



# Analysis of Guangdong-Hong Kong-Macao Greater Bay Area's Economic Growth Trend Based on Big Data Mining

Chao-ping Ma<sup>1</sup>(✉) and Xiao-yun Lin<sup>2</sup>

<sup>1</sup> Department of Economics and Trade, Guangzhou College of Technology and Business, Guangzhou 510850, China  
machaoping369@163.com

<sup>2</sup> Library, Guangzhou College of Technology and Business, Guangzhou 510850, China

**Abstract.** For the sake of improving the optimal management and dispatching ability of Guangdong-Hong Kong-Macao Greater Bay Area's economy, it is essential to optimize and predict the growth trend of the Greater Bay Area's economy, put forward the optimization prediction method of the Greater Bay Area economic growth trend based on 3500 mining, and construct the economic growth model of statistical sequence distribution. Big data mining method is chosen to model the big data statistical information of the area's economic growth, extract the characteristic quantity of the association rules of the big data economic growth trend, use the fuzzy fusion clustering method to carry on the automatic clustering processing to the economic growth trend, and establish the optimal iterative model of the prediction of the economic growth trend. Combined with adaptive optimization algorithm, the Greater Bay Area's economic growth trend is optimized and predicted. The simulation outputs show that the method has good adaptability to predict economic growth trend of the area we talked about, and has high accuracy in predicting growth trend, which improves the adaptive scheduling and management ability of the economy in the bay area.

**Keywords:** Guangdong-Hong Kong-Macao Greater Bay Area · Economic growth · Big data mining · Trend analysis

## 1 Introduction

It is of great significance to study the economic growth trend of the bay area of Guangdong, Hong Kong and Macao [1, 2]. The prediction algorithm of the economic growth trend and the prediction algorithm based on the Markov model are proposed. However, the fuzzy degree of the method is large, and the adaptability of the prediction process is not good enough. To consider the above problems, based on big data mining, this paper suggests a prediction method to the economic growth trend of bay area, and constructs a statistical sequence distribution model for the growth trend of bay area economic [3, 4]. Using big data mining method, this paper builds a big data statistical information model for the economic growth trend of the Greater Bay Area, extracts the characteristics of the relevant laws of the economic growth trend and uses the fuzzy big

data fusion clustering method to automatically cluster the economic growth, and establishes the big bay optimal iteration model for the prediction of the economic growth trend, and the economic trending of the bay area is predicted. Combined with the adaptive optimization algorithm, the optimization of Guangdong and Macao region is implemented. Finally, the performance of improving the prediction ability of economic growth trend of Guangdong, Hong Kong and Macao region is verified by simulation experiment.

## 2 Analysis on the Trending of Regional Economy in the Area of Hong Kong and Macao

This paper sets up an information management model of the economy of the Hong Kong, Macao and Guangdong area, and the economic growth trend of that region is optimized, and the automatic dispatching and information management capability of the economy in the Hong Kong and Macao is improved. On the basis of the optimization and prediction of the economic trend in the big bay area, the paper analyses and analyzes the big data of the economy in the district. It is greatly significant to study the trend of the bay area's economic growth. The distribution is shown in Fig. 1.



**Fig. 1.** Guangdong-Hong Kong-Macao Greater Bay Area's distribution

### 2.1 Shortcomings of Guangdong-Macao-Hong Kong Greater Bay Area's Regional Economy

The Greater Bay Area of Guangdong, Macao and Hong Kong gave full play to the geographical advantages of coastal areas and made great progress in the early stage of reform and openness. At the same time, it also led to the improvement of the economy in the surrounding areas to varying degrees. However, due to the limitations of innate regional conditions, the development of the bay district also has many deficiencies,

mainly manifested in: first, the industrial structure of different areas is increasingly convergent. Repeat the low level of construction and other problems, resulting in waste of resources, resulting in low economic benefits. The second is the regional poverty [5]. Although the economy of that region is relatively developed, there are still some areas where the economy is backward and the infrastructure construction is seriously lagging behind, which leads to the poverty of the whole region, the imperfect industrial structure, and the lack of motivation for development. Third, the gap between different regions is further widened, which leads to different geographical regions driven by interests to form different forms of local protection barriers, which increases the difficulty of forming the overall market of economic development.

