# Economic Forecast of the Southern China on BP Neural Network--- Taking Chongqing as an Example

Du Wenke[1*], Ge Jing[2], Sun Shidong[3]

[1] *Renmin University of China, Beijing, 100872*
[2] *Insititute of Economics, Zhejiang University of Technology, Hangzhou, Zhejiang, 310014*
[3] *School of Public Finance & Taxation, Northeast University of Finance and Economics, Dalian, Liaoning, 116000*
*\*3417617790@qq.com*

**ABSTRACT**

Based on the good economic development trends in Chongqing in recent years, the large changes in GDP increments of the subordinate districts and counties, and the gradual attention paid to the characteristics of regional economic regions. This article used the idea of sampling to establish comprehensive economic development indicators. The back propagation neural network model of comprehensive economic development can accurately predict the overall economic development of Chongqing and southern regions in the next five years, and explore its correlation with the overall economic development of the country. Through the application of principal component analysis and dimensionality reduction time series analysis, it fits the relationship between the comprehensive indicator data of Chongqing area from 2000 to 2019 obtained from the National Bureau of Statistics, financial reports and annual reports so as to carry out model training, analysis and prediction. The conclusion from this passage has strong universality in the current situation of regional economic differentiation and commonality. After the promotion and establishment of a mechanism to obtain a regional economic overview by sampling representative cities in various regions, it will greatly promote the formulation and modification of macroeconomic policies.

*Keywords: Economic Forecasting, BP Neural Network, Principal Component Analysis, Regional Economic Differentiation*

## 1. INTRODUCTION

In the understanding of traditional economic society, economic development is mainly reflected in three major indicators: employment rate, inflation rate, and economic growth rate. The supply-side structural reforms emphasized by General Secretary Xi Jinping have also repeatedly demonstrated that the three carriages that stimulate domestic economic demand such as investment, consumption and net exports have important reference value for regional economic development forecasts. With the popularization of the concept of the five major economic theaters, regional differentiation and inter-regional cooperation mechanisms will also have a great synergy effect on the national economy. Predicting the development status of the regional economy in the next 5 years is conducive to the establishment of a comprehensive national economic development system. By comparing the different factors that affect the economic development of the southern theater and the northern theater, the differences in geography and

humanities can be better combined and make a more accurate forecast of economic development in the changing era.

In the relevant forecasts of existing government agencies, the volatility index of the index is often within 5 percentage points up and down, resulting in very obvious characteristics of local inaccuracy. At present, most relevant literature studies are aimed at micro-regions, such as Beijing and micro-economic objects. For example, research on the sale of eggs, although the data collection is convenient and the difference is small, it is also not scalable.

In the existing literature, BP neural network has been criticized repeatedly because of its poor generalization ability. The reasons behind it are the incorrect extension of incomplete induction, which causes the limitation of the local optimal solution to be enlarged, and the second is the error data and the impact of missing data. This article intends to solve this problem by increasing the data dimension through data collection in regional statistical

yearbooks and other aspects and comparisons between regions.

The research in this article is based on historical data and adopts the method of sampling observation, represented by the emerging city of Chongqing in the southern region in recent years, which greatly reduces the workload. The four major opportunities that Chongqing faces in terms of economic development also represent the normal state of many regions in the south, including Guangxi, Xi'an and Yunnan, which are actively responding to external economic stimulation in the new era with Chinese characteristics. Research on economic activities can draw out the "obstacle indicators" and "excellent indicators" of Chongqing's economic resource carrying capacity in the future social and economic development, which are also of reference value for northern and national cities.

