Beijing-Tianjin-Hebei Energy Demand Combination Forecast Analysis

Jian Li * and Xiaoqian Zhang *
School of Management, Tianjin University of Technology, Tianjin, China

*Corresponding author e-mail: zhangxq7426@stud.tjut.edu.cn, lijian631219@163.com

Abstract. Energy demand forecasting is the basis for responding to high-quality economic development requirements and targeted adjustments to the Beijing-Tianjin-Hebei energy structure. This paper selects five main factors that affect energy demand, constructs a combined forecasting model of a combination of multi-factor gray neural network and ARIMA-BP neural network, and introduces the idea of chaos optimization on this basis to simulate and analyse data from 2012 to 2016, and predict the energy demand in the Beijing-Tianjin-Hebei region in 2020 and 2025. The results show that: 1. Compared with the CGA-ARIMA-BP model and the CGA-GNN model, the CGA-GNN-ARIMA-BP model has higher prediction accuracy; 2. It is estimated that in 2020 and 2025, the energy demand in the Beijing-Tianjin-Hebei region will reach 493 and 552 million tons of standard coal.

1. Introduction
With the acceleration of China's industrialization and urbanization, China's economic development level has been continuously improved, and environmental pollution has become increasingly prominent. The Beijing-Tianjin-Hebei region is one of the regions with severe air pollution in China, and its main source of air pollution is energy consumption. At present, China's economy has shifted from the "high-speed development" stage to the "high-quality development" stage. Adhering to green sustainable development and gradually getting rid of dependence on non-renewable energy has become the only way to change the way of economic development and resolve the contradictions of economic development. The prediction of the total energy demand and trend in the future is of great significance to the formulation and implementation of energy strategies in the Beijing-Tianjin-Hebei region and to guarantee the green and sustainable development of the economy and society.

2. Literature Review
At present, experts and scholars in related fields have conducted in-depth research on influencing factors of energy demand and energy demand forecasting methods. The research basis is mainly economics theory and engineering technology theory. The main forecasting methods used are system dynamics methods [1, 2], Grey System Theory [3], bottom-up method [4], neural network model [5], combined prediction, etc [6, 7]. However, the above prediction methods have their own limitations. They cannot accurately describe the nonlinear relationship of complex giant systems, and the prediction results are non-robust. In recent years, due to the rapid development of computer technology and the abundant
amount of stored data, the combined prediction model with multiple model structures can make better use of computer software's non-linear mapping capabilities to present more accurate prediction results. Therefore, this paper makes use of the combination forecasting model to effectively predict the future energy demand of Beijing-Tianjin-Hebei while effectively overcoming the limitations of traditional forecasting models.

This article is based on the historical data of Beijing-Tianjin-Hebei energy demand. Based on the analysis of the factors affecting energy demand, this paper introduces the idea of chaos optimization and constructs a combined forecasting model to predict the energy demand of Beijing-Tianjin-Hebei from 2020-2025. The chaotic genetic algorithm was used to optimize the neural network, and then the energy demand system was regarded as a sequence with time regularity. The gray neural network and the ARIMA-BP neural network were used to make predictions respectively, and the dual nonlinear combination prediction was realized.

3. Research Method

3.1. Chaos genetic algorithm
The chaos genetic algorithm (CGA) is proposed based on the inversion of the optimization of the genetic algorithm and the ergodicity of the chaos optimization. The basic idea is to first use the chaos method to generate the initial value of the population and adjust the individual's fitness. Then, based on the idea of genetic algorithm, select, cross, and mutate the initial population to achieve the evolution of the population. It has the advantages of simple calculation process, fast calculation speed, high accuracy, etc., and can make up for the shortcomings of traditional neural networks in terms of low prediction accuracy and slow convergence speed.

