. During the operation, the central air-conditioning cold source system can't adapt to the dynamic changes of environment and load, which causes the energy waste. The energy consumption of the system can be adjusted to reduce if the energy consumption of cold source system can be accurately predicted under different operating conditions. Therefore, there have been many studies on energy consumption prediction for cold source system operation. At present, the cold source system energy consumption prediction model is mainly divided into white box model, gray box model and black box model. Among them, Meng Hua[2] established a white box model of central air conditioning cold source system for system control, and its modeling process required many equipment parameters related to system. Yao[3] established a central air-conditioning cold source system gray box model based on the experiment data, but lack of analysis to equipment model. Andrew[4] established energy consumption ANN model of the central air conditioning system and provided the system energy saving algorithm. In the petrochemical energy field, hybrid model is often applied to the chemical reaction process for higher accuracy[5]. This paper establishes and analyzes grey box model and ANN model for cold source system energy consumption prediction. Then, hybrid model combined with two models is proposed to predict the cold source system energy consumption accurately.

## 2. Model building

### 2.1. Modeling Data Description

In this paper, the ANN model and the gray box model are established with the cold source system



operation data, which is generated by the simulation program of central air conditioning cold source system. The simulated equipment parameters of cold source system are shown in Table 1. In the simulation program, the used models have been verified by the actual data, and the relative errors between simulation program data and actual data is within 10% [6,7,8]. With the simulation program, the cold source system operation data in the range of 45%-95% cold load is obtained under different operating conditions. At each 5% cold load, 240 sets of data are randomly selected, and there are totally 2640 sets of operation data for modeling.

Table 1. Central air conditioning cold source system component parameters

| Component     | Related Parameters                                      | Units |
|---------------|---|-------|
| chiller       | Cooling capacity 1758kw, power rating 322kw             | 1     |
| Chilled pump  | power rating 37kw, Water flow 300m <sup>3</sup> /h      | 1     |
| Cooling pump  | power rating 45kw, Water flow 400m <sup>3</sup> /h      | 1     |
| Cooling Tower | power rating 16.5kw, Air volume 419634m <sup>3</sup> /h | 1     |

## 2.2. Modeling with Gray Box Model

Considering the central air-conditioning system operation and adjustment, the energy consumption prediction model is established with gray box model. The model output is the energy consumption of the cold source system, and the model input variables choose principle are the impact on system energy consumption and equipment operation control. There are two categories of model input variables, uncontrollable input variable and controllable input variable. This paper considers the outdoor ambient wet bulb temperature and cold load demand as uncontrollable variables. The chilled water supply temperature, the chilled water supply and return temperature difference, the chiller cooling water inlet temperature and the cooling water inlet and outlet temperature difference are chosen as controllable input variables. The energy consumption gray box model of the central air conditioning cold source system is established with the input and output variables, as shown in equations (1)-(11). The unknown coefficients in energy consumption gray box model are identified by the model identification algorithm with the system operation data.

The energy consumption model of the chiller based on operating COP is as follows:

$$COP = a_1 + a_2 Q_e + a_3 T_{chws} + a_4 T_{cwi} + a_5 Q_e^2 + a_6 Q_e T_{chws} + a_7 Q_e T_{cwi} + a_8 T_{chws} T_{cwi} \quad (1)$$

$$P_{chiller} = \frac{Q_e}{COP} \quad (2)$$

Where  $Q_e$  is cold load,  $T_{chws}$  is chilled water supply temperature,  $T_{cwi}$  is the chiller cooling water inlet temperature,  $P_{chiller}$  is chiller energy consumption,  $a_1 - a_8$  are model coefficient.

The chilled pump energy consumption model can be modeled by a polynomial of chilled water flow:

$$P_{chp} = b_1 + b_2 G_{chp} + b_3 G_{chp}^2 \quad (3)$$

$$G_{chp} = \frac{3.6 Q_e}{c \Delta T_{chw}} \quad (4)$$

Where  $P_{chp}$  is chilled pump energy consumption,  $G_{chp}$  is chilled pump water flow,  $c$  is specific heat capacity of water,  $\Delta T_{chw}$  is chilled water supply and return temperature difference,  $b_1 - b_3$  are model coefficient.