## 2. DATA COLLECTION AND PROCESSING

### 2.1. Key economic indicators and data interpretation

First introduce three indicators: economic growth rate, inflation rate (price index), and unemployment rate. RGDP is the comparison between the final GDP and the base GDP. The final GNP is calculated at the current price at the final period, and the resulting growth rate is the nominal economic growth rate. The final GNP calculated at the constant price is the actual growth rate. Economic growth rate. Inflation rate refers to the rate of increase in the general price level in a certain period of time. Reflect the degree of inflation. It is usually expressed by the increase in the price index and the decrease in the purchasing power of the currency. The unemployment rate refers to the number of unemployed laborers among the employed population who meet all the employment conditions in a certain period of time. It aims to measure the idle labor capacity and is the main indicator reflecting the unemployment status of a country or region.

Investment in fixed assets is expressed in currency, the amount of work an enterprise has to construct and purchase fixed assets in a certain period of time, and the changes in related expenses. Including real estate, buildings, machinery, machinery, transportation tools, and enterprises used for capital construction, renovation, overhaul and other fixed asset investments.

Per capita consumption expenditure refers to all expenditures used by residents to meet the daily consumption of households, including expenditures on purchases of goods and services. Per capita consumption expenditure is the main body of social consumption demand, a direct factor driving economic growth, and an important indicator that reflects the living standards and quality of residents. The current statistics are divided into two indicators: per capita consumption expenditure of urban residents and per capita consumption expenditure

of rural residents. The per capita consumption expenditure of urban residents is calculated monthly, and the per capita consumption expenditure of rural residents is calculated quarterly.

Total import and export refer to the total amount of goods actually entering and leaving my country's borders. Total imports and exports are used to observe the total scale of a country's foreign trade. China stipulates that export goods are counted at FOB prices, and imported goods are counted at CIF prices.

### 2.2. Determination of data indicators

This article invests in data on economic indicators over the past 20 years, including 8 major indicators such as fixed asset investment, credit scale, currency supply, gross national product, savings, exchange rate, total import and export, and interest rates. With the aid of its permanent population and other data with social and human factors, the relationship between it and the next year's inflation rate, unemployment rate and economic growth rate is trained through BP neural network fitting to complete the whole process.

Under the general framework, this article will also subdivide the added value of industries into the primary, secondary, and tertiary industries, and observe the impact of high-tech manufacturing, commodity services, and even the subdivision of the real estate industry on Chongqing's economy. The ratio of each parameter in the neural network is completed.

After obtaining the development trend of Chongqing's economic segmentation field and the development model of Chongqing's overall economy, we cross-validate it, and use stochastic gradient descent to iteratively optimize the data, and finally get the BP neural network for 2020-2025. Based on the prediction results of the Chongqing Economic Index, we finally tried to use the model to predict the data of Beijing in 2018 and 2019, judge the generalization ability of the training model in Chongqing, and draw the conclusion of regional economic differentiation. The model is used to predict the national economic data in 2018 and 2019, and to determine the correlation between Chongqing as a key economic region in the South and national economic development. Through the joint comparison of multiple regions, the Mutually Exclusive Collectively Exhaustive rule is used to establish an index map of China's economic development model as much as possible.

In the specific selection of data used, the first step is to continue the principle of principal component analysis and perform factor analysis to reduce the complexity of training. By testing the correlation between variables and economic results, the conclusion is drawn: After Chongqing's GDP, total imports and exports, investment in fixed assets, and savings in the coefficients in previous years, the follow-up indicators are added, and the impact

on economic growth has not changed much. This does not mean that subsequent indicators, including interest rates and exchange rates, are not closely related to economic growth. However, through subsequent experiments, it is found that the accuracy of the training model using 8 indicators and the accuracy of training the model using 4 indicators is only 0.00001. To shorten the iteration time and propose a more concise prediction model, the preprocessing stage chooses to remove the following four Indicators and only the first four indicators are considered. In the long-term economic cycle, saving is equivalent to consumption, and there are fewer missing data on consumption and a large number of data samples.

