Volume 1, Issue 12, pp. 48-53, 2017.
International Journal of Scientific Engineering and Science
ISSN (Online): 2456-7361

A Population Prediction Model Based on Variable Weight Coefficients

Xiang Wan, Bing-Xiang Liu, Xing Xu*
School of Information Engineering, Jingdezhen Ceramic Institute, Jiangxi, China, 333000

Abstract— *To deal with the lack of accuracy and consistency in some single models, it took advantage of different algorithms to optimize the BP neural network respectively. Regarding predicting outcomes of basic BP model, GA -BP model and PSO-BP model as the input cells of new BP neural network's learning data, the combination predicting outcome can be got finally. The novel of this combination model lies in the collection of different predicting outcomes of single models together, and the input of neural network together with a combination forecast of the total population. Entering corresponding years simply, combination forecast outcomes can be got at once. It would have no impact because of inaccurate data and unknown corresponding factors. And weight coefficients of combination model change over time is very scientific. Taking advantage of this combination model to predict the total population of China from 2014 to 2020, the results show that total population of China in 2020 is within 1.4 billion. The trend of China's population growth has been effectively controlled by the government and it is of great possibility for the government to control the total population within 1.45 billion in 2020.*

Keywords— *Combination model; BP neural network; population of China.*

I. INTRODUCTION

Since the reform and opening up policy, the economic development speed of China isn't only rapid, but also sustained and steady, which has become a "miracle" of the world now. When why China can create such a great "miracle" of economic growth is discussed, lots of people insist that "demographic dividend" is a vital reason, therefore the "demographic dividend" is combined with the prospects of China's continuous economic growth together. The demographic problem is related to long-term development of the country, therefore the accurate and effective prediction has a very important guiding significance for policies of making and formulating development plans. Lots of researches have been done by scholars at home and abroad.

On abroad, mainstream population development models are Malthus model and Logistic model [1]. In recent years, there has been a new progress on population models. Lee and Tuljapurkar (1994) used ARIMA model to predict the birth rate in American [2]. Matysiak and Nowok (2007) used demographic probabilistic forecast method to predict the population of Poland from 2005 to 2050 [3]. Malthus model and Logistic model only apply to constantly increasing population. ARIMA model is a kind of steady-models, so it just decribes the statistics characteristic of steady array. However, demographic probabilistic forecast method needs some knowledge and experience in statistics background. The population of China will fluctuate according to demographic composition in the future, traditional models can hardly deal with it, which severely handicap their application.

Most researches on demographic projections at home have focused on using GM (1, 1) model and intelligent algorithms to fit data. Ke-Feng DUAN (2012) built a complex model base on logistic retarded growth model, then complex model got parameters by monadic linear regression method [4]. Zhi-Gang ZHOU et al. (2012) built a gray PGM (1, N) combined with chaos neural network model to predict the population of Hubei province during the Twelfth Five Year Plan period [5]. Qiu-Ying WANG et al. (2014) used the Song Jian equation model of discrete population development to predict the population of Heilongjiang province in the next 50 years [6]. These models can predict the population of China in a short-term, but as time goes on, errors of prediction will multiply. Because of restrictions on the availability of information or data, traditional models must make some approximates and assumption in corresponding factors. Errors of models are maldistribution and precision of models can't be effectively guaranteed.

In this paper, a demographic prediction model was built based on BP neural network. With the help of Genetic Algorithm (GA) and Particle Swam Optimization (PSO) algorithm to optimize the BP neural network's connection weights and thresholds, the networks got their best weights and thresholds to build GA-BP model and PSO-BP model respectively. Regarding these results of basic BP neural network model, GA -BP model and PSO-BP model as the input cells of new BP neural network's learning data, finally results modified by combination model were given. Author only took a quantitative analysis. Fitting process was conducted on total population and corresponding years by intelligent algorithms. With the help of small amounts of data, an accurate forecast of the short-term was made by combination model. It solved problems of traditional models restriction on the availability of data or scope.

