International Journal of Biology and Genetics
Research Article

The Forecast of The Number of Inbound Tourists and The Analysis of The Source Market During The Epidemic of Coronavirus Disease

Xia Jiang[1], Bin Zhao[1,2*]

[1]Hospital, Hubei University of Technology, China.
[*2]School of Science, Hubei University of Technology, China.

Received: 25 August, 2020; **Accepted**: 27 August, 2020; **Published:** 05 September, 2020

***Corresponding Author:** Dr. Bin Zhao, School of Science, Hubei University of Technology, China. E-mail address: zhaobin835@nwsuaf.edu.cn

Abstract

With the rapid development of economy, the competition of inbound tourism market is more and more fierce. The key point of sustainable development of inbound tourism is to ensure a certain number of tourists. Therefore, it is an important step to predict the number of inbound tourists and study the market of inbound tourists. As a leading tourism city in China, how to attract more tourists is not only related to the development of inbound tourism in Shanghai, but also provides some inspiration for other cities during the epidemic of Coronavirus Disease.

In this paper, an improved grey markov (GM) model is used to predict the number of inbound tourists in Shanghai during the epidemic of Coronavirus Disease, and then the market changes of inbound tourists are studied by the deviation-share analysis method. Finally, the tim-scale characteristics and trends of inbound tourists in Shanghai are analyzed by ensemble empirical mode decomposition.

GM (1,1) model is one of the most widely used grey dynamic prediction models in grey system theory, which is composed of a first order differential equation with a single variable. The initial value correction improves the gray GM (1,1) model, and introduces the center point triangle albino weight function in the state division to improve the Markova model. Comparing with the results of traditional GM (1,1), initial value modified GM (1,1) and traditional grey markov prediction models, the prediction effect of this model is verified to be better. These models are better than linear regression and time series.

Deviation-share analysis explores the changes in the inbound tourist market, and the results show that from 2004 to 2017, the inbound tourist market in Shanghai developed faster than that in the whole country, with a more reasonable and competitive structure. In addition to Japan, the number of inbound tourists from each country to the whole country and Shanghai has increased and increased greatly.

The time-scale characteristics and trends of inbound tourists in Shanghai are analyzed by ensemble empirical mode decomposition. The results show that: first, the total number of inbound tourists and the number of foreign tourists mainly change within 3 or 6 months, while that of Hong Kong, Macao and Taiwan fluctuates between high and low frequency. Second, the main cyclical fluctuations and no significant trend of the source countries. The fluctuation

period of Japan, Thailand, Britain, France and Germany is 3 months; Macau is 3, 6, 12, 60, 180 months; Singapore is 3, 6, 180 months. Third, there is a clear trend and cycle fluctuations as a supplement to the source countries. The fluctuation periods in Hong Kong are 3, 6, 90 and 180 months; In Taiwan, Canada and Russia it is 3 , 6 months; In Indonesia, the United States, Italy and New Zealand it is 3, 6 and 12 months; In Malaysia it is 3, 180 months; In South Korea it is 3 ,45 months; In Australia it's four or seven months. Taiwan, Canada, Russia and New Zealand showing the most significant upward trend.

From the above research results, specific Suggestions and strategies of market structure competition can be put forward to the inbound tourism industry in Shanghai according to the predicted number of inbound tourists in Shanghai, the structure of the source market and the cyclical fluctuation and trend of the source country.

Keywords: Gray Arkov; Initial Value Correction; Albino Weight Function; Deviation-Share; Ensemble Empirical Mode Decomposition.

Introduction

Research Background and Significance

Research Background

Due to the great development of the world economy and the rising standard of people's living, people gradually began to choose tourism to relax themselves and broaden their horizons. China's tourism industry is a newly developed industry. It started late, but its rapid development is unimaginable. It can be divided into international tourism and domestic tourism, and international tourism can be divided into outbound tourism and inbound tourism.

In fact, tourism is a kind of labor service, the utility of which is invisible to the naked eye. As a special value to society, it can meet many other requirements such as tourists' travel, leisure and visiting relatives and friends, especially inbound tourism, which attracts tourists from all over the world due to the attraction of different cultures and lifestyles and transnational business activities.

The competition of international tourism is more and more intense. At present, China can be called a tourism power, but there are still many distances between the tourism power and the tourism power. It is urgent to carry out inbound tourism and increase its international status in order to transform itself from a tourism power to a tourism power. Shanghai, located in the southeast coast, is one of the entries and exit ports in China and the gateway to the world in the future. There are great advantages in developing inbound tourism. In 2014, Shanghai tourism authorities wanted to build Shanghai into a world-famous tourist city, and the Shanghai International Resort project relying on Shanghai Disneyland has been progressing steadily, which also shows that Shanghai authorities attach great importance to inbound tourism, which has obviously become a new vitality to promote development. Based on the analysis of the number of inbound tourists, the market structure of inbound tourists and the multi-scale characteristics of inbound tourism time, on the one hand, it can understand the current situation of Shanghai inbound tourism, on the other hand, it can also provide new ideas for the development of Shanghai inbound tourism. Based on the above background, the title of this paper is determined during the epidemic of Coronavirus Disease.

Research Meaning

Studying the data of Shanghai's inbound tourism in the past years and predicting inbound tourism in Shanghai during the epidemic of Coronavirus Disease and its future development will help us to understand Shanghai's inbound tourism market from another perspective. It can also give some useful suggestions to Shanghai's tourism departments and enterprises. Shanghai's inbound tourism market has a good prospect, so it is of great significance to study it. It is also helpful to seize the opportunity, seize the foreign tourism market, produce the best-selling tourism products, and determine the optimal plan.

Global tourism is becoming more and more popular, and the competition for tourism market is becoming more and more fierce. The focus of competition is that the tourist market should have a certain scale, so it is necessary to discuss the inbound tourist market. Whether a country or a region's tourism industry is internationalized and immature depends on its inbound tourism development, and the inbound tourism market structure is very important to the scale and development of the inbound tourism industry. Deviation share analysis method is comprehensive and non-static, which can well analyze the good and bad situation of regional inbound tourist market structure and competition, reflect the change of tourist market structure, and better study the tourist market.

Inbound tourism is one of the main markets of China's tourism industry, which can obtain income, increase the communication between China and foreign countries, and increase China's influence in foreign countries. Its rapid development makes our country's tourism industry flourish, and makes our country's position in the world further rise. Therefore, it is of great significance to analyze China's inbound tourism market and understand the time scale characteristics and basic laws of the inbound tourism market for planning our country's inbound tourism development strategy.

Research Purpose, Ideas and Contents

Research Objective

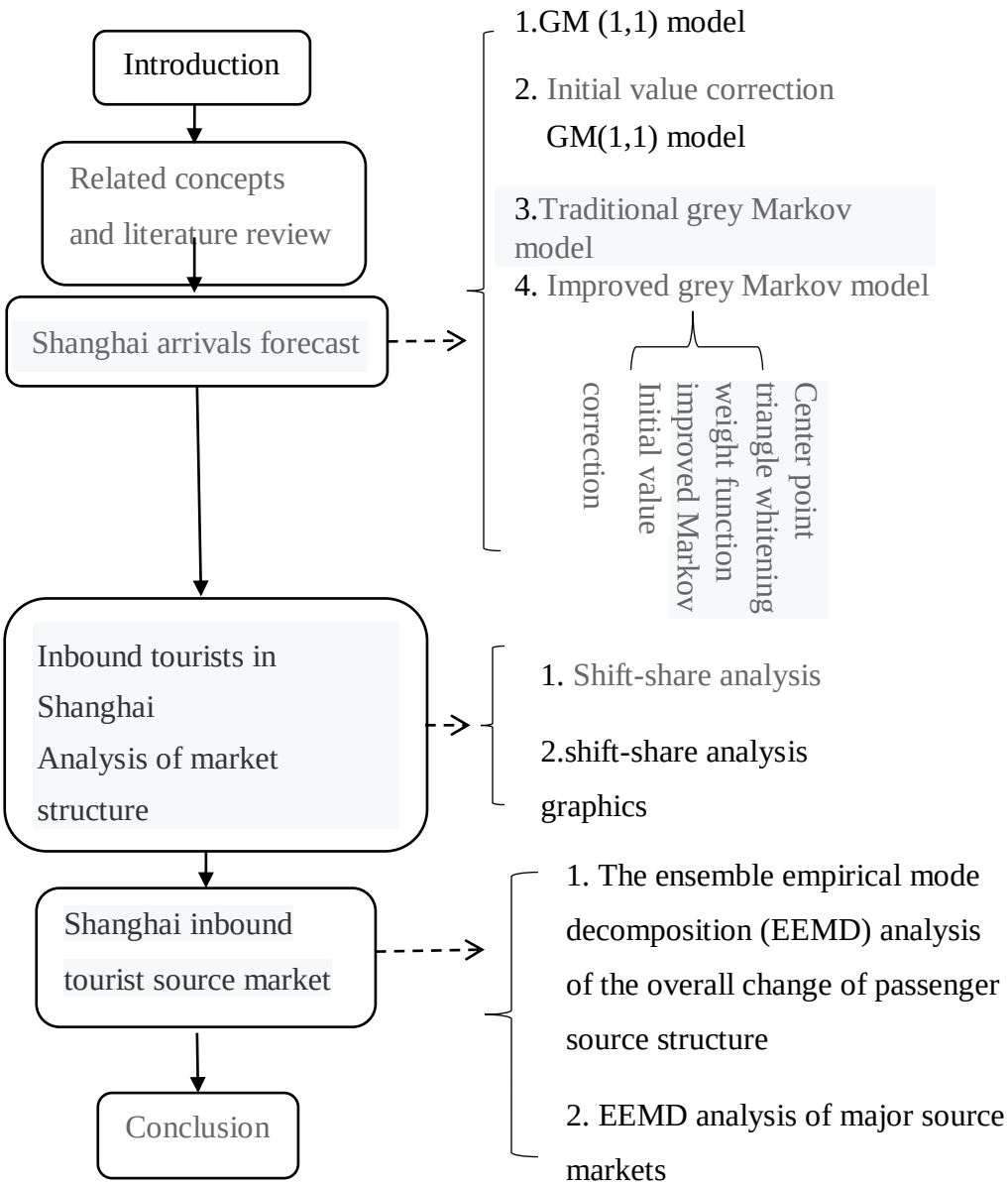
The purpose of this paper is to forecast the number of tourists in Shanghai, to provide some reference for the next year's plan of Shanghai Tourism Administration, and then to study the structure of Shanghai inbound tourist market. We can see whether Shanghai's tourist source countries are on the rise or on the decline, the structure is not reasonable, and the competitiveness. Finally, we study the multi-scale characteristics and trend analysis of Shanghai inbound tourist market to get the periodicity of tourism from the source country to Shanghai. All of these provide good suggestions for Shanghai's tourism industry and can better develop Shanghai's tourism industry.

Research Ideas

The research idea of this paper is first the introduction part, which elaborates the research background, research significance, research purpose, research idea, research content, research method and innovation of this article. Then, the domestic and foreign literature review of forecasting model, tourist market structure analysis, deviation share analysis and set empirical mode decomposition are described in detail. Next, the paper forecasts the number of Shanghai inbound tourists, analyzes the structure of Shanghai inbound tourism market, and analyzes the multi-scale characteristics and trends of Shanghai inbound tourism market. At last, it summarizes the article, describes the innovation and shortcomings of the article.

The specific research ideas of this paper are as follows:

Figure (1.1) Flow chart of research ideas



Research Contents

Our study is divided into six parts to study the prediction of the number of inbound tourists in Shanghai, the changes of inbound tourist market and the multi-scale characteristics and trends of inbound tourists

Part 1, introduction. This part gives a comprehensive overview of the research background and significance, research purpose, content and ideas as well as research methods of the paper, and puts forward innovative places, which plays an overall role in the research of this paper.

Part 2, related concepts and literature review. This Part mainly introduces the related concepts of inbound tourism and the research status of relevant literature at home and abroad, including the prediction of inbound tourism number, the analysis of inbound tourism market structure and the literature review of empirical mode decomposition. A comprehensive summary of the relevant literature plays a theoretical role in the study of the main body.

Part 3, the forecast and analysis of the number of inbound passengers in Shanghai. In this part, the initial value of GM

(1,1) is modified first, then the central point triangular whitening weight function improves Markov model and constructs the improved grey Markov prediction model. By comparing with the results of linear regression, time series, traditional GM (1,1), initial value modified GM (1,1) and traditional grey Markov model, the optimal number of passengers is obtained.

Part 4 analyzes the structure of Shanghai inbound tourist market. This part uses the deviation share analysis method to analyze the changes of inbound tourist market, its overall structure and competitiveness, according to the results, which tourist source countries will go up or down to Shanghai in the future.

In part 5, the time multi-scale characteristics of Shanghai's inbound tourists are analyzed. The periodic change and growth trend of tourist source countries in Shanghai are analyzed by using ensemble empirical mode decomposition (EEMD).

Part 6, conclusion. The conclusion of the above parts is summarized, and the shortcomings are put forward. Based on the results, some suggestions for improving Shanghai's tourism industry are provided.

Research Method

Literature Analysis

Literature analysis is the basic method of paper writing. After careful reading of domestic and foreign literature, we can understand the results and latest research trends in literature.

According to the papers written by our predecessors, we can learn some ideas and research directions of writing articles, and master the latest frontier science. Based on the analysis of domestic and foreign literature, this paper grasps the current situation of the research on the prediction of the number of inbound tourists, the structure of inbound tourists' market and the time scale change characteristics of inbound tourism market. Refer to a large number of domestic and foreign documents about inbound tourism, and finally elaborate these documents in detail to provide theoretical support for the study.

Statistical Analysis

Statistical analysis is to analyze the quantitative relationship of things, understand and express the relationship between research objects, explain the law of change and development trend, so as to accurately explain and predict the research objects. All things in the world have quality and quantity. To understand the essence, we need to grasp the law of quantity.

Statistical analysis uses mathematical methods to establish mathematical models, and makes statistical analysis on the data and data obtained from the survey, so as to obtain quantitative conclusions that are easy to understand. This method is a scientific, accurate and objective evaluation method which is often used now. In this paper, the collected data are analyzed by R language, MATLAB, Excel and so on. The number of inbound tourists in Shanghai is predicted, and the structure of Shanghai inbound market and the time scale change characteristics of the inbound market are found out to analyze.

Innovation

This paper not only uses the grey Markov model, but also uses the improved grey Markov model to predict the number of inbound passengers in Shanghai. First of all, the grey model is improved by modifying the initial value. The traditional GM (1,1) model $x^{(1)}(1) = x^{(0)}(1)$ For the initial condition, the information brought by the new data is lost, so the initial condition $x^{(1)}(1)$ Generate last item with new accumulation $x^{(1)}(n)$ Predict together to improve accuracy. Then, the central point triangular whitening weight function is introduced to improve the Markov model. The traditional state division of Markov model does not reflect the preference degree of each fluctuation index, while the central point triangular whitening weight function comprehensively considers the preference degree of each fluctuation index in two adjacent intervals, indicating the possibility that the object belongs to a certain state, which can be compensated Subjective division of states. At last, the improved Gray Markov prediction model is

constructed, which greatly improves the prediction accuracy.

Shift share method (SSM) and ensemble empirical mode decomposition (EEMD) have been used in many fields, but they have not been used to study a problem together and draw a conclusion. The shift-share method can obtain the rationality of the market structure of inbound tourists in Shanghai and which source countries have competitive advantages. The EEMD method can obtain the cyclical changes and trend changes of some source countries to Shanghai. The two results are linked together to provide some suggestions for the tourism industry in Shanghai.

Related Concepts and Literature Review

Related Concepts

Inbound Tourism

Inbound tourism refers to the tourism activities of foreign residents or residents of other countries coming to our country. Inbound tourism is a part of international tourism. The market of major tourist source countries to China and Chinese cities is divided into two parts: first, compatriots and overseas Chinese from Hong Kong, Macao and Taiwan; second, foreigners (including Chinese who already have foreign nationality).

Inbound tourists

Inbound tourists are foreigners, compatriots in Hong Kong, Macao and Taiwan who come to China (mainland) for tourism, holidays, visits relatives and friends, medical treatment, clothes and cosmetics, attend meetings or work related to economic, cultural, sports, religious and other activities within the reporting period, that is, the number of inbound tourists. When the relevant personnel make statistics, the inbound tourists are considered as one person at a time. Both inbound tourists and one-day tourists belong to inbound tourists.

Inbound tourists are foreigners and compatriots of Hong Kong, Macao and Taiwan who stay in the tourist accommodation facilities in China (mainland) for more than or equal to one night. Inbound passengers do not include:

1. Officials who are invited to visit China at or above the level of government ministers and accompanying persons;
2. Staff, diplomats of foreign embassies in China and their accompanying family service personnel and dependents;
3. Foreign experts, students, journalists and business personnel who have lived in China for at least one year;
4. Transit passengers who do not need to enter the international flight ports of mainland China through passport inspection;
5. Border residents in and out of border areas;
6. Compatriots from Hong Kong, Macao and Taiwan who have returned to settle in the motherland;
7. Foreigners settled in China and those who return to settle in China after leaving the country;
8. Mainland Chinese citizens returning home.

One day inbound tourists refer to foreigners, Hong Kong compatriots, Macao compatriots and Taiwan compatriots who do not stay overnight in China (mainland).One day inbound tourists include automobile, train, boat and yacht inbound tourists, overnight tourists on board or on board, and service personnel on board, but excluding overseas (domestic) Chinese (mainland) living in China (mainland) but working in China (mainland), compatriots from Hong Kong, Macao and Taiwan who come back on the same day, as well as border people from neighbouring countries.