$$\begin{cases} X_t = \emptyset_1 + \emptyset_1 X_{t-1} + \cdots + \emptyset_p X_{t-p} + \varepsilon_{t-1} - \theta_1 \varepsilon_{t-1} - \cdots \theta_q \varepsilon_{t-q} (\varphi_p \neq 0, \theta_q \neq 0) \\ E(\varepsilon_t) = 0 \qquad Var(\varepsilon_t) = \sigma_\varepsilon^2 \qquad E(\varepsilon_1 \varepsilon_t) = 0 \qquad (s \neq t) \\ E(\varepsilon_t X_s) = 0 \qquad (\forall s < t) \end{cases}$$

Where $X_t$- random process $(\varepsilon_t)$-white noise process. In particular, at that time $\varphi_0 = 0$, it was called the centralized ARMA(p,q) model, and the delay operator b was introduced, which was recorded as follows.

$$\begin{cases} \varphi(b)X_t = \theta(b)\varepsilon_t \\ \varphi(b) = 1 - \varphi_1 b - \varphi_2 b^2 - \cdots - \varphi_2 b^p \\ \varphi(b) = 1 - \theta_1 b - \theta_2 b^2 - \cdots - \theta_q b^q \end{cases}$$

Where $\varphi(b)$ - p-order autoregressive coefficient polynomial. $\theta(b)$ -q-order moving average coefficient polynomial. Therefore, when establishing a model for a stationary time series, you can first fit an ARMA(p,q) model, then find a more reasonable model based on its parameter values $\phi_1, \phi_2, \ldots, \phi_p$, $\theta_1, \theta_2, \theta_3, \ldots \theta_q$ and whether it is significantly 0, and then fit the corresponding model to test its adaptability.

Need to order the model, that is, determine the values of p and q. The autocorrelation function and partial autocorrelation function of stationary series generally have relatively standardized statistical characteristics. For the simple autoregressive ARMA model and the moving average MA model, p and q can be judged according to the tailing of the partial correlation function PACF and the autocorrelation function ACF. As for the ACF and PACF of the ARMA model, both have tailing properties, so the statistical characteristics of ACF and PACF cannot be used to determine the order of the ARMA model. AKAIKE [20] proposed a joint ranking of AIC and FPE criteria. Different p and q are selected to fit the time series, and the corresponding AIC value and FPE value are calculated, and make the model with the minimum value as the best model. After selecting the best p and q, the model residuals are calculated and used for white noise testing to check whether the model is suitable. In addition, it is necessary to normalize the data in order to judge the error function equivalently through any one

Therefore, in this experiment, per capita consumption expenditure is used to replace the index of savings. The final index is regional production value, fixed asset investment, total import and export, and per capita consumption expenditure.

### 2.3. Data filling based on the ARMA model

For the missing data in historical data, 20 items of per capita disposable income in the Jiangbei area are missing 6 items, corresponding to a missing 30%, which is filled by the autoregressive moving average model. The key formulas of the ARMA model are as follows.

of Pearson's correlation coefficient, Cosine similarity, and the square of Euclidean distance to improve the accuracy of the error function value.
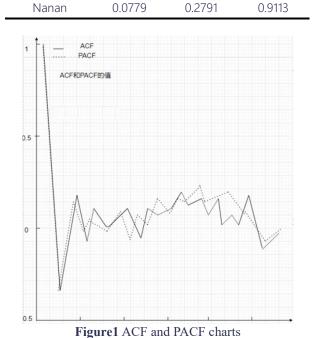
### 2.4. Data Stationarity Test

As shown in Figure 1, the logarithm of the changes in the main economic indicators of Chongqing is taken and differential processing is performed to test the stationarity. After the ADF test, the sequence after the difference is a white noise sequence, which is a stationary sequence. As shown in Figure1, through the 20-order calculation of the autocorrelation function ACF and the partial correlation function PACF values, there is no obvious truncation, that is, tailing, so the ARMA model can be used. The establishment of the ARMA model and the testing of the fitting accuracy of the ARMA model in all regions. The errors of training the ARMA (1, 5) model using the changes in the main economic indicators of each city are shown in the table 1. In the table 1, MSE, RMSE and R2 are mean square error, root mean square error and coefficient of determination, respectively. For the changes in the main economic indicators, the ARMA (1,5) model predicts more accurately, the mean square error MSE is below 0.15, and the coefficient of determination of goodness of fit R2 is above 0.88. It shows that the ARMA (1,5) model predicts the economic growth rate, inflation rate, and unemployment rate of Chongqing area more accurately, and can be used to fill in missing data.