II. BP NEURAL NETWORK MODEL

2.1 BP Neural Network Theory

BP neural network was first proposed by a group of scientists headed by Rumelhart and McCelland in 1986. It is a multilayer feedforward learning network and training according to error backward propagation. The strength of network's connections between neurals changes by the size of the weights and thresholds. The biggest advantage of BP neural network is that network can approximate any nonlinear function with arbitrary precision without to known these mathematical expressions exactly. With the help of training

http://ijses.com/

samples back propagation and weights, thresholds change to adjust the network, it can achieve the goal of network.

2.2 BP Neural Network Design

It was proved by Hecht-Nielsen that a three-layer neural network with enough nodes can generate arbitrary complex mapping to solve common problems. In this paper, relying on China National Bureau of Statistics' national total population data from 1978 to 2013, a three-layer BP neural network was built with a single hidden layer [7]. Data was classified from 1978 to 2008 for training, then data from 2009 to 2013 was taken as test data for network. Input layer neurons were corresponding years, the output was the total population, so input layer neurons number of the network was 1, the output layer node was 1.

Determining neural network's hidden layer is a very complex problem, it is the key to build a neural network successfully. According to kolmogorv theorem and experiences, the number of hidden layer nodes were tested one by one, finally network got the best number of hidden layer nodes was 5. Transfering function and training function were logsig-purelin-trainlm, training steps was 1000, the goal of training was 0.0001 and network learning rate was 0.01. A 1-5-1 BP neural network was been built. As seen in the figure 1, when the training step was 6, error was less than 1×10^{-4}, training was over.

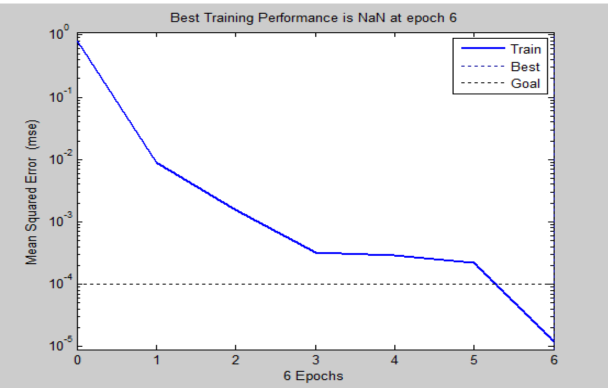

Fig. 1. Correlation of network between training steps and error curve.

2.3 Comprehensive Performances Index of Models

Precision of model is closely related to errors, in order to known relative merits between different models, some comprehensive performance index functions were used in this paper.

$$SSE=\sum_{t=1}^{N}(X_t-\hat{X}_t)^2 \tag{1}$$

$$\mathrm{MSE}=\sqrt{\sum_{t=1}^{N}(X_t-\hat{X}_t)^2}\Big/N \tag{2}$$

$$MAE=\sum_{t=1}^{N}\left|X_t-\hat{X}_t\right|\Big/N \tag{3}$$

$$\text{2-norm}=\sqrt{\sum_{t=1}^{N}(X_t-\hat{X}_t)^2} \tag{4}$$

2.4 Analysis of the Performances of BP Neural Network

The 2-norm of training data was 354.2, MSE= 11.4; the 2-norm of testing data was 887.1, MSE =177.4. Through contrast analysis, it was found that the network was short of stability and consistency, and the generalization ability was weak. So the BP neural network model is not perfect, it must be improved.

Weights and thresholds of BP neural network are random numbers initialized in [0.5, 0.5] interval, but weights and thresholds have a great influence on network training. Weights and thresholds of network without optimize by algorithms will drop into local extremum. These problems may lead to the inability of finding the global optimal point of network, it also causes a decline in network generalization ability, low precision, no practical value and so on [8].

