Port Container Throughput Forecast Based on ABC Optimized BP Neural Network

Huang Fucheng1*, Liu Dexin1, An Tiansheng1 and Cao Jie1

1Navigation College, Dalian Maritime University, Dalian, Liaoning, 116026, China

*Corresponding author’s e-mail: 13204070386@163.com

Abstract. In order to improve the accuracy of port container throughput prediction, an improved ABC-BP prediction model based on Artificial Bee Colony (ABC) was proposed. With the advantage of global search ability and difficulty in local optimization, the weights and thresholds of Back Propagation (BP) neural network are optimized, and the optimal weights and thresholds in the network are finally determined. Referring to the existing research results, Qingdao 2014-2019 (24 quarters) GDP, port throughput, total foreign trade import and export volume were selected as network input, and Qingdao port 2014-2019 port container throughput statistics were used as network output, a total of 24 sets of data were constructed, so as to build a BP neural network prediction model. The first 20 groups of data were used as the training set, and the last 4 groups of data were used as the test set for instance verification. The results show that, compared with the traditional BP neural network and the BP neural network prediction model optimized by Genetic Algorithm (GA), the ABC-BP model can significantly improve the prediction accuracy of port container throughput.

1. Introduction
The accurate and scientific forecast of the container throughput of the port is the theoretical basis of the port’s overall scientific planning for the port authorities, and also has a guiding role in port construction, investment and development direction.

There are many research achievements by domestic and foreign scholars on the forecast of container throughput of the port. For example, Yang Jinhua[1] used grey forecasting methods to predict the container throughput of Shanghai Port in the next three years, and judged whether Shanghai Port is facing overcapacity through the capacity utilization index; Du Gang [2] compared the prediction results of the seasonal time series model with the whole autoregressive moving average model, and concluded that the former has a higher prediction accuracy for container throughput; Tang Simin[3] used the typical factors of port throughput as independent variables and established it using multiple linear regression analysis. Forecasting model of typical factors of port throughput; Wang Zhenzen[4] combined the three-dimensional exponential smoothing method with Markov model and applied it to the forecast of the quarterly throughput of container ports.

In recent years, Back Propagation (BP) neural network has also been widely used in port container throughput forecasting. For example, Zhang Shukui[5] combined the grey model and BP neural network to construct a grey neural network model to predict port container throughput; Chen Jinwen[6] used the improved BP neural network model with stronger generalization ability to improve the prediction accuracy of port container throughput.
However, due to the shortcomings of Back Propagation (BP) neural network, such as poor global search ability, slow convergence and easy to fall into the local minimum, it is inevitable to improve it. The Artificial Bee Colony (ABC) proposed by the Karaboga team in 2005 is an algorithm based on the intelligent behaviour of the bee colony, and has been proven to be better than Genetic Algorithm (GA), Particle Swarm Algorithm (PSO), and differential Evolution Algorithm. (DE) and Particle Swarm Evolution Algorithm (PS-EA) have better performance [7-9]. It mainly has the advantages of being able to jump out of the local minimum, avoiding local optimization, and having fewer setting parameters.

Therefore, this paper introduces the Artificial Bee Colony (ABC), through optimizing the weights and thresholds of the Back Propagation (BP) neural network to obtain the best weights and thresholds, constructs the ABC-BP ship traffic flow prediction model, and the original BP neural network model and Comparison of prediction models optimized by Genetic Algorithm (GA) to verify the accuracy of ABC-BP prediction models.

2. Model building

2.1. BP neural network model
Back Propagation (BP) neural network is a typical multi-layer feedforward network, which is trained through two processes of signal forward propagation and error back propagation. It consists of an input layer, a hidden layer, and an output layer. Each layer contains one or more neurons, and the hidden layer may contain multiple layers. Its basic network structure is shown in Figure 1.

![Figure 1. BP neural network structure.](image)

The training set passes through the input layer as an input signal, enters the network structure, and passes to the hidden layer after weighted summation. The hidden layer contains the activation function required for network training. Under this function mapping, it is passed to the weighted summation again to The output layer, the output layer eventually generates an output signal. The output signal is compared with the expected output value, and the generated error will be compared with the preset error range. If it exceeds the range, the error will be back-propagated, and the weights and thresholds of the network will be corrected to obtain the ideal output signal.