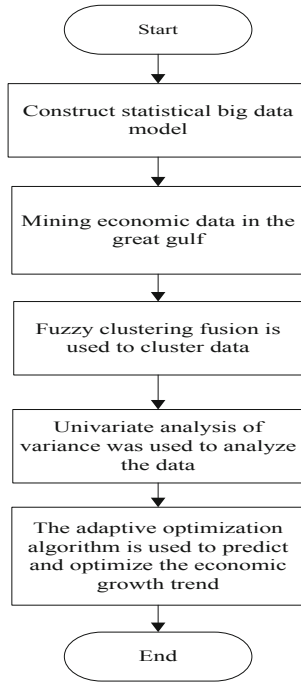
## **2.2 The Influences of the Economic Development in the Region of Hong Kong and Macao**

By analyzing the bay region's economic development, we can conclude that the main influences of the regional economic development are as follows: first, the coordination mechanism between different geographical regions still has great disadvantages, including the ecological compensation mechanism and the resource development mechanism, in which the ecological mechanism is not perfect, and the compensation theory is not systematic, coupled with the lack of unity of the standards of ecological compensation. As a result, there are great differences in many aspects of compensation, so that the consensus of compensation cannot be formed, the mechanism of resource development is not perfect, the overall value of resource products is low, there is great uncertainty in price, and the abundance of resources cannot be transformed into the driving force of economic development. Second, there is a lack of overall planning in the management policies of different regions, and there is no pertinence in formulating the economic policies of the region. Even when the specific regional planning is issued, there is a lack of clear development direction and development goals, which affects the regional economic development [6].

## **3 Construction and Feature Analysis of Statistical Big Data Model**

### **3.1 Statistical Big Data Modeling of the Economic Growth in the Bay Area of Guangdong, Hong Kong and Macao**

To predict the trending of regional economic growth, the statistical sequence analysis and adaptive big data mining are used to design the prediction algorithm. Firstly, it is critical to construct the statistical sequence analysis model, and to use the interactive information scheduling model to sample the growth trend [6]. The specific process is shown in Fig. 2.



**Fig. 2.** The regional economic trending analysis flow chat based on big data mining

According to the results of information collection, the automatic prediction is carried out, and the statistical sequence distribution model is constructed [7]. The big data mining is used to predict the big data economic growth trend adaptively. Combined with the parameter aggregation analysis way, the nonlinear statistical analysis model nd is described by statistical time series.

$$x_n = x(t_0 + n\Delta t) = h[z(t_0 + n\Delta t)] + \omega_n \tag{1}$$

In the formula,  $h(\cdot)$  is the scalar distribution sequence according to the previous statistical results, which is a multivariate quantitative value function, and  $\omega_n$  is the measurement error of Guangdong-Hong Kong-Macao Greater Bay Area’s economic growth trend prediction. It is assumed that the time series monitored by a given amount is expressed as follows:

$$U = \{U_1, U_2, \dots, U_N\} \tag{2}$$

In which,  $U_i$  is a random variable whose dimension is d-dimension, a regression analysis model of prediction in the region is constructed by using a statistical analysis method, The statistical data of the economic growth trend in the large-bay area of Guangdong-Hong Kong-Macao Greater Bay Area has formed a fuzzy rough set in the three-dimensional space [8].

The above formula represents the randomly sampled scalar sequence. Combined with the fuzzy information fusion method, the sampling results of statistical sequence are obtained as follows:

$$X = K[s_1, s_2, \dots, s_K]_{\mathbf{n}} = K(x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau}) \tag{3}$$

Wherein,  $K = N - (m - 1)\tau$  represents the subspace clustering dimension of statistical sequence, which is a time delay. According to the above information flow model, it provides an accurate data analysis basis optimization prediction [9].

### 3.2 Statistical Characteristic Analysis

The statistical sequence distribution model of the economic growth trend is built, the large data statistical modeling is carried out by adopting a large data mining method, the economic growth trend prediction model is established, the discrete characteristic component of the economic demand information is  $s_i = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau})^T$ , the above formula is a group of short-time discrete information distribution set, and in the embedded space of the economic growth trend prediction [10], the distribution function of the state-set distribution function is shown as follows:

$$\frac{dz(t)}{dt} = F(z) \tag{4}$$

Combined with the above formula, the effective probability density is higher by selecting the appropriate  $m$  and  $\tau$ . At this time, the statistical regression analysis probability density characteristics are expressed as follows:

$$R_1 = \{X_1, X_2, X_3, \dots, X_d\}^T \tag{5}$$

By using the continuous descriptive statistical analysis method, multiple comparative analysis was carried out to obtain the correlation function:

$$R_1^T R_1 = \{X_1, X_2, \dots, X_m\} \{X_1, X_2, \dots, X_m\}^T \tag{6}$$

Set up a meter to calculate the best characteristic decomposition value. The characteristic decomposition expression of the prediction is as follows:

$$R_1^T R_1 = V_1 \sum_1 V_1^T \tag{7}$$

The fuzzy analogy of the economic forecast is expressed as follows:

$$R_2^T R_2 = V_2 \sum_2 V_2^T \tag{8}$$

$$R_2 = \{X_{d+1}, X_{d+2}, \dots, X_{d+m}\}^T \tag{9}$$

$$R_2^T R_2 = \{X_{d+1}, X_{d+2}, \dots, X_{d+m}\} \{X_{d+1}, X_{d+2}, \dots, X_{d+m}\}^T \quad (10)$$

In the formula, the wide area characteristic components of the prediction are expressed as follows:

$$V = [V_1, V_2, \dots, V_m] \in R^{m \times m} \quad (11)$$

On the basis the analysis of variance of univariate, according to the evolution characteristics of load value  $X_m$  and  $X_k$ , the predicted load is  $X_{m+1}$  and  $X_{k+1}$ .

## 4 Optimization of Guangdong-Hong Kong-Macao Greater Bay Area's Economic Growth Trend Prediction Algorithm

### 4.1 Extracting the Features from Relevance Rules

The mean of big data mining is used to model the economic development by big data statistical big data, and the association rules' characteristic quantity is extracted [11]. The association characteristics of the growth trend at the level of salience are represented as:

$$W_x(t, v) = \int_{-\infty}^{+\infty} X(v + \xi/2) X^*(v - \xi/2) e^{i2\pi\xi t} d\xi \quad (12)$$

Taking the quantitative average value of the persistent statistical value, the following are obtained:

$$E_x = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W_x(t, v) dt dv \quad (13)$$

Wherein,  $W_x(t, v)$  represents the decisive factor of the economic demand forecast:

$$\begin{cases} \int_{-\infty}^{+\infty} W_x(t, v) dt = |X(v)|^2 \\ \int_{-\infty}^{+\infty} W_x(t, v) dv = |X(t)|^2 \end{cases} \quad (14)$$

On the basis of a comprehensive analysis, the correlation rule feature extraction is carried out, and the automatic prediction of the economic growth trend is carried out according to the feature extraction results [12].

### 4.2 Fuzzy Clustering and Adaptive Prediction

On the basis of extracting the characteristics of association rules, the fuzzy data fusion clustering method is set to automatically cluster the regional economic growth [13, 14],

and the univariate variance analysis is made use of analyzing the orthogonal eigenvector solution [15, 16]. It is assumed that the single component of the growth trend is unknown, and the results are as follows:

$$VV^T = I_M \tag{15}$$

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_m) \in R^{m \times m} \tag{16}$$

Under the constraint of surplus probability, the order of eigenvalues is as follows:

$$\sigma_1 > \sigma_2 > \sigma_3 > \dots > \sigma_{s+1} > \sigma_m \tag{17}$$

Under the evolutionary game, the new characteristic sequence of trending forecast is:

$$X_{m+1}(m) = X_{k+1}(m) \pm \sqrt{(d_m(0)e^{\lambda_1} +)^2 - \sum_{i=1}^{m-1} [X_{m+1}(i) - X_{k+1}(i)]^2} \tag{18}$$

A new group of time series  $y_k$ , which replaces regional economic growth trend, is produced, and the forecast results are as follows:

$$D_k = D_{k-1} - N_{k-1}^2 / D_{k-1} \tag{19}$$

$$\phi_{kk} = N_k / D_k \tag{20}$$

$$\phi_{kj} = \phi_{k-1,j} - \phi_{kk} \cdot \phi_{k-1,k-j} \tag{21}$$

An initial value  $N(0, 1)$ , is generated in the Gaussian distribution  $N(0, 1)$  to realize the improved design of prediction algorithm [17, 18].