The cooling pump energy consumption model can be modeled by a polynomial of cooling water flow:

$$P_{cp} = c_1 + c_2 G_{cp} + c_3 G_{cp}^2 \quad (5)$$

$$Q_{con} = Q_e + P_{chiller} \quad (6)$$

$$G_{cp} = \frac{3600 Q_{con}}{\rho_w c \Delta T_{cw}} \quad (7)$$

where  $P_{cp}$  is cooling pump energy consumption,  $G_{cp}$  is cooling pump water flow,  $\Delta T_{cw}$  is cooling water

inlet and outlet temperature difference,  $Q_{con}$  is condensation heat,  $\rho_w$  is water density,  $c_1 - c_3$  are model coefficient.

The cooling tower heat transfer model is as follows:

$$Q_{con} = e_1 + e_2(T_{cwi} + \Delta T_{cw} - T_{wb}) + e_3(T_{cwi} + \Delta T_{cw} - T_{wb})^2 + e_4m_{cw} + e_5m_{cw}^2 + e_6(T_{cwi} + \Delta T_{cw} - T_{wb})m_{cw} + e_7m_{air} \quad (8)$$

Cooling tower fan energy consumption can be modeled by a polynomial of fan air volume:

$$G_{fan} = \frac{3600m_{air}}{\rho_{air}} \quad (9)$$

$$P_{fan} = d_1 + d_2G_{fan} + d_3G_{fan}^2 \quad (10)$$

Where  $T_{wb}$  is outdoor air wet bulb temperature,  $m_{cw}$  is cooling water mass flow,  $m_{air}$  is cooling tower fan air mass flow,  $\rho_{air}$  is air density,  $P_{fan}$  is cooling tower fan energy consumption,  $G_{fan}$  is cooling tower fan air flow,  $e_1 - e_7$ ,  $d_1 - d_3$  are model coefficient.

The total energy consumption of cold source system is calculated as equation (11):

$$P_{total} = P_{chiller} + P_{chp} + P_{cp} + P_{fan} \quad (11)$$

Where,  $P_{total}$  is total energy consumption of cold source system.

### 2.3. Modeling with ANN Model

Artificial neural network is a data-driven black box model that abstracts and analyzes the information between input and output data like brain neurons. Artificial neural network has excellent adaptive learning and nonlinear mapping capabilities for complex system. It is widely applied in the fields of pattern recognition, automatic control and predictive estimation. Therefore, the central air conditioning cold source system energy consumption prediction model has been established with ANN model, and compare the performance of cold source system gray box model.

The ANN in this paper is a multi-layer feedforward network, which adopts error back propagation algorithm. It consists of input layer, hidden layer and output layer. In the network, the input layer and output layer are one layer, and the hidden layer is one or more layers. The three-layer ANN can map bounded continuous function with arbitrary precision. Therefore, the three-layer ANN applied to establish the energy consumption prediction model of cold source system. The structure of the three-layer ANN is shown in Fig.1. The basic components of each layer are neurons in ANN. The layers are connected by weights, thresholds and transfer functions. The connection equation are as follows:

$$n = \sum_{i=1}^R w_i p_i + b \quad (12)$$

$$a(n) = f(\sum_{i=1}^R w_i p_i + b) \quad (13)$$

Where  $p_i$  is input variable,  $w_i$  is input variable weight,  $b$  is input variable threshold,  $f$  is transfer function,  $a$  is output variable.

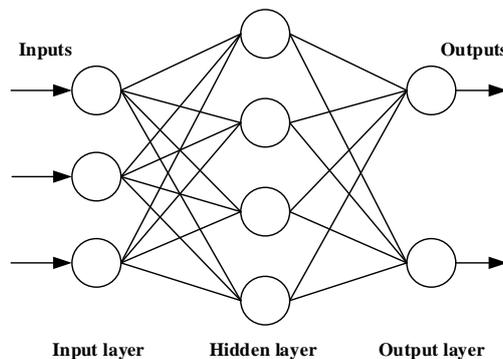


Figure 1. Schematic diagram of three-layer ANN

Different network structures have a significant impact on the prediction accuracy of ANN model. The structure of ANN model is analyzed in paper. It can be found that input parameters, output parameters, transfer functions, learning function and number of hidden layer neuron influence the prediction accuracy of ANN model. For comparing with gray box model, the same input variable is used as the input neuron in the ANN model of cold source system. The influence of different transfer function, learning function and number of hidden layer neuron are analyzed on prediction performance of the ANN model. The ANN model of cold source energy consumption prediction is established based on component energy consumption prediction ANN model, the related parameters of component ANN model are shown in Table 2.