III. The Improvement and Optimization of BP Neural Network Model

3.1 Genetic Algorithm to Optimize Weights and Thresholds of BP Neural Network

GA was first proposed by Bagley J.D in 1967 and it was a simple simulation of the evolution theory of Darwin and followed the rule of "natural selection, survival of the fittest". In order to search solutions of spaces, these solutions are encoded by GA. The algorithm generates a number of individuals called "chromosome" and chooses some chromosomes in a random way. Through generations of selection, crossover and mutation processes, it can get the best chromosomes and converges to global optimal solution eventually. The processes of genetic algorithm is shown in figure 2:

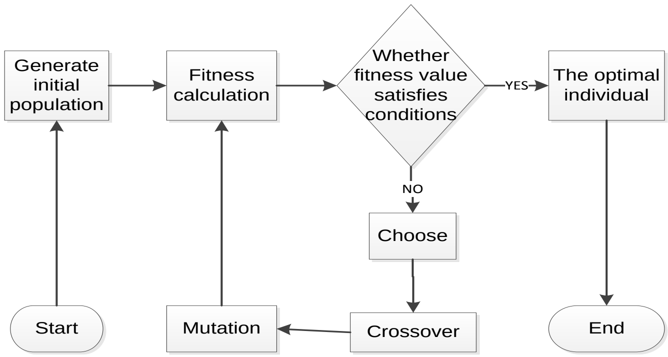

Fig. 2. Processes of genetic algorithm.

GATBX toolbox of Sheffield University in England was used to optimize weights and thresholds of BP neural network to get optimal weights and thresholds, and then these weights and thresholds were given to the network learning and training data. As seen in figure 3, when population evolution to the 60th generation, 2-norm of GA - BP model was very small, it reached predetermined accuracy standard.

International Journal of Scientific Engineering and Science
Volume 1, Issue 12, pp. 48-53, 2017. ISSN (Online): 2456-7361

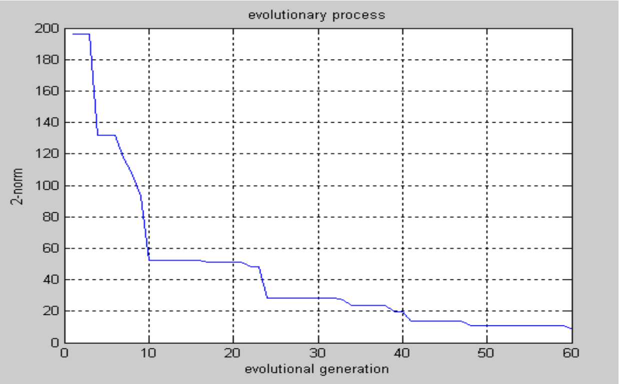

Fig. 3. The 2-norm of model.

There was an obvious improvement of GA-BP neural network in the precision of testing data. Using genetic algorithm to optimize the BP neural network has certain feasibility [9]. The 2-norm of testing data was 8.2529, MSE = 1.65, MAE = 3.06. It showed that GA- BP model was used to predict future population with a high accuracy. Unfortunately, errors of training data was too large, MSE = 51.07, MAE = 193.19. Obviously, GA -BP model was lack in stability, precision of training and consistent prediction and the generalization ability needed to be further improved.

For these problems, the main reason remains in GA-BP model itself. The original population is generated in a random way, it may produce some overfitness individuals. It causes "premature" problem and leads to algorithm drop into local extreme, the global optimal point can't be found finally [10].

3.2 Particle Swarm Optimization Algorithm to Optimize Weights and Thresholds of BP Neural Network

PSO algorithm was first proposed by James Kennedy and Russell Eberhart in 1995, the guiding thoughts of PSO was simulating the foraging behavior of birds.

PSO generates an initial population in a random way and iterating is used to search the optimal solutions. The optimal point of each particle is Pbest and the global optimal point of population is Gbest. In iteration processes, each particle updates its speed and position with the help of Gbest and Pbest parameters. As shown below:

$$V_{id}^{k+1} = wV_{id}^{k} + c_1 r_1 (P_{id}^{k} - X_{id}^{k}) + c_2 r_2 (P_{gd}^{k} - X_{id}^{k}) \tag{5}$$

$$X_{id}^{k+1} = X_{id}^{k} + V_{(k+1)id} \tag{6}$$

w is the inertia weight; d =1, 2, ……, D; i = 1, 2, ……, n; k is the current iteration times; V_{id} is the speed of particles; c_1 and c_2 are non-negative constants, called acceleration factors; r_1 and r_2 are random numbers in [0, 1] interval.

