

Urban cold-chain logistics demand predicting model based on improved neural network model

Ying Chen^{1,3,*}, Qiuming Wu¹, and Lijin Shao²

¹ School of Economic and Management, Fuzhou University, Fuzhou 350001, Fujian, PR China

² Fuzhou Melbourne Polytechnic, Fuzhou 350001, Fujian, PR China

³ School of business administration, Fujian Business University, Fuzhou 350001, Fujian, PR China

Received: 9 November 2019 / Accepted: 15 February 2020

Abstract. With the popularity of the Internet and mobile terminals, the development of e-commerce has become hotter. Therefore, e-commerce research starts to focus on the statistics and prediction of the cargo volume of logistics. This study briefly introduced the back-propagation (BP) neural network model and principal component analysis (PCA) method and combined them to obtain an improved PCA-BP neural network model. Then the traditional BP neural network model and the improved PCA-BP neural network model were used to perform the empirical analysis of the cold chain logistics demand of fruits and vegetables in city A from 2010 to 2018. The results showed that the main factors that affected the local cold chain logistics demand were the growth rate of GDP, the added value of primary industry, the planting area of fruits and vegetables, and the consumption price index of fruits and vegetables; both kinds of neural networks model could effectively predict the cold chain logistics demand, but the predicted value of the PCA-BP neural network model was more fitted with the actual value. The prediction error of the BP neural network model was larger, and the fluctuation was obvious within the prediction interval. Moreover, the time required for the prediction by the PCA-BP neural network model was less than that by the BP neural network model. In summary, the improved PCA-BP neural network model is faster and more accurate than the traditional BP model in predicting the cold chain logistics demand.

Keywords: Back-propagation neural network / principal component analysis / cold chain logistics / demand prediction

1 Introduction

With the popularity of the Internet and mobile terminals, e-commerce has developed rapidly. At the same time, people's living standards have also improved, and the demand for food begin to diversify with the development of the economy [1]. Fresh fruits and vegetables, as well as seafood, which require transportation with cryopreservation, have gradually become part of goods in e-commerce. For e-commerce companies, the prediction of logistics demand is an important basis for strategic decision-making and market analysis [2], especially for fruits, vegetables and seafood which require cold chain logistics. Fruits, vegetables, seafood, etc. are limited by their own expiration date and cannot be stored as long as other products. Once the judgment of cold chain logistics is wrong, it will cause huge losses. Therefore, the prediction of the demand for products

which require cold chain logistics is important [3]. Cheng et al. [4] proposed an improved simulated annealing particle swarm optimization (SAPSO) algorithm and space-time adjustment strategy for establishing a combined prediction model of logistics traffic and then set up one based on SAPSO-back-propagation (BP) neural network. The simulation test verified that the SAPSO which was proposed had better convergence performance and higher stability than Particle Swarm Optimization (PSO)-BP or BP neural network models. By using the comprehensive prediction theory, Xia et al. [5] combined with many factors which reflected the level of regional economic development, established a logistics demand index system, proposed a continuous gray neural network prediction model, and empirically analyzed the logistics demand in Guangxi. The results showed that the predicted value was close to the actual value, which had a good predictive effect. Ya [6] designed and applied multiple regression and AW-BP prediction methods based on order parameters of system for cold chain logistics of food.

* Corresponding author: cyingyc@yeah.net

Table 1. Factors that affect the cold chain logistics demand of fruits and vegetables.