**Table 2. Energy consumption prediction ANN model parameters of component**

| Parameter       | Chilled Pump          | Chiller   | Cooling Pump             | Cooling Tower                             |
|-----------------|-----------------------|---|--------------------------|---|
| ANN structure   | 2-9-1                 | 5-7-1   | 2-9-1                    | 4-14-1                                    |
| Input           | $Q_e, \Delta T_{chw}$ | $Q_e, T_{chws}, T_{cwi}, \Delta T_{chw}, \Delta T_{cw}$ | $Q_{con}, \Delta T_{cw}$ | $Q_{con}, T_{wb}, T_{cwi}, \Delta T_{cw}$ |
| Output          | $P_{chp}$             | $P_{chiller}$   | $P_{cp}$                 | $P_{fan}$                                 |
| Trans Function  | Tansig-Purelin        |   |                          |   |
| Learn algorithm | Levenberg-Marquardt   |   |                          |   |
| Learning rate   | 0.05                  |   |                          |   |
| epochs          | 1000                  |   |                          |   |

**3. model training and verification**

For comparing the energy consumption prediction performance of gray box model and ANN model, different training samples are divided by cold load range to train the gray box model and ANN model respectively, and the same test sample is applied in model test. The training samples and test sample are shown in Table 3. Each component energy consumption ratio is different in the central air conditioning cold source system. Therefore, the total energy consumption prediction accuracy of cold source system is considered as the standard to compare the performance of gray box model and ANN model comprehensively. In this paper, the performance is evaluated by relative error (RE) and average relative error (ARE), the mathematical expressions of them are shown as (14)-(15).

$$RE = \left| \frac{\hat{y} - y}{y} \right| \tag{14}$$

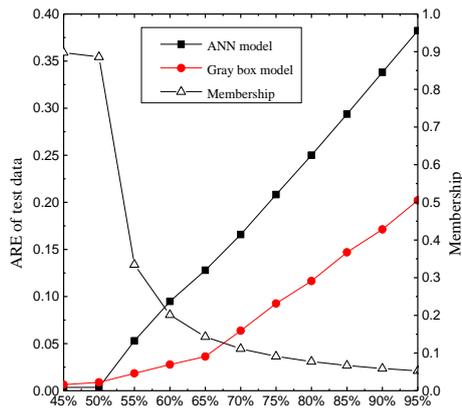
$$ARE = \frac{\sum_{i=1}^N RE(i)}{N} \tag{15}$$

Fig.2 shows the cold source system energy consumption prediction ARE of ANN model and gray box model under different training samples. The x-coordinate value indicates 240 sets input test data at the corresponding cold load, and the left ordinate indicates the model energy consumption prediction ARE corresponding to 240 sets input test data. It can be seen from Fig.2(a)-(e) that energy consumption prediction ARE curve of two models have a unique intersection point at the same test sample. The gray box model energy consumption prediction ARE is smaller than the ANN model at the right range of the point. At the left range of the point, the gray box model prediction ARE is larger than the ANN model. With the training sample data increase, the range continuously expands until the ANN model prediction ARE is smaller than the gray box model in the whole test data range, as shown in Fig.2(f)-(i). The phenomenon is accounted for the mapping and generalization ability difference between two models at the same training sample data.

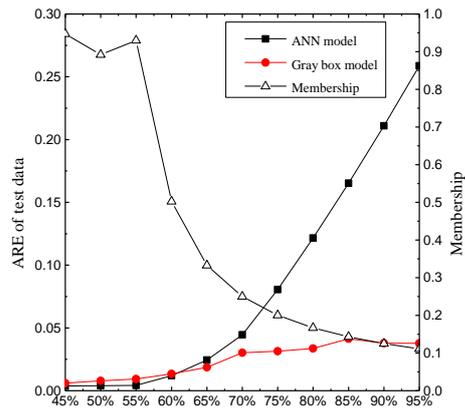
**Table 3. Model training and test sample data description**

| NO | Train Sample | Train Sample Range | Test Sample | Test Sample Range |
|----|--------------|--------------------|-------------|-------------------|
| 1  | 480          | 45%-50%            | 2640        | 45%-95%           |
| 2  | 720          | 45%-55%            | 2640        | 45%-95%           |