3.1.1. Initial population chaos generation. In the standard CGA, the initial population of individuals is randomly generated, and there are characteristics such as large dispersion, low correlation, and high repetition rate among individuals, which cannot effectively identify the convergence speed. Therefore, in order to further improve the efficiency of the algorithm, a chaotic ergodic idea is introduced for searching to obtain a higher quality initial population. The chaotic optimization system selected in this paper is in the form of Logistic mapping, the specific calculation formula is shown in formula 1:

\[ p_i^{(u+1)} = \lambda p_i^{(u)} \left( 1 - p_i^{(u)} \right) \]  

\[ PI \] is the chaotic variable; \( i = 1, 2, \ldots, n \) is the chaotic variable number; \( u = 1, 2, \ldots, N \) is the population number; \( \lambda \) is the chaotic attractor, where \( \lambda = 4 \), and the number of chaotic variables and population reflects the length of the individual chromosome and the population size.

3.1.2. Fitness Function. The BP neural network is used to train the chaotic initialization data, and the absolute difference between the output result and the expected value is used as the individual fitness \( F \), the specific calculation formula is shown in formula 2:

\[ F = \mu \left( \sum_{i=1}^{n} |Y_i - O_i| \right) \]  

In the formula, \( i = 1, 2, \ldots, n \) is the output node number of the BP neural network, \( Y_i \) is the expected value of the \( i \)-th node; \( O_i \) is the predicted value of the \( i \)-th node; the coefficient is \( \mu = 1 \).

3.1.3. Select operation. There are many methods for genetic algorithm selection operations, such as roulette and tournament methods. In this paper, roulette is selected, the specific calculation formula is shown in formula 3 and formula 4:

\[ f_i = k_i F_i \]  

\[ k_i \]
In the formula, \( f_i \) is the fitness value of the individual, \( k \) is the coefficient; \( S_i \) is the selection probability of any individual, \( N \) is the number of individuals in the population.

3.1.4. Cross operation. In this paper, individuals are coded with real numbers, so the method of cross-operation is the real-number cross method, that is, the cross-operation method of the \( m \)-th chromosome \( A_m \) and the \( n \)-th chromosome \( A_n \) at the \( j \) position is shown in formula 5:

\[
A_{mj} = A_{mj} (1 - B) + A_{nj} B \\
A_{nj} = A_{nj} (1 - B) + A_{mj} B
\]  

(5)

3.1.5. Mutation operation. Select the \( i \)-th chromosome and \( j \)-th gene for mutation. The mutation operation method is shown in formula 6 and formula 7:

\[
f(g) = r \left( 1 - \frac{g}{G_{\text{max}}} \right)^2
\]  

(6)

\[
A_{ij} = \begin{cases} 
A_{ij} + (A_{ij} - A_{\text{max}}) \cdot f(g), & r \geq 0.5 \\
A_{ij} + (A_{\text{min}} - A_{ij}) \cdot f(g), & r < 0.5 
\end{cases}
\]  

(7)

In the formula, \( A_{\text{max}} \) is the upper bound of the gene \( A_{ij} \); \( A_{\text{min}} \) is the lower bound of the gene \( A_{ij} \); \( g \) is the current number of iterations; \( G_{\text{max}} \) is the maximum number of evolutions; \( r \) is a random number between \([0,1]\).

3.1.6. Chaotic optimization of outstanding individuals. The specific calculation formula is shown in formula 8:

\[
\delta_k' = (1-\varphi) \delta^* + \varphi \delta_k
\]  

(8)

Among them, \( \delta^* \) is the optimal chaotic vector; \( \delta_k \) is the chaotic vector after \( k \) iterations; \( \delta_k' \) is the corresponding chaotic vector after random perturbation; where \( 0 < \varphi < 1 \), and an adaptive selection method is adopted. The method of determining \( \varphi \) in this paper is shown in formula 9:

\[
\varphi = 1 - \left( \frac{k - 1}{k} \right)^t
\]  

(9)

In the formula, \( t \) is an integer and is determined according to the optimization objective function; \( k \) is the number of iterations.