For the problem of inertia weight, analysis of Shi. Y indicates that a larger inertia weight is beneficial to global search, small inertia weight has a better local search ability. In order to balance the differences between global search ability and local search ability, a linear decreasing inertia weight was introduced into the PSO algorithm by Shi. Y [11]. The linear decreasing inertia weight formula in this paper is:

$$w(k) = w_{start} - (w_{start} - w_{end})(\frac{k}{T_{\max}})^2 \tag{7}$$

w_{start} is the initial inertia weight; w_{end} is the termination of evolution to maximum generation; k is the current evolution generation; $T_{\max}$ is the maximum generation. w_{start}=0.9, w_{end}=0.4. As seen in figure 4, when population evolution to the 300th generation, average fitness of PSO-BP model tends to be stable, error reaches predetermined accuracy standard.

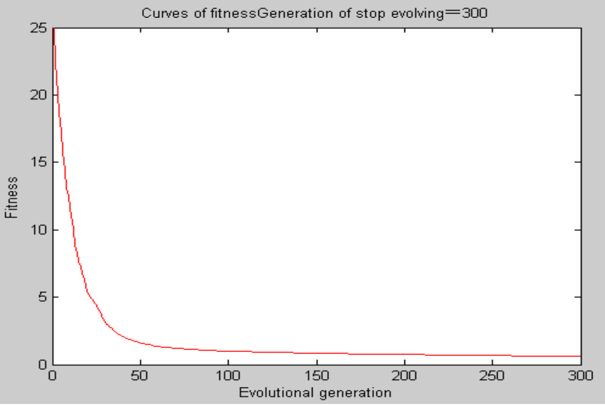

Fig. 4. The average fitness curves of population.

By calculating, there was obvious improvement of PSO-BP neural network than BP network in stability and consistency. It also had a higher precision than BP network. For testing data, the 2-norm of PSO-BP model was 141.7251, MSE=28.35, MAE=52.40. While the 2-norm of training data was 550.6115, MSE=17.76, MAE=70.81. Although stability and consistency of PSO-BP model was higher than BP network, prediction accuracy of training data was declining. This is a shortcoming of PSO-BP model.

3.3 Performances Comparison between GA-BP Neural Network and PSO-BP Neural Network

Compared with PSO algorithm and GA, PSO algorithm is easier for realization, it saves a lot of complex genetic encoding and crossover mutation processes. It also has memory ability and particles can search current situation and dynamically adjust the searching strategies in-time. Although PSO-BP model lacks in precision, the PSO-BP neural network has an obviously better improvement than GA-BP network in stability and consistency.

This paper found through comparison that no matter use what kind of global optimization algorithm to optimize weights and thresholds of BP neural network, there must be a certain probability leading BP neural network drops into local extremum. Is there a model that both has the features of stability and consistency of PSO-BP model and has a high prediction accuracy of GA-BP model? A combination model based on variable weighting coefficients, which combine the advantages of GA-BP model, PSO-BP model and BP neural network will be discussed in the next section, and it can make the combination model more reliable, stable and accurate.

IJSES

International Journal of Scientific Engineering and Science
Volume 1, Issue 12, pp. 48-53, 2017. ISSN (Online): 2456-7361

IV. A Combination of Prediction Model Based on Variable Weight Coefficients of BP Neural Network

4.1 The Theory and Overview of Combination Forecast

Since 1969, Bates. J. M. and Granger. C. W. J. put forward the combination forecast in the journal of *Operations Research* for the first time, the researches of combination forecast are in full swing at home and abroad [12]. The combination forecast is using information provide by different forecasting models together to build a combination model in the form of weighted average. Generally speaking, it can be expressed as a mathematical programming problem as follows:

$$\max(\min)\, f = f(l_1, l_2, \cdots, l_m)$$
$$S.T\begin{cases} \sum_{i=1}^{m} l_i = 1 \\ l_i \geq 0, i = 1, 2, \cdots, m \end{cases} \tag{8}$$

$f(l_1,l_2,\cdots,l_m)$ is the objective function, $l_1,l_2,\cdots,l_m$ are the weighted coefficients of different methods.