Number	Unit	Factor
X_1	%	GDP growth rate
X_2	Hundred million yuan	Added value of primary industry
X_3	Ten thousand/hm ²	Planting area of fruits and vegetables
X_4	/	Consumer price index
X_5	km	Total transportation distance
X_6	Ten thousand households	Number of mobile phone users
X_7	Vehicle	Transport vehicle
X_8	Ten thousand tons	Annual transportation volume
X_9	Hundred million yuan	Transportation facility investment

By introducing error function correction and dynamic adaptive weight, it avoided some disadvantages of being slow convergence and easily trapped in the situation of local optimization which happened in the general BP neural network. Experiments showed that the convergence speed, prediction accuracy and avoidance of local extremum of the new prediction method were greatly improved. For cold chain logistics, the new prediction method was adaptable, simple, practical and efficient. This study briefly introduced the BP neural network model and the principal component analysis (PCA) method, and combined the two to obtain the improved PCA-BP neural network model. The traditional BP neural network model and the improved PCA-BP neural network model were used to perform the empirical analysis of the cold chain logistics demand for fruits and vegetables in city A from 2010 to 2018.

2 BP neural network model for predicting the cold chain logistics demand of fruits and vegetables

For fruits and vegetables, one of their edible values is freshness, but because of their perishable characteristic and the demand of consumers for higher quality, merchants need to transport fruits and vegetables in the environment of cryopreservation all the way, which is called cold chain logistics. For merchants, the implementation of cold chain logistics can ensure the freshness of fruits and vegetables and increase their sales in a certain extent, however, the cost is increased, too. Thus, merchants must have an accurate judgment on the cold chain logistics demand of fruits and vegetables. Also, there are many models that can be used to predict logistics demand. In order to predict the logistics demand, the factors that influence it need to be analyzed firstly.

For factors affecting cold chain logistics, different researchers will choose different factors due to different concerns. After analysis, this study finally selected the relevant factors from aspects of the importance of the influencing factors, whether the relevant data are easy to obtain and whether the factors can reflect the objectivity and operability of cold chain logistics. The selected factors

are shown in Table 1. Nine influencing factors were selected from three aspects: macroeconomic level, logistics traffic information level and sustainable development level. Among them, GDP growth rate, added value of primary industry, planting area of fruits and vegetables and consumer price index belong to macroeconomic level; total transportation distance, number of mobile phone users and transport vehicles belong to logistics transportation information level; annual transportation volume and transportation facility investment belong to sustainable development level.

After analysis, it is known that the changes affected by the above factors in the cold chain logistics demand are nonlinear changes, and most of the above prediction models are more suitable for data prediction which belongs to linear variation or changes into linear variation after preprocessing. For the non-linear variation prediction of the cold chain logistics demand in this study, a nonlinear large-scale adaptive BP neural network model is adopted.

The basic structure of BP neural network [5] is divided into input layer, hidden layer and output layer, as shown in Figure 1. The BP neural network model is a three-layer structure. x_1, x_2, \dots, x_n is input vector X which represents all kinds of factors that affect the cold chain logistics demand in this study, i, j, k are relative dimensions of input layer, hidden layer, and output layer, ω_{ij} is the weight of the input layer to the hidden layer, ω_{jk} is the weight of the hidden layer to the output layer, and d_1, d_2, \dots, d_n is output vector D , which indicates the predicted cold chain logistics demand. The algorithm used in the BP neural network is the error back propagation algorithm whose training principle for the neural network [7] is: the input and the predetermined output are firstly set, and then the actual output is calculated layer by layer forwardly and compared with the predetermined output. When there is an error, the weight is adjusted in the opposite direction of the network to keep the error between the actual output and predetermined output in the specified range.

As shown in Figure 2, the weight is firstly initialized by setting it to the minimum value, then the input vector (the collected data of influence factors related to cold chain logistics which is collected in recent years) and expected output (the cold chain logistics demand in recent years), and then they are calculated layer by layer in the forward

direction of the network. The calculation formula in the hidden layer [8] is:

$$o_j = f\left(\sum_{i=1}^n \omega_{ij}x_i - \alpha_j\right), \quad (1)$$

where o_j is the output vector of hidden layer; α_j is the adjustment item of hidden layer; $f(\bullet)$ is the activation function of the hidden layer, either linear or non-linear. The calculation formula of the output layer [9] is:

$$y_k = g\left(\sum_{j=1}^n \omega_{jk}o_j - \beta_k\right), \quad (2)$$

where y_k is the output vector of output layer; β_k is the adjustment item of output layer; $g(\bullet)$ is the activation function of the output layer, either linear or non-linear.