In the figure, $X = (x_1, x_2, \cdots, x_m)$ is the input signal, $Y = (y_1, y_2, \cdots, y_n)$ is the output signal, $w_{ij}$ is the connection weight from the input layer neuron to the hidden layer neuron, and $v_{ij}$ is the connection weight from the hidden layer to the output layer.
The weighted input from the input layer to the hidden layer is set to \( T = (t_1, t_2, \cdots, t_p) \). The threshold between the input layer and the hidden layer is \( \theta_j \). The calculation formula for \( t_j \) is as follows:

\[
 t_j = \sum_{i=1}^{n} w_{ij} x_i - \theta_j
\]

(1)

The characteristic of the S-type activation function is that the function itself and its derivatives are continuous, so data processing is very convenient in the training process. The activation function in the hidden layer uses the Sigmoid function, namely \( f(x) = \frac{1}{1 + e^{-x}} \), we set the actual output of the hidden layer to \( L = (l_1, l_2, \cdots, l_q) \), so the actual output of the hidden layer is \( l_j = f(t_j) \).

Suppose the threshold of the output layer is \( \alpha_o \), and the weighted input from the hidden layer to the output layer is set to \( R = (r_1, r_2, \cdots, r_n) \), which is \( r_i = \sum_{j=1}^{q} v_{ij} l_j - \alpha_j \). From this, the actual output \( y_t = f(r_t) \) can be obtained by the following training error:

\[
 e_k = \frac{1}{2} \sum_{k=1}^{n} (d_k - y_k)^2
\]

(2)

Among them, \( d_k \) represents the expected output value of the \( k \) time, and \( e_k \) is the training error of the \( k \) time. The whole process is to continuously iteratively obtain the training error, thereby adjusting the weights and thresholds of the input and output layers, so that the training error is continuously reduced, so that the actual output is close to the expected output.

2.2. Artificial bee colony algorithm

Artificial bee colony algorithm is an optimization method proposed by Karaboga to solve the multivariate function optimization problem [8-9].

The artificial bee colony algorithm is divided into the following 4 stages.

1) Initialization phase

Use the reconnaissance bee to initialize all vectors \( x_m \) \((m = 1 \cdots SN)\) representing the honey source, \( SN \) is the size of the population (the number of feasible solutions), and set the control parameters. Since each honey source \( x_m \) is the solution vector of the problem to be optimized, each \( x_m \) contains \( n \) variables \((x_{mi}, i = 1 \cdots n)\). These vectors will be optimized to minimize the objective function. The formula for randomly generating a feasible solution is as follows:

\[
 x_{mi} = l_i + rand(0,1) * (u_i - l_i)
\]

(3)

\( l_i \) and \( u_i \) are the lower limit and upper limit of \( x_{mi} \) respectively.

2) Employed bee phase

According to the position in the memory, the employed bees will find the best honey source among the neighbouring honey sources. When they find a honey source, they will evaluate its fitness. They use the following formula to determine the neighbouring new honey source \( v_{mi} \):

\[
 v_{mi} = x_{mi} + \phi_{mi} (x_{mi} - x_{ki})
\]

(4)

\( x_{ki} \) is a randomly selected honey source, \( i \) is a randomly selected position index, and \( \phi_{mi} \) is a random number between \([-a, a]\). When a new honey source \( v_{mi} \) is generated, it will calculate its fitness
value, and will choose between $x_m$ and $v_m$ according to the greedy algorithm. The fitness value of the solution can be calculated by the following formula:

$$fit_m(x_m) = \begin{cases} 
\frac{1}{1 + f_m(x_m)} & \text{if } f_m(x_m) \geq 0 \\
\frac{1}{1 + \text{abs}(f_m(x_m))} & \text{if } f_m(x_m) < 0 
\end{cases}$$

(5)

$f_m(x_m)$ is the objective function value of the problem $x_m$ to be solved.