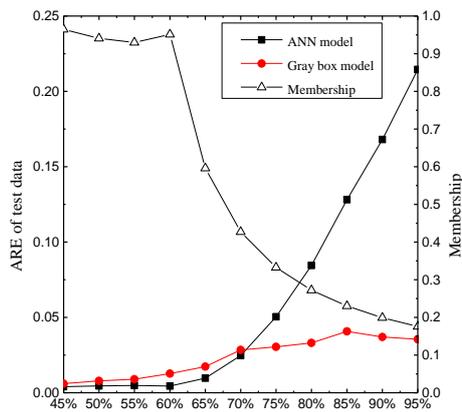
|   |      |         |      |         |
|---|------|---------|------|---------|
| 3 | 960  | 45%-60% | 2640 | 45%-95% |
| 4 | 1200 | 45%-65% | 2640 | 45%-95% |
| 5 | 1440 | 45%-70% | 2640 | 45%-95% |
| 6 | 1680 | 45%-75% | 2640 | 45%-95% |
| 7 | 1920 | 45%-80% | 2640 | 45%-95% |
| 8 | 2160 | 45%-85% | 2640 | 45%-95% |
| 9 | 2400 | 45%-90% | 2640 | 45%-95% |



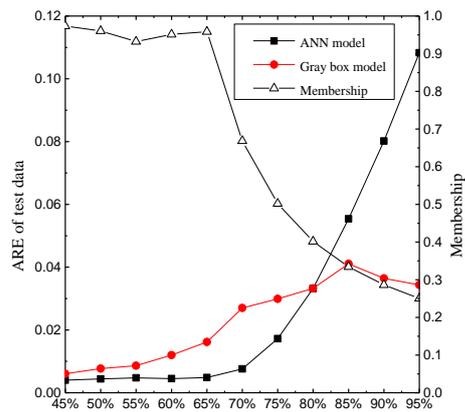
(a) Training sample NO.1



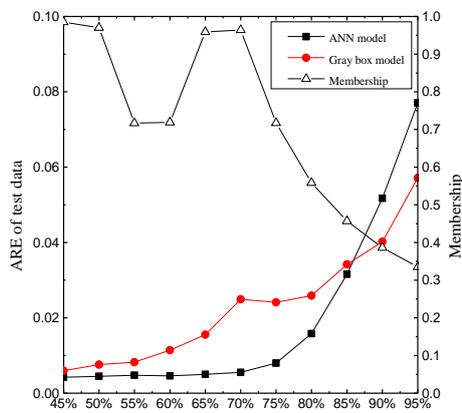
(b) Training sample NO.2



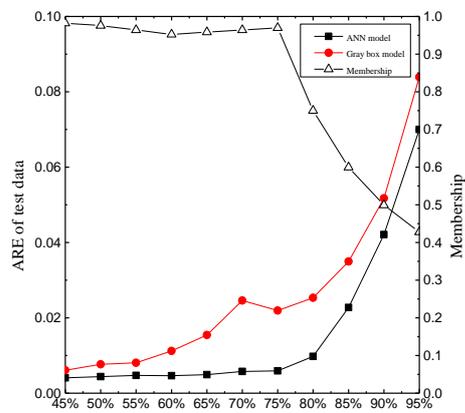
(c) Training sample NO.3



(d) Training sample NO.4



(e) Training sample NO.5



(f) Training sample NO.6

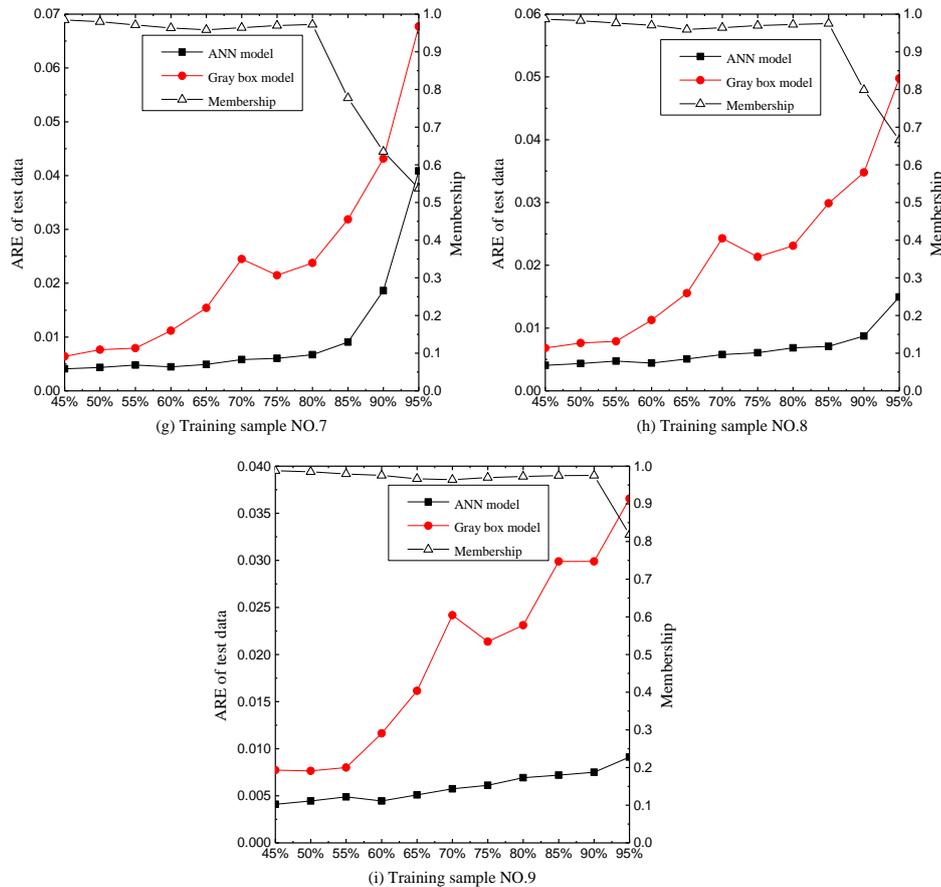


Figure 2. ARE and membership of test sample data under different training samples

#### 4. Model hybrids based on K-means cluster algorithm

In the analysis above, the energy consumption prediction accuracy with two models is related to the proximity of the test input data and the trained input data at the same training sample. K-means cluster algorithm is applied to further analyze the influence of the proximity on the model prediction accuracy. The membership is considered to measure the proximity of the test input data and the trained input data under different training samples after clustering center of the trained input data is obtained by K-means cluster algorithm. The basic clustering progress with K-means cluster algorithm is as follow. The first step is to determine the number of cluster centers. The second step is to randomly select k sets of data as the initial cluster center. The third step is to calculate the distance of each sample data to each cluster center, and sample data is divided into the nearest class based on distance. The last step is to recalculate the cluster centers of each class until the criterion function converges. The criterion function is shown in equation (16). The mathematical expressions of membership calculation are shown in equation (17)-(18).