3.2. Grey neural network theory

Let the time series be \( A \), where \( S \) is the initial time series, and \( D \) is the first-order accumulation generation sequence, expressed by the continuous function \( y(t) \), then the differential expression of the grey neural network system is shown in formula 10:

\[
\frac{dy_1}{dt} + Ay_1 = B_1y_2 + B_2y_3 + \cdots + B_{n-1}y_n
\]  

(10)

In the formula: \( y_1 \) is the output data; \( y_2, y_3, \ldots, y_n \) is the input data; \( A, B_1, B_2, \ldots, B_{n-1} \) is the equation parameter. The expression of the solution of formula (11) is:

\[
d = \frac{B_1}{A} y_2(t) + \frac{B_2}{A} y_3(t) + \cdots + \frac{B_{n-1}}{A} y_n(t)
\]  

(11)
Where Map to the extended BP neural network to get the gray neural network, and its model structure is shown in Figure 1.

![Figure 1. Schematic diagram of the gray neural network model](image)

3.3. ARIMA-BP neural network theory

ARIMA model is a typical time series prediction method, the specific calculation formula is shown in formula 12:

\[
\omega_t = \phi_1 \omega_{t-1} + \cdots + \phi_p \omega_{t-p} + \varepsilon + \mu_t + \theta_1 \mu_{t-1} + \cdots + \theta_q \mu_{t-q} \tag{12}
\]

In the formula, \( \omega_t \) is a stationary sequence. The residual sequence of the ARIMA model contains non-linear laws that cannot be explained by the ARIMA model. As a nonlinear dynamic system modeling and prediction, BP neural network can theoretically predict the residual well. The process of ARIMA-BP neural network prediction is as follows:

1. Establish the ARIMA model for the historical data of the variables that need to be predicted, and use the ARIMA model to predict the variables. Let the original data be \( y_t \) and the ARIMA predicted value be \( \hat{Y}_t \), then the residual \( e_t = y_t - \hat{Y}_t \).

2. The BP neural network is used to predict the obtained residual sequence \( e_t \). First determine the factors that have a major influence on the variables, and use it as the input layer of the neural network. Use the residual sequence \( e_t \) as the output layer of the neural network. Then use the trial and error method to determine the hidden layer of the neural network. The residual sequence is predicted to obtain a predicted value \( \hat{e}_t \).

3. The final predicted value is \( \hat{y}_t \). Its calculation formula is \( \hat{y}_t = \hat{Y}_t + \hat{e}_t \).

3.4. Combining prediction steps

Step 1 Specifically analyze the influencing factors of energy demand, and use the obtained influencing factors as the input variables of the gray neural network and the ARIMA-BP neural network, take the energy demand as the output variable of the gray neural network, and the residual predicted by ARIMA as the output variable of the network of the ARIMA-BP neural.

Step 2 Divide the data into training data and test data, use the training data to train the gray neural network and ARIMA-BP neural network, and use the chaotic genetic algorithm to optimize the gray neural network and ARIMA-BP neural network, respectively. The trained gray neural network and ARIMA-BP neural network are simulated and predicted separately to obtain the predicted energy demand of the two.

Step 3 Use the obtained energy demand forecast value of the gray neural network and ARIMA-BP neural network as the input variables of the new BP neural network, and take the actual energy demand as the output variable. Use the chaotic genetic algorithm to optimize the new BP neural network.
neural network for training. After the training, the new BP neural network is simulated and predicted, and the predicted value of the new BP neural network is used as the final predicted value of energy demand.