**Table1** ARMA model error table

| Distinct | MSE | RMSE | R2 |
|---|---|---|---|
| Jiangbei | 0.0541 | 0.2326 | 0.9269 |
| Jiulongpo | 0.0821 | 0.2865 | 0.8807 |
| Shapingba | 0.077 | 0.2775 | 0.9041 |
| Wanzhou | 0.1106 | 0.3326 | 0.9312 |
| Yubei | 0.1202 | 0.3467 | 0.9119 |
| Fulin | 0.0419 | 0.2047 | 0.9534 |

| Nanan | 0.0779 | 0.2791 | 0.9113 |
|---|---|---|---|



**Figure1** ACF and PACF charts

## 3. Model selection and establishment

### 3.1. The introduction of BP neural network

The relationship between economic indicators and comprehensive economic conditions shows nonlinear characteristics. Considering the complexity of ensemble learning, logistic regression, polynomial regression and other fitting methods cannot achieve the desired results. Therefore, the BP neural network is used to fit the set of data. BP neural network is suitable for system problems of high-order, nonlinearity, multiple feedback and time-varying characteristics, and is one of the most widely used neural network models.

The BP neural network model is composed of four parts: input and output model, action function model, error calculation model and self-learning model.

### 3.2. Model solution and prediction

After model optimization, the BP neural network model with a 1-9-11 layer structure is selected. The number of nodes in the input layer is 1, corresponding to important economic indicators of input; the number of nodes in the output layer is 1, corresponding to the comprehensive situation of the output economy; the number of neurons in the hidden layer is 9. The hidden layer transfer function selects the LOGSIG type function. The transfer function type of the output layer is the PURELIN linear function, the softmax function is used as the activation function, the training algorithm uses the LM algorithm, and the adaptive learning function uses the gradient descent backpropagation learning rule (LEARNED). Taking the data of Jiangbei District as an example, the BP neural network has been trained through 240 iterations, and the SSE is only 0.0475, and the error value is maintained at a relatively stable level. Through the idea of cross-validation, the ratio of the training set to the validation set is changed, and the most suitable training and simulation ratios are finally determined by comparing the prediction error functions of 7:3, 9:1, and 5:5.

Using the data of 26 districts with a time span of 20 years for training, the error is shown in the figure. Except for a few cities, where the error is obvious, the overall situation of fixed asset investment, total import and export volume, regional economic total, per capita consumption expenditure and the correlation between unemployment rate, inflation rate, and economic growth rate in most other cities can be determined by nerves. The network fits well. And the mean square error MSE is below 0.12, the root mean square error is below 0.3464, and the coefficient of goodness of fit determination R2 is above 0.80. It shows that the BP neural network has a very significant effect on this nonlinear correlation fitting. The evaluation function of the model after training is shown in the figure, the predicted point and the real point are basically close to the 1:1 curve, and R² reaches 0.9562, indicating that the fitting is more accurate. The final prediction result is partially intercepted as shown in the table below.

**Table 2** Future comprehensive economic situation of Chongqing

| year | Inflation rate | Unemployment rate | GDP growth rate |
|---|---|---|---|
| 2020 | 5.5% | 3.5% | 10.7% |
| 2021 | 5.4% | 3.4% | 11.5% |
| 2022 | 5.3% | 3.3% | 10.9% |
| 2023 | 5.2% | 3.35% | 11.4% |
| 2024 | 5.2% | 3.28% | 12.5% |