## 5 Analysis of Experiment Results

In the sake of testing the application performance in the prediction of the economic growth, the experimental analysis is carried out, and the trending prediction simulation is carried out with SPSS statistical analysis software and Matlab simulation software. The descriptive statistical analysis is used to compare and analyze the linear proportion. The analysis results are shown in Table 1.

**Table 1.** Results of the multiple comparative analysis based on the annual

| Year      | 2009–2012 | 2013–2016 | 2017–2019 |
|-----------|-----------|-----------|-----------|
| 2009–2012 | 1         | –         | –         |
| 2013–2016 | 0.043     | 1         | –         |
| 2017–2019 | 0.056**   | 0.018*    | 1         |

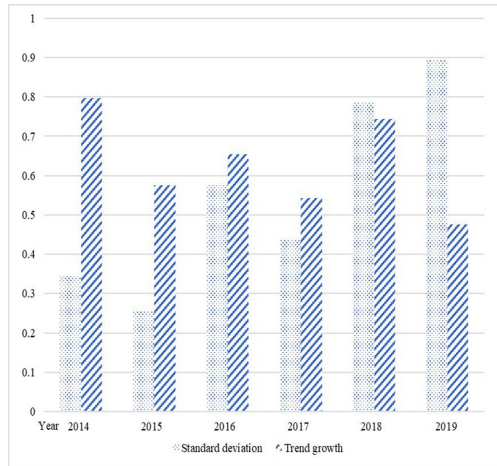
Note: \* indicates significant at 0.05 level.

According to the results of the above comparative analysis, it is found that the annual surplus has a significant impact on the economic, but it is not a decisive factor. On this basis, the descriptive statistical analysis is carried out, and the results are shown in Table 2.

**Table 2.** Results of continuous descriptive statistics

| Year | Sample number | Standard deviation | Max value | Min value | Average value |
|------|---------------|--------------------|-----------|-----------|---------------|
| 2014 | 478           | 0.343              | 467.89    | 12.86     | 42.32         |
| 2015 | 346           | 0.256              | 547.86    | 24.56     | 44.56         |
| 2016 | 578           | 0.578              | 326.45    | 27.54     | 76.76         |
| 2017 | 865           | 0.437              | 655.68    | 45.65     | 76.78         |
| 2018 | 546           | 0.785              | 544.67    | 35.32     | 65.98         |
| 2019 | 489           | 0.893              | 543.45    | 43.45     | 56.75         |

According to the above descriptive statistical analysis results, the statistical analysis value is shown in Fig. 3.



**Fig. 3.** Statistical Analysis Results of the prediction of the Economic Growth Trend in the district.

According to the statistical analysis results of Fig. 3, the prediction is realized, and the prediction accuracy is tested. The comparative results are shown in Fig. 3. The analysis of Fig. 4 shows that the prediction of the Greater Bay Area economy by using this method is more accurate.

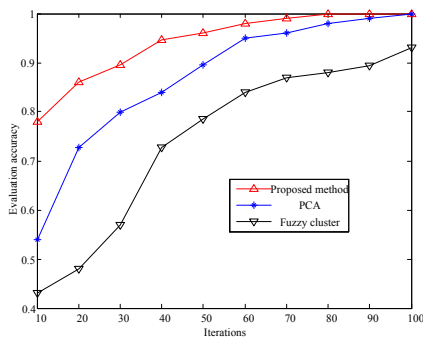


Fig. 4. Comparison of prediction accuracy

## 6 Conclusions

Based on the optimization and prediction of the economic growth trend of Guangdong-Hong Kong-Macao Greater Bay Area, the automatic dispatching and information management ability of the economy in Guangdong and Macao is improved. In this paper, the economic growth trend optimization and prediction with big data mining is put forward. Through the phase space reconstruction, the correlation statistical sequence distribution of the economic growth trend prediction of the Greater Bay Area is obtained, and the large-data statistical large-data modeling of the economic growth trend of the big bay region is carried out by adopting a large data mining method. With the fuzzy large-data fusion clustering way, the trending is automatically clustered, and the orthogonal feature vector solution analysis of the economic growth trend is carried out by using the variance analysis method of the single variable. And by using fusion clustering, the trend of economic in the large-bay area of the three districts is automatically cluster-treated by means of the fuzzy large-data fusion clustering method, and the optimization and prediction of the economic growth trend of the region of Hong Kong and Macao are realized by using the self-adaptive optimization algorithm. The outcomes show that the adaptive scheduling and management ability of the economic growth trend of the bay region of Guangdong, Hong Kong and Macao is better, the prediction accuracy is higher, and the self-adaptive scheduling and management ability of the large-bay district is improved.

