**ANN Methods for COP Prediction of Supermarket Refrigeration System**

Wei-le OUYANG and Rui-qing KANG*

School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China

*Corresponding author

**Keywords:** BP neural network, Genetic algorithm, RBF neural network, COP prediction, Refrigeration system.

**Abstract.** Refrigeration system is one of the main energy consumers. Since COP prediction is the base of optimization analysis of energy saving, ANN models were developed by using on-site testing data to predict COP of the refrigeration system in this paper. BP and RBF neural network were applied to the prediction models respectively, and genetic algorithm was used to optimize the BP network. The prediction results of the models show that the GA-BP model has higher accuracy and better stability than the BP network only. And the RBF model has the smallest error and shortest time on calculation. Both BP and RBF models can apply in COP prediction of the refrigeration systems, but performance of the RBF model is better.

**Introduction**

Energy saving and emission reduction are significant missions for Chinese government. Refrigeration units for selling and storing foods is one of the main energy consumers in supermarkets. COP for a refrigeration system is defined as the ratio of the cooling capacity output to the total electric power input. It is used to show the efficiency of the entire system. Workers in the supermarket usually adjust parameters manually to improve efficiency based on experience. In order to use intelligent control system, COP prediction is need to be done before the optimization analysis of energy saving. How to build an accurate COP prediction model is a primary issue for saving energy.

ANN methods provide new solutions for complex nonlinear problems to predict the performance of refrigeration system. Tian et al. put forward a BP neural network to predict the COP of a variable frequency screw chiller in a cinema, which achieved high accuracy [1]. Xue et al. compared different black-box prediction models and found out BP is the proper model in predicting COP of chiller accurately. And by adopting GA-based operating variables identification, optimal values of the variables were found in specified ranges as well [2]. Lazrak et al. also developed an ANN approach to evaluate the energy performance of absorption chillers, but the reliable information with the long-term performance of the systems was obtained from a laboratory [3]. Arzu et al. used ANNs for performance analysis of a single-stage vapor compression refrigeration system. And mathematical formulations had been obtained from the summation and activation functions and weights of neurons [4]. However, most of researches fixed on air conditioners or a single unit, few studies are on a whole system. Additionally, most of studies used experimental data to simulate systems. In this paper, different ANN models were proposed to predict COP of the refrigeration system in a supermarket. And all data we used were on-site testing data, which showed different conditions from the performance shown in laboratory. This would provide more theoretical basis for optimizing energy-saving control system.

**Data Preparation**

**Variables of Models.** The on-site operating data needed for the ANN models had been obtained from the refrigeration system of a large supermarket in Beijing (see Table 1). To obtain the actual
COP of the system as the output of the ANN models, this paper used software REFLIBfxl. With the help of this tool thermodynamic data and transport properties of refrigerants can be calculated within EXCEL simply. Although REFLIBfxl can calculate the COP instead of complex thermodynamic equations, it cannot be applied to the real-time intelligent control system, so machine learning is applied to predict the COP.

<table>
<thead>
<tr>
<th>Table 1. Part of operating data samples.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>Outdoor temperature[°C]</td>
</tr>
<tr>
<td>Discharge temperature  [°C]</td>
</tr>
<tr>
<td>Suction temperature [°C]</td>
</tr>
<tr>
<td>Condensing pressure [bar]</td>
</tr>
<tr>
<td>Condensing temperature [K]</td>
</tr>
<tr>
<td>Evaporating pressure [bar]</td>
</tr>
<tr>
<td>Evaporating temperature [°C]</td>
</tr>
<tr>
<td>Flow rate [m³/h]</td>
</tr>
<tr>
<td>Compression ratio</td>
</tr>
<tr>
<td>COP</td>
</tr>
</tbody>
</table>

Based on the physical laws and theoretical calculation of COP, this paper presents correlation analysis of all variables from Table 1 to determine the input variables of ANN models. We found that all variables are related to COP, while the outdoor temperature, discharge temperature, condensing pressure are significantly related to COP. Since the evaporating pressure and temperature are highly correlated, the condensing pressure and temperature are highly correlated as well, temperatures were omitted in order to make the model more concise. Hence, seven variables of the rest in Table 1 were selected as the input of the ANN models eventually.

**Data Preprocessing.** On-site measurement data always come with uncertainties that bring about lower accuracy and long training time when a model is developed. Except for purchases, big change in weather or any other uncertain factors, the data should fluctuate within a small range. In this paper, the 3σ rule was adopted. Besides, the actual COP should be less than ideal COP. Ideal value ε for each point were calculated by Eq. 1, and points where the actual COP is bigger than the ideal value should be eliminated. Finally, 36 outliers of original data were eliminated, and the percentage was 1.4%.

\[
\varepsilon = \frac{T_e + 273}{T_e - T_c}
\]

Where \( T_e \) is evaporating temperature [°C], \( T_c \) is condensing temperature [°C].

Due to the different units and ranges of input variables, the neural network may converge slowly and results may be wrong. Therefore, the original data need to be normalized before network training and the output of network should be anti-normalized before analysis.

**COP Prediction**

The total number of measurement points after preprocessing is 2560 (see Table 1), the first 2510 points are selected for neural network training, and the last 50 are used for testing.