After the layer-by-layer calculation, the actual output vector is obtained and compared with the output vector that is set. The error calculation formula [10] is:

$$E = \frac{\sum_{k=1}^t (d_k - y_k)^2}{2}, \quad (3)$$

where E is the error between the actual output vector and the expected output vector; t is the dimension of the output vector. After judging the error, if the result is within the specified range, the result is directly output. If not within the specified range, the weights of the calculation formulas in the hidden layer and the output layer are inversely adjusted, and the weight adjustment formula of the output layer to the hidden layer [11] is:

$$\omega_{jk}(t + 1) = \omega_{jk}(t) + \eta(1 - y_k)o_j, \quad (4)$$

where η is the learning rate that adjusts the weight. The weight adjustment formula [12] of the hidden layer to the input layer is:

$$\omega_{ij}(t + 1) = \omega_{ij}(t) + \eta o_j(1 - o_j) \sum_{k=1}^t \sigma_k \omega_{jk} o_j, \quad (5)$$

where σ_k is the error between the actual vector and the expected vector in the k -dimensional output layer. After the weight adjustment, the error of the actual output vector and the expected output vector is recalculated, and the above steps are repeated until the error reaches the specified range.

3 PCA

PCA [13] recombines the original variables into new sets of aggregate variable in the practical application, the aggregate variables are unrelated, and aggregate variables can reflect original variable information as much as possible with relatively small amount. Combining PCA with the BP neural network, it is possible to greatly reduce the amount of calculation and improve the computational efficiency while ensuring the accuracy.

The way in which PCA reduces the original data is as follows. Firstly, the original data matrix of the dimension $m \times n$ is transformed into a matrix form by

$$Y = (Y_1, Y_2, \dots, Y_n)^T = W^T X. \quad (6)$$

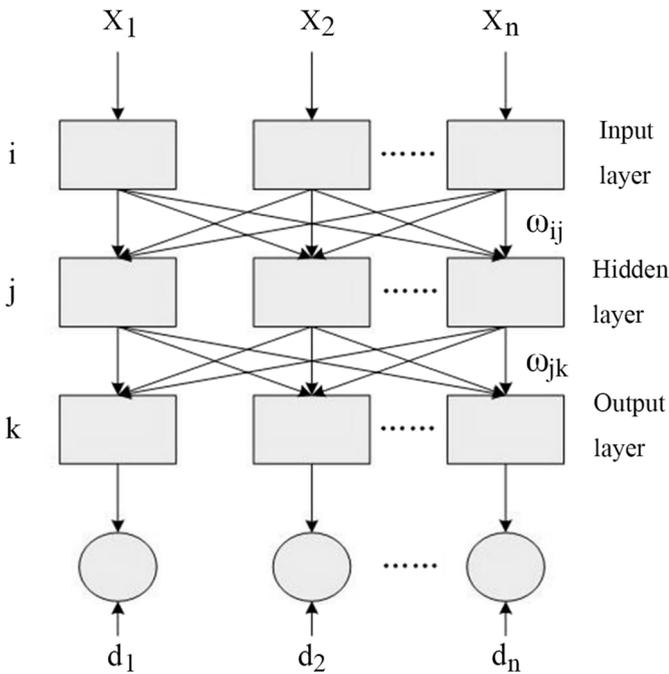


Fig. 1. Structure of BP neural network model.