Tourist Source Market

Tourist market refers to the actual and hidden buyers of a specific designated tourist product in the tourist area. From the perspective of economics, it is the sum of the supply and demand of tourism products.

Geographically, it is the tourism economic activity centre of the tourism market. It is a general category of commodity market, with the basic characteristics of commodity market, including tourists and tourist destinations, as well as the relationship

between tourism operators and buyers. There are differences between the general commodity market and the tourism market. The tourism market does not sell specially designated related material products, but planned routes in advance.

Tourist source market is the sum of tourism supply and demand market, which embodies the economic relations among countries, countries and tourism operators, tourism operators and tourists. The formation and development of tourism market is the inevitable result of the coordinated development of these relations.

Tourist market is a huge market with many people and complicated structure. These people come from different places, have different personalities, different cultures and different travel requirements. Therefore, the demand of tourist market is diversified. The staff of tourist destination should make detailed investigation instead of relying on their imagination to set down the requirements of tourist source, and adopt efficient methods to achieve the demand diversification of tourist source market.

Journals Reviewed

This section carefully clarifies the relevant references of the prediction of the number of inbound tourists at home and abroad, the market structure analysis of inbound tourists and the time multi-scale characteristic analysis of inbound tourists, as well as the research status of Grey Markov model, deviation share analysis method and empirical mode decomposition method.

Literature Review on The Prediction of The Number of Inbound Tourists

With the rapid growth of the National People's economy and the continuous improvement of people's income, people often choose to travel in their spare time to improve their quality of life. At present, tourism has been gradually integrated into our life, people like to go out to play. The world has stepped into the trend of tourism. As a very important part of tourism demand forecasting, the theory and application of this field are being studied step by step at home and abroad. There are many methods to predict the number of Inbound Tourists: Jincheng Tang, Songsak sriboonchitta and Xinyu yuan [1] use the time series model combined with belief function to predict the demand of China's international tourism. Veloce W [2] uses error correction model (ECM) and traditional regression model to predict Canadian inbound tourists. Witt S F and Turner L W [3] use the integrated time series econometric analysis (Sitea) method to predict the number of Chinese inbound tourists. Ying y, Yirui W and Shanghai G [4] put forward a comprehensive moving average of seasonal trend autoregression using tree neural network model (SA-D model) to forecast tourism demand. Hu Lijuan selects the prediction index of Qinhuangdao's inbound tourism demand, and constructs the prediction model of inbound tourism demand based on BP neural network. Li Naiwen and Han Jingjing combined ARIMA model with RBF neural network to predict the number of inbound tourists in China. Wang Xiaoshan established a gravity model to adapt to China's inbound tourism, and can measure the impact of various factors on China's inbound tourism. Qiao Rui uses index and logistic curve model to predict the number of inbound tourists in Shanghai. In Xiong Liang's model of Shanghai inbound tourists, the time series model has the highest accuracy and the best effect, which is often used in the random series changing with time; the quartic curve model has the better effect in the one-way model, which is suitable for the medium and short-term prediction; in the case of incomplete indicators, the multiple regression model is more suitable for the ranking analysis of main factors. Chen Peng predicted the number of inbound tourists in Anhui Province Based on GM (1,1) model, and put forward some suggestions for the development of inbound tourism market. It can be seen from the above literature that the prediction methods of inbound tourists include time series, error correction model (ECM), regression model, tree neural network model (SA-D model), SARIMA, BP neural network, differential autoregressive

moving average model (ARIMA), gravity model, index and logistic curve, time series model, GM (1,1) model, etc.

There are many ways to modify the initial value of GM (1,1) model. He Xia and Xu Hongwei[11]prove that the initial value of GM (1,1) model is modified by $x^{(0)}(1)+\alpha$, $\beta x^{(0)}(1)$ The equivalence of the grey prediction model is only related to the exponential structure of the solution of the model, but not to the determination of the background value and the estimation of the model parameters. Shao Hongmei, Yang Jianhua and LAN Yuexin choose two different initial conditions [12] $x^{(1)}(n)$ and $x^{(1)}(n)+\beta$ The grey model prediction is carried out respectively. The non-equidistant grey prediction model of wangche effect based on the initial value correction is used[13] $\alpha x^{(0)}(k_1)$ and $x^{(0)}(k_1)+\alpha_1 k_1+\alpha_2$ Two methods are used to correct the initial value. Heng Yali and Wang Bo[14] think that the initial value of the original GM (1,1) model $x^{(0)}(1)$ Can affect the solution of differential equation $\hat{x}^{(0)}(k+1)$. Consider using $\hat{x}^{(0)}(1)=x^{(0)}(1)+\sigma$ to correct the initial value. Wang Zhongtao, Peng Xin and Dai Qi[15] in order to avoid the limitation of solving equations by using ${x_1}^{(1)}={x_0}^{(1)}$, so they use $x^{(1)}(1)=b_i x^{(1)}(m)$ to correct, to determine the undetermined parameters by minimizing the sum of squares of errors b_i. When m taking different values, we can get different prediction models after correction, and then, the model error sum of squares is minimized, and n kinds of prediction results are combined with optimal weighting to produce a new model. Liu Muxiao[16] set the initial value of the model as $\beta x^{(0)}(1)$, substituted this into the traditional GM(1,1) model, and used the minimum gradient method to find the modified parameter with the minimum residual index function β.

Yao T, Gong Z and Xie N studied the growth rate of simulation values under different initial values of discrete GM (1,1) model, and optimized to obtain the initial values[17-24].

As can be seen from the above literature, the gray GM (1, 1) model for the initial value correction method also has a lot of, such as $x^{(0)}(1)+\alpha$、$\beta x^{(0)}(1)$、$x^{(1)}(n)$、$x^{(1)}(n)+\beta$、${x_1}^{(1)}={x_0}^{(1)}$、$x^{(1)}(1)=b_i x^{(1)}(m)$、the growth rate of the simulation values under different initial value to optimize the initial value and so on the many kinds of modified method of initial value.

Zhang Rui used the improved grey Markov model to predict the grain yield of Shaanxi Province and studied the influencing factors from the traditional input and agricultural policy. Shi Chaoyang proposes to improve the grey Markov chain prediction model, which solves the disadvantages of the former Grey Markov chain model in terms of the subjective experience of the modeler. On the other hand, it improves the efficiency of judging whether the divided state is stable through the Markov test. The improved grey Markov model studies and forecasts the data of China's consumer price index. Hu Xiaoyong mainly constructs the grey system prediction model, the weighted Markov chain prediction model and the grey Markov chain prediction model, and compares the effect of the three models.

Zhanli and Jinhua set up grey Markov model to predict the fire. Yushui Geng's grey Markov model based on the improved GM (1,1) algorithm is used for sales forecasting. This model has achieved good results in the sales forecast and has carried on the comparative analysis to the forecast result. Zongqian Jia, Zhifang Zhou, Hongjie Zhang, Bo Li and Youxian Zhang established the GM (1,1) prediction model of coal consumption in Gansu Province, and then modified the GM (1,1) model with Markov chain prediction method, and tested the accuracy of the modified model. Hongyan Huan and Qingmei Tan put forward the gray Markov model, which has the advantages of dealing with bad information and long-term fluctuation sequence, and forecast the cultivated land scale of Jiangsu Province. It can be seen from the above literature that the combination of grey model and Markov model is widely used, which also shows that the combination of the two models can be realized very well.

Literature Review on the Analysis of Market Structure of Inbound Tourists

The steady and normal development of inbound tourism market is an inevitable requirement for a city to become an important tourism destination in the world. Therefore, the analysis of inbound tourist market structure is helpful to get the market expansion strategy. Only when we understand the structure change and future development trend of the source market, can we give a better strategy to improve. There are many articles about the analysis of tourist market structure at home and abroad.

Liu fajian, Chen Dongdong, Zhu Jianhua, Qian Zang and Li Binbin build the market subordinate network of China's inter provincial tourist source countries based on the 2-model network analysis. Ma Lijun, sun Gennian and he Jingru used market competition and pro scene degree to establish a comprehensive evaluation model to analyze the inbound tourist market and its changes in the two periods. Xiao Lai uses SSM, pro scene model and Boston matrix theory model to explore the situation of Hangzhou inbound tourist market, and tries to find ways to summarize the characteristics and disadvantages of each tourist market, and provides some suggestions for it. Yu Tong used geographic concentration index, clustering and pro scene degree to analyze the market development of Shenzhen's inbound tourist source countries in 2008-2017.Shi bin and Ma Yaofeng use the DSSM model to explore the evolution of Shaanxi inbound tourism market structure in 2011-2015.It can be seen from the above literature that there are many models for the study of inbound tourist market, such as 2-mode network analysis, comprehensive evaluation model, Boston matrix theory model, pro scene analysis model, geographic concentration index, cluster analysis, transfer share analysis and dynamic deviation share method[25-29].

The different application cases of deviation share method to the analysis of inbound tourism market are as follows. Li Cuilin and Qin Hao used the deviation share method to analyze the development structure of Xinjiang's inbound tourism industry, and selected the 2001-2015 foreign exchange income of Xinjiang and the whole country for empirical analysis. Lin Longfei and Jiang Yan analyzed the market structure of inbound tourists in Zhengzhou by using the data of inbound tourists in Zhengzhou from 2001 to 2012 and the statistical data of all ethnic groups in China. Tang Dai established the model of deviation share analysis of Shantou City, compared Shantou City with the entry market of the whole country and Guangdong Province, and obtained the development status, different characteristics and future development trend of Shantou City compared with the whole country and Guangdong Province. Zhang Kai and Wang Yuqin constructed the SSM model to analyze the inbound tourism market of each province and city in the Yangtze River Basin, and briefly analyzed the competitiveness of 17 major tourist source countries to these places. Yu mengke's dynamic SSM method is used to analyze the entry development of the main tourist market in Guizhou, and get the market basis and competitiveness of the main tourist countries, so as to get the development status of each tourist market in Guizhou. Caiping Z [35] uses the deviation share method to analyze the structure of Zhejiang inbound tourism market. Peiji S [36] divides the changes of inbound tourism market in Gansu Province into two stages: 1996-2000 and 2000-2005, and analyzes them with deviation share method. Wang Li and Meng duo selected overseas tourists in Liaoning Province as samples, based on some inbound tourism statistics, using the deviation share analysis method, to analyze the development status and structure of the tourism market in Liaoning Province. Huang [38] used the deviation share analysis method to study the inbound tourism market structure of Hainan. Yasin m, Alavi J and Sobral F [39] use the deviation share method to study the characteristics and competitive position of Portugal in the tourism market [30-34].

Some cases of deviation share analysis in other fields are as follows. Based on the empirical analysis of the structural

adjustment of planting industry by the deviation share method, Li Yan used the deviation share method to analyze the competitiveness of different crops. It is measured by growth share bias, structure bias and competitiveness bias. Yu Xiaoyang, Tian Shuai, Lu Yi and Liu Shuai used the method of deviation share to compare the production structure and competitive advantage of grain crops. Luoyuanhong and Pingying aim to analyze the four marine pillar industrial structures of Jiangsu Province in recent years with the method of deviation share. According to the unique marine economy of Jiangsu Province, they provide some suggestions for the development of marine pillar industries of Jiangsu Province. The four pillar industries of coastal tourism, marine shipping industry, marine transportation and marine fishery continue to tap potential, improve competitiveness and have good industrial foundation advantages. Guo Xiaojie and Tian Huiping use the deviation share method to explore the industrial structure and competitiveness of various industries in Guangdong service industry [40-43].

Literature Review of Empirical Mode Decomposition Method

Empirical mode decomposition (EMD) is a great breakthrough of linear and steady-state spectrum analysis based on Fourier transform. According to the characteristics of data time scale, EMD can decompose the signal without setting the basis function in advance. Because of this characteristic, EMD can be used to decompose all kinds of signals in essence, so it has obvious advantages in solving non-stationary and non-linear data, and the signal-to-noise ratio is also high. Since the EMD method was proposed, it has been applied in different fields rapidly and efficiently. Yu Xiangyang, Sha run, Zhu Guoxing and Hu Shanfeng used empirical mode decomposition (EMD) method to explore the characteristics of passenger flow fluctuation in Huangshan Scenic Area, and predicted by EMD and LSSVM. Li Xiaoxuan, LV Benfu, Zeng Pengzhi and Liu Jinjie put forward clsi-emd-bp prediction model based on network search to improve prediction accuracy on the basis of noise interference prediction. At first, we use CLSI method to synthesize the index of Internet search data, then use EMD to process the noise sequence, separate the high-frequency noise from the original sequence, and then use the denoised web search data to predict the passenger flow. Based on the emd-arima-bp composite model of network search index, Lu Lijun studies the tourism behavior of tourism consumers in the Internet era, and further improves the accuracy of the prediction of tourist volume. Based on BP and Elman neural network model, Lu Lijun and Liao Xiaoping introduced EMD method to improve the traditional BP prediction model to study the new characteristics of the passenger parade in the Internet era. Li Xiaolong, Xu Baoguang and Shi Biao use EMD to analyze air passenger flow considering that air passenger flow is affected by macroeconomic, seasonal, competitive and other factors. Zhao Junyuan, Gao Zhanyu, Li Xiaoli and Zhang Jiayong used empirical mode method to analyze the monthly fluctuation and reasons of Sichuan tourism foreign exchange income since August 2002.Xu Biwei, Su Chengli, Yang Wei and Cao Jiangtao proposed an isolated word recognition algorithm based on EMD and DTW considering the large noise interference in speech recognition. The emd-bpn method, which combines empirical mode decomposition with neural network, is used to predict the short-term passenger flow of high-speed railway. Xiong Tao's hybrid model based on EMD is used for nonlinear time series analysis and prediction. At the same time, aiming at the problems in the field of enterprise and finance, the corresponding prediction model is designed. It can be seen from the above literature that empirical mode decomposition (EMD) can not only be used in the field of tourism, but also be widely used in other fields. It can also be combined with other models to study problems [44-52].

Set EMD is an improved method of EMD, which improves the phenomenon of mode aliasing in EMD.EEMD also has a lot of research in various fields. Hen Lingling, he Liang and Li Yuxia used the EEMD method

to analyze the time, multi-scale change rule and the relationship between inbound tourism and economic growth by using the monthly data of inbound tourism foreign exchange income, inbound tourism population and GDP. From the perspective of tourism demand forecast, Zhang Muzi uses the combined forecast model of EEMD and ARIMA to predict the occupancy rate of hotels in Charleston area and its east Cooper area and North Charleston area. Sun Jianbo based on the integrated empirical mode decomposition (EEMD) and variable weight combination to predict the photovoltaic power, which improves the accuracy of the model prediction. Zhao Hua Wu and Norden E. Huang [56] think that EEMD is a time-space analysis method. After many experiments, the increased white noise is averaged out. The only persistent part that survives in the average process is the component of signal (original data), and then it is regarded as a real and more physical answer. Hongchao Wang, Jin Chen and Guangming Dong [57] use EEMD method to decompose the early weak fault signals of rolling bearing, obtain some intrinsic modal functions, and then select IMF with the largest kurtosis index to be processed by tqwt. Combining the advantages of EEMD and WNN, Yaguo Lei, Zhengjia he and Yanyang Zi put forward an effective method of automatic fault diagnosis for rolling bearing of locomotive. Hua Li, Tao Liu, Xing Wu and Qing Chen proposed a sensitive mode function selection method based on FBE for the problem of multiple intrinsic mode functions generated by EEMD, which can better reflect the fault characteristics. Said GACI [60] proposed a denoising method based on integrated empirical mode decomposition (EEMD) and compared it with the thresholding method of DWT.As can be seen from the above literature, EEMD is often used in signaling, but slowly, it is also used in other fields, such as our tourism industry. EEMD can also be combined with other models for research [53-55].

Forecast and Analysis of The Number of Inbound Passengers in Shanghai:

Grey GM (1,1) Model

With the development of science and technology, information communication becomes very important in people's economic life and scientific teaching and research activities. How to extract, screen and process information effectively has been paid more and more attention. Therefore, the grey prediction is not just an unprovoked occurrence, but a subject emerged at the historic moment.

The uncertainty research object increases the research difficulty, but the research method of uncertainty object needs probability statistics, fuzzy mathematics and grey system. Probability statistics mainly studies random events, finds some laws that look like random events, and then forecasts them, and needs a lot of data. Fuzzy mathematics mainly studies the clear regularity of data in essence, but there is no law to find in data itself. The research of grey system needs little data, and the analysis of few data can get some rules. But probability statistics and fuzzy mathematics need a lot of data, which is the biggest advantage of grey system.

From the beginning of grey system to now, many scholars have participated in the research and improvement. There are many models based on grey system, such as grey equation, grey sequence, grey matrix, grey model and so on. The differential equation used in the grey system is to get the function by using the sequence generating operator after obtaining the data of different time, which provides the premise for the establishment of the model. In this process, the operator enlarges the law of scattered data, improves the degree of data association to increase the certainty. This is the key to using less data in grey system.

The Principle of Grey GM (1,1) Model

The very important module of grey system principle is grey prediction, in which GM model has a very important role, which makes grey prediction go on smoothly and has a very accurate accuracy. The advantage of grey model is that it is not only the difference model of short-term prediction, but also the differential equation of system change. It does not need a lot of data, but only a small amount of discrete data can be used, and then the accumulated generated data will be generated to eliminate a large part. Random errors, find the regular changes in the data, and complete the state prediction. The highlights of grey GM (1,1) model are fewer sample data, easy to understand method, simple algorithm, high accuracy of short-term prediction and easy to test [61-62].