**Table 3** Comprehensive future economic situation in the southern region

| year | Inflation | Unemployme | GDP | | rate | nt rate | growth rate |
|---|---|---|---|---|---|---|---|
| | | | | 2020 | 5.4% | 3.6% | 12.7% |
| | | | | 2021 | 5.35% | 3.8% | 13.5% |

| 2022 | 5.38% | 3.3% | 12.9% |
| 2023 | 5.29% | 3.1% | 12.6% |
| 2024 | 5.25% | 3.0% | 11.9% |

### 3.3. Extensive discussion of model results

Through the training of some data, it is found that when the total fixed investment increases by 40%, the total economic volume will not increase by 40%, but calculated according to the ratio of fixed investment, total import and export, per capita investment and total economic volume of 2:3:4:5. After the comprehensive calculation index increases by 40%, the economic aggregate increases by 40% and becomes a high probability event. However, when a certain index grows too fast in the short term, the correlation will be invalid, and it is effective in the stable economic development stage.

Bringing the training model of Chongqing area into the economic data prediction of Beijing area, it is found that the correlation coefficient between the two is 0.6956, which does not have a very strong correlation, but they can influence each other. After bringing into the national economic situation data prediction, it is found that the two are related. The coefficient of sex is equal to 0.7983, which is greater than the correlation with the Beijing area, but less than the correlation within the Chongqing area. From this, it can be explained that the differences in regional economy have caused the southern and northern regions to formulate different fiscal and monetary policies in order to achieve the same level of economic growth. In Chongqing, the economic industries that are supported by the real estate industry and the emerging cultural and creative industries are very different from the prevailing economic status of Beijing's established industries and residents' consumption propensities are also different. At the same time, there is also a gap in economic status, and Chongqing residents are more inclined Consumers of entertainment products, and compared with Beijing's high housing prices, the enthusiasm for real estate investment in Chongqing is higher.

### 4. CONCLUSIONS

This paper selects a total of 569 economic indicator data and corresponding 478 economic growth data from 26 districts and counties in Chongqing from 2000 to 2019 to explore the correlation between the main indicators of economic changes and comprehensive economic strength, and uses BP neural network for fitting. After model optimization, the BP neural network model with a 1-9-11 layer structure is selected. After training, the mean square error MSE is below 0.12. Through data fitting, it can be found that the relationship between the main economic indicators and the comprehensive economic strength has a relatively general law under the stable state of the

economy. Through the analysis of economic data in Chongqing, the interaction mechanism between Chongqing and Beijing is determined, and the corresponding national policies are formulated through the highly overlapped parts, and the national economic policy recommendations are completed through time series analysis. The establishment of the regional economic influence mechanism, through the analysis of the economic data of Chongqing area to determine the mechanism of interaction with the Beijing area, and formulate corresponding national policies through the highly overlapping parts of the two, and complete the country through the advantages of time series analysis Economic policy advice project.

This paper proposes to combine the characteristics of regional differentiation and apply the model results of BP neural network more widely to fully improve the ability to deal with complex factors (such as geographical environment, humanistic choices, etc.) and predict economic phenomena when the local economy is variable. By comparing regional commonalities and differences, when studying the factors that affect the economy by market segmentation, the principal component analysis under exploratory factor analysis is used to simplify the BP neural network model. The combined application of the two quantitative models is an innovative attempt in scientific research. However, due to the complexity of regional economic factors, how to collect important influencing factors in the first time and eliminate interference factors or too small factors will be the next optimization goal of this article.

### REFERENCES

[1] Amir Abbas Shojaie, Ali Dolatshahi Zand, Shahrzad Vafaie. Calculating production by using short term demand forecasting models: a case study of fuel supply system[J]. Evolving Systems,2017,8(4).

[2] Hualing Shi. Application of Back Propagation (BP) Neural Network in Marine Regional Economic Forecast[J]. Journal of Coastal Research, 2019, 98(sp1).

[3] Xiaofang Luo. Construction of artificial neural network economic forecasting model based on the consideration of state transition diagram[J]. Neural Computing and Applications,2019,31(12).