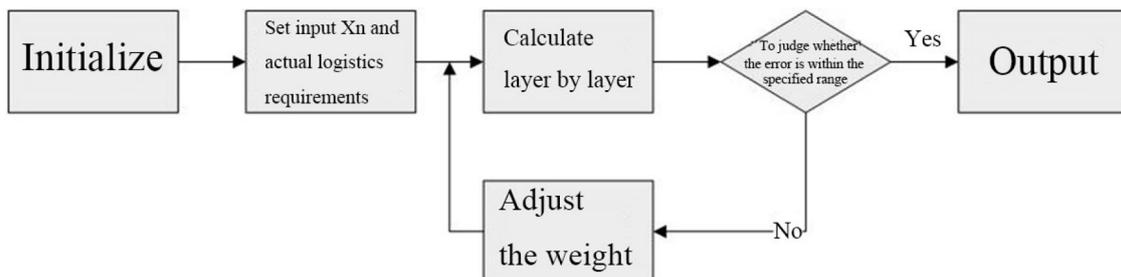


Fig. 2. Algorithm flow for predicting the cold chain logistics demand of fruits and vegetables.

Table 2. Statistics required for the prediction of the cold chain logistics demand of fruits and vegetables in city A from 2010 to 2018.

Number	Unit	2010	2011	2012	2013	2014	2015	2016	2017	2018
X_1	%	14.1	12.4	11.2	10.2	9.6	9.2	9.1	8.8	8.5
X_2	Hundred million yuan	1206	1391	1520	1588	1683	1772	1904	1953	1988
X_3	Million hm^2	52.0	53.4	54.9	56.2	57.3	58.6	60.5	61.8	62.3
X_4	/	105.5	111.3	105.1	104.6	103.6	103.2	104.5	99.4	100.1
X_5	km	101 454	114 450	120 333	122 674	128 262	119 949	134 026	138 075	139 654
X_6	Ten thousand households	1812	2364	2636	2805	2940	3055	3141	3450	3555
X_7	Vehicle	301 678	467 560	370 112	545 622	584 036	599 586	603 301	635 001	641 255
X_8	Ten thousand tons	100 340	111 577	107 021	135 035	151 772	130 280	138 119	15 354	16 356
X_9	Hundred million yuan	305	576	472	489	702	818	959	722	701

Then a new set of variables $Z = (Z_1, Z_2, \dots, Z_d)$ $d \leq n$ is calculated. The new variables of the set are independent of each other and can reflect the information in the original variable set X to the greatest extent. Its calculation formula [14] is:

$$\begin{cases} D(Y_i) = W_i^T \sum W_i \\ \text{cov}(Y_i, Y_k) = W_i^T \text{cov}(X, X) W_k = W_i^T \sum W_k. \end{cases} \quad (7)$$

$i, k \in [1, n]$

The principal component index is extracted from new variable Z as an input vector of the BP neural network.

4 Case analysis

4.1 Experimental environment

In this study, BP neural network model algorithm was written by MATLAB software, and the principal component analysis of the original information was carried out by SPSS software. The experiment was performed on a lab server with configurations of Windows 7 system, I7 processor, and 16 GB memory.

4.2 Experimental data

This study used the statistical data of cold chain logistics of fruits and vegetables in city A from 2010 to 2018.

As shown in Table 2, the products, such as fruits and vegetables, in the logistics transportation of commodities need maintain a low temperature environment in the logistics process because of their perishable characteristics. Therefore, it was necessary to consider the above characteristics together in the prediction of the cold chain logistics demand of fruits and vegetables products.

4.3 Experimental steps

Data preprocessing: it was seen from Table 1 that the units used in different influence factors were not the same. The

calculation result would inevitably become a large error if the original data was used directly. Thus, the “standardized method” in the SPSS software was used to deal with original data to eliminate dimensional differences between different factors.

Total variance analysis based on PCA: the total variance analysis was performed on the preprocessed original data by using the PCA model, and the main influence factors were selected according to the contribution rate of cumulative variance.

The prediction of the cold chain logistics demand of fruits and vegetables: the preprocessed statistics from 2010 to 2014 were taken as the training set, and the preprocessed statistics from 2015 to 2018 were taken as the testing set. The BP neural network model and PCA-BP neural network model were used to predict the cold chain logistics demand of fruits and vegetables.