Let the nonnegative original sequence be: $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$ The 1-AGO sequence is as follows $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\}$.

Among them, $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 1,2,\cdots,n$. generating sequence for accumulation $X^{(1)}$ to construct whitening differential equation:

$$\frac{\mathrm{D}x^{(1)}}{\mathrm{dt}} + ax^{(1)} = b\#(1)$$

By solving the equation, we can get: $x^{(1)}(t) = \frac{b}{a} + Ce^{-at}$

with$x^{(1)}(1) = x^{(0)}(1)$Get the time response function of grey GM (1,1) model for the initial condition

$$x^{(1)}(t) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(t-1)} + \frac{b}{a}\#(2)$$

Time series of GM (1,1) model

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak} + \frac{b}{a}, k = 1,2,\cdots,n-1\#(3)$$

Parameter estimation by least square method

$$\hat{\alpha} = \left(\hat{a}, \hat{b}\right)^T = (B^T B)^T B^T Y\#(4)$$

Among them,

$$\mathrm{B} = \begin{pmatrix} -\left(x^{(1)}(2) + x^{(1)}(1)\right)/2 & 1 \\ -\left(x^{(1)}(3) + x^{(1)}(2)\right)/2 & 1 \\ \vdots & \vdots \\ -\left(x^{(1)}(n) + x^{(1)}(n-1)\right)/2 & 1 \end{pmatrix}, Y = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{pmatrix}\#(5)$$

The predicted value of the original data column is reduced by one step

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), k = 1,2,\cdots,n-1\#(6)$$

The restore value of the original sequence is

$$\hat{x}^{(0)}(k+1) = (1 - e^{a})\left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak}, k = 1,2,\cdots,n-1\#(7)$$

Data Sources

Due to the small amount of demand data, this paper selects the number of inbound passengers in Shanghai from 2004 to 2016 as the basis for prediction. In order to facilitate the calculation of the number of inbound passengers, the unit is set to million. The specific data comes from Shanghai Tourism Bureau.

Grade Ratio Test

2004Initial sequence of inbound passengers in Shanghai from 2016 $\{x^{(0)}\}$ as follows:
{4.9192, 5.7135,6.0567,6.6559,6.4037,6.2892,8.5112,8.1757,8.004,7.574,7.913,8.0016,8.5437}

The order ratio of the sequence is calculated by the formula as follows:

$$\sigma(k) = x^{(0)}(k-1)/x^{(0)}(k) \#(8)$$

$$\sigma(k) = \{0.861,0.943,0.91,1.039,1.018,0.739,1.041,1.021,1.057,0.957,0.989,0.93\}$$

Only satisfied$\sigma(k) \in (e^{(-2/n+1)}, e^{(2/n+1)}) = (0.726,1.154)$In order to model and forecast GM (1,1).It can be seen from the order ratio sequence, $\{\sigma(k)\} \subset (0.726,1.154)$Therefore, the initial sequence satisfies the conditions of GM (1,1) model.

Forecasting the Number of Inbound Passengers in Shanghai by Traditional GM (1,1) Model

The traditional GM (1,1) model is used to predict the number of inbound passengers in Shanghai from 2004 to 2016 using R software, as shown in the following table:

Table (3.1): Prediction value of grey model for the number of inbound passengers in Shanghai

particular year	Actual value (million)	Forecast value (million)	Relative error (%)	Average relative error (%)	Posterior difference ratio	Accuracy (%)
2006	6.0567	6.3156	-4.2745	-0.6443	0.3087	99.3557
2007	6.6559	6.5181	2.0697			
2008	6.4037	6.7272	-5.0516			
2009	6.2892	6.9429	-10.3946			
2010	8.5112	7.1656	15.8097			
2011	8.1757	7.3954	9.5439			
2012	8.0040	7.6326	4.6402			
2013	7.5740	7.8774	-4.0057			
2014	7.9130	8.1300	-2.7427			
2015	8.0016	8.3908	-4.8637			
2016	8.5437	8.6599	-1.3598			

According to **(Table 3.1)**, the average relative error is very small, the posterior error is less than 0.35, and the accuracy is 99.3557%. It can be seen from the relative error, the posterior error ratio and the accuracy that the prediction effect of the traditional GM (1,1) model is very good.

According to the traditional GM (1,1) model, the parameters can be estimated

Development coefficient: $-a = 0.0316$ Grey action: $b = 5.868$

Therefore, the prediction formula of grey model is:

$$\hat{x}^{(0)}(k+1) = 5.9292e^{0.0316k}, k = 1,2,\cdots,12\#(9)$$

According to the above formula, we can get the 2017 forecast value of 8.9376 million people. Similarly, adding the 2017 forecast value to the original data, we can get the 2018 forecast value of 9.1753 million people.

Grey Markov Model

Because the development trend of grey model prediction value is very single, it cannot predict the change characteristics of data columns, and Markov model can predict the development trend of data columns. The fitting effect of grey model is not good and the prediction accuracy is low. However, Markov probability matrix can predict the future development situation by the transition probability between States, and can predict the data series with large fluctuation. Therefore, combining the highlights of the two models, the traditional gray Markov model is used to predict the number of inbound passengers in Shanghai.

Markov Model Principle

Markov process was first proposed by mathematician a.a.markov, and then studied by many scholars. Up to now, the theoretical knowledge of Markov process is very rich and sound. Markov forecast is a random change system. The random model is constructed by probability mathematics theory. The change rule of variable state is obtained by the state transition matrix and initial probability of variable. In fact, many processes in reality belong to Markov processes, such as the change process of factory personnel, the process of population growth, the change process of the number of animals in the forest, etc. This process was studied by A.H. Kolmogorov in 1931. Differential equation was proposed to study the process, which laid a solid foundation for Markov process analysis. If there are many states in a sequence, and there is no rule for the transition between the states of the sequence, it is considered that the sequence is in a certain state at present, then it may be transferred to another state at the next moment, which is called Markov process. The following describes the principle of Markov model.

In Markov model, states are divided by relative error [63]

$$v_k = \frac{\hat{x}^{(0)}(k) - x^{(0)}(k)}{x^{(0)}(k)} . 100\% \ (k = 1,2,\cdots,n) \#(10)$$

With $\hat{x}^{(0)}(k)$ is the trend value of item, $x^{(0)}(k)$ is the actual value of item. Obviously, the v_k larger the actual value deviates from the trend value, the worse the effect is; The v_k smaller the value is, the closer the actual value is to the trend value, the better the effect is.

According to v_k , relative error sequence $\mathrm{V} = (v_1, v_2, \cdots, v_n)$It can be divided into s states, if any $v_k \in [a_{1i}, a_{2i}], i = 1,2,\cdots,s$, the state i of the k TH term is E_i, where a_{1i}, a_{2i} represent the lower bound and upper bound of E_i respectively. So, the state set $\mathrm{E} = (E_1, E_1, \cdots, E_s)$ it is generated according to the relative errors. In order to make full use of the latest data and reduce the influence of random errors, this paper constructs the state transition probability matrix and adopts the method of multi-step transition.

w status after step i to state j the transfer probability of is

$$p_{ij}{}^{(w)} = \frac{m_{ij}{}^{(w)}}{M_i} \#(11)$$

$m_{ij}{}^{(w)}$ is the number of states i to state j after w step is moved; M_i is the total number of states i.

The transition probability matrix after $w$is made up of the transition probability $p_{ij}{}^{(w)}$ after w

$$P^w = \begin{bmatrix} p_{11}{}^{(w)} & p_{12}{}^{(w)} & \cdots & p_{1s}{}^{(w)} \\ p_{21}{}^{(w)} & p_{22}{}^{(w)} & \cdots & p_{2s}{}^{(w)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{s1}{}^{(w)} & p_{s2}{}^{(w)} & \cdots & p_{ss}{}^{(w)} \end{bmatrix}, i = 1,2,\cdots,s$$

Select the last s state as the initial state and transfer step w is the distance between the forecast item and selected item. Consider the initial state of s state respectively $(i_1, i_2, \cdots, i_s)$, w Step transition probability $p_i{}^{(w)} = \left(p_{i1}{}^{(w)}, p_{i2}{}^{(w)}, \cdots, p_{is}{}^{(w)}\right), i = 1,2,\cdots,s$, forming the transfer probability matrix $P^{(w)}$. Therefore, we can get the State transition probability matrix of prediction data:

$$\mathrm{R} = \begin{bmatrix} p_{i_1 1}{}^{(w)} & p_{i_1 2}{}^{(w)} & \cdots & p_{i_1 s}{}^{(w)} \\ p_{i_2 1}{}^{(w)} & p_{i_2 2}{}^{(w)} & \cdots & p_{i_2 s}{}^{(w)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{i_s 1}{}^{(w)} & p_{i_s 2}{}^{(w)} & \cdots & p_{i_s s}{}^{(w)} \end{bmatrix}, i = 1,2,\cdots,s$$

By selecting $\max\{p_i = \sum_{w=1}^{s} p_{.i}{}^{(w)}, i = 1,2,\cdots,s\}$ Determine which state the forecast data belongs to, and the upper and lower bounds of this state can be obtained respectively a_{1*} , a_{2*} . Finally, the formula of the predicted value is

$$\hat{y}(t) = \hat{x}^{(0)}(t) \cdot [1 + 0.5(a_{1*} + a_{2*})] \#(12)$$

For section $\mathrm{t} + 1$ The state transition probability matrix is added to the t the status of the item. Then, by reconstructing the probability matrix to achieve equal dimension processing, the next modified prediction value can be obtained. Repeat this method until all predictions are obtained.

State Division

In this paper, the data are divided into five state intervals, and the traditional GM (1,1) model is used to calculate the corresponding prediction value of the original sequence. Find out the corresponding relative error sequence, and determine five state intervals according to the relative error, taking the interval measure as 6%. The status interval is specifically

$$E_1 = [-12\%, -6\%], E_2 = [-6\%, 0], E_3 = [0,6\%],$$
$$E_4 = [6\%, 12\%], E_5 = [12\%, 18\%]$$

These five state intervals are called relative error state intervals, and the state of the initial data is determined based on the initial data and relative error intervals. The specific division is as follows:

Table (3.2): State table based on relative error

particular year	Relative error (%)	state	particular year	Relative error (%)	state
2004	0		2011	9.5439	E_4
2005	-7.1031	E_1	2012	4.6402	E_3
2006	-4.2745	E_2	2013	-4.0057	E_2
2007	2.0697	E_3	2014	-2.7427	E_2
2008	-5.0516	E_2	2015	-4.8637	E_2
2009	-10.3946	E_1	2016	-1.3598	E_2
2010	15.8097	E_5			

Mahala Nobis Test

Using MATLAB software, the transfer frequency matrix and probability transfer matrix are as follows:

$$f^{(1)} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 3 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad P^{(1)} = \begin{bmatrix} 0 & 0.5 & 0 & 0 & 0.5 \\ 0.2 & 0.6 & 0.2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Marginal probability formula $P_{.j} = \sum_{j=1}^{5} f_{ij} / \sum_{i=1}^{5} \sum_{j=1}^{5} f_{ij}$ Therefore, the marginal probability values are $P_{.1} = 1/11, P_{.2} = 6/11, P_{.3} = 2/11, P_{.4} = 1/11, P_{.5} = 1/11$. According to the obtained probability transfer matrix, the statistical values are calculated as follows:

Table (3.3): Statistic $\chi^2 = 2\sum_{i=1}^{5}\sum_{i=1}^{5} f_{ij} \left|\ln \frac{P_{ij}}{P_{.j}}\right|$ Calculation table

state	$f_{i1}\left\|\ln \frac{P_{i1}}{P_{.1}}\right\|$	$f_{i2}\left\|\ln \frac{P_{i2}}{P_{.2}}\right\|$	$f_{i3}\left\|\ln \frac{P_{i3}}{P_{.3}}\right\|$	$f_{i4}\left\|\ln \frac{P_{i4}}{P_{.4}}\right\|$	$f_{i5}\left\|\ln \frac{P_{i5}}{P_{.5}}\right\|$	total
E_1	0	0.087	0	0	1.705	1.792
E_2	0.788	0.286	0.095	0	0	1.170
E_3	0	1.212	0	0	0	1.212
E_4	0	0	1.705	0	0	1.705
E_5	0	0	0	2.398	0	2.398
total	0.788	1.585	1.800	2.398	1.705	8.276

Calculate the statistics according to the formula $\chi^2 = 2 \times 8.276 = 16.553$ The significant level was selected $\alpha = 0.5$ And since the number of States is 5, it can be seen from the level table ${\chi^2}_{0.5}((m-1)^2) = {\chi^2}_{0.5}(16) = 15.338$ Statistics can be obtained by comparison $\chi^2 > {\chi^2}_{0.5}(16)$ Therefore, the number of inbound tourists in Shanghai from 2004 to 2016 is Markov and Markov model can be established[64].

Application of Grey Markov Model

The transfer probability matrix corresponding to the step size of 2, 3, 4, 5 is calculated [19]

$$P^{(2)} = \begin{bmatrix} 0.1 & 0.3 & 0.1 & 0.5 & 0 \\ 0.12 & 0.66 & 0.12 & 0 & 0.1 \\ 0.2 & 0.6 & 0.2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$P^{(3)} = \begin{bmatrix} 0.06 & 0.33 & 0.56 & 0 & 0.05 \\ 0.132 & 0.576 & 0.132 & 0.1 & 0.06 \\ 0.12 & 0.66 & 0.12 & 0 & 0.1 \\ 0.2 & 0.6 & 0.2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$P^{(4)} = \begin{bmatrix} 0.066 & 0.788 & 0.066 & 0.05 & 0.03 \\ 0.115 & 0.544 & 0.215 & 0.06 & 0.066 \\ 0.132 & 0.576 & 0.132 & 0.1 & 0.06 \\ 0.12 & 0.66 & 0.12 & 0 & 0.1 \\ 0.2 & 0.6 & 0.2 & 0 & 0 \end{bmatrix}$$

$$P^{(5)} = \begin{bmatrix} 0.158 & 0.572 & 0.208 & 0.03 & 0.033 \\ 0.109 & 0.599 & 0.169 & 0.066 & 0.058 \\ 0.115 & 0.544 & 0.215 & 0.06 & 0.066 \\ 0.132 & 0.576 & 0.132 & 0.1 & 0.06 \\ 0.12 & 0.66 & 0.12 & 0 & 0.1 \end{bmatrix}$$

According to the above transition probability matrix and the last five years' state, the new transition probability matrix in 2017 is predicted as follows:

$$R^{(2017)} = \begin{bmatrix} 0.2 & 0.6 & 0.2 & 0 & 0 \\ 0.12 & 0.66 & 0.12 & 0 & 0.1 \\ 0.132 & 0.576 & 0.132 & 0.1 & 0.06 \\ 0.115 & 0.544 & 0.215 & 0.06 & 0.066 \\ 0.115 & 0.544 & 0.215 & 0.06 & 0.066 \end{bmatrix}$$

In $R^{(2017)}$, the sum of the numbers in the second column is the largest, so 2017 belongs to E_2 and the forecast value in 2017 is:

$$\hat{y}(2017) = \hat{x}^{(0)}(2017) \cdot [1 + 0.5(a_{12} + a_{22})] = 8.8035\#(13)$$

Similarly, adding the 2017 forecast to the raw data gives the 2018 status E_2 and the forecast for 2018 is 9.0377.

Improved grey Markov Model

GM (1,1) Model Principle of Initial Value Correction

Tradition GM (1,1) model $x^{(1)}(1) = x^{(0)}(1)$ As the initial condition, the information brought by the new data is lost, so the final prediction is generated by the initial condition and the new accumulation to improve the accuracy. The GM (1,1) model can be improved by modifying the initial value method, which can greatly improve the model fitting and prediction accuracy. The model is as follows [65]:

Separate order $t = 1, n$, It can be concluded that:

$$\begin{cases} x^{(1)}(1) = C_1 e^{-a} + \dfrac{b}{a} \\ x^{(1)}(n) = C_2 e^{-an} + \dfrac{b}{a} \end{cases} \#(14)$$

Order $\mathrm{C} = \frac{1}{2}(C_1 + C_2)$, get The expression is:

$$\mathrm{C} = \frac{1}{2}\left[\left(x^{(1)}(1) - \frac{b}{a}\right)e^{a} + \left(x^{(1)}(n) - \frac{b}{a}\right)e^{an}\right] \#(15)$$

The restore value of the original sequence is

$$\hat{x}^{(0)}(k+1) = \frac{1}{2}(e^{-a} - 1)\left[\left(x^{(1)}(1) - \frac{b}{a}\right)e^{a} + \left(x^{(1)}(n) - \frac{b}{a}\right)e^{an}\right]e^{-ak}$$
$$k = 1,2,\cdots,n-1\#\#(16)$$

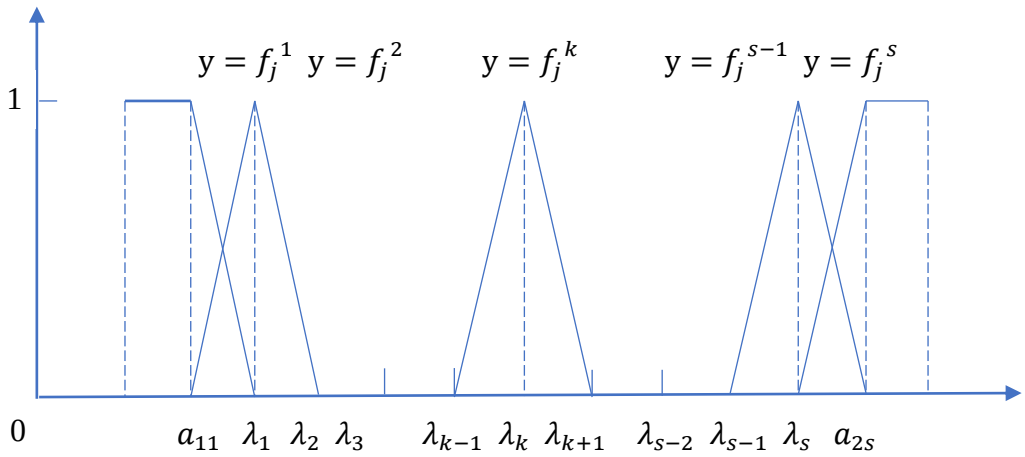
The Principle of Improved Markov Model with Centre Point Triangular Whitening Weight Function

According to equation (8), the relative error sequence $\mathrm{V} = (v_{1,} v_{2,} \cdots, v_n)$, by The relative error sequence is divided into s states. For the division of each state interval, factors such as the range of relative error, the distribution of each item in the range, the division between different states and the convenience of calculation must be considered.

