Research on Regional Logistics Prediction Based on Genetic Neural Network

Futou Li

School of Management, Tianjin Polytechnic University, Tianjin, 300387, China
E-mail: futou_li@163.com

Abstract: As one of the important areas of modern logistics industry development, regional logistics, whether it is coordinated with the level of regional economic development, has become an important factor affecting the sustainable, rapid, stable and healthy development of the regional economy. In-depth study of the relationship between regional logistics demand and regional economic development level, for the development of logistics business of third-party logistics enterprises, to enrich and improve regional logistics theory, and to improve the scientific and practical development of government departments to develop modern logistics industry development plans Targeted, both have important theoretical and practical value. Based on the investigation of the status quo of domestic and international regional logistics development and current problems, this paper points out the main factors affecting regional logistics demand through the theoretical analysis of the relationship between regional logistics and regional economic development, and builds regional logistics demand on this basis. The predictive index system, and then through the use of artificial intelligence modeling methods, established a genetic neural network model, and achieved a good prediction effect. Finally, an empirical study of the model is carried out in Tianjin, and relevant suggestions and countermeasures are proposed.

Keywords: regional logistics; regional economy; neural network; genetic algorithm

1. Introduction

With the acceleration of economic globalization and regional economic integration, the continuous advancement of modern information network technology, and the increasingly perfect supply chain management theory, the logistics industry as an emerging service industry with broad prospects and value-added functions, on a global scale Developed rapidly. At present, all countries in the world attach great importance to the development of the logistics industry, especially in the related theoretical research, logistics practice, logistics technology innovation and other aspects have made great progress, but also make the logistics industry become a new growth point and pull of the national economy. As one of the important areas of modern logistics industry development, regional logistics, whether it is coordinated with the level of regional economic development, has become an important factor affecting the sustainable, rapid, stable and healthy development of the regional economy. Eksoz et al. [1] applied grey model, neural network model and combination model, combined with demand factors, predicted the short-term demand of railway cold chain logistics in a region. Bahram et al. [2] proposed a generalized autoregressive conditional heteroscedasticity model for fitting, and the prediction effect is slightly better than that of the autoregressive model.

2. Selection of Predictive Indicators

The regional economic indicators used in this paper refer to the economic indicators used in the regional logistics demand forecasting process that have a significant impact on regional logistics demand. From the analysis of the economic impact factors of regional logistics demand, it can be seen that the regional economic aggregate, regional economic structure and its distribution are important economic factors affecting regional logistics demand. In addition, intra-regional trade and per capita income and consumption levels in the region are also important aspects that affect regional logistics needs. Therefore, when setting regional economic indicators, it is necessary to include as many relevant indicators as possible in order to effectively use regional logistics demand, and the data of these aspects are relatively easy to obtain from the statistical yearbooks of various provinces and cities. Therefore, economic indicators can be set from the following aspects:

(1) Regional economic aggregate indicators: gross domestic product (GDP) and per capita GDP;

(2) Regional industrial structure indicators: total industrial output value, total agricultural output value, total output value of the tertiary industry, etc.;

(3) Trade indicators within and outside the region: total retail sales in the region and total regional foreign trade;

(4) Regional consumption level indicators: per capita income and per capita consumption level. Considering the availability of data, this paper only uses the freight volume Y (10,000 tons) as the logistics demand scale indicator when forecasting the logistics demand.

Through the above analysis of the economic factors that affect the regional logistics demand, the corresponding indicators for the forecast of logistics demand scale are: Primary industry X_1 (100 million yuan); Second industry output value X_2 (100 million yuan); Third industry output value X_3 (100 million yuan);Regional retail sales X_4 (100 million yuan); Regional foreign trade X_5 (100 million US dollars); Household consumption level X_6 (yuan).

3. Establishment of Genetic Neural Network Model

In 1975, the genetic algorithm was inspired by the theory of biological evolution by Professor J. Holland of the University of Michigan in the United States. It is an efficient global search method that effectively uses existing information to automatically acquire and accumulate relevant searches during the search process. Knowledge of space, and adaptive control direction to the optimal solution. Genetic algorithms are particularly well suited for solving complex and nonlinear problems. Neural network is an abstract mathematical model based on the results of modern biological research on human brain tissue to simulate the structure and behavior of human brain neural network. Since the American psychologist W. McCullch and the mathematician W. Pitts proposed the abstract mathematical model of formal neurons in 1943, the artificial neural network has undergone a tortuous development for more than half a century. In 1989, Robert Hecht-Nielson proved that any continuous function in a closed interval can be approximated by a hidden BP neural network, so a three-layer BP neural network can perform arbitrary n-dimensional to m-dimensional mapping. The commonly used neural network generally has three layers: the input layer, the hidden layer and the output layer. For an arbitrary time series of a set $\{(x(t), y(t)) x \in R^m, y \in R^n, t = 1,2,\cdots,N\}$ of N samples, the three-layer neural network can completely perform any highly nonlinear input and output. That is, we can find the relevant mappings to make: $F : R^m \to R^n$. Dividing N sample sets into training sample Φ_1 and detecting sample Φ_2:

$$\Phi_1 = \begin{Bmatrix} (x(t), y(t)) x \in R^m, y \in R^n, \\ t = 1,2,\cdots, N_1, N_1 \le N \end{Bmatrix} \tag{1}$$

$$\Phi_2 = \begin{Bmatrix} (x(t), y(t)) x \in R^m, y \in R^n, \\ t = N_1 + 1, N_1 + 2,\cdots, N \end{Bmatrix} \tag{2}$$

First use Φ_1 to establish a mapping relationship, and then see if the network can give the correct input and output, if so, the model can be predicted in practice. For such a time series, it is possible to implement a three-layer BP neural network with m input nodes, *n* output nodes, and P hidden layer nodes. The nonlinear relationship between input and output in a neural network is described by the following function:

$$\varphi(t) = \sum_{j=1}^{p} v_{jk} f\left[\sum_{i=1}^{m} w_{ij} \cdot x_i(t) + \theta_j\right] + r_k \tag{4}$$

where $f(x) = \dfrac{1}{1+e^x}$, $k = 1,2,\cdots n$, $t = 1,2,\cdots N_1$ x_i is the input to the network, $\varphi(t)$ is the output of the network, and w_{ij} is the weight of the input layer to the output layer node. v_{jk} is the weight of the hidden layer node to the output layer node, θ_j is the threshold at the hidden layer node, r_k is the threshold of the output node, and f is the activation function. Let the overall error of the system be less than δ_1, then there are:

$$E_1 = \frac{1}{2}\sum_{t=1}^{N_1}\sum_{k=1}^{n}[\varphi(t) - \varphi_1(t)]^2 \le \delta_1 \tag{5}$$

Let the average error of the test sample be less than a, then there are:

$$E_2 = \frac{1}{N - N_1}\sum_{t=N_1}^{N}\sum_{k=1}^{n}[\varphi(t) - \varphi_1(t)]^2 \le \delta_2 \tag{6}$$

The connection weight of BP neural network mainly affects the performance of the network. Its acquisition is gradually adjusted in training by giving a set of initial weights, and finally a better weight distribution is obtained, but the initial point selection is blind. It is difficult to select a global initial point, and the solution to the actual problem is often a complex multi-dimensional surface. There are many local extremum points that make the neural network into a local extremum. In addition, the advantages and disadvantages of the network structure have a great impact on the processing power of the network. In practice, the incremental or declining heuristics are used to determine the structure of the network. So far, there is no effective solution. Using the global search ability of the genetic algorithm to optimize the connection weight of the neural network can better overcome this difficulty and improve the performance of the neural network. Genetic algorithm optimizes the basic idea of neural network: changing the neural network relies on the guidance of gradient information to adjust the network weight, but using the global search characteristics of genetic algorithm to find the most suitable connection and network structure. Since the three-layer neural network is composed of the input layer, the output layer and the hidden layer, and the input layer and the output layer are determined by the modeling itself, the number of hidden layer nodes is mainly optimized when optimizing the neural network structure.

The theory of genetically optimized neural networks is as follows:

$$\begin{cases} \min E(w, \ v, \ \theta, \ r) = \dfrac{1}{2}\sum_{t=1}^{N_1}\sum_{k=1}^{n}[\varphi(t)-\varphi_1(t)]^2 \\ s.t \quad w \in R^{m\times p}, \quad v \in R^{n\times p}, \theta \in R^p, r \in R^n \end{cases} \tag{7}$$

The genetic algorithm is mainly used to solve the optimization problem of the maximum value. Based on the above problems, this paper adjusts the fitness function as follows:

$$F(w, \ v, \ \theta, \ r) = \frac{1}{\sum_{t=1}^{N_1}\sum_{k=1}^{n}[\varphi(t)-\varphi_1(t)]^2} \tag{8}$$

Then the above formula can be transformed into the following plan:

$$\begin{cases} \max F(w, \ v, \ \theta, \ r) = \dfrac{1}{\sum_{t=1}^{N_1}\sum_{k=1}^{n}[\varphi(t)-\varphi_1(t)]^2} \\ s.t \quad w \in R^{m\times p}, \quad v \in R^{p\times n}, \theta \in R^p, r \in R^n \end{cases} \tag{9}$$

Using matlab software to solve the above planning problem, the corresponding optimal weight w_1, w_2 of the neural network and the optimal threshold B_1, B_2 can be obtained.