3) Follow bee phase

Non-employed bees contain two types of bee groups: follower bees and scout bees. The employed bees will share the honey source information with the following bees waiting in the nest. Afterwards, the following bees will randomly select the food source based on the information. In the artificial bee colony algorithm, the following bees select the honey source according to the probability calculated by the fitness value provided by the employed bees. Therefore, a selection method based on fitness is needed, such as the roulette selection method. The probability $p_m$ that $x_m$ is selected by the following bee can be calculated by the following formula:

$$p_m = \frac{fit_m(x_m)}{\sum_{m=1}^{n}fit_m(x_m)}$$

(6)

When honey source $x_m$ is selected by the following bee, the neighbouring honey source is generated by equation (4), and its fitness value is calculated. In the stage of hiring bees, the choice between $x_m$ and $v_m$ is based on the greedy algorithm. Therefore, more following bees are attracted to a richer honey source and give positive feedback behaviours.

4) Scout bee phase

In order to prevent the algorithm from falling into the local optimum, if the employed bee passes the iteration limit and the quality of the solution has not improved, the employed bee will become a scout bee and the owned solution will be abandoned. After the conversion, the scout bee starts to randomly search for new solutions. For example, if the solution $x_m$ is abandoned, the scout bee converted from the employed bee will generate a new solution according to equation (3). Therefore, honey sources that are initially poor or become poor due to collection are discarded, and negative feedback behaviours are generated to balance positive feedback behaviours.

2.3. A Predictive Model of BP Neural Network Based on Improved ABC

There are many factors that affect the prediction accuracy of Back Propagation (BP) neural network, such as the number of hidden layer neurons, the choice of activation function type, learning rate, initial weights and thresholds. This paper introduces an artificial bee colony algorithm for optimizing the initial weights and thresholds of the Back Propagation (BP) neural network. The specific optimization process is shown in Figure 2.
The specific steps are as follows:
1) Create a new Back Propagation (BP) neural network and initialize
2) Artificial bee colony algorithm parameter initialization. The number of employed bees is $N_e$; the number of following bees is $N_o$; The number of honey sources is the number of solutions is $N_s$; And in the artificial bee colony algorithm, these three numbers are the same, namely $N_e = N_o = N_s$. The maximum number of iterations is $MCN$ (max cycle number); The optimization limit is $limit$. The $D$-dimensional vector $X_i (i = 1, \cdots, N_s)$ is the weight and threshold of the group BP neural network, and the dimension of $D$ satisfies the following formula:

$$D = (I + 1) \times H + (H + 1) \times O$$ (7)

Where $I$ is the number of neurons in the input layer, $H$ is the number of neurons in the hidden layer, and $O$ is the number of neurons in the output layer. Initially randomly generated honey source.
3) Calculate the fitness value corresponding to each solution:

$$f(X_i) = \begin{cases} 1 & MSE_i = 0 \\ \frac{1}{MSE_i + 1} & MSE_i > 0 \end{cases}$$ (8)

Figure 2. ABC optimizes BP neural network process.
 Among them, \( i = 1, \ldots, N_x \), \( MSE_i \) represents the mean square error of the Back Propagation (BP) neural network of the \( i \) solution.

4) The employed bee searches for a new solution according to equation (3) and uses a greedy algorithm. If the fitness value of the new solution is greater than the old solution, the old solution is updated and the new solution is recorded. Otherwise, the number of failed updates of the old solution is increased by 1.

5) Calculate the selected probability of each solution according to equation (6), and then follow the bee to find a new solution from the neighbourhood of the existing solution according to the calculated selected probability

6) If the number of times that the solution \( X_i \) fails to update exceeds the preset optimization limit number \( \text{limit} \), it means that the optimization of this solution has reached the upper limit and can no longer be optimized. The reconnaissance bee will abandon this solution and generate a new solution to replace this solution according to equation (4), and save the optimal solution

7) When the number of iterations is greater than the maximum number of iterations \( MCN \), end the training, otherwise return to step (4) to continue searching

8) Give the optimal solution to the BP neural network for example verification

3. Example verification and analysis

In order to compare the prediction accuracy of the ABC-BP prediction model more clearly, this article also selects the BP neural network prediction model, GA-BP prediction model and ABC-BP prediction model, based on the data for experimental verification and analysis.

For the prediction accuracy, we choose the average absolute error \( MAE \), the average absolute percentage error \( MAPE \), and the root mean square error \( RMSE \) in the statistical prediction method for evaluation.

\[
MAE = \frac{1}{N} \sum \left| Y_{\text{pre}}(t) - Y_{\text{real}}(t) \right| 
\]

\[
MAPE = \frac{1}{N} \sum \left| \frac{Y_{\text{pre}}(t) - Y_{\text{real}}(t)}{Y_{\text{real}}(t)} \right| 
\]

\[
RMSE = \sqrt{\frac{1}{N} \sum \left( Y_{\text{pre}}(t) - Y_{\text{real}}(t) \right)^2} 
\]

Among them, \( Y_{\text{pre}}(t) \) is the predicted value, \( Y_{\text{real}}(t) \) is the actual value.