**Acknowledgements.** Supported by Guangdong Provincial Philosophy and Social Sciences “13th Five-Year Plan” 2018 Project: “Guangdong-Hong Kong-Macao Greater Bay Area ‘One Country, Two Systems, Three Levels and Four Cores’ Fusion Development Study” (Approval No.: GD18XYJ22).

## References

1. Aslanidis, N., Christiansen, C.: Smooth transition patterns in the realized stock-bond correlation. *J. Empir. Finance* **19**(4), 454–464 (2012)
2. Parvin, H., Mirnabibaboli, M., Alinejad-Rokny, H.: Proposing a classifier ensemble framework based on classifier selection and decision tree. *Eng. Appl. Artif. Intell.* **37**, 34–42 (2015)
3. Hwang, Y., Yu, T.Y., Lakshmanan, V., et al.: Neuro-fuzzy gust front detection algorithm with S-band polarimetric radar. *IEEE Trans. Geosci. Remote Sens.* **55**(3), 1618–1628 (2017)
4. Bin, M.A., Xianzhong, X.I.E.: PSHO-HF-PM: an efficient proactive spectrum handover mechanism in cognitive radio networks. *Wireless Pers. Commun.* **79**(3), 1–23 (2014)
5. Chaowen, C.: Discussions on the reform of Tanching practice for computer science majors. *J. Nanning Teach. Coll.* **22**(3), 75–77 (2005)
6. Tao, W.U.: Research on the infiltration of tea painting art in art education in colleges and universities. *Tea Fujian* **24**(4), 45–47 (2018)
7. Maheshwari, P., Singh, A.K.: A survey on spectrum handoff techniques in cognitive radio networks. In: 2014 International Conference on Contemporary Computing and Informatics (IC3I), pp. 996–1001. IEEE (2014)
8. Ding, X.M., Fuh, J.Y., Lee, K.S.: Interference detection for 3-axis mold machining. *Comput. Aided Des.* **33**(8), 561–569 (2001)
9. Lasemi, A., Xue, D., Gu, P.: Recent development in CNC machining of freeform surfaces: a state-of-the-art review. *Comput. Aided Des.* **42**(7), 641–654 (2010)
10. Wu, Y.X., Wen, X.: Short-term stock price forecast based on ARIMA model. *Stat. Decis.* **(23)**, 83–86 (2016)
11. Li, X., Peng, L., Hu, Y., Shao, J., et al.: Deep learning architecture for air quality predictions. *Environ. Sci. Pollut. Res.* **23**(22), 22408–22417 (2016)
12. Zaho, X., Wu, Y., Huang, H.Y., et al.: New generation of migrant workers housing purchasing power and the housing sales to inventory constraints relationship model – taking Zengcheng area as an example. *Manag. Eng.* **22**(1), 8–13 (2017)
13. Wang, Q., Bai, X., Feng, N.: The systemic risk measurement and influencing factors of commercial banks—an analysis based on CCA-POT-Copula Method **02**, 1–9 (2016)
14. Yao, L., Yao, W.X.: Research on the policy effect of incremental expansion of margin and securities lending: based on the multi period DID model and Hausman’s test. *Int. Financ. Res.* **349**(5), 85–96 (2016)
15. Ning, C., You, F.: Data-driven adaptive nested robust optimization: general modeling framework and efficient computational algorithm for decision making under uncertainty. *AICHe J.* **63**(9), 3790–3817 (2017)
16. Feng, W., Heng, Z., Kangshun, L., et al.: A hybrid particle swarm optimization algorithm using adaptive learning strategy. *Inf. Sci.* **36**(4), 162–177 (2018)
17. Laizhong, C., Genghui, L., Xizhao, W., et al.: A ranking-based adaptive artificial bee colony algorithm for global numerical optimization. *Inf. Sci.* **417**(10), 169–185 (2017)
18. Konstantinos, A.T., George, N.: Optimization of a fully adaptive quality and maintenance model in the presence of multiple location and scale quality shifts. *Appl. Math. Model.* **54**(6), 64–81 (2017)