4. Beijing-Tianjin-Hebei energy demand combination forecast analysis

4.1. Analysis of influencing factors
Energy demand (recorded as Y, unit is 10,000 tons of standard coal) is affected by many factors such as economy, society, technology, and industrial structure. The selection of influencing factor indicators is shown in Table 1:

| Influencing factors          | Evaluation index                                      | unit          |
|-----------------------------|-------------------------------------------------------|---------------|
| Economic growth (X1)        | Gross domestic product                                | 100 million yuan |
| Industrial structure (X2)   | Proportion of secondary industry to GDP               | %             |
| Total population (X3)       | Total population                                      | Ten thousand people |
| Technical progress (X4)     | Energy consumption per unit of GDP                    | 10,000 tons of standard coal / 100 million yuan |
| Household consumption level | Household consumption level                            | yuan          |

4.2. Comparison of prediction accuracy
First, the energy demand influencing factors are used as the input variables of the gray neural network and the ARIMA-BP neural network, and the energy demand is correspondingly used as the output variable of the gray neural network. The residual predicted by the ARIMA model is used as the output variable of the ARIMA-BP neural network. Use 1990-2010 data as the training samples of the neural network to continuously train the neural network, optimize the performance of the neural network, and use the predicted value of energy demand obtained from the simulation of the gray neural network and ARIMA-BP neural network as the input variables of the new neural network to train the new neural network; Finally, the values of the variables that affect the energy demand from 2011 to 2016 are substituted into the neural network. The trained gray neural network and the ARIMA-BP neural network are used for simulation and prediction to obtain the 2011-2016 energy demand forecast value. The predicted value of the new neural network is used as an input variable to obtain the predicted value of the energy demand of the new neural network. In order to illustrate the validity and accuracy of the methods, this article lists the comparison between the predicted values and the actual values of each method. Among them, CGA-GNN represents the gray neural network model based on chaotic genetic algorithm, CGA-ARIMA-BP represents the ARIMA-BP neural network model based on chaotic genetic algorithm, and CGA-GNN-ARIMA-BP represents the gray neural network and ARIMA based on chaotic genetic algorithm - BP neural network intelligent combination prediction model. The results are shown in Table 2:

| year | Actual value | Predictive value | Error (%) | Predictive value | Error (%) | Predictive value | Error (%) |
|------|--------------|------------------|-----------|------------------|-----------|------------------|-----------|
| 2012 | 42652.13     | 42298            | -0.83     | 42362            | -0.68     | 42426            | -0.53     |
| 2013 | 44270.11     | 44766            | 1.12      | 44027            | -0.55     | 44602            | 0.75      |
| 2014 | 44296.47     | 44717            | 0.95      | 43738            | 1.26      | 44571            | 0.62      |
| 2015 | 44508.09     | 44219            | -0.65     | 44828            | 0.72      | 44321            | -0.42     |
| 2016 | 45001.40     | 45541            | 1.2       | 45379            | 0.84      | 45343            | 0.76      |
From Table 2, we can conclude that the average error predicted by the CGA-GNN model is 0.95%, the average error predicted by the CGA-ARIMA-BP model is 0.81%, and the average error predicted by the CGA-GNN-ARIMA-BP model is 0.62%. Obviously the CGA-GNN-ARIMA-BP model has the highest prediction accuracy of the three, followed by the CGA-ARIMA-BP model, and the worst is the CGA-GNN model. From the perspective of prediction accuracy, the prediction results of the three models are obviously very good. This is because the gray neural network model and ARIMA-BP neural network adopted in this paper are deeply integrated combined prediction models rather than simple combined prediction models. A genetic algorithm based on chaotic thought optimization is used to optimize the performance of the neural network, which can deeply integrate the prediction methods into one. Its advantages are exerted while ensuring the overall consistency.

4.3. Beijing-Tianjin-Hebei Energy Demand Forecast

In order to accurately predict China's future energy demand, it is necessary to accurately analyze the future changes in the factors affecting energy demand. This article adopts a prediction method combining ARIMA model and grey system theory prediction. The ARIMA model is used to predict the linear part, the GM (1, 1) model is used to predict the non-linear part (ie, the residual part), and the linear + non-linear prediction method is used. You can get more accurate predictions. After the future values of the variables of the energy influence factors are predicted, the input to the neural network can predict the future energy demand values.