Considering the condition of traditional Markov chain S-state partition, the preference degree of each fluctuation index is not reflected. For example, there are three state intervals: [2,5], [5,10], [10,20].5.01 belongs to the second state, 9.99 also belongs to the second state. However, compared with the central value of this state (7.5), 5.01 is closer to the first state interval, and 9.99 is closer to the third state interval than 7.5. In the traditional state division process, it is not objective that they are in the same state. Considering the preference degree of each fluctuation index synthetically in two adjacent intervals, the triangle whitening weight function of the center point represents the possibility that the object belongs to a certain state, which can compensate for the subjective division of the state. λ_i it's No i The center point of each state means the maximum possibility of each state interval. Generally speaking, let's $\lambda_i = 0.5(a_{1i} + a_{2i}), i = 1,2,\cdots,s$. On the basis of the weight function of the centre point triangle whitening, the left and

right endpoints should be respectively from the λ_1 and λ_2 Extend horizontally to both sides. For the left endpoint, extend the line to [66] $x = a_i$ Axis; and the right endpoint should extend to $x = b_i$ Axis. This means that the relative error of a certain data is less than λ_1. It will completely belong to the first state. Similarly, the relative error of a certain data is greater than λ_2 will be completely in s state. Therefore, the centre triangle whitening weight function has more normative and practical significance (as shown in the figure below) [67].

Figure (3.1): Weight function of triangle whitening at center point



For section i the center point triangle whitening weight function is defined as follows:

$$f^1(V_k) = \begin{cases} 0, v'_k \notin [a_{11}, \lambda_{i+1}] \\ 1, v'_k \in [a_{11}, \lambda_i] \\ \dfrac{\lambda_{i+1} - v'_k}{\lambda_{i+1} - \lambda_i}, v'_k \in [\lambda_i, \lambda_{i+1}] \end{cases}, i = 1$$

$$f^i(V_k) = \begin{cases} 0, v'_k \notin [\lambda_{i-1}, \lambda_{i+1}] \\ \dfrac{v'_k - \lambda_{i-1}}{\lambda_i - \lambda_{i-1}}, v'_k \in [\lambda_{i-1}, \lambda_i] \\ \dfrac{\lambda_{i+1} - v'_k}{\lambda_{i+1} - \lambda_i}, v'_k \in [\lambda_i, \lambda_{i+1}] \end{cases}, i = 2,3, \cdots, s-1 \#(17)$$

$$f^s(V_k) = \begin{cases} 0, v'_k \notin [\lambda_{i-1}, a_{2s}] \\ \dfrac{v'_k - \lambda_{i-1}}{\lambda_i - \lambda_{i-1}}, v'_k \in [\lambda_{i-1}, \lambda_i], i = s \\ 1, v'_k \in [\lambda_i, a_{2s}] \end{cases}$$

According to formula (16), in different states k function values are $\sigma_k = \left(f^1(v_k), f^2(v_k), \cdots f^s(v_k)\right), k = 1,2, \cdots, n$ .Therefore, the state probability matrix is

$$\sigma = \begin{bmatrix} f^1(v_1) & f^2(v_1) & \cdots & f^s(v_1) \\ f^1(v_2) & f^2(v_2) & \cdots & f^s(v_2) \\ \vdots & \vdots & \cdots & \vdots \\ f^1(v_n) & f^2(v_n) & \cdots & f^s(v_n) \end{bmatrix}$$

w step size from state i to state j the transfer probability of ${p_{ij}}'^{(w)}$

$$ {p_{ij}}'^{(w)} = \sum_{k=1}^{n-w} \left[\frac{f^i(v_k)}{\sum_{k=1}^{n-w} f^i(v_k)} \cdot f^j(v_{k+w}) \right] \#(18) $$

w step transition probability matrix is determined by ${p_{ij}}'^{(w)}$ form

$$ P'^{(w)} = \begin{bmatrix} {p_{11}}'^{(w)} & {p_{12}}'^{(w)} & \cdots & {p_{1s}}'^{(w)} \\ {p_{21}}'^{(w)} & {p_{22}}'^{(w)} & \cdots & {p_{2s}}'^{(w)} \\ \vdots & \vdots & \cdots & \vdots \\ {p_{s1}}'^{(w)} & {p_{s2}}'^{(w)} & \cdots & {p_{ss}}'^{(w)} \end{bmatrix}, i = 1,2,\cdots,s $$

Choose the last s the transition step size is $1,2,\cdots$ s。 Considering the state bias of each initial term $(f^1(v_k)E_1 + f^2(v_k)E_2 + \cdots + f^s(v_k)E_s, k = n-s+1, n-s+2, \cdots, n)$, the kth term corresponds to the state bias vector multiplied by the transition probability vector ${p_i}'^{(w)} = \left({p_{i1}}'^{(w)}, {p_{i2}}'^{(w)}, \cdots, {p_{is}}'^{(w)}\right), i = 1,2,\cdots s$, and then sum. w prediction after step length t the transfer probability vector of data is

$$ {r_k}'^{(w)} = \left({r_{k1}}'^{(w)}, {r_{k2}}'^{(w)}, \cdots, {r_{ki}}'^{(w)}\right) = f^1(v_k){p_1}'^{(w)} + f^2(v_k){p_2}'^{(w)} + \cdots + f^j(v_k){p_s}'^{(w)} $$

Therefore, it is predicted that t the transition probability matrix under the new state is

$$ R' = \begin{bmatrix} {r_{n-s+1}}'^{(1)} \\ {r_{n-s+2}}'^{(2)} \\ \vdots \\ {r_n}'^{(s)} \end{bmatrix} $$

To eliminate cumulative effects, states i the standardized probability prediction data of can be described as follows:

$$ q_i = \frac{\sum_{n-s+1}^{n} {r_{ki}}'^{(w)}}{s}, i = 1,2,\cdots,s \#(19) $$

Finally, no t the predicted value of data is

$$ \hat{y}'(t) = \hat{x}'^{(0)}(t) \cdot [1 + \sum_{i=1}^{m}(q_i \cdot \lambda_i)] \#(20) $$

For the prediction of (t+1), the state of item t is added to the state transition probability matrix. Then, the probability matrix is reconstructed to realize the equal dimension processing, and the next modified prediction value is obtained. Repeat this method until all predictions are obtained.

Application of Improved Grey Markov Model

The GM (1,1) model is improved by the initial value modification method, and the Markov model is improved by the central point triangular whitening weight function. Then the two models are combined into the improved grey Markov model to predict the number of inbound passengers in Shanghai.

Using R software and GM (1,1) model of initial value correction to predict the number of inbound passengers in Shanghai from 2004 to 2016, as shown in the following table:

Table (3.4): Forecast value of initial value modified grey model of Shanghai inbound passengers

particular year	Actual Value (million)	Forecast (million)	Relative error (%)	Average relative error	Posterior difference ratio	Accuracy (%)
2004	4.9192	4.9192	0	-0.6221	0.3089	99.3779
2005	5.7135	6.1078	-6.9010			
2006	6.0567	6.3152	-4.2683			
2007	6.6559	6.5178	2.0755			
2008	6.4037	6.7268	-5.0454			
2009	6.2892	6.9425	-10.3881			
2010	8.5112	7.1652	15.8146			
2011	8.1757	7.3950	9.5492			
2012	8.0040	7.6322	4.6458			
2013	7.5740	7.8769	-3.9996			
2014	7.9130	8.1296	-2.7367			
2015	8.0016	8.3903	-4.8575			
2016	8.5437	8.6594	-1.3538			

According to table 3.4, the average relative error is very small, the posterior error is less than 0.35, and the accuracy is 99.3779%. From the three aspects of relative error, posterior difference ratio and accuracy, the prediction effect of initial value modified GM (1,1) model is better than that of traditional model GM (1,1).

According to the GM (1,1) model modified by initial value, the estimated value of parameters can be obtained:

Development coefficient: $-a = 0.0316$ Grey action amount: $b = 5.868$

Therefore, the prediction formula of grey model is:

$$\hat{x}'^{(0)}(k+1) = 5.9289e^{0.0316k}, k = 1,2,\cdots,12\#(21)$$

According to the above formula, we can get the 2017 forecast value of 8.9371 million people. Similarly, adding the 2017 forecast value into the original data, we can get the 2018 forecast value of 9.135 million people.

According to the initial value, the relative error of grey model prediction is modified to determine the initial data state. The specific state interval is

$$E_1 = [-12\%, -6\%], E_2 = [-6\%, 0], E_3 = [0,6\%],$$
$$E_4 = [6\%, 12\%], E_5 = [12\%, 18\%]$$

According to the interval range of each state, the center point is determined as $\lambda_1 = -9\%, \lambda_2 = -3\%, \lambda_3 = 3\%, \lambda_4 = 9\%, \lambda_5 = 15\%$. The center point triangle whiteness weight function is as follows:

$$f^1(V_k) = \begin{cases} 0, v'_k \notin [-12\%, \lambda_2] \\ 1, v'_k \in [-12\%, \lambda_1] \\ \dfrac{-3\% - v'_k}{-3\% - (-9\%)}, v'_k \in [\lambda_1, \lambda_2] \end{cases}, i = 1$$

$$f^i(V_k) = \begin{cases} 0, v'_k \notin [\lambda_{i-1}, \lambda_{i+1}] \\ \dfrac{v'_k - \lambda_{i-1}}{\lambda_i - \lambda_{i-1}}, v'_k \in [\lambda_{i-1}, \lambda_i] \\ \dfrac{\lambda_{i+1} - v'_k}{\lambda_{i+1} - \lambda_i}, v'_k \in [\lambda_i, \lambda_{i+1}] \end{cases}, i = 2,3,4\#(22)$$

$$f^5(V_k) = \begin{cases} 0, v'_k \notin [\lambda_4, 18\%] \\ \dfrac{v'_k - 9\%}{15\% - 9\%}, v'_k \in [\lambda_4, \lambda_5], i = 5 \\ 1, v'_k \in [\lambda_5, 18\%] \end{cases}$$

According to the relative error state division in table 3.2, the clustering coefficient matrix is obtained

$$\sigma = \begin{bmatrix} 0.65 & 0.35 & 0 & 0 & 0 \\ 0.211 & 0.789 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0.726 & 0.274 & 0 & 0 \end{bmatrix}$$

In order to express the different real meaning of the elements in the cluster coefficient matrix, it is shown in the following table:

Table (3.5): State partition of Improved Grey Markov model

particular year	state	particular year	state
2005	$0.65E_1$+$0.35E_2$	2011	$0.908E_4$+$0.092E_5$
2006	$0.211E_1$+$0.789E_2$	2012	$0.726E_3$+$0.274E_4$
2007	$0.154E_2$+$0.846E_3$	2013	$0.167E_1$+$0.833E_2$
2008	$0.341E_1$+$0.659E_2$	2014	$0.956E_2$+$0.044E_3$
2009	E_1	2015	$0.31E_1$+$0.69E_2$
2010	E_5	2016	$0.726E_2$+$0.274E_3$

According to the triangle whitening weight function of the center point of each state, the 1-5 step transition probability matrix is as follows:

$$P'^{(1)} = \begin{bmatrix} 0.179 & 0.347 & 0.101 & 0 & 0.373 \\ 0.244 & 0.554 & 0.202 & 0 & 0 \\ 0.262 & 0.738 & 0 & 0 & 0 \\ 0.039 & 0.193 & 0.557 & 0.211 & 0 \\ 0 & 0 & 0.061 & 0.855 & 0.084 \end{bmatrix}$$

$$P'^{(2)} = \begin{bmatrix} 0.052 & 0.15 & 0.232 & 0.383 & 0.183 \\ 0.182 & 0.493 & 0.149 & 0 & 0.176 \\ 0.524 & 0.449 & 0.027 & 0 & 0 \\ 0.128 & 0.862 & 0.01 & 0 & 0 \\ 0.014 & 0.07 & 0.665 & 0.251 & 0 \end{bmatrix}$$

$$P'^{(3)} = \begin{bmatrix} 0.183 & 0.232 & 0.326 & 0.247 & 0.013 \\ 0.326 & 0.3 & 0.082 & 0.215 & 0.077 \\ 0.143 & 0.319 & 0 & 0 & 0.538 \\ 0.072 & 0.894 & 0.034 & 0 & 0 \\ 0.153 & 0.844 & 0.004 & 0 & 0 \end{bmatrix}$$

$$P'^{(4)} = \begin{bmatrix} 0.371 & 0.378 & 0.112 & 0.043 & 0.096 \\ 0.179 & 0 & 0.245 & 0.164 & 0.411 \\ 0 & 0.335 & 0.127 & 0.489 & 0.049 \\ 0.238 & 0.699 & 0.064 & 0 & 0 \\ 0.026 & 0.934 & 0.04 & 0 & 0 \end{bmatrix}$$

$$P'^{(5)} = \begin{bmatrix} 0.026 & 0.563 & 0.02 & 0.087 & 0.304 \\ 0.056 & 0.281 & 0.057 & 0.389 & 0.216 \\ 0 & 0 & 0.726 & 0.274 & 0 \\ 0 & 0.726 & 0.724 & 0 & 0 \\ 0.284 & 0.693 & 0.023 & 0 & 0 \end{bmatrix}$$

According to the above transition probability matrix and the last five years' state, the new transition probability matrix in 2017 is predicted as follows:

$$R'^{(2017)} = \begin{bmatrix} 0.201 & 0.589 & 0.153 & 0.058 & 0 \\ 0.16 & 0.435 & 0.163 & 0.064 & 0.177 \\ 0.318 & 0.301 & 0.079 & 0.206 & 0.097 \\ 0.239 & 0.117 & 0.204 & 0.127 & 0.314 \\ 0.041 & 0.204 & 0.241 & 0.357 & 0.157 \end{bmatrix}$$

2017The forecast values for the year are as follows:

$$\hat{y}'(2017) = \hat{x}'^{(0)}(2017) \cdot \left[1 + \sum_{i=1}^{5} (q_i \cdot \lambda_i)\right] = 8.6944\#(23)$$

Similarly, adding the 2017 forecast to the original data gives a forecast of 8.9135 million people in 2018.

Comparative Analysis of Results

Using the linear regression prediction model, this paper forecasts the number of inbound tourists in Shanghai from 2004 to 2016

$$y = 0.2662x + 5.2718\#(24)$$

On the test of equations, the standard deviation of residuals $\hat{\sigma} = 0.5984$, the square of the correlation coefficient$R^2 = 0.7661$, the P value of F distribution is 8.895×10^{-5} . Therefore, it is very significant.

The prediction model of time series is used to predict the number of inbound tourists in Shanghai from 2004 to 2016. The prediction model formula of time series is

$$Y_t = 0.3984 + 0.8469Y_{t-1} + \varepsilon_t - \varepsilon_{t-1}\#(25)$$

The p value of the unit root statistic ADF is 0, which means that there is no unit root, that is, the data is stable. The P values of AR (1) and MA (1) are both less than 0.05, and the parameters are significant. Using linear regression and time series prediction models, the number of inbound tourists in Shanghai from 2004 to 2016 is predicted. The number of inbound tourists in 2017 is 8.9991 million, that in 2018 is 9.2653 million, that in 2017 is 9.1005 million, and that in 2018 is 9.6697 million. The prediction results are compared with our four models.