For the same set of data, all kinds of useful information are explained by these models from different angles. Although methods and principles of models are different, models need not be mutually exclusive and in some case should even be mutually reinforcing. Essential of combination forecast is integrating information from single models in order to improve the stability, accuracy and reliability of combination model [13]. Theories and practices show that the combination model has a higher stability than single models, prediction accuracy is greatly increased. Combination model greatly expands the application fields and takes more agility.

4.2 Population Prediction Model Based on Variable Weight Coefficients

4.2.1 Constant Weight Coefficients Model and Variable Weight Coefficients Model

According to whether weight coefficients of combination model is a function of time series, the combination model can be divided into constant weight coefficients model and variable weight coefficients model. Most of combination models have been constant weight coefficients models so far, they get weight coefficients of each single model through linear programming or other traditional methods [14-16]. Unfortunately, in practical issues, some possible problems are far more complex than constant weight coefficients model to be dealt with, then ideal results can hardly be got on the basis of the existing methods. Weight coefficients of combination model which don't change with time are too idealistic. Dealing with complex problems, constant weight coefficients model is much more unstability and lower precision than any single models.

4.2.2 Population Prediction Model Based on Variable Weight Coefficients of BP Neural Network

Variable weight coefficients model refers to weight coefficients of combination model change with time. Variable weight combination model is much more complex, the most difficulty problem is getting weight coefficients of single models at different points of time. In general, the combination model is built by evaluation functions, which are constructed by using the minimum sum of squares of errors. The best weight coefficients of each models is figured out in combination model one by one according to this method [17]. It is a cumbersome processes and must be combined with constraint conditions. Due to the rigorous constraint conditions, application of the combination model is greatly restricted.

In this paper, with the help of BP neural network can approximate any nonlinear function with arbitrary precision to build "Population Prediction Model Based on Variable Weight Coefficients". It can collect different total population prediction outcomes of GA-BP model, PSO-BP model, BP model and corresponding years together, regarding them as the input of neural network, the output is the actual total population of corresponding years. Weight coefficients of GA-BP model, PSO-BP model and BP model can be got by network with the help of learning and training.

4.2.3 Analysis on the Performances of the Population Prediction Model Based on Variable Weight Coefficients of BP Neural Network

As seen in figure 5, the comparison between the simulative results of different models and the actual population of China from 2009 to 2013. It can be seen from the picture that simulative results of combination model have the advantage of high goodness of fit and satisfactory effect. Although simulative results of GA-BP model also achieved high conformity with actual population, there was still have some volatility. There is obvious improvement of combination model than GA-BP model in stability, it can be reflected in comprehensive performance index throughout the rest of this paper.

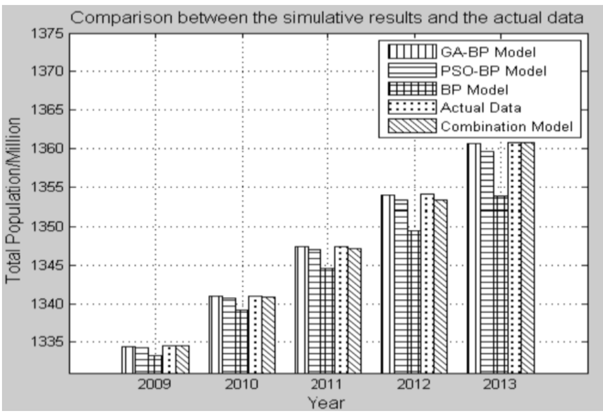

Fig. 5. Comparison between simulative results and actual data.

As can be seen in Table I, there is a greater improvement of variable weight coefficients combination model than single models in precision, stability and consistency. The combination model not only has the advantage of PSO-BP model in stability and consistency, but also has the high precision of GA-BP model in prediction. It has a stronger generalization ability than traditional BP neural network. "Population Prediction Model Based on Variable Weight Coefficients" is stable and reliable.