The reasons for why the logistics cold chain data from 2010 to 2018 were selected as the training set and testing set are as follows. On the one hand, these data are the cold chain logistics data in recent years, so the training model can be more close to the actual situation in recent years. On the other hand, these data are known data (including influencing factor data and logistics demand data), which is convenient to use accurate results to adjust the internal parameters of the model during training and evaluate the accuracy of the model prediction results with accurate data during testing. Taking the data from 2010 to 2014 as the training set is to analyze the law of the past cold chain logistics data, while taking the data from 2015 to 2018 as the testing set is to use the trained model containing the law of logistics to predict and analyze the future data; as the data of 2015 ~ 2018 are known, the prediction results can be evaluated.

4.4 Experimental results

As shown in Table 3, the eigenvalue of influence factor X_1 was 7.745, and its rate of variance contribution was 85.72%; the eigenvalues of influence factors $X_2 \sim X_4$ were 0.752, 0.256, and 0.125 respectively, and their rates of variance contribution were 8.34, 2.82, and 1.42%

Table 3. PCA based total variance analysis.

Influencing factor	Eigenvalues	Variance contribution rate/%	Cumulative variance contribution rate/%
X_1	7.745	85.72	85.72
X_2	0.752	8.34	94.06
X_3	0.256	2.82	96.88
X_4	0.125	1.42	98.30
X_5	0.089	0.98	99.28
X_6	0.042	0.45	99.73
X_7	0.019	0.20	99.93
X_8	0.007	0.06	99.99
X_9	0.001	0.01	100

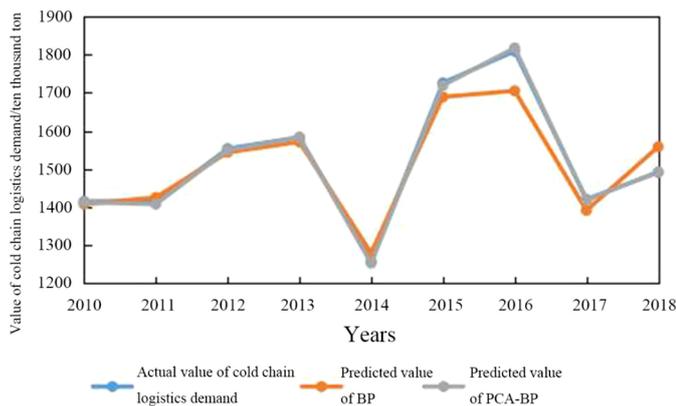


Fig. 3. Actual values and BP and PCA-BP neural network model prediction values of cold chain logistics demand of fruits and vegetables.

respectively; the eigenvalues of influence factors $X_5 \sim X_9$ were 0.089, 0.042, 0.019, 0.007, and 0.001 respectively, and their rate of variance contribution were 0.98, 0.45, 0.20, 0.06, and 0.01% respectively, among which the rates of variance contribution of influence factors $X_1 \sim X_4$ were above 1%, and the rate of cumulative variance contribution of the four influence factors was 98.30%, which was larger than 95%; therefore, the four influence factors $X_1 \sim X_4$ could be chosen to represent information for the entire sample and input into the PCA-BP neural network model.

As shown in Figure 3, the cold chain logistics demand of fruits and vegetables in city A showed a volatility growth between 2010 and 2018. The predicted value of the PCA-BP neural network model and the actual value of cold chain logistics basically coincided, while the predicted value of the BP neural network model after 2015 clearly deviated from the actual value. But the trend of the three curves was basically the same: in 2014, the actual cold chain logistics demand was the least, 12.57 million tons, the predicted value of the BP neural network model was 12.79 million

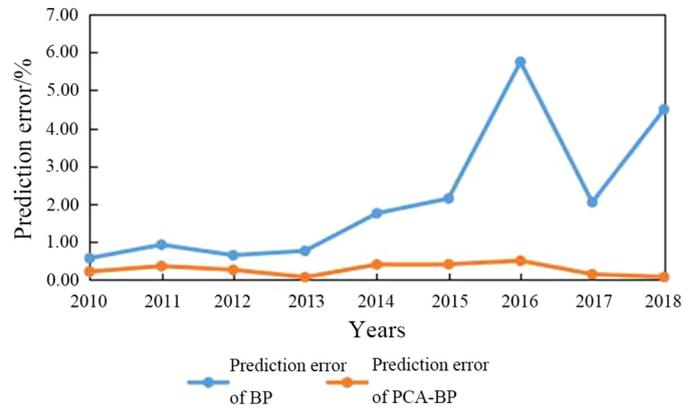


Fig. 4. Prediction errors of BP and PCA-BP neural network models.