3.1. BP neural network prediction model

1) Data selection

There are many factors that affect the container throughput of the port. This article draws on the existing research results [10-11] and selects the GDP, port throughput, and total foreign trade import and export of Qingdao City during 2014-2019 (24 quarters) as the network input. Qingdao Port 2014-2019 (72 Month) The statistical data of the container throughput of Qingdao City during 2014-2019 (72 months) is output as the network. After preliminary processing, 24 sets of statistical data are obtained, as shown in Table 1. Take the first 20 sets of data as the training set, and the last 4 sets of data as the test set.

| Serial number | Training set input | Training set output | Serial number | Testing set input | Testing set output |
|---------------|--------------------|--------------------|--------------|------------------|-------------------|
| 1             | 13813,3190,1,1363.7| 493                | 13           | 12469,1937,53,921.4| 443.04            |
| 2             | 14339,3362.62,1441.2| 537                | 14           | 12606,2583.77,1059.7| 450.37            |
Model parameter setting

The BP neural network adopts the original three-layer structure, the number of neurons in the input layer $m$ is 3, the number of neurons in the output layer $n$ is 1, and for the determination of the number $h$ of hidden layer neurons, we draw on existing experience Formula[12], as shown in equations (12) and (13), the range of $h$ is obtained by calculation. Before optimization, multiple experiments are performed to select the number of hidden layer neurons.

$$h = \sqrt{m + n + a}$$  \quad (12)

$$\frac{(m+n)}{2} \leq h \leq (m+n) + 10$$  \quad (13)

$a$ is an adjustment constant between 1-10. Through calculation, we get that when the number of hidden layer neurons $h$ ranges between $[2, 12]$, the prediction effect is better. After many experiments, we finally chose $h$ 10 as the optimal choice. S-shaped function as the activation function, trainlm as the training method, 1000 training times, $10^{-3}$ training target, and a learning rate of 0.01. The prediction results are shown in table 2.

| Serial number | Actual value | Predictive value | Residual |
|---------------|--------------|------------------|----------|
| 21            | 415.36       | 504.37           | 89.01    |
| 22            | 420.99       | 440.92           | 19.93    |
| 23            | 415.3        | 429.71           | 14.41    |
| 24            | 421.21       | 418.12           | -3.09    |

3.2. GA-BP prediction model

The optimization of BP neural network by Genetic Algorithm (GA) is also aimed at the initial weights and thresholds. The basic idea is to use the prediction error of the first BP neural network as the fitness value of GA individuals, and the individuals in GA are to be optimized. Initial weights and thresholds, and then through selection, crossover, mutation operations to find the best individual, that is, the optimal initial weights and thresholds.

Among them, we set the initial population of the genetic algorithm to 50, the number of iterations to 100, and set other operating parameters. The parameters of the BP neural network still use the previously set values for training. BP neural network prediction model and the prediction results of this model are shown in table 3.

| Serial number | Actual value | Predictive value  | Residual |
|---------------|--------------|-------------------|----------|
|               | BP GA-BP     | BP GA-BP          |          |
| 21            | 415.36       | 504.37 378.07     | 89.01 -37.29 |
| 22            | 420.99       | 440.92 414.17     | 19.93 -6.82  |
| 23            | 415.3        | 429.71 381.01     | 14.41 -34.29 |
| 24            | 421.21       | 418.12 413.41     | -3.09 -7.8  |
3.3. ABC-BP prediction model

The network is initialized according to the parameters of the previous Back Propagation (BP) neural network prediction model. Then, the parameters of ABC-BP prediction model are set. Where, the initial population size is 50, the maximum number of iterations $MCN$ is 100, the limit number is 50, and the dimension of $D$ is calculated as 50 according to Equation (7) for initialization and training. The prediction model of GA-BP neural network is compared with the results of the model, as shown in table 4.