Compare the predicted value of 2017 and 2018 of linear regression, time series and the four models in this paper with the real value and calculate the relative error, which is (predicted value real value) / real value. See table 3.6 for the results:

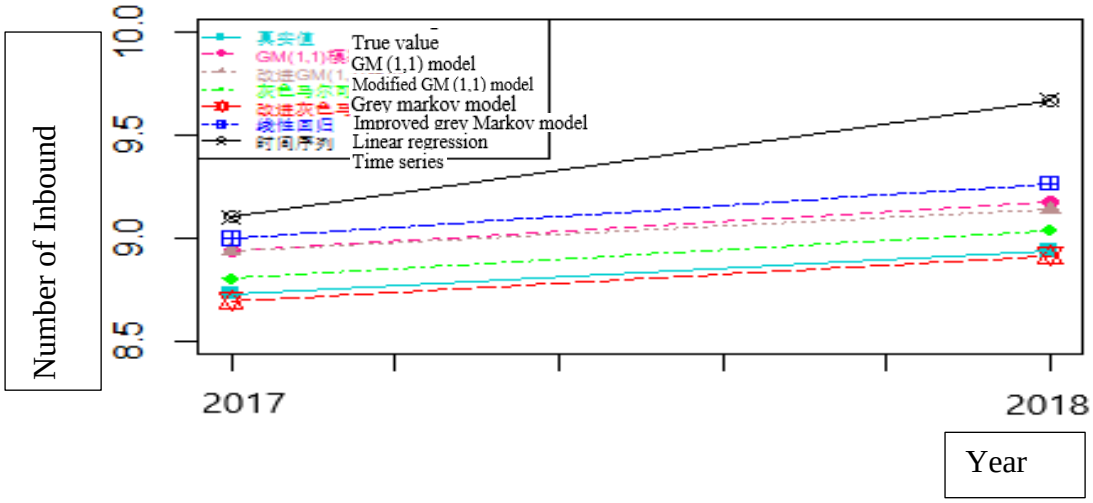
Table (3.6): Comparison of prediction results

particular year	True value	time series	relative error	linear regression	relative error
2017	8.7301	9.1005	4.243%	8.9991	3.081%
2018	8.9371	9.6697	8.197%	9.2653	3.672%
Average relative error	-	-	6.220%	-	4.901%
particular year	True value	GM (1,1) model	Relative error	Improved GM (1,1) model	Relative error
2017	8.7301	8.9376	2.377%	8.9371	2.371%
2018	8.9371	9.1753	2.665%	9.135	2.214%
Average relative error	-	-	2.521%	-	2.293%
particular	True value	grey markov	relative	Improved grey	relative

year		model	error	Markov model	error
2017	8.7301	8.8035	0.841%	8.6944	-0.409%
2018	8.9371	9.0377	1.126%	8.9135	-0.264%
Average relative error	-	-	0.983%	-	-0.336%

Linear regression, time series and the four models in this paper are used to compare the predicted and real values in 2017 and 2018, as shown in **(Figure 3.2)**

Figure (3.2): Comparison between the predicted value and the real value in 2017 and 2018



According to (Table 3.6 and Figure 3.2), the prediction results of linear regression and time series are larger than the prediction results of a series of gray models in this paper, and are the least close to the true value. The relative error values of GM (1,1) model, improved GM (1,1) model, gray markov model and improved gray markov model in turn with the real value are smaller and smaller, and it is also closer and closer to the real value in turn from the comparison diagram. Therefore, it is concluded that the prediction effect of the model is getting better and better in turn, while the prediction result of the improved grey Markov model is the best and closest to the real value.

Summary of This Part

To sum up, this paper focuses on two key improvements of Grey Markov model. In grey model prediction, the initial value is used to modify GM (1,1) model to reduce the error. In the process of Markov model, the possibility of state is determined by the center point triangle whitening weight function, which makes up for the subjective defect of state division. State possibility is also called preference degree in two adjacent intervals, which evaluate different possibilities in an interval in a more objective way. In addition, by calculating the average value of each interval as the center point, it is easy to get the triangle whitening weight function of the center point, thus establishing the function. In the case study, according to the relative error, the grey Markov model based on the initial value correction and center point triangle whitening weight function is superior to the traditional GM (1,1), initial value modified GM (1,1) and the traditional grey Markov model, and a series of grey models in this paper have better prediction results than linear regression and time series. Therefore, the model in this paper is verified in theory and practice.

An Analysis of The Market Structure of Inbound Tourists in Shanghai:

Deviation Share Analysis

The shift share analysis method regards the regional economy as a changing process, takes the regional or national economic development as a reference, and divides the regional economic aggregate in a certain period of time into share component, structure deviation component and competitiveness deviation component. It is used to evaluate the advantages and disadvantages of regional economic structure as well as the size of competitiveness, and obtain relatively competitive sectors, so as to specify an optimal direction for future economic development and provide methods for industrial structure adjustment.

The shift share analysis method is used to analyse the structure of tourist market. A region is selected as the sample and the upper level of the whole country is taken as the reference frame. The share component, structure transfer component and competitiveness transfer component are used to analyse the trend of tourism market change. It is used to explain why the regional economic development and decline, and to evaluate the advantages and disadvantages of the regional inbound tourism market structure Market competitiveness of inbound tourists. Objective and quantitative to find a competitive tourism market, for the region's future development of inbound tourism market points out good ideas [68].

Model Principle of Deviation Share Analysis

According to the PRINCIPLE of SSM model [69], after a period of $[0, t]$ between region i and country, the total amount of inbound tourism and the tourism market structure change. The total number of inbound tourists in the initial region i is b_{i0}, and the total number of inbound tourists in the final region is b_{it}. At the same time, region $i$divides into $j$markets according to the inbound tourists. $b_{ij,0}$, $b_{ij,t}$ $(j = 1,2,\cdots,n)$ are used to represent the number of inbound tourists from the jth inbound tourist source market to region I in the early and late stages, and B_{j0} and B_{jt} are used to represent the total number of tourists from the jth inbound tourist source market corresponding to the reference area of region i in the early and late stages.

The rate of change at $[0, t]$ of the jth tourist source in region i is:

$$r_{ij} = \left(b_{ij,t} - b_{ij,0}\right)/b_{ij,0} \ (j = 1,2,\cdots,n,) \tag{26}$$

The rate of change at $[0, t]$ of the jth source in the reference area is:

$$R_j = (B_{j,t} - B_{j,0})/B_{j,0} \tag{27}$$

The standardized value of passenger volume of the jth source market in region i based on the share of the reference region:

$$b'_{ij,0} = (b_{i,0} \times B_{j,0})/B_0 \tag{28}$$

The growth amount of inbound tourists in region i in the jth source market in the time period $[0, t]$ G_{ij} can be decomposed into three components, namely the share component N_{ij}, the source structure deviation component P_{ij} and the market competition deviation component D_{ij}, which can be expressed as:

$$G_{ij} = N_{ij} + P_{ij} + D_{ij} \tag{29}$$

Among them:$N_{ij} = b'_{ij,0} \times R_j$; $P_{ij} = (b_{ij,0} - b'_{ij,0}) \times R_j$;

$D_{ij} = (r_{ij} - R_j) \times b_{ij,0}$.

$$PD_{ij} = P_{ij} + D_{ij} \tag{30}$$

Suppose the total growth amount of inbound tourism market passengers in the ith region is G_i, then the total share component N_i, structure offset component P_i and competitiveness offset component D_i can be expressed as follows:

$$G_i = b_{i,t} - b_{i,0} = N_i + P_i + D_i \tag{31}$$

Among them:

$N_i = \sum_{j=1}^{n} b'_{ij,0} \times R_j$; $P_i = \sum_{j=1}^{n} (b_{ij,0} - b'_{ij,0}) \times R_j$; $D_i = \sum_{j=1}^{n} b_{ij,0} \times (r_{ij} - R_j)$.

Passenger Flow In region i at the next higher level:

$$\mathrm{L} = \frac{b_{i,t}}{b_{i,0}} \Big/ \frac{B_t}{B_0} \tag{32}$$

Introduction:

$$M_{j,0} = b_{ij,0} / B_{j,0} \tag{33}$$

$$M_{j,t} = b_{ij,t} / B_{j,t} \tag{34}$$

Respectively represent the proportion of the passenger flow from initial source j to region i in the corresponding national passenger flow from corresponding source in the same period, divide L into structural effect index W and competitiveness effect index U, and get:

$$\mathrm{L} = \mathrm{W} \times \mathrm{U} \tag{35}$$

Among them: $\mathrm{W} = \frac{\sum_{j=1}^{n} M_{j,0} \cdot B_{j,t}}{\sum_{j=1}^{n} M_{j,0} \cdot B_{j,0}} \Big/ \frac{\sum_{j=1}^{n} B_{j,t}}{\sum_{j=1}^{n} B_{j,0}}$; $\mathrm{U} = \frac{\sum_{j=1}^{n} M_{j,t} \cdot B_{j,t}}{\sum_{j=1}^{n} M_{j,0} \cdot B_{j,t}}$.

According to the above formula, if G_i is larger, $\mathrm{L} > 1$ means that the growth rate of tourists in region i is faster than that in China. On the contrary, the smaller G_i is, $\mathrm{L} < 1$ means that the growth rate of tourists in region i is slower than that in China.

If the P_i is larger, $W > 1$, it indicates that the number of tourists in region i is in a good growth stage for the whole country, while the number of tourists in region I is in a growth stage for the whole country. On the contrary, the P_i is smaller, $W < 1$, indicating that the growth situation is generally poor, and the source structure needs to be adjusted appropriately.

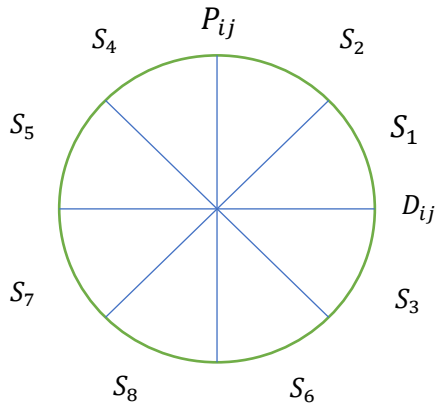
If D_i is larger, $U > 1$ indicates that region i has a fast development speed and strong competitiveness. On the contrary, the smaller D_i is, $U < 1$ indicates that the tourist volume of each source area of region I does not have a strong growth momentum and competitive advantage.

Draw Shift Share Analysis Chart

According to the data in the analysis table, the comparison and classification of the source areas are carried out, and the deviation share analysis chart is made, so that the results are clearer and the types of each source area can be determined. By analyzing two bisectors with 45 degrees of inclination angle, the coordinate system is divided into eight sectors.

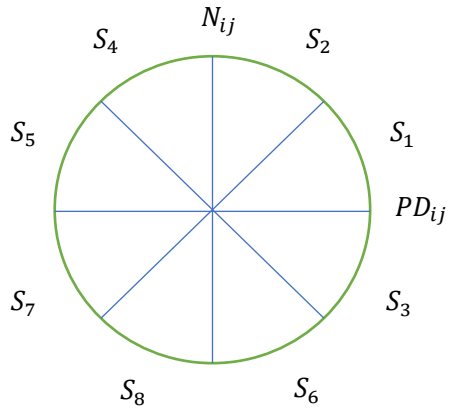
Each tourist source area and the whole area are marked on the coordinate system. According to their sectors, they can be divided into several types to evaluate the overall structure and competitiveness of the region, and then determine which regions are competitive advantages. At this time, the analysis chart can also be used to compare various regions to find out the good and bad structure and competitiveness level [70].

Figure (4.1): Deviation component analysis of inbound tourist source structure



For the analysis of the deviation components of the structure of inbound tourists, firstly, the D_{ij} and P_{ij} the values are plotted in the coordinate system and classified according to the sector of each inbound tourist destination in the coordinate system. Sector 1 in **(Figure 4.1)** refers to the better source areas with good original foundation and strong competitiveness; sector 2 refers to the better source areas with strong competitiveness and good original foundation; sector 3 refers to the better or general source areas with poor foundation and rapid development; sector 4 refers to the better or general source areas with better foundation but declining status; sector 5 refers to the poor source areas with good foundation but poor competitiveness; sector 6 refers to the base poor and fast developing tourist source areas; sector 7 and 8 mainly refer to the worst tourist source areas with poor foundation and lack of competitiveness.

Figure (4.2): Advantage analysis of inbound tourist source structure



For the advantage analysis chart of inbound tourist source structure, the N_{ij} and PD_{ij} In the numerical coordinate system, it is classified according to the sector of each inbound tourist source in the coordinate system. In **(Figure 4.2)**, those belonging to sectors 1 and 2 are better; those belonging to sectors 3 and 4 are general; those belonging to sectors 5 and 6 are poor; and those in sectors 7 and 8 are the worst.

Analysis on The Market Structure of Inbound Tourists in Shanghai

Division of Data Sources and Tourist Sources

The data is from the Shanghai Statistical Yearbook. The number of inbound tourists from different countries in Shanghai from 2004 to 2017 was collected from 12 sample markets of source countries, including Japan, Singapore, Germany, France, the United Kingdom, Italy, Canada, the United States and Australia, as well as Hong Kong, Macao and Taiwan compatriots.

As a sample of national inbound tourist market, the upper level region uses Excel to process data and compare the data at the early stage (2004) and the late stage (2017) to study the development and structure of Shanghai inbound tourist market, as shown in **(table 4.1)**.

Table (4.1): Structure data of inbound tourist market in Shanghai and China

Tourist destination	National inbound passenger volume / 10000		Shanghai inbound passengers / 10000			
	2004year $B_{j,0}$	2017year $B_{j,t}$	2004year $b_{ij,0}$	Structure proportion/%	2017year $b_{ij,t}$	Structure proportion /%
Japan	333.43	268.30	120.67	24.53	116.68	13.37
Singapore	63.68	94.12	12.74	2.59	23.36	2.68
Germany	36.53	63.55	16.42	3.34	33.65	3.85
France	28.11	49.47	8.98	1.83	22.17	2.54
Britain	41.81	59.18	9.52	1.94	24.33	2.79
Italy	12.24	28.05	4.81	0.98	12.35	1.41
Canada	34.80	80.60	6.08	1.24	25.08	2.87
U.S.A	130.86	231.29	35.46	7.21	95.88	10.98
Australia	37.63	73.43	7.75	1.58	25.34	2.90
Hong Kong and Macao compatriots	8842.05	10444.59	48.95	9.95	73.64	8.44
Taiwan compatriots	368.53	587.13	71.75	14.59	128.16	14.68
Other countries	974.15	1968.29	148.79	30.25	292.37	33.49
total	10903.82	13948.00	491.92	100.00	873.01	100.00

Shift Share Analysis of Shanghai Inbound Tourist Market

According to the above formula, using the data in **(table 4.1)** and SSM principle to calculate, the shift share analysis table of Shanghai inbound tourist source market **(table 4.2, table 4.3)** is obtained, and the calculation results of different stages are obtained. The analysis table consists of three parts [71].

(1) Raw data: $b_{ij,0}$, $b_{ij,t}$, $B_{j,0}$, $B_{j,t}$

(2) Intermediate results: r_{ij}, R_j, $b'_{ij,0}$, $b_{ij,0} - b'_{ij,0}$, $r_{ij} - R_j$

(3) Final analysis results: G_{ij}, N_{ij}, P_{ij}, D_{ij}, PD_{ij}

Table (4.2): SSM analysis of Shanghai inbound tourist market

Source of tourists	National change rate R_j	Rate of change in Shanghai r_{ij}	Shanghai standardization scale $b'_{ij,0}$	Share component N_{ij}	Mechanism deviation component P_{ij}	Competitiveness deviation component D_{ij}	Proportion of base period in China $M_{j,0}$	Proportion in China at the end of the period $M_{j,t}$
Japan	-0.20	-0.03	15.04	-2.94	-20.63	19.58	0.3619	0.4349
Singapore	0.48	0.83	2.87	1.37	4.72	4.53	0.2001	0.2482
Germany	0.74	1.05	1.65	1.22	10.93	5.08	0.4495	0.5295
France	0.76	1.47	1.27	0.96	5.86	6.37	0.3195	0.4482
Britain	0.42	1.56	1.89	0.78	3.17	10.85	0.2277	0.4111
Italy	1.29	1.57	0.55	0.71	5.50	1.33	0.3930	0.4403
Canada	1.32	3.13	1.57	2.07	5.94	11.00	0.1747	0.3112

U.S. A	0.77	1.70	5.90	4.53	22.68	33.21	0.2710	0.4145
Australia	0.95	2.27	1.70	1.62	5.76	10.22	0.2060	0.3451
HK and Mc compatriots	0.18	0.50	398.90	72.30	-63.43	15.82	0.0055	0.0071
Taiwan compatriots	0.59	0.79	16.63	9.86	32.70	13.85	0.1947	0.2183
Other countries	1.02	0.96	43.95	44.85	106.99	-8.26	0.1527	0.1485

Table (4.3): overall effect of Shanghai inbound tourism market structure on China

project	Total passenger growth G_{ij}/Ten thousand people	Relative growth rate L	Structure effect index W	Competitiveness effect index U	Total score component N_{ij}	Total structural offset component P_{ij}	Total competitiveness offset component D_{ij}	Offset component PD_{ij}
numerical value	381.09	1.39	1.19	1.16	137.34	120.18	123.57	243.75

General Analysis

As can be seen from **(table 4.1)**, from 2004 to 2017, the structure of sample units in Shanghai's inbound tourist source market has undergone obvious changes, basically showing a downward or upward trend. The number of passengers from other countries has always been the first, and the proportion has increased by 3.24% from 30.25%; the number of Japanese tourists has decreased from the second to the third, from 24.53% to 13.37%; the number of Taiwan compatriots has increased from the third to the second, from 14.59% to 14.68%; the number of Hong Kong and Macao compatriots has decreased from the fourth to the fifth, from 9.95% to 8.44%; the number of American tourists has increased from the fifth to the third.

The proportion has increased significantly, from 7.21% to 10.98%; the number of German tourists has remained at the sixth place, and the proportion has increased from 3.34% to 3.85%; the number of Singapore tourists has slipped from the seventh to the tenth, but the proportion has increased from 2.59% to 2.68%; the number of British tourists has dropped from the eighth to the ninth, but the proportion has increased from 1.94% to 2.79%; the number of French tourists has dropped from the ninth to the eleventh, but the proportion has dropped from 1.83%.

The number of tourists from Australia jumped from 10th to 7th, from 1.58% to 2.9%; the number of tourists from Canada jumped from 11th to 8th, from 1.24% to 2.87%; the number of tourists from Italy remained the 12th, but the proportion increased from 0.98% to 1.41%.Hong Kong and Macao compatriots, other countries, Taiwan compatriots, Japan and the United States are in the top five of China's and Shanghai's inbound tourist source markets. The number of tourists from Hong Kong and Macao has always been the first in China, while it has been in the fourth and fifth places respectively in Shanghai. There is a big difference in the ranking of inbound tourists between China and Shanghai.