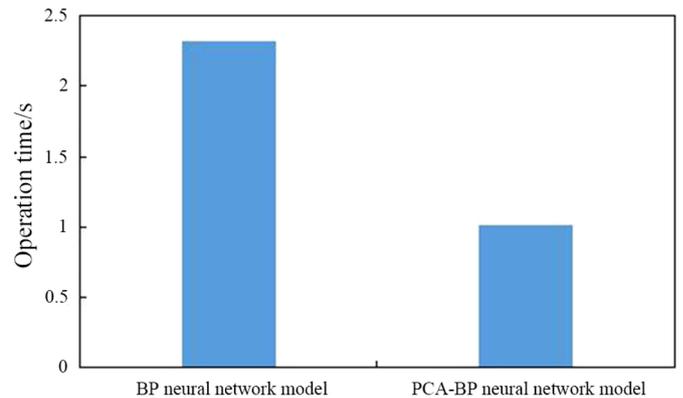


Fig. 5. Operation time of BP neural network model and PCA-BP neural network model.

tons, and the predicted value of the PCA-BP neural network model was 12.52 million tons; in 2016, the actual cold chain logistics demand was the largest, 18.09 million tons, the predicted value of the BP neural network model was 17.05 million tons, and the prediction of the PCA-BP neural network model was 18.18 million tons.

It was clearly seen from Figure 4 that the prediction error of the BP neural network model was not only higher than that of the PCA-BP neural network model, but also had a large fluctuation. From 2013, the prediction error of the BP neural network model increased significantly, and the prediction error fluctuated greatly between 2015 and 2018; the prediction error of the PCA-BP neural network model was relatively stable and remained at a low level. The prediction error of the BP neural network model was 5.75%, and the prediction error of the PCA-BP neural network model was 0.50%. It could be seen that the PCA-BP neural network model was more accurate than the BP neural network model in predicting demand for the cold chain logistics of fruits and vegetables.

As shown in Figure 5, the BP neural network model took 2.316s to predict the value of cold chain logistics demand of fruits and vegetables, and the PCA-BP neural

network model took 1.011 s. The computation time of the PCA-BP neural network model was significantly less than that of the BP neural network model. The reason was that the significantly reduced input vector of the whole neural network model after PCA processing reduced computation amount, which made the PCA-BP neural network model compute faster and more accurate than the BP neural network model although the PCA-BP model had PCA which BP did not have.

5 Discussion

In this study, the demand of urban cold chain logistics was predicted by BP neural network, and BP neural network was improved using PCA method. The prediction principle of BP neural network for cold chain logistics demand is to gradually approach the non-linear law between the influencing factors and logistics demand by adjusting the weight parameters using the internal non-linear function. Therefore, for the prediction model, the input influencing factors are critical. Different researchers will choose different influencing factors for different concerns. Based on the analysis, this study selected nine influencing factors from three aspects: macroeconomic level, logistics transportation information level and sustainable development level considering the objectivity, accessibility and importance of the influencing factors, as shown in Table 1. The macroeconomic level reflects the external environment for the development of cold chain logistics in urban areas. After all, both the consumption activities of logistics and the transportation of logistics need sufficient economic support. The logistics transportation and information level reflect the infrastructure conditions of logistics. For the logistics, the higher the level of transportation and communication, the higher the cost and speed of logistics. Sustainable development level reflects the development potential of cold chain logistics.