TABLE I. Comprehensive performance index of models.

Performance \ Model	GA-BP Model	PSO-BP Model	BP Model	Combination Model
Total SSE	2506701	323259	916019	40161
Total MSE	43.98	15.79	26.59	5.57
Total MAE	166.79	68.25	93.81	18.42
Total 2-norm	1583	569	957	200

4.3 Weight Coefficients of Combination Model

In this paper, weight coefficients of combination model can be got with the help of learning and training by network, then it can make a combination forecast of the total population. Putting minimum of total prediction absolute error as the object function of the linear programming, it can reverse to resolve variable weight coefficients of single models at different points of time in the combination model. It can avoid not only complex convex quadratic programming problems, but also negative weight coefficients with no sense, this method greatly simplifies the processes of solving the weight coefficients. The solving processes are as follows:

For the same object, the prediction results of GA-BP are A1, A2, A3……, weight coefficients of combination model are w_{A1}, w_{A2}, w_{A3}……; the prediction results of PSO-BP model are B1, B2, B3……, weight coefficients of combination model are w_{B1}, w_{B2}, w_{B3}……; the prediction results of BP model are C1, C2, C3……, weight coefficients of combination model are w_{C1}, w_{C2}, w_{C3}……; the prediction results of "Combination Model Base on Variable Weight Coefficients of Population Prediction" are T1, T2, T3……; The formula is as follows:

$$\min f=\sum_{n=1}^{3}\left|w_{Ai}(Ti-Ai)+w_{Bi}(Ti-Bi)+w_{Ci}(Ti-Ci)\right|$$
$$S.T\begin{cases}w_{Ai}\times Ai+w_{Bi}\times Bi+w_{Ci}\times Ci=Ti\\ w_{Ai}+w_{Bi}+w_{Ci}=1\\ w_{Ai},w_{Bi},w_{Ci}\in(0,1),i=1,2,3\cdots\end{cases}\tag{9}$$

V. TOTAL POPULATION PREDICTED BY COMBINATION MODEL BASED ON VARIABLE WEIGHT COEFFICIENTS

Based on quantitative analysis method, combination model was used to predict the total population of China. For qualitative analysis, demographic structure wasn't considered in the paper yet. Due to combination model only suitable for short-term population prediction, the total population from 2014 to 2020 can be predicted. As seen in table II, weight coefficients of each model.

TABLE II. Total population predicted by combination model from 2014 to 2020 and weight coefficients of each models.

Year	Combination Model/Million	W_{Ai}	W_{Bi}	W_{Ci}
2014	1364.09	0.48895343	0.30155780	0.20948877
2015	1369.12	0.44670186	0.31186939	0.24142875
2016	1373.78	0.40172668	0.32969497	0.26857835
2017	1378.01	0.36708184	0.33213017	0.30078799
2018	1381.83	0.33320926	0.33334050	0.33345024
2019	1385.19	0.30217030	0.33474757	0.36308213
2020	1388.13	0.27184897	0.33608201	0.39206901

VI. CONCLUSION

Through analyzing and comparing performances of different models, the results showed that combination model based on BP neural network was more simple and accurate. "Population Prediction Model Based on Variable Weight Coefficients", which contains three kinds of useful information and advantages of three different models, in addition, it also can overcome special defects of single model. Weight coefficients of combination model, which is a function of time series, robustness and accuracy, is greatly better improved than single model and it is scientific, reasonable and reliable for short-term prediction. As it is known, combination model can get weight coefficients of single models at different points of time more effectively and reasonablely, it can also avoid negative weight coefficients and unreasonable weight coefficients greater than 1.