In 2017, the national market is followed by Hong Kong and Macao compatriots, other countries, Taiwan compatriots, Japan, the United States, Singapore, Canada, Australia, Germany, the United Kingdom, France and Italy, while the Shanghai market is followed by other countries, Taiwan compatriots, Japan, the United States, Hong Kong and Macao compatriots, Germany, Australia and Canada, UK, Singapore, France and Italy. The number of Hong Kong and Macao compatriots in China is very large, but their ranking in Shanghai is relatively low. This is

what Shanghai needs to emphasize in the future development of inbound tourist market. From 2004 to 2017, the number of inbound tourists from every country except Japan to the whole country and Shanghai increased significantly.

As the number of inbound tourists from Japan decreased, Shanghai should focus on the Japanese market later. As can be seen from **(table 4.3)**, the total growth of Shanghai inbound passenger market in the base period and the end period reached 3.8109 million, compared with the growth rate of the whole China L 39, more than 1, indicating that the inbound passenger market in Shanghai is developing faster than that in China. The total structural offset component of Shanghai is 120.18, and the structural effect index W 19, greater than 1, indicating that Shanghai's inbound tourist market contains a certain proportion of high-speed growth market, the structure is relatively reasonable; the total competitiveness offset component is 123.57, the competitiveness effect index U it is 1.16, more than 1, which indicates that Shanghai's inbound tourist market competitiveness is strong.

While maintaining the current inbound tourism market strategy, we should also enhance tourism marketing and publicity in developed European and American countries, such as Japan, the United States, Canada, Germany, the United Kingdom, France, Russia, Russia and Australia, so as to increase the market share of inbound tourism resources in these countries.

Shift-Share Analysis chart

Market Structure Deviation Component Analysis Chart

According to (table 4.1), the horizontal axis is the deviation component of market competitiveness D_{ij} , the vertical axis is the structural offset component P_{ij} To complete the structural deviation component analysis chart of Shanghai inbound tourism market, and mark the scatter points of each source market in the coordinate system **(Figure 4.3).**Note: in **(Figure 4.3, Figure 4.4, Figure 4.5 and Figure 4.6)** [72] 45° The bisector is corrected according to the scale of vertical and horizontal coordinates.

Figure (4.3): Analysis chart of Shanghai inbound tourism market structure deviation

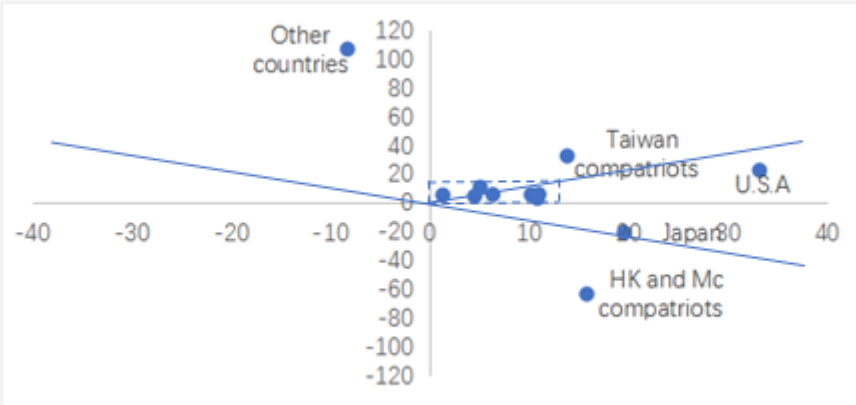
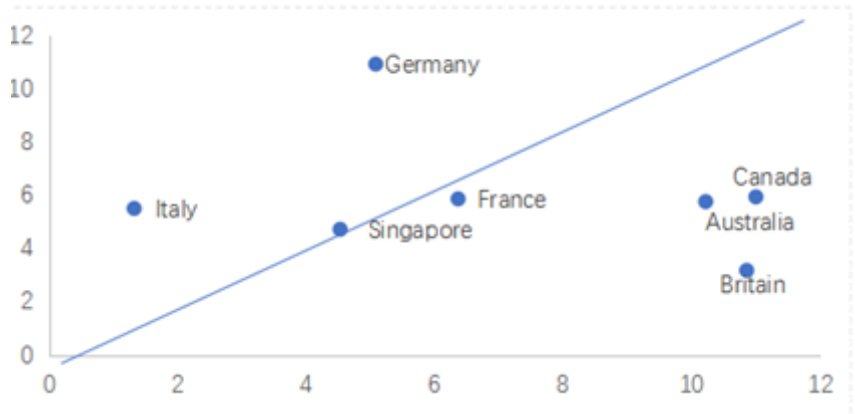


Figure (4.4): Dotted part of figure 4.3



As can be seen from **(figure 4.3)** and **(figure 4.4)**, taking the whole country as the upper level reference system, Taiwan compatriots, Germany and Italy have good market foundation and strong competitive advantages, among which Taiwan compatriots have the best market base; Singapore, France, Australia, Britain, Canada and the United States have better market base and competitive ability, and the US has a better market base It is the best; Japan's market base is poor, As can be seen from (**figure 4.3 and figure 4.4)**, taking the whole country as the upper level reference system, Taiwan compatriots, Germany and Italy have good market foundation and strong competitive advantages, among which Taiwan compatriots have the best market base; Singapore, France, Australia, Britain, Canada and the United States have better market base and competitive ability, and the US has a better market base It is the best; Japan's market base is poor, but its development is fast, it is a better or general source; other countries with a better foundation but declining status are better or general source areas; Hong Kong and Macao compatriots' tourist source areas are poor in foundation, but fast in development.

Market Structure Advantage Analysis Chart

The horizontal axis is the component of deviation PD_{ij}, the vertical axis is the share component N_{ij} , make an analysis chart of the market advantages of inbound tourists in Shanghai, and mark the source market of each scattered point in the coordinate system **(Figure 4.5)**.

Figure (4.5): analysis of structural advantages of Shanghai inbound tourism market.

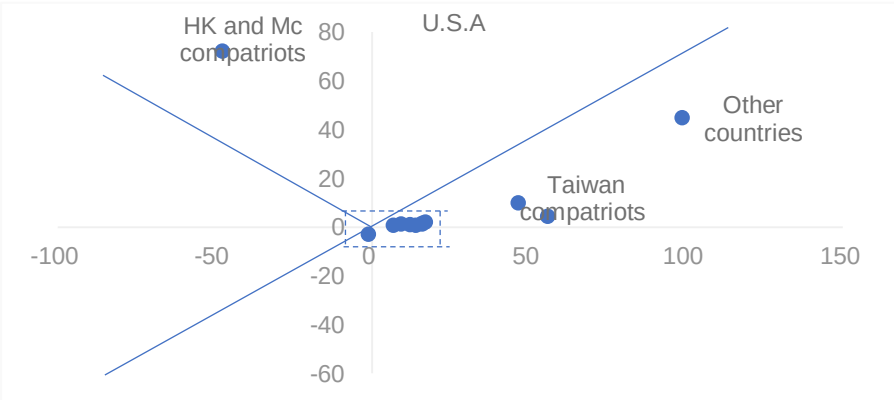
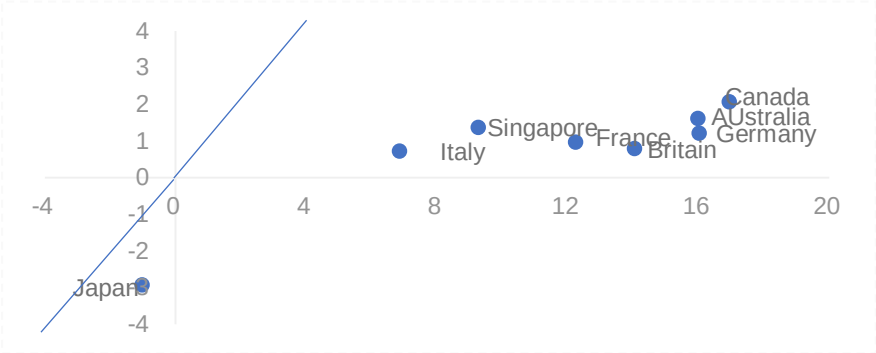


Figure (4.6): dashed part of figure 4.5



It can be seen from **(figure 4.5 and figure 4.6)** that the deviation and share components of other countries, Taiwan compatriots, the United States, Canada, Australia, Germany, Britain, France, Singapore and Italy in the inbound tourist source market of Shanghai are all positive. The regional source market advantage of these places, namely the deviation component, contributes more to the total growth of Shanghai's tourists than the share; Both of them deviate from the market share of Hong Kong and Macao, which shows that the contribution of the total volume of tourists from that of Japan is relatively large. The better source countries include other countries, Taiwan compatriots, the United States, Canada, Australia, Germany, the United Kingdom, France, Singapore and Italy; the compatriots of Hong Kong and Macao belong to the general source areas; and Japan is the worst source area.

It can be seen that most of the inbound tourist market in Shanghai has developed well and belongs to better and general market types.

Summary of This Part

After exploring the deviation share of Shanghai's inbound market structure, its overall structure is relatively reasonable, with a certain overall competitiveness, and the growth rate of tourists is faster than that of China, but its advantages are not obvious. We should seize the time to take effective measures, adhere to the tourism policy, develop a new tourism resource market, give full play to its own characteristics and potential tourism resources, and continue to add the structural and competitive advantages of Shanghai's inbound tourist market.

Time multi-scale characteristics of inbound tourists in Shanghai:

Data source and description

In order to show the current situation of inbound tourism development in Shanghai, this paper selects the monthly data from January 2004 to December 2018, including the number and composition of inbound tourists (i.e. Hong Kong, Macao, Taiwan and foreigners), as well as the number of inbound tourists from 18 major source countries in Shanghai. The length of each time series is 180 months. All inbound passenger data are from Shanghai Tourism Bureau.

The principle of set empirical mode decomposition (EEMD)

Empirical mode decomposition (EMD) is a new adaptive decomposition method found by Huang et al.Compared with wavelet transform, the basis function of EMD is determined by the data itself, which is more suitable for the analysis of non-linear and non-stationary data. EMD is used to decompose a troublesome signal into a set of complete and almost orthogonal signal components (IMF) and a residual (res). However, one of the main disadvantages of EMD is that mode mixing is easy to occur, which may blur the physical meaning of IMFs. In order to solve the inherent pattern mixing problem in EMD, Wu and Huang [56] proposed an effective EEMD analysis method. In fact,

EEMD is a noise aided EMD program. It makes full use of the statistical characteristics of Gaussian white noise uniform distribution, improves the distribution of extreme points in the original signal, and solves the mode mixing situation. In this method, the average value of the comprehensive test is regarded as a real and more meaningful IMF component.

It is based on the decomposition of a high frequency EMD signal from a transient to a high frequency. Data of different sizes keep their physical meaning and are well extracted and expressed. The lowest frequency component (res) represents the average value or overall trend of the original signal [73].

After the improvement of EMD, EEMD inherits the advantage of EMD adaptability and avoids the instantaneous noise that may be carried by the original data, which makes it difficult to carry out scale mixing. The results of signal decomposition are more stable and consistent, and the nonlinear and non-stationary data processing is more accurate.

The following is a brief introduction to EEMD algorithm [53]

(1) Initialize the number of integration m and add the amplitude of white noise, the first test m= 1.

(2) The mth test is carried out on the signal containing white noise.

(a) A white noise sequence with a given amplitude is added to the target signal to generate a new signal,

$x_m(t) = x(t) + n_m(t)$ (36)

Among them, $n_m(t)$ represents the mth white noise sequence added, $x(t)$ indicates the studied signal, $x_m(t)$ represents M signal sequences added with noise.

(b) Using EMD method to add noise signal sequence $x_m(t)$ it is decomposed into a series of IMFs, $C_{i,m}(i = 1,2,\cdots,I)$

$x_m(t) = \sum_{i=1}^{I} C_{i,m} + r_n$(37)

Among them, $C_{i,m}$ indicates the mth test i IMF, I represents the number of IMF in each trial, r_n represents a time series that represents the trend or mean of the original data series.

(c) If $m < M$, go to step (a), and $m = m + 1$. Steps (a) and (b) are repeated each time in a different white noise sequence until $m = M$.

(3) Calculate the overall mean value of M tests corresponding to each IMF in the decomposition C_i.

$C_i = \frac{1}{M}\sum_{m=1}^{m} C_{i,m}$, $i = 1,2,\cdots,I, m = 1,2,\cdots,M$(38)

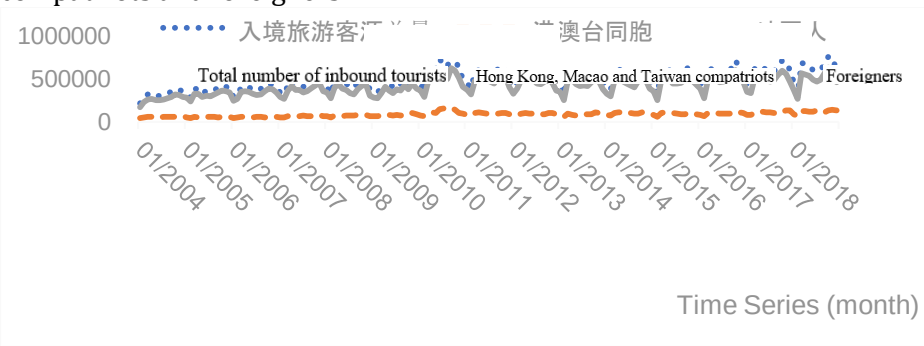
(4) Each mean of IMFs $C_i(i = 1,2,\cdots,I)$ the final IMFs represent the simple oscillation modes embedded in the studied signal.

Results and analysis

Analysis of EEMD decomposition results of the overall change of inbound tourist source structure

The total number of inbound tourists, the number of compatriots from Hong Kong, Macao and Taiwan and foreigners from January 2004 to December 2018 in Shanghai are compared in a broken line chart, and the monthly changes of the data can also be seen.

Figure (5.1): changes of total inbound tourist source, Hong Kong, Macao and Taiwan compatriots and foreigners



From **(Figure 5.1)**, we can see the actual changes of inbound tourist sources and their components from January 2004 to December 2018, that is, the number of compatriots and foreigners in Hong Kong, Macao and Taiwan. It can be seen that the data is not linear, but also non-static, which is the multi-scale characteristics of the data, so it is very suitable to restore the vibration information of different scales with EEMD.

The total amount of inbound tourists in Shanghai, the IMF components of Hong Kong, Macao and Taiwan compatriots and foreigners decomposed by EEMD, and the change information and maximum points of trend items are as follows[74]:

Figure (5.2): Location of IMF, trend value and maximum value of total inbound tourists after EEMD decomposition

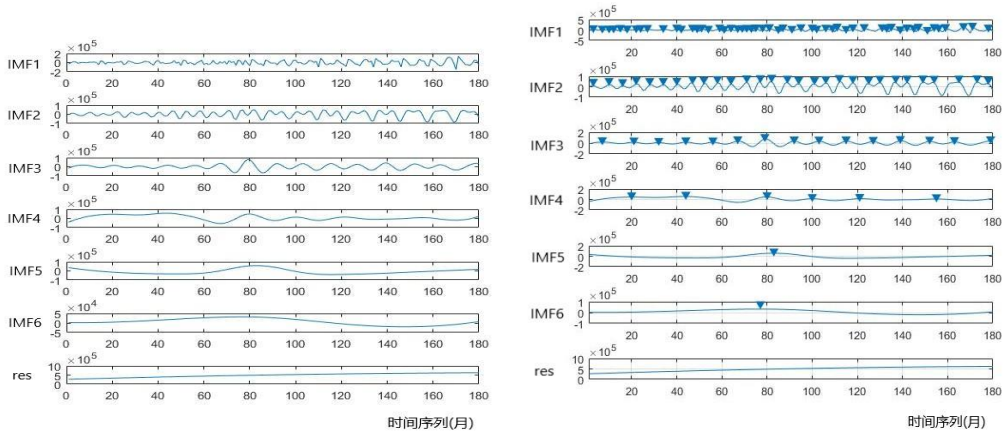


Figure (5.3): position of IMF, trend value and maximum value after EEMD decomposition of Hong Kong, Macao and Taiwan compatriots