4. Empirical Analysis

All data in this paper are derived from China's relevant statistical yearbook. According to the index system stipulated in the foregoing, the data is processed appropriately. To ensure the accuracy of the forecast and eliminate the influence of time factors, the data is averaged into a ring ratio. The 14 samples from 1996-2006 were used as learning samples for the model, and 2007-2014 was used as the test sample for the model. The relevant parameters of the genetic optimization neural network algorithm are set as follows: The number of input nodes is set to 3, the number of output nodes is set to 1, the population size is set to 50, the algebra of evolution is set to 150, the number of learning times of the network is set to 3000, and the training error rate of the network is set to 0.0001, The relevant parameters of the network model can be obtained by using the genetic algorithm toolbox as follows.

$$w_1 = \begin{pmatrix} 0.2892 & 0.1520 & 0.6273 \\ 0.8083 & 0.1613 & -0.1342 \\ -0.7505 & 0.3547 & 0.7705 \end{pmatrix} \tag{10}$$

$$w_2 = (0.0282 \quad -0.0173 \quad -0.5187) \tag{11}$$

$$B_1 = (-0.5292 \quad -0.0412 \quad -0.8715) \tag{12}$$

$$B_2 = (0.8105) \quad N = 3 \tag{13}$$

where w_1, w_2 is the weight of each neuron from the input layer to the hidden layer and the hidden layer to the output layer, B_1, B_2 are the thresholds of the hidden layer and the output layer neurons, respectively, N is the number of interneurons. Using the optimized neural network, 14 data from 1996-2006 were used as learning samples, and samples from 2007-2015 were used as test samples. The training results are shown in Table 1. It can be seen that the reliability of the training results of the genetic neural network model is very high, and the training fitting error is only 0.67%.

As the model training is completed, the parameters related to the model are determined. The next step is to test the credibility of the model. The results of the test samples are shown in Table 2.

Comparison of genetic neural networks with multiple regression models and neural networks, as shown in Tables 3 and 4.

Table 1. Training sample set

NO.	Input index						True value	Predicted value	Error
	x_1	x_2	x_3	x_4	x_5	x_6			
1	0.8865	1.0674	1.2202	1.1798	1.1618	1.0802	1.0287	1.0415	0.0116
2	1.0474	1.0347	1.0633	1.1033	0.9208	1.0578	1.0316	1.0416	0.0098
3	1.2581	1.0203	1.0657	1.0122	0.9381	0.9845	1.0205	1.0263	0.0057
4	1.0156	1.0242	1.1147	1.1213	1.0635	1.0678	1.0244	1.0232	-0.0012
5	1.2766	1.0875	1.1838	1.2288	1.0628	1.1804	1.0207	1.0236	0.0028
6	1.1314	1.1822	1.2377	1.4014	1.1758	1.4188	1.0705	1.0393	-0.0311
7	1.0083	1.0318	1.1114	1.1353	1.0057	1.1554	1.0486	1.0368	-0.0114
8	1.0971	1.084	1.1803	1.1442	1.1521	1.0907	1.0673	1.0698	0.0028
9	1.2666	1.1886	1.1782	1.3134	1.2082	1.2944	1.0531	1.0633	0.0103
10	1.0831	1.0764	1.0682	1.1201	1.0833	1.1072	1.0216	1.0242	0.0023
11	1.1555	1.0845	1.2046	1.0073	0.9461	1.0802	1.0154	1.0217	0.0062
12	0.9946	1.0892	1.2782	1.1445	1.0826	1.2052	1.0498	1.0517	0.0021
13	1.0027	1.2303	1.3033	1.2167	1.2131	1.1738	1.0365	1.0386	0.0022
14	1.1072	1.3322	1.4375	1.4543	1.3047	1.4644	1.0593	1.0598	0.0006

Table 2. Test sample set

NO.	Input index						True value	Predicted value	Error
	x_1	x_2	x_3	x_4	x_5	x_6			
1	1.0252	1.0553	1.1648	1.1101	1.2656	1.0733	1.0467	1.0414	-0.0053
2	1.0087	1.0603	1.1479	1.0811	1.2317	1.0886	1.0788	1.0681	-0.0108