$$E = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2 \quad (16)$$

Where  $E$  is the squared error sum of all points in sample data,  $p$  is the point in sample data,  $m_i$  is cluster center of  $C_i$ .

$$d_i = \sqrt{\sum_{j=1}^n (x_j - m_{ij})^2} \quad (17)$$

$$I = \frac{Max(d_i) - Min(d_i)}{Max(d_i) + Min(d_i)} \quad (18)$$

Where  $x$  is the sample data point,  $m_j$  is cluster center point,  $d_i$  is Euclidean distance from the sample point to the cluster center  $C_i$ ,  $I$  is the membership of the sample point to sample data.

The calculated membership of test input data to the trained input data under different training sample is shown in Fig.2, the right ordinate indicates the average membership corresponding to 240 sets input test data. It can be seen from Fig.2 that the membership of test input data to the trained input data with the training sample size increasing. The energy consumption prediction accuracy of cold source system with the ANN model and gray box model is different under different membership. The prediction accuracy with ANN model exceeds gray box model when the membership is greater than a specific value. The prediction accuracy with gray box model exceeds ANN model when the membership degree is lower than a specific value. It is found that the membership value of 0.4 can be considered to judge which model is more accurate at the same test input data. A hybrid model is proposed to combine the advantage of two models based on the membership of input data to the trained input data. The ANN model is applied to predict the cold source system operation energy consumption when the membership value is greater than 0.4. The gray box model is applied to predict when the membership value is lower than 0.4. Under different training samples, the energy consumption prediction ARE with hybrid model is shown in Fig.3, comparing with ANN model and gray box model.