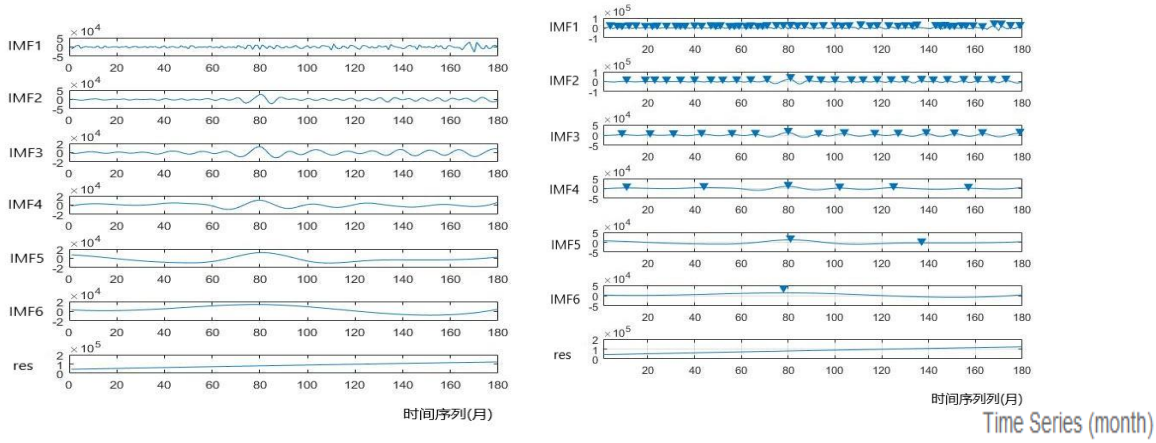
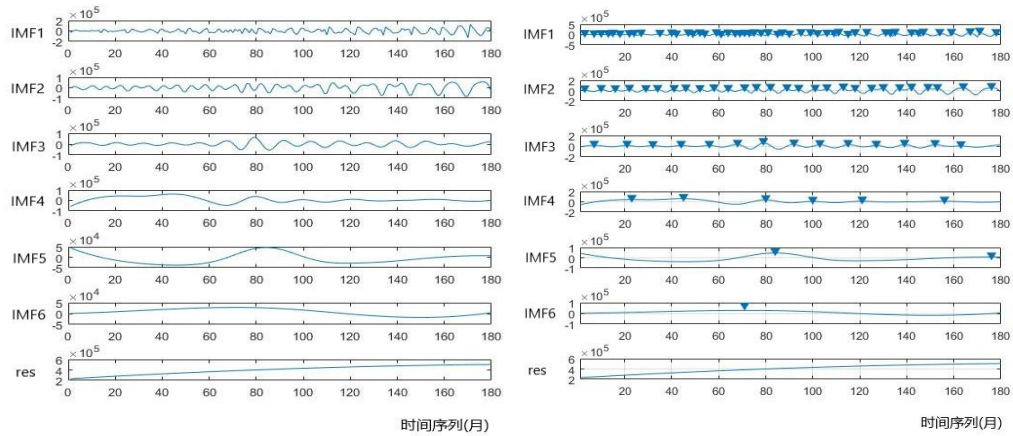


Figure (5.4): location of IMF, trend value and maximum value of the number of foreigners after EEMD decomposition



As can be seen from **(Figure 5.2, Figure 5.3 and figure 5.4)**, the data is divided into six IMF components. These IMFs reflect the information of oscillation from high frequency to low frequency in different time scales, and reflect that each IMF component is very strong in data reduction. According to the distribution of each IMF maximum point, we can see the relatively stable quasi periodic change of each IMF component. Therefore, the average period of each IMF component change and the corresponding variance contribution rate are obtained by MATLAB as follows:

Table (5.1): Contribution rate of period and variance corresponding to different IMF components

	Inbound tourists	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	Res
Average period (month)	Total inbound tourists	3.46	6.21	12	30	180	180	-
	Compatriots of Hong Kong, Macao and Taiwan	3.21	6.67	12	30	90	180	-
	foreigners	3.67	6.43	12.86	30	90	180	-
Variance contribution rate (%)	Total number of inbound tourists	12.96	6.24	3.78	4.43	4.46	1.74	66.38
	Compatriots of Hong Kong, Macao and Taiwan	7.33	6.21	2.47	1.93	4.73	6.43	70.90
	foreigners	15.83	8.29	3.43	5.97	5.21	2.16	59.10

From **(table 5.1)**, it can be seen that the total number of inbound tourists and the data of inbound tourists from compatriots of Hong Kong, Macao and Taiwan and foreigners have a period of about 3 months, 6 months, 12 months and 30 months, which is very consistent with the seasonal and interannual changes of tourism.

The high frequency periodic fluctuation of the three types of tourists is the largest in the total number of inbound tourists, all of which are 180. The high frequency fluctuation of compatriots from Hong Kong, Macao and Taiwan is the same as that of foreign tourists, with the cycle of 90and 180.

From the variance contribution rate of the three IMF components, it can be seen that the variance contribution rate of imf1 component of the total number of inbound tourists is the largest, which is 12.93%, followed by the variance contribution rate of imf2 component is 6.24%, and the variance contribution rate of imf3-imf6 is slightly small and negligible; the variance contribution rate of imf1, imf2 and imf6 components of Hong Kong, Macao and Taiwan compatriots is very close, which are 7.33%, 6.21% and 6.43%, respectively; the contribution rate of variance is slightly small and can be ignored; the largest contribution rate of variance of imf1 component of the number of foreigners is 15.83%, followed by the contribution rate of variance of imf2 component is 8.29%, and the contribution rate of variance of imf3-imf6 is slightly small and can be ignored.

According to the results and analysis in table 5.1, the change of the total number of inbound tourists in Shanghai is mainly three-month high-frequency

periodic fluctuations, supplemented by six-month fluctuations. The high-frequency and low-frequency periodic fluctuations of inbound tourists from Hong Kong, Macao and Taiwan occur alternately in three, six and 180 months; the fluctuation frequency of foreign tourists is high in three and six months. The variance contribution rate of the three trend items is 66.38%, 70.9% and 59.1% from Hong Kong, Macao and Taiwan compatriots, total inbound tourists and foreigners, respectively. It can be seen that the three trends are the most important changes in the future.

Analysis of EEMD decomposition results of Main Inbound Tourist Market

In addition to the compatriots of Hong Kong, Macao and Taiwan in China, China's inbound tourist market also includes 15 major source countries, including 6 Asian countries, 2 American countries, 5 European countries and 2 Australian countries. Using EEMD to decompose the market data of each inbound passenger, the average period and variance contribution rate are shown in **(table 5.2)**.

Table (5.2): The change cycle and variance contribution rate of Main Inbound Tourist Market

	country		IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	Res
China	Hong Kong	Average period	3.33	6	13.85	30	90	180	-
		Variance contribution rate	13.12	8.56	3.21	2.94	7.75	9.02	55.40
	Macao	Average period	3	6.67	12	25.71	60	180	-
		Variance contribution rate	14.06	10.61	9.80	5.07	10.48	17.17	32.79
	Taiwan	Average period	2.90	6.67	12	30	90	180	-
		Variance contribution rate	5.77	4.94	2.96	1.68	3.35	3.85	77.44
Asia	Japan	Average period	2.90	6.92	12	25.71	180	180	-
		Variance contribution rate	30.27	11.78	5.24	7.54	21.62	12.49	11.06
	Singapore	Average period	3.16	6.21	12.86	30	60	180	-
		Variance contribution rate	20.54	11.26	2.04	1.63	3.49	17.57	43.47
	Thailand	Average period	3.75	6.43	16.36	25.71	90	180	-
		Variance contribution rate	26.55	7.62	5.29	3.80	18.49	0.71	37.55
	Indonesia	Average period	2.86	6	12	22.50	60	90	-
		Variance contribution rate	21.07	10.79	8.01	2.65	1.42	1.28	54.78
	Malaysia	Average period	3.60	6.67	12.86	36	60	180	-
		Variance	19.93	6.12	5.30	2.91	3.00	10.83	51.91

		contribution rate							
	the republic of Korea	Average period	2.81	5.63	13.85	45	60	180	-
		Variance contribution rate	11.57	4.90	6.72	13.08	2.93	1.22	59.58
America	U.S.A	Average period	3.33	6.21	12	30	90	180	-
		Variance contribution rate	15.94	10.99	8.70	2.13	3.19	1.99	57.06
	Canada	Average period	3	6	12.86	25.71	180	180	-
		Variance contribution rate	14.71	10.23	3.35	3.51	5.82	0.71	61.66
Europe	Britain	Average period	3.83	6.67	12.86	36	60	180	-
		Variance contribution rate	28.71	13.32	5.96	1.15	0.91	1.31	48.63
	France	Average period	4.19	6.92	12.86	36	60	180	-
		Variance contribution rate	25.15	10.58	7.15	5.53	2.08	6.56	42.96
	Germany	Average period	3.83	6.92	13.85	36	60	180	-
		Variance contribution rate	40.36	16.37	5.98	1.79	2.43	0.47	32.60
	Italy	Average period	3.10	6.67	12	25.71	45	180	-
		Variance contribution rate	16.47	16.39	12.62	2.73	1.02	0.70	50.07
	Russia	Average period	3.27	6.43	12.86	36	90	180	-
		Variance contribution rate	14.24	5.88	1.50	2.29	3.12	0.27	72.69
Australia	Australia	Average period	3.75	6.67	13.85	30	90	180	-
		Variance contribution rate	23	9.44	2.62	2.73	4.62	0.45	57.14
	New Zealand	Average period	3	6	12	36	60	180	-
		Variance contribution rate	6.90	7.16	5.94	2.06	0.66	0.40	76.88

According to the results of EEMD decomposition of six source countries in Asia, the data changes can be divided into three categories. One is represented by Singapore, Thailand, Malaysia and South

Korea. The trend change is the main source market, supplemented by alternating high and low frequency cycles. In Singapore, the high and low frequency cycles change alternately in March, June and 180 months, Thailand in three and 90 months, Malaysia in three and 180 months, and South Korea in three and 45 months The second category is represented by Indonesia, and the main source market is trend change, supplemented by low-frequency cycle change in 3, 6 and 12 months; the third category is represented by Japan, supplemented by trend change, and the main source market is high-frequency and low-frequency cycle change in 3, 6 and 180 months.

According to the results of the period and variance contribution rate of the number of inbound tourists from the main source countries of the United States, Europe and Australia, we can divide these countries into three categories. The first is represented by the United States, Italy and New Zealand, which are mainly characterized by trend changes, supplemented by periodic fluctuations of 3, 6 and 12 months, and the sum of variance contribution rates of imf1-imf3 is 35.63%, 45.48% and 20% respectively; the second is represented by the United Kingdom, France, Germany and Australia, which are mainly characterized by periodic fluctuations of 4 and 7 months, and the variance contribution of imf1-imf2 of the four source countries The sum of contribution rates is 42.03%, 35.73%, 56.73% and 32.44% respectively; the three categories are represented by Canada and Russia, which are mainly characterized by trend change, supplemented by 3 and 6-month periodic fluctuations, and the sum of variance contribution rates of imf1-imf2 of these two source countries are 24.94% and 14.06% respectively. In the future, all the 15 source countries are on the rise.

Summary of this part

The EEMD method is used to decompose the total number and composition of inbound tourists in Shanghai, and the cycle and variance contribution rate are obtained. It can be seen that the total number of inbound tourists, compatriots of Hong Kong, Macao and Taiwan, and the number of foreigners entering Shanghai all have three-month and six-month cycle changes, and the variance contribution rate of the three-month and six-month changes of the total number of inbound tourists and the number of foreigners respectively reaches 19.2% and 24.12%, basically Master the scale change of time. There is also a significant variance contribution rate in 180 months, which is mainly characterized by high and low frequency periodic changes. The variance contribution rates of the three trend items are very high, which indicates that there is a great possibility of continuous increase in the future. Among them, Hong Kong, Macao and Taiwan account for the highest variance contribution rate, which indicates that our country's inbound tourists will come from our compatriots in Hong Kong, Macao and Taiwan in the future [75].

From the perspective of time multi-scale characteristics, the three-month cycle of Japan, Thailand, Britain, France and Germany is more obvious in Shanghai inbound tourist source countries; the high-frequency and low-frequency cycles occur alternately in three, six, twelve, sixty and 180 months in Macao, China, and three, six and 180 months in Singapore. The common characteristics of these market changes are obvious cyclical characteristics, the trend change is slightly not significant.

The trend changes are obvious in Hong Kong, Taiwan, Indonesia, Malaysia, South Korea, the United States, Canada, Italy, Russia, Australia and New Zealand, among which Taiwan, Canada, Russia and New Zealand are the most obvious. In Hong Kong, 3, 6, 90 and 180 months of seasonal high-frequency and low-frequency fluctuations are supplemented; Taiwan, Canada and Russia are supplemented by 3-month and 6-month seasonal changes; Indonesia, the United States, Italy and New Zealand are supplemented by 3-month, 6-month and 12-month low-frequency cycles; Malaysia is supplemented by 3-month and 180-month high-frequency and low-

frequency cycles; and in Korea, 3-month and 45-month high-frequency and low-frequency cycles alternate In Australia, 4 and 7 months are the secondary cyclical fluctuations[76].

In a word, most of the inbound tourism market in Shanghai is dominated by trend items, while a few are dominated by cyclical fluctuations. This shows that the inbound tourism market in Shanghai has grown steadily in the past 15 years, but there are still some tourists from some source countries that have grown slowly. Therefore, it is necessary to improve the development of Shanghai's inbound tourism industry. For the source countries with periodic fluctuations, we cannot rely on the rapid growth of tourists to realize the increase of inbound tourism income, but can only increase the per capita tourism consumption. We can improve the supply of inbound tourism, strengthen the attraction of inbound tourism products and commodities to tourists, so as to improve the tourism consumption of everyone and promote development of inbound tourism, and make it a pillar industry in the process of Shanghai tourism development [77].

Conclusion:

Summary

This paper takes Shanghai as the research object, forecasts the number of inbound tourists in Shanghai by improving the grey Markov model, and then analyzes the market structure of inbound tourists in Shanghai by using the deviation share method. Finally, based on the EEMD method, this paper studies the multi-scale characteristics and trends of inbound tourists in Shanghai, and obtains the following conclusions:

The improved GM (1,1) with initial value modification and Markov model with center point triangle whitening weight function are constructed. Compared with the results of traditional GM (1,1), initial value modified GM (1,1) and traditional grey Markov forecasting model, the prediction results of this model are more accurate, and the four models are better than linear regression and time series model.

Using the method of deviation share analysis to analyze the changes of inbound tourist market, the results show that the inbound tourist market is relatively reasonable and competitive, but for each inbound tourist country, the specific situation is different.

The ensemble empirical mode decomposition (EEMD) method is used to analyze the time multi-scale characteristics and trends of inbound tourists in Shanghai. The results show that the total number of inbound tourists and the number of foreign tourists entering Shanghai fluctuate periodically; the source countries and specific periodic fluctuations are obtained, which are mainly periodic fluctuations and have no significant trend change; the source countries with obvious trend change are obtained, and the detailed weekly fluctuations and trend changes are known.

According to the above conclusions, accurate prediction is of great significance to the decision-making of tourism departments and the travel of tourists, which is conducive to the rational allocation of resources and avoid congestion and tourists' detention. For the vigorous development of Shanghai's inbound tourism industry, some suggestions are provided to further strengthen and enhance the competitiveness of Shanghai's inbound tourism market.

Deficiencies

Because the author's theoretical knowledge and writing time are limited, there are still some areas that are not good enough in the paper, and some areas need further research and discussion, such as the initial value correction of the grey model may have better methods, which can make the prediction results more accurate; the deviation share analysis method is used to study the changes of Shanghai inbound tourist market, because many tourist sources China’s data cannot be found, only 12 known tourist source countries can be used for analysis. Moreover, the data of Hong Kong and Macao are only combined, and there is no separate data of Hong Kong and Macao, which makes many conclusions impossible to show. The time multi-scale characteristics and trends of inbound tourist sources are analyzed by using the

method, and the data of some tourist source countries cannot be found, only the data of main source countries, Therefore, for some countries, we cannot predict the situation of tourism in Shanghai and so on.

Conflict of Interest

We have no conflict of interests to disclose and the manuscript has been read and approved by all named authors.

Acknowledgement

This work was supported by the Philosophical and Social Sciences Research Project of Hubei Education Department (19Y049), and the Staring Research Foundation for the Ph.D. of Hubei University of Technology (BSQD2019054), Hubei Province, China.

References

1. Tang J, Sriboonchitta S, Yuan X. (2015) Forecasting Inbound Tourism Demand to China Using Time Series Models and Belief Functions [J]. Studies in Computational Intelligence, 583(03): 329-341.
2. Veloce W. (2004) Forecasting Inbound Canadian Tourism: An Evaluation of Error Corrections Model Forecasts [J]. Tourism Economics, 10(3): 263-280.
3. Witt SF, Turner LW. (2002) Trends and Forecasts for Inbound Tourism to China [J]. Journal of Travel & Tourism Marketing, 26 (01): 97-107.
4. Ying Y, Yirui W, Shangce G. (2017) Statistical Modeling and Prediction for Tourism Economy Using Dendritic Neural Network [J]. Computational Intelligence and Neuroscience, 36(03): 1-9.
5. Lijuan H. (2013) Research on influencing factors and prediction of inbound tourism demand in Qinhuangdao [D]. Yanshan University.
6. Naiwen L, Jingjing H. (2015) Prediction of China's inbound tourists based on time series correction algorithm [J]. Resource development and market, 31 (01): 126-128.
7. Wang Xiaoshan. (2014) Research on the influencing factors of China's inbound tourism based on gravity model [D]. Hailan University.
8. Rui Q. (2010) Prediction of the trend of inbound tourists in Shanghai under the influence of World Expo [J]. Modern commerce and industry, 22(05): 127-128.
9. Liang X. (2004) Research on the statistical analysis and prediction model of Shanghai inbound tourism market [D]. East China Normal University.
10. Peng C. (2014) Prediction of the number of inbound tourists in Anhui Province Based on the model [J]. Journal of Suzhou University, 29 (09): 37-40.
11. Xia H, Hongwei X. (2013) Equivalence of initial value modified grey prediction model [J]. Journal of Xihua University (NATURAL SCIENCE EDITION), 32 (04): 54-57.
12. Hongmei S, Jianhua Y, Yuexin L. (2015) Combined GM (1,1) model based on initial value modification and its application [J]. Statistics and decision making, 34(02): 89-90.
13. Xiao WC. (2006) Non equidistant grey prediction model based on initial value correction [J]. Journal of Chongqing Normal University (NATURAL SCIENCE EDITION), 15 (03): 42-44.
14. Yali C, Bo W. (2014) Prediction and analysis of China's rural residents' consumption level index: an improved GM (1,1) model based on initial value modification [J]. Income and consumption, 25 (04): 113-115
15. Zhongtao W, Xin P, Qi D. (2007) Improvement and application of grey prediction model based on initial value correction [J]. Journal of Chongqing Institute of Technology (NATURAL SCIENCE EDITION), 21(10):81-84.
16. Muxiao L. (2010) Application of improved GM (1,1) model in ship maintenance cost prediction [J]. Ship Electronics Engineering, 36 (12): 151-154.
17. Yao T, Gong Z, Xie N. (2010) The Discrete Grey Prediction Model Based on Optimized Initial Value[M]. Springer Berlin Heidelberg.
18. Rui Z. (2012) Grey Markov prediction of grain yield in Shaanxi Province [D].