In this study, the nine influencing factors were screened using PCA method, from which the factors that had the greatest impact on logistics demand were selected. The cumulative variance contribution rate of the first four factors was 98.30%, i.e., the effective information provided by the first four factors accounted for 98.30% of the total sample, especially the first GDP growth rate accounted for 85.72% of the total effective information. The added value of primary industry, planting area of fruit and vegetable and consumer price index accounted for 8.34, 2.82 and 1.42% respectively, which showed that the most important thing for the local cold chain logistics was the local economic level.

The models were tested using the testing set and compared with the actual logistics demand. It was seen from the final error results that the prediction errors of the two models were at a relatively low level in 2010 ~ 2014, while the improved PCA-BP neural network remained at a low level in 2015 ~ 2018, but the error of BP neural network

had great fluctuations. The error of the improved PCA-BP neural network was smaller than that of BP neural network. Data of 2010 ~ 2014 were training set data, and the weight parameters of the models followed the standard of training set; therefore the error remained at a relatively low level. The data of 2015 ~ 2018 were testing set data, and the model predicted results through the laws obtained by the training set; therefore the error changed. BP neural network calculated the nine factors at the same time, which reduced the calculation efficiency, but also produced interference to the accuracy because of data containing few effective information. The improved PCA-BP neural network screened factors using PCA method, which not only reduced calculation amount, but also eliminated the interference caused by data containing few effective information.

6 Conclusion

This study briefly introduced the BP neural network model and the PCA method and combined them to obtain the improved PCA-BP neural network model. Then the traditional BP neural network model and the improved PCA-BP neural network model were used for the example analysis of the cold chain logistics demand of fruits and vegetables in city A from 2010 to 2018. The results are as follows. The main influence factors that affected the cold chain logistics demand were the growth rate of GDP, the added value of primary industry, the planting area of fruits and vegetables, and the price index of fruits and vegetables consumption. Both PCA-BP and BP neural network models had good predictive ability, and the predicted value trend was basically the same as the actual value; but the prediction error of the PCA-BP neural network model was smaller and more stable than the BP neural network model, and the prediction error of the BP neural network model fluctuated to increase year by year. The time required for the prediction by the PCA-BP neural network model was 1.011 s, and the time required for the BP neural network model was 2.316 s; thus, the prediction of the cold chain logistics demand by the PCA-BP neural network model was faster and more accurate.

References

1. J. Nie, *NeuroQuantology* **16**, 686–691 (2018)
2. J.X. Zhu, Z.Z. Liao, *Sensor. Lett.* **13**, 1112–1117 (2015)
3. Y. Li, S. Lu, in *2015 IEEE International Conference on Grey Systems and Intelligent Services (GSIS)*, 2015
4. C. Zhou, J.C. Tao, *Kybernetes* **44**, 707–710 (2015)
5. G.E. Xia, et al., in *International Conference on Systems & Informatics*, 2017
6. B. Ya, *J. Comput. Theor. Nanos.* **13**, 4019–4024 (2016)
7. J. Wang, et al., in *Control & Decision Conference*, 2015

8. S. Wang, et al., *Renew. Energy* **94**, 629–636 (2016)
9. W. Jia, et al., *Appl. Intell.* **43**, 176–191 (2015)
10. X. Li, et al., *Int. J. Bifurc. Chaos* **25**, 1540030 (2015)
11. J. Wang, et al., *Trans. Chin. Soc. Agric. Eng.* **33**, 116–122 (2017)
12. Z.M. Lv, Z. Wang, Z.Y. Wang, *Appl. Mech. Mater* **713-715**, 7 (2015)
13. Y. Zhang, et al., in *Power & Energy Engineering Conference*, 2016
14. A. Satija, J. Caers, *Adv. Water Resour.* **77**, 69–81 (2015)

Cite this article as: Ying Chen, Qiuming Wu, Lijin Shao, Urban cold-chain logistics demand predicting model based on improved neural network model, *Int. J. Metrol. Qual. Eng.* **11**, 5 (2020)