There are large numbers of population growth factors to deal with, the combination model only takes a quantitative analysis, it is just a fitting model has the relationship between total population and time. With the help of small amounts of population data, the total population of China can be got by combination model. For qualitative analysis, such as the birth rate, mortality, national policy and other factors weren't considered in the paper yet. It is very suitable for short-term population prediction model without data of corresponding factors. The combination model has a higher precision and stability than single model or simple data fitting. The thinking method of this model can be used into other unstability data that single model can hardly deal with. However, the advantages of combination model based on variable weight coefficients can be more obviously than growing data.

Although some beneficial exploration of population model had been made in this paper, there are still some questions that need future research. For example, the combination model only took a quantitative analysis, corresponding factors weren't considered. Prospected applications of the subject and proposed theoretical problems in this paper, showed the future research direction.

ACKNOWLEDGMENTS

This paper was supported by the Science Foundation of Jiangxi Provincial Depart- ment of Education, the National Natural Science Foundation of China (No.61702239).

REFERENCES

[1] T. Malthus, An Essay on the Principle of Population, Cambridge: Cambridge University Press, 1992.
[2] R. Lee and S. Tuljapurkar, “Stochastic population forecasts for the United States: Beyond high, medium and low,” *Journal of the American Statistical Association*, vol. 89, issue 428, pp. 1175-1189, 1994.
[3] A. Matysiak and B. Nowok, “Stochastic forecast of the population of Poland, 2005-2050,” *Demographic Research*, vol. 17, pp. 301-338, 2007.
[4] Ke-Feng DUAN, “A population prediction model of China based on complex models,” *Statistics & Decision*, vol. 20, pp. 30-32, 2012.
[5] Zhi-Gang ZHOU et al., “Gray-Chaos neural network model and its application to short-term population prediction,” *Systems Engineering*, vol. 30, issue 10, pp. 118-122, 2012.
[6] Qiu-Ying WANG et al., “A model based on Song Jian discrete population evolution equations applied in population forecast and analysis of Heilongjiang province,” *Mathematics in Practice and*

Volume 1, Issue 12, pp. 48-53, 2017. *ISSN (Online): 2456-7361*

Theory, vol. 44, issue 10, pp. 203-207, 2014.

[7] China National Bureau of Statistics, China Statistics Yearbook [J]. Beijing: China Statistics Press, 1979-2014.

[8] Lei Yu, Feng Shi, Hui Wang, et al.. *Analysis 30 Examples of MATLAB Intelligent Algorithm*, Beijing: Beijing University of Aeronautics and Astronautics Press, 2015.

[9] Chun-Yan LIU et al., "Performance comparison between GA- BP Neural network and BP neural network," *Chinese Journal of Health Statistics*, vol. 30, issue 2, pp. 173-176, 181, 2013.

[10] Shi-Hong Chen and Ying-Ming Zhang, "Research of population prediction based on GA-BP neural network," *Journal of Strategy and Decision-Making*, vol. 3, issue 4, 48, 2012.

[11] Shi Y, Eberhart R. "Fuzzy adaptive particle swarm optimization," *Proc. Congress on Evolutionary Computatio*n, Seoul, Korea, pp. 101-106, 2001.

[12] Bate. J. M, Granger. C. W. J. "The combination of forecasts," Operations Research Quarterly, vol. 20, issue 4, pp. 451-468, 1969.

[13] Hua-Juan Dai, Combination Prediction Model and Its Application Research, 2007.

[14] Shao-Feng TAN, "A optimal combination prediction model of dam based on GM (1,1)-ARIMA," *Bulletin of Surveying and Mapping*, (S1), pp. 66-69, 2014.

[15] Hui-Zi SUN et al., "Drought period stream-flow forecasting based on optimal weighted combination model," *Journal of Northwest A & F University (Natural Science Edition)*, vol. 39, issue 11, pp. 201-208, 2011.

[16] Yang Song et al., "An improved optimal combinative model for population forecast and its application in land use planning," *Resources and Environment in the Yangtze Basin*, vol. 17, issue 3, pp. 346-350, 2008.

[17] Yong-Gang WAN, Hong-Yun ZHENG, "Prediction of air transportation incidents based on combined model with optimal variable weights," *China Safety Science Journal*, vol. 23, issue 4, pp. 26-31, 2013.