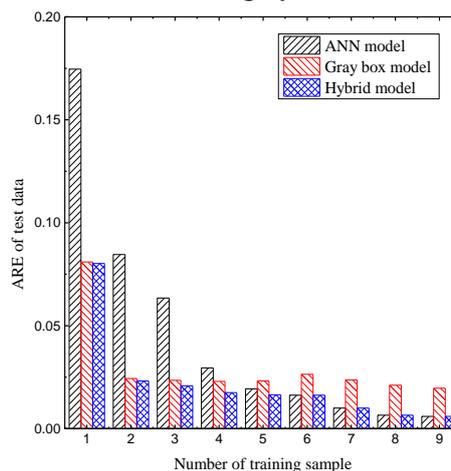


Figure 3. ANN model, gray box model and hybrid model energy consumption prediction ARE

It can be seen from Fig.3 that the hybrid model can improve the cold source operation energy consumption prediction accurate on average, comparing with ANN model and gray box model. In training sample No.1-5, the membership of test input data to the trained input data is partly greater than 0.4, and part of membership is lower than 0.4. Therefore, hybrid model conditionally selects ANN model or gray box model to predict the cold source operation energy consumption, its prediction accuracy exceeds ANN model and gray box model. In training sample NO.6-9, all the membership is greater than 0.4, hybrid model only selects ANN model to predict, its prediction accuracy exceeds gray box model but equals to ANN model. Compared with gray box model and ANN model, the prediction ARE of hybrid model averagely decreased by 27.7% and 33.85% under different training sample.

## 5. Conclusion

The grey box model and ANN model of operation energy consumption prediction are established respectively in the central air conditioning cold source system. Grey box model and ANN model are trained to predict the cold source system operation energy consumption under different training sample data which is obtained by simulation, and the prediction accuracy of models is evaluated by average relative error. It can be found that ANN model has higher prediction accuracy than gray box model when the membership of input variables to the trained input data exceeds a specific value. Otherwise, the gray

box model has higher prediction accuracy. Combining ANN model and gray box model, a hybrid model is proposed to enhance prediction accuracy of central air conditioning cold source system operation energy consumption based on K-means cluster algorithm. Hybrid model conditionally selects ANN model or gray box model to predict by judging the membership. Hybrid model selects gray box model at the lower membership. Hybrid model selects ANN model at the higher membership. Compared with ANN model and gray box model, hybrid model can effectively enhance the operation energy consumption prediction accuracy of central air conditioning cold source system.

### **Acknowledgments**

This research has been supported by the 13th Five-Year National Science and Technology Support Key Project of China (Grant No. 2017YFC0702501-3).

### **Reference**

- [1] Zhou X Q, Liu F, Chen W Q. (2007) Investigation and Study on Energy Consumption of Public Buildings in Guangzhou. *Building Science*, 23(12):76-80.
- [2] Men H, "Study on simulation of central chilled water system and online optimal supervisory control," Ph.D. thesis, Tongji University, 2004.
- [3] Yao Y, Lian Z W, Hou Z J. (2004) Optimal operation of a large cooling system based on an empirical model. *Applied Thermal Engineering*, 24:2303-2321.
- [4] Kusiak, A., Li, M., Tang, F. (2010) Modeling and optimization of HVAC energy consumption. *Applied Energy*, 87(10):3092-3102.
- [5] Sohrab, Z., Nima, R., Ali, L. (2018) Applications of hybrid models in chemical, petroleum, and energy systems: a systematic review. *Applied Energy*, 228:2539-2566.
- [6] Zhou Y, "Theory and application research on operation energy efficiency optimization for air conditioning cooling sources of a shopping mall," Ph.M. thesis, South China University of Technology, 2014.
- [7] Liu Z B, "Study of operation efficiency and Parameter optimization for chilled water system of central air-conditioning of the large-scale shopping malls building," Ph.M. thesis, South China University of Technology, 2016.
- [8] Zhang Q, "Study on energy saving control strategies of central air conditioning system," Ph.M. thesis, Southeast University, 2016.