Table 3 lists the energy demand values of the Beijing-Tianjin-Hebei region in 2020 and 2025. In the process of using intelligent combined neural network prediction, the relationship between economic growth and energy demand is used to determine that China's economy has entered a new normal situation. The energy demand growth rate in the Beijing-Tianjin-Hebei region has also shifted from high-speed growth to medium- and high-speed growth. According to the forecast results, it is estimated that in 2020, the energy consumption per unit of GDP in the Beijing-Tianjin-Hebei region will decrease by 28.57% compared to 2015, far higher than the expected target. The total energy demand can reach 55.21 million tons of standard coal in 2025. It can be seen that although the growth rate of energy demand in the Beijing-Tianjin-Hebei region has slowed down, there is still a certain degree of pressure. Therefore, in order to change this situation, the Beijing-Tianjin-Hebei region should accelerate the process of industrial structure transformation and upgrading, and continue to make breakthroughs in the development of new energy and new technologies in order to better alleviate the pressure on Beijing-Tianjin-Hebei energy needs.

5. Conclusion

In the prediction method, this paper adopts the gray neural network model based on chaotic genetic algorithm and ARIMA-BP neural network model, and establishes an energy demand prediction model based on influencing factors. GDP, industrial structure, technological progress, population and residents' consumption levels are used as the input variables of the network and energy demand are used as the output variable of the network. The data from 1990 to 2016 are used for simulation and simulation. The simulation results show that the intelligent combined neural network model has higher accuracy and better prediction results.

By analyzing and forecasting the results, we can see that in the context of China’s economic development entering a high-quality development stage, the economic growth rate of the Beijing-Tianjin-Hebei region is shifting from high-speed growth to medium- and high-speed growth, but the overall energy demand is still showing a growing trend. The relationship between economic growth and energy demand revises the forecasting model. It is found that by 2025, the total energy demand in the
Beijing-Tianjin-Hebei region will reach 5.52 billion tons of standard coal. Therefore, it is necessary to accelerate the transformation and upgrading of the industrial structure in the Beijing-Tianjin-Hebei region, and continue to make breakthroughs in the development of new energy and new technologies in order to achieve the transformation and upgrade to a "low-carbon economy."

References
[1] Fang D B, Shi S S, Yang J P. Research on Forecast and Early Warning of China’s Energy Demand under the New Normal [J]. Resources Development and Market, 2017, 33 (01): 8-13 + 26.
[2] Li J, Guo J, Yuan Q M. Research on Energy Demand Forecast and Policy Impact in the Background of Beijing-Tianjin-Hebei Coordinated Development [J] .Journal of Arid Land Resources and Environment, 2018,32 (05): 5-11.
[3] Li J, Yu Y P. Secondary nonlinear energy demand prediction based on error correction [J]. Journal of Arid Land Resources and Environment, 2016, 30 (11): 13-18.
[4] Chen R, Rao Z H, Liu J X, Tong Y Y, Liao S M. Research on Energy Demand Forecast and Countermeasures of Changsha City Based on LEAP Model [J] .Resources Science, 2017,39 (03): 482-489.
[5] Li J B, Xian X F. Forecast of Chongqing Energy Demand Based on Grey Neural Network Model [J]. Journal of Southwest University (Natural Science Edition), 2016, 38 (06): 136-141.
[6] Wang Y. Forecast of China’s Energy Consumption Demand Trend Based on Model Combination Method [J]. Statistics and Decision, 2018, 34 (20): 86-89.
[7] Xiao J, Sun H Y, Liu D H, Cao H W, Wang S Y. Research on Energy Consumption Forecasting Based on GMDH Hybrid Model [J] .China Management Science, 2017,25 (12): 158-166.