| Serial number | Actual value | Predictive value | Residual |
|---------------|--------------|------------------|----------|
|               |              | GA-BP            | ABC-BP   | GA-BP   | ABC-BP   |
| 21            | 415.36       | 378.07           | 422.54   | -37.29  | 7.18     |
| 22            | 420.99       | 414.17           | 423.03   | -6.82   | 2.04     |
| 23            | 415.3        | 381.01           | 421.25   | -34.29  | 5.95     |
| 24            | 421.21       | 413.41           | 423.17   | -7.8    | 1.96     |

3.4. Error comparison analysis

We calculated through formulas (9) (10) (11), and obtained the average absolute error $MAE$, the average absolute percentage error $MAPE$, and the root mean square error $RMSE$ of the three models.

| Model       | MAE  | MAPE% | RMSE  |
|-------------|------|-------|-------|
| BP          | 31.61| 7.5918| 46.1984|
| GA-BP       | 21.55| 5.1766| 25.8539|
| ABC-BP      | 4.28 | 1.0278| 4.8723 |

It can be seen from table 5 that the average percentage error of the ABC-BP neural network prediction model is 4.11488% lower than that of the GA-BP prediction model, while it is 6.564% lower than that of the original BP prediction model. The comparison of prediction results of the data of the last four test sets is shown in figure 3. It can be seen that compared with GA-BP prediction model and the original BP prediction model, the prediction curve of the ABC-BP prediction model is more consistent with the actual value and the fitting accuracy is higher.

![Figure 3. Comparison of the prediction results of the three models.](image)
4. Conclusion
In this paper, weights and thresholds, one of the factors affecting the accuracy of Back Propagation (BP) neural network prediction model, are introduced, and the Artificial Bee Colony (ABC) with fewer parameters and strong global optimization ability is used to optimize the weights and thresholds of BP neural network. The example shows that the Back Propagation (BP) neural network prediction model optimized by ABC has the highest fitting accuracy and the lowest corresponding error compared with BP prediction model and GA-BP prediction model, and it is feasible to predict port container throughput.

Acknowledgments
Authors would like to acknowledge my teacher for his guidance on this thesis and my classmates for their help in data collection.

References
[1] Yang J.H., Yang Y. Container throughput forecast of Shanghai Port based on Grey Model[J]. Journal of Shanghai Maritime University, 2014, 35(02): 28-32.
[2] Du G., Liu Y.N. Forecasting of Shanghai Port container throughput under seasonal variation influence[J]. Journal of East China Normal University (Natural Science), 2015(01): 234-239.
[3] Tang S.M., Lan P.Z., Zhu J.J. Typical factor prediction model of port throughput[J]. Journal of Shanghai Maritime University, 2015(01): 234-239.
[4] Wang Z.Z., Chang D.F., Zhu Z.L., Luo T. Analysis and Prediction of Port Quarterly Container Throughput with ES-Markov Model[J]. Navigation of China, 2019, 42(04): 125-130.
[5] Zhang S.K., Xiao Y.J., Lu Z.A. Prediction Model of Port Container Throughput Based on Grey Neural Network[J]. Journal of Chongqing Jiao tong University (Natural Science), 2015, 34(05): 135-138.
[6] Chen J.W., Lan P.Z. Port Throughput Prediction Based on Improved BP Neural Network[J]. Journal of Jimei University (Natural Science), 2015, 34(05): 135-138.
[7] Karaboga D. Akay B. A comparative study of artificial bee colony algorithm [J]. Applied Mathematics and Computation, 2009, 214(1): 108-132.
[8] Karaboga D. Basturk B. On the performance of artificial bee colony (ABC) algorithm[J]. Applied Soft Computation, 2008, 8(1): 687-697.
[9] Karaboga D. Basturk B. A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm[J]. Journal of Global optimization,2007,39(3):459-71.
[10] Zhu J., Yu S.Q. Influence factors of foreign trade cargo throughput of China’s coastal ports based on trade gravity mod-el[J]. Journal of Shanghai Maritime University, 2011,32(02):14-19+24.
[11] Jiang J., Wang H.Y, Yang Z. Econometric analysis based on the throughput of container and its main influential factors[J]. Journal of Dalian Maritime University, 2007(01):83-86+106
[12] Wang R.B., Xu H.Y., Li B., Feng Y. Research on Method of Determining Hidden Layer Nodes in BP Neural Network[J]. Computer Technology and Development,2018,28(04):31-35.