**BP Network Model.** BP (Back Propagation) neural network is composed of three layers: input layer, hidden layer and output layer. It optimizes the weighted connections by allowing error to spread from output layer towards lower layers. BP models need to determine the number of hidden layer nodes before training. In this paper, we used Eq. 2 to obtain a range for hidden layer nodes \( l \).

\[
l = \sqrt{n + m + a}
\]

Where \( m \) and \( n \) are output and input layer nodes respectively, and \( a \) is a constant within [0,10].

Since \( m = 1 \) and \( n = 7 \), \( l \) should be a constant within [2,13]. We compared prediction error of networks for each possible values of \( l \), and the results are shown in Fig. 1. Although error at \( l = 9 \) is
smaller than that at \( l=3 \), considering the problem of learning time and over-fitting, the best number of the hidden layer nodes was chosen to be 3 in this paper. Therefore, we can finally determine the BP network structure (see Fig. 2). Log-sigmoid as the hidden layer transfer function \( f(x) \) was used. The max iteration number was set to 1000, the network learning rate was 0.01, and the training goal error was 0.001.

There are two criteria for the network performance, mean relative error (MRE) and coefficient of determination \( (R^2) \). They defined as Eq. 3 and Eq. 4, respectively.

\[
MRE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{\hat{y}_i - y_i}{y_i} \right|
\]  

\[
R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2}
\]  

Where \( N \) is the number of testing samples, \( \hat{y}_i \) is predicted value of \( i \) set, \( y_i \) is actual value of \( i \) set, and \( \bar{y} \) is the mean value of testing points. The smaller the MRE is and the closer the \( R^2 \) is to 1, the better performance the model achieves.

Fig. 3 shows the comparison of the BP neural network prediction results, \( R^2 \) is 0.9518 and MRE is 0.0182. However, the initial weights and thresholds of the BP network training are not fixed, which makes results unrepeatable. This paper introduced genetic algorithm (GA) to optimize the BP
network by determining the optimal initial weights and thresholds before training, so that the predicted results will not be affected by the initialization of parameters.

**Figure 3. Prediction result of BP model.**

**GA-BP Model.** Genetic algorithm is a method to search for optimal solutions by simulating natural evolutionary processes. The training starts with GA, which performs a global search on weight ranges and finds the best initial weights for the BP neural network. Then, the BP algorithm begins training with the best initial weights provided by GA and approaches the optimum solution. The flowchart is shown in Fig. 4. The BP network structure in this paper is 7-3-1, so the number of parameters that the genetic algorithm needs to optimize is 28.

Result of GA-BP model is shown in Fig. 5. It can be seen that the performance after adding the genetic algorithm is significantly improved. The predicted values and the actual values of the model are closer than those of the single BP model. Except for the 38th sample, the relative error is in the range of -3% to 3% (Fig. 6). The MRE of the model after combining with GA reduces from 0.0182 to 0.0062, and the $R^2$ increases from 0.9518 to 0.9889. It proved that the genetic algorithm improves the accuracy of the prediction model and can ensure the stability of the training results.
Figure 4. The flowchart of GA-BP algorithm.

Figure 5. Prediction result of GA-BP model.
RBF Network Model. RBF (Radial Basis Function) neural network and BP neural network are basically the same in structure. The difference is that the RBF network transfer function of hidden layer is a radial basis function, which means that only the functions with their centers close to the input patterns will give a response [5]. These basis functions, also called kernel, can be selected among several types of functions, but for most applications they are chosen to be Gaussian functions, so was this paper. Training the RBF network involves determining the number of RBF units, the width of RBF units and the output layer weight values. The RBF network can converge rapidly and determine the number of hidden layer nodes automatically. The parameters spread and training goal error was set to 0.3 and 0.001, respectively. The prediction result of the RBF model is shown in Fig. 7. It can be seen that the error of the 38th sample is significantly improved. The result proved that RBF has higher approximation property.

Table 2 compares prediction performances of the three models above. Every ANN model obtained good quality performance, but the RBF neural network model approximates the high nonlinear process with the smallest error, and its calculation speed is faster as well. Besides, the genetic algorithm have improved the BP network model with higher accuracy. However, when the RBF neural network model has more training samples, it needs more hidden layer neurons than the BP neural network to achieve the goals, which is prone to over-fitting.
Table 2. Comparison of prediction models.

<table>
<thead>
<tr>
<th></th>
<th>BP</th>
<th>GA-BP</th>
<th>RBF</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRE[%]</td>
<td>1.82</td>
<td>0.62</td>
<td>0.21</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.9518</td>
<td>0.9889</td>
<td>0.9996</td>
</tr>
<tr>
<td>Training time[s]</td>
<td>6.35</td>
<td>-</td>
<td>1.17</td>
</tr>
</tbody>
</table>

**Conclusion**

In this paper, COP prediction models of the supermarket refrigeration system were studied. The input of the models were selected by correlation analysis. The BP neural network and RBF neural network were trained respectively, and genetic algorithm was used to optimize the BP network model. The results indicate that both BP neural network and RBF neural network can be used to predict the COP of the refrigeration system, and the prediction results are very close to the actual data. The mean relative error of BP model and GA-BP model are 1.82% and 0.62%, and $R^2$ of them are 0.9518 and 0.9889 respectively, which demonstrates that genetic algorithm can optimize the BP network. Besides, performance of RBF model was the best among three methods. It has shortest training time and higher prediction accuracy, where mean relative error is 0.21% and $R^2$ is 0.9996.

**References**