3	1.0295	1.1125	1.1679	1.0831	1.4171	1.1281	1.0354	1.0343	-0.0016
4	1.0173	1.0885	1.0974	1.0811	1.1132	1.0811	1.0899	1.0747	-0.0153
5	1.0214	1.0913	1.1137	1.0931	1.1933	1.1127	1.0843	1.0741	-0.0103
6	1.0167	1.2236	1.1203	1.0911	1.5467	1.0846	1.0298	1.0246	-0.0051
7	1.0301	1.2128	1.2037	1.1051	1.4237	1.1561	1.0844	1.0737	-0.0108
8	1.0817	1.1442	1.1277	1.1191	1.1647	1.1168	1.0544	1.0595	0.006
9	1.0817	1.1442	1.1277	1.1191	1.1647	1.1168	1.0544	1.0595	0.0062
10	1.0215	1.0914	1.1136	1.0831	1.1733	1.1027	1.0853	1.0751	-0.0113

Table 3. Comparison results with multiple regression models.

Sample	Genetic-neural network			Multivariate nonlinear regression model		
	Absolute error	Predictive value	Relative error	Absolute error	Predictive value	Relative error
1	290.77	1.0414	0.0103	1099.78	3.0076	0.3765
2	300.56	1.0681	0.0028	899.98	3.5673	0.3644
3	285.93	1.0747	0.0062	980.37	2.9978	0.2988
4	389.43	1.0737	-0.0067	996.45	4.6734	0.0985
5	276.55	1.0741	0.0034	999.67	3.7764	0.0764

Table 4. Comparison results with neural networks.

Sample	Genetic-neural network			Neural Networks		
	Absolute error	Predictive value	Relative error	Absolute error	Predictive value	Relative error
1	290.77	1.0414	0.0103	463.22	1.9988	0.1164
2	300.56	1.0681	0.0028	388.45	2.0043	0.0934
3	285.93	1.0747	0.0062	409.88	1.9967	0.0099
4	389.43	1.0737	-0.0067	499.43	1.7843	-0.0779
5	276.55	1.0741	0.0034	379.44	1.6932	0.1009

5. Conclusions and Suggestions

From the perspective of Tianjin's economic development and logistics demand, this paper believes that the development of Tianjin's logistics industry needs to provide more room for development, and on this basis, strive to optimize the supply capacity of logistics services. Therefore, when the local government formulates policy planning for logistics, it can proceed from the following aspects: (1) Develop regulations and policies related to logistics. Regulations and policies are a kind of protection and promotion of the development of logistics industry in the system, mainly including government support, taxation finance and other related preferential policies. (2) Formulate a coordinated development plan for logistics and economy. Tianjin's fast-growing economy plays an important role in China, and the logistics industry has an irreplaceable role in economic development, creating huge profits for the economy. Therefore, we must grasp the coordination of logistics and economic development as a whole, and we can use the corresponding forecasting tools and methods to predict the future flow of goods, and use this as a basis to formulate corresponding policies and long-term planning strategies. (3) Improve Tianjin's logistics facilities planning. From the actual situation, the logistics industry has strong dependence on facility planning. In other words, the infrastructure largely determines the effective logistics service capabilities. As the economic center of the north, Tianjin has gradually increased its influence on the internal economy, and Tianjin's foreign trade is increasing year by year. Therefore, its logistics service capacity is also bound to increase. However, the existing infrastructure construction can no longer meet the current fast-growing economy, so it is necessary to speed up the construction of logistics infrastructure to ensure the coordination of logistics and economic development. (4) Improve logistics service levels and optimize structure. Advancedization and softening are the main trends in the economic structure of Tianjin. Products are shifting towards high and fine. With the change of consumer consumption concept and the rapid development of e-commerce, the logistics service and structure in Tianjin have changed. Therefore, the government should adopt corresponding means to rationally adjust the service structure of logistics to better promote the development of the logistics industry. (5) Develop logistics technical standards. The local government should speed up the construction of logistics information technology in accordance with the existing standards of the country. At the same time, it must be in line with the international community to ensure that the logistics service system platform operates normally in a unified, orderly and efficient environment.

References

[1] Eksoz C., Afshinmansoun S., Michael B. Collaborative forecasting in the food supply chain: A conceptual frame-work. *INT. J. PROD. ECON.*, **2014**, 158: 120-135.

[2] Bahram Adrangi, Arjun Chatrath, Kambiz Raee. The demand for US air transport service: A chaos and nonlinearity investigation, *TRANSPORT RES. E-LOG.*, **2001**, 37: 337-353.