Lanzhou University of Commerce.
19. Chaoyang S. (2014) Improvement and application of Grey Markov chain [D]. Lanzhou University.
20. Xiaoyong H. (2014) Research on the application of Grey Markov chain in the prediction of occupational safety training number [D]. Southwest Jiaotong University.
21. Zhanli M, Jinhua S. (2011) Application of Grey-Markov Model in Forecasting Fire Accidents[J]. Procedia Engineering,314-318.
22. Geng YS. (2011) The Research of Improved Grey-Markov Algorithm [C]. Intelligent Information Technology Application Association, 36(02):121-128.
23. Jia Z, Zhou Z, Zhang H, (2020). Forecast of coal consumption in Gansu Province based on Grey-Markov chain model [J]. Energy, 52 (01): 199.
24. Huan H, Tan Q. (2015) The forecast of cultivate land quantity based on Grey-Markov model[J]. 5(1): 127-136.
25. Fajian L, Dongdong C, Jianhua Z. (2016) Market structure and interactive pattern of China's inter provincial inbound tourism: Based on 2-mode network analysis [J]. Progress in Geographical Sciences, 35 (08): 932-940.
26. Lijun M, Gennian S, Jingru H. (2014) An analysis of China's inbound tourist market based on dual perspective competition and pro scenic degree [J]. Yunnan geographic environment research, 26 (06): 6-11 + 19.
27. Lai X. (2010) Research on the development strategy of inbound tourism in Hangzhou Based on the analysis of tourist source market [D]. Zhejiang Normal University.
28. Tong Y. (2019) Empirical analysis on the market characteristics of Shenzhen inbound tourists [J]. Special Economic Zone,21(08): 123-125.
29. One belt, one road, Bin S, Feng MY. (2017) The development of Shaanxi's inbound tourist market under the background of "one belt and one road": the perspective of the evolution of the structure of the tourist market based on DSSM [J].36 (09): 80-86.
30. Cuilin L, Hao Q. (2018) Analysis and optimization of Xinjiang inbound tourism industry structure based on SSM [J]. Journal of Guangxi Vocational and technical college, 11 (01): 64-70.
31. Longfei L, Yan J. (2015) Research on the market structure of Zhengzhou inbound tourism based on shift share analysis method [J]. Journal of Hunan Institute of Engineering (Social Science Edition), 25 (02): 11-15.
32. Dai T. (2018) Research on the inbound tourist market of Shantou City [D]. Xiamen University.
33. Kai Z, Yuqin W. (2008) Based on the shift share analysis of inbound tourism market of provinces and cities in the Yangtze River Basin[J]. Tourism Forum, 1 (06): 421-425
34. Mengke Y. (2018) Research on inbound tourism market in Guizhou Province [D]. Guizhou University.
35. Caiping Z. (2008) An Analysis of Tourist Structure in the Inbound Tourism Market Based on SSM Method-Taking Zhejiang Province as An Example [J]. Tourism Tribune,2008.
36. Peiji S. (2009) SSM Analysis on the Structure of Inbound Tourism Market in Gansu Province [J]. Journal of Arid Land Resources and Environment.
37. Li W, Duo M. (2011) SSM analysis of inbound tourism market structure in Liaoning Province [J]. Resource development and market, 27 (11):1047-1050.
38. Huang. (2011) A Shift-share analysis on the structure of tourism market-A case study on Hainan in-bound tourism market [C]. International Conference on E-business & E-government.
39. Yasin M, Alavi J, Sobral F. (2004) A Shift-Share Analysis Approach to Understanding the Dynamic of the Portuguese Tourism Market [J]. Journal of Travel & Tourism Marketing,17(4): 11-22.
40. Yan L. (2019) Analysis on adjustment of planting structure in Liaoning Province Based on shift share method [D]. Shenyang Agricultural University.
41. Xiaoyang Y, Shuai T, Yi LV, Shuai L. (2017) Analysis of grain yield increase

structure change in "sickle bend" region based on shift share model and Discussion on corn planting structure adjustment [J]. China Agricultural Resources and regionalization, 38 (05): 14-20.
42. Yuanhong L, Ying P. (2020) Analysis of Jiangsu marine industrial structure based on shift share analysis [J]. Marine development and management, 37 (01): 92-96.
43. Xiaojie G, Huiping T. (2019) Analysis of competitiveness of service industry in Guangdong Province Based on shift share analysis [J]. Journal of Changsha University, 33 (05): 93-97.
44. Xiangyang Y, Run S, Guoxing Z, et al. (2012) EMD Based passenger flow fluctuation characteristics and its combination prediction in Huangshan scenic spot [J]. Progress in Geographical Sciences,31 (10): 1353-1359.
45. Xiaoxuan L, Benfu LV, Pengzhi Z, et al. (2017) Research on tourist flow prediction based on network search and clsi-emd-bp [J]. System engineering theory and practice,37 (01): 106-118.
46. Lijun L. (2019) Prediction of tourist volume based on network search index and emd-arima-bp combination model: a case study of Zhangjiajie [J]. Journal of Jishou University (SOCIAL SCIENCE EDITION,40 (01): 138-150.
47. Lijun L, Xiaoping L. (2019) Research on tourist volume prediction based on emd-bp neural network [J]. Statistics and decision making, 35 (04): 85-89.
48. Xiaolong L, Baoguang X, Biao S. (2015) Analysis of airline passenger flow based on empirical mode decomposition [J]. Systems engineering, 2015, 33 (04): 142-148.
49. Junyuan Z, Zhanyu G, Xiaoli L, (2013). Multi time scale analysis on the fluctuation of tourism foreign exchange income in Sichuan Province and its causes based on the application of EMD method [J]. Western economic management forum,2013,24(02): 1-5.
50. Biw X Chengl S Wei Y, et al. (2018) (Isolated word speech recognition based on DTW and EMD [J]. Journal of Liaoning University of petroleum and chemical technology,38 (01): 74-78
51. Cheng C. (2015) Short term passenger flow forecast of high-speed railway based on emd-bpn method [D]. Lanzhou Jiaotong University.
52. Tao X. (2014) Research on hybrid modeling technology and application of time series prediction based on EMD [D]. Huazhong University of science and technology.
53. Lingling C, Liang H, Yuxia L. (2015) A multi-scale comparative study of China's inbound tourism and economic growth based on EEMD [J]. Journal of Nanjing Normal University (NATURAL SCIENCE EDITION)38 (3).
54. Muzi Z. (2016) Prediction of hotel occupancy rate based on eemd-arima [D]. Shaanxi Normal University.
55. Jianbo S. (2019) Photovoltaic power generation forecasting model and Application Research Based on EEMD and variable weight combination forecasting [D]. North China Electric Power University (Beijing).
56. Wu Z, Huang NE. (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv. Adapt. Data Anal,1(01): 1-41.
57. Wang H, Chen J, Dong G. (2014) Feature extraction of rolling bearing's early weak fault based on EEMD and tunable Q-factor wavelet transform. Mech. Syst. Signal Process, 48(1-2):103-119.
58. Lei Y, He Z, Zi Y. (2011) EEMD method and WNN for fault diagnosis of locomotive roller bearings[J]. Expert Systems with Applications, 38(6): 7334-7341.
59. Li H, Liu T, Wu X, et al. (2019) Application of EEMD and improved frequency band entropy in bearing fault feature extraction[J]. ISA Transactions, 88: 170-185.
60. Gaci S. (2016) Decomposition (EEMD) Denoising Method for Seismic Signals[J]. Energy Procedia, 97: 84-91.
61. Youwen W. (2014) Mathematical model based on GM (1, l) Yangquan tourist number prediction [J]. Journal of Shanxi Normal University (NATURAL SCIENCE EDITION), 28 (03): 14-17.
62. Sifeng L. (2004) Grey system theory and its application[M]. Beijing Science

Press.
63. Chen Z，Xu AM. (2012) Prediction of gas flow-rate from Boreholes based on Grey Markov model [J]. China Safety Sci J, 22(3): 79-85.
64. Weidong F, Huaibo. (2011)Markov test and prediction of Shanghai Shenzhen 300 index [J]. Science and technology and engineering, 11 (20): 4833-4835.
65. Xia H, Weifeng L. (2011) Equivalence of initial value modified grey GM (1,1) model [J]. Journal of Xinxiang University (NATURAL SCIENCE), (05): 391-393.
66. Wang HZ，Qiang FJ，He BC. (2014) A new grey evaluation method based on the improved center-point triangular whitenization weight function [J]. Stat Decis, 8: 69-72.
67. Sifeng L, Naiming X. (2011) A new grey evaluation method based on improved trigonometric whiteness weight function [J]. Systems engineering, 26 (02): 244-250.
68. Liwei W, Zhengxian X, Jinfeng W. (2009) An analysis of the structural changes of Guangxi inbound tourism market: Based on the perspective of SSM [J]. Social scientist, (11):95-96.
69. Kedan Z, Yongping G, Wenzhen W. (2011) An analysis of the market structure of Guangxi inbound tourism based on the shift share analysis [J]. Guangxi Social Sciences, (05): 34-38.
70. Fuyuan B, Huizi Z, Huixian D. (2018) Research on the deviation share of inbound tourist market structure -- Taking Sanya City as an example [J]. Economic perspective, (04): 27-37.
71. Longfei L, Yan J. (2015) Research on the market structure of Zhengzhou inbound tourism based on shift share analysis method [J]. Journal of Hunan Institute of Engineering (Social Science Edition), 25 (02): 11-15.
72. Yu Z, Zhibao D, Xujia C. (2014) Research on the market structure of inbound tourists in Guangzhou [J]. Henan science, 32 (04): 673-678.
73. Lingling C, Zhenshan L, Jie G, et al. (2009) Research on China's food security based on EMD [J]. Chinese Journal of Agricultural Sciences,42 (1): 180-188.
74. Lingling C, Wei Y. (2017) Multi scale analysis and prediction of China's inbound tourist market based on EEMD [J]. Business economic research, 17(10): 189-192.
75. Xiaorong H, Kunjie P. (2020) Impact of COVID-19 on China's Tourism industry: Prediction and Countermeasures [J]. Journal of Sichuan Tourism University, 32(04): 65-71.
76. Xiaoqiong P. (2020) Quantitative Analysis of the impact of COVID-19 epidemic on China's Tourism industry [J]. Market Modernization,51(12): 147-149.
77. Jing X, Rui L, Linqing L. (2020) Impact mechanism and countermeasures of covid-19 on Shanghai's economic and social development [J]. Shanghai urban management, 29(03): 4-10.

Appendix

1. Prediction of 2017 by Traditional Grey Markov Model

(1). Traditional Grey Prediction Model (Forecasting 2017)

```
#1. GM (1,1) model with x0 (1) as initial condition
gm11 < - function (x0, t) {#x0 is input column, t is forecast number
A sequence of < 0-xmscum is generated
b<-numeric(length(x0)-1)
n<-length(x0)-1
For (I in 1: n) {#) generates the nearest neighbor mean generating sequence of x1
b[i]<--(x1[i]+x1[i+1])/2
b} The sequence B is the nearest mean generating sequence of x1
D<-numeric(length(x0)-1)
D[]<-1
B<-cbind(b,D)
```

```
BT < - t (b) ◌ do inverse matrix
M<-solve(BT%*%B)
YN<-numeric(length(x0)-1)
YN<-x0[2:length(x0)]
The least squares estimation parameters of alpha < - M% *% BT% *% yn ◌ model satisfy alpha
alpha2<-matrix(alpha,ncol=1)
a<-alpha2[1]
u<-alpha2[2]
Cat ("GM (1,1) parameter estimated value: ',', development coefficient - a =", - A, "," "grey action quantity u =", u, '', '") # the parameter estimation value a, u is obtained by using the least square method
y<-numeric(length(c(1:t)))
y[1]<-x1[1]
For (W in 1: (t-1)) {substitute the estimated value of a, u into the time response sequence function to calculate the X1 fitting sequence y
y[w+1]<-(x1[1]-u/a)*exp(-a*w)+u/a
}
Cat ("analog value of X (1): ', n', y ',' \ n ')
xy<-numeric(length(y))
xy[1]<-y[1]
For (o in 2: T) {#, the model input sequence x0 prediction sequence is obtained by using the post subtraction operation
xy[o]<-y[o]-y[o-1]
}
Cat ("analog value of X (0):", ', XY,' \ n ',' \ n ')
#Calculate the residual error E
e<-numeric(length(x0))
for(l in 1:length(x0)){
Residual error of E [l] < - x0 [l] - XY [l]
}
Cat ("residuals: ', n', e, '\ n')
#Calculate the relative error
e2<-numeric(length(x0))
for(s in 1:length(x0)){
Relative error of E2 [S] < - (E [S]) / x0 [S] ◌
}
Cat ("relative error: ', n', E2 ',' \ n ')
Cat ("sum of squares of residuals =", sum (e ^ 2), '\ n')
Cat ("average relative error =", sum (E2) / (length (E2) - 1) * 100, "%", '\ n')
Cat ("relative precision =", (1 - (sum (E2) / (length (E2) - 1)) * 100, "%", '\ n', '\ n')
#Posterior difference ratio test
Avge < - mean (ABS (E)); esum < - sum ((ABS (E) - avge) ^ 2); EVAR = esum / (length (E) - 1); SE = sqrt (EVAR) ๑ calculate the variance of residuals
Avgx0 < - mean (x0); x0sum < - sum ((x0-avgx0) ^ 2); x0var = x0sum / (length (x0)); SX = sqrt (x0) # calculate the variance of the original sequence x0
CV < - Se / SX ◌
Cat ("posterior difference ratio test: ', n'," C value = ", CV, '\ n'). The posterior difference ratio is tested and compared with the general standard to judge whether the prediction result is good or not.
if(cv < 0.35){
Cat ("C value < 0.35, GM (1,1) prediction accuracy grade: good", '\ n')
}else{
if(cv<0.5){
```

```
Cat ("C value belongs to [0.35,0.5), GM (1,1) model prediction accuracy grade: qualified", ', n', '\ n')
}else{
if(cv<0.65){
Cat ("C value belongs to [0.5,0.65), and the prediction accuracy grade of GM (1,1) model is: barely qualified", '\ n'
}else{
Cat ("C value > = 0.65, prediction accuracy grade of GM (1,1) model: unqualified", '\ n')
}
}
}
#Draw the prediction sequence of input sequence x0 and the comparison image of x0
Plot (XY, col = blue ', type = B', PCH = 16, xlab ='time series', ylab ='value ')
points(x0,col='red',type='b',pch=4)
Legend ('topleft ', C ('forecast price', 'original price'), CEX = 0.6,text.width=1.4,pch=c(16,4),lty=l,col=c('blue','red'))
}
a<-
c(4.9192,5.7135,6.0567,6.6559,6.4037,6.2892,8.5112,8.1757,8.004,7.574,7.913,8.0016,8.5437)
gm11(a,length(a)+2)
```

(2).Find F and P Matrix Matlab

```
format rat
CLC
a=[1 2 3 2 1 5 4 3 2 2 2 2];
for i=1:5
for j=1:5
f(i,j)=length(findstr([i j],a));
End
End
F
ni=(sum(f'))'
for i=1:5
p(i,:)=f(i,:)/ni(i);
End
P
```

(3). R Language to Find 5-Step Transition Probability Matrix

```
A=array(c(0.0,0.5,0.0,0.0,0.5,0.2,0.6,0.2,0.0,0.0,0.0,1.0,0.0,0.0,0.0,0.0,0.0,1.0,0.0,0.0,0.0,0.0,0.0,1.0,0.0),dim=c(5,5))
A=t(A)
A
B2=A%*%A
B2
B3=A%*%A%*%A
B3
B4=A%*%A%*%A%*%A
B4
B5=A%*%A%*%A%*%A%*%A
B5
```

2. Improved Grey Markov Prediction (2017)

(1). Improved Grey Prediction

```
#GM (1,1) model with X1 (1) and X1 (n) as initial conditions
Gm11 < - function (x0, t) {x0 is input column, t is forecast number
A sequence of < 0-xmscum is generated
b<-numeric(length(x0)-1)
n<-length(x0)-1
For (I in 1: n) {#) generates the nearest neighbor mean generating sequence of x1
b[i]<--(x1[i]+x1[i+1])/2
b} The sequence B is the nearest mean generating sequence of x1
D<-numeric(length(x0)-1)
D[]<-1
B<-cbind(b,D)
BT < - t (b) ◌ do inverse matrix
M<-solve(BT%*%B)
YN<-numeric(length(x0)-1)
YN<-x0[2:length(x0)]
The least squares estimation parameters of alpha < - M% *% BT% *% yn ◌ model satisfy alpha
alpha2<-matrix(alpha,ncol=1)
a<-alpha2[1]
u<-alpha2[2]
Cat ("GM (1,1) parameter estimated value: ',', development coefficient - a =", - A, "," "grey action quantity u =", u, '', '") # the parameter estimation value a, u is obtained by using the least square method
y<-numeric(length(c(1:t)))
f=length(x1)
y[1]<-x1[1]
For (W in 1: (t-1)) {substitute the estimated value of a, u into the time response sequence function to calculate the X1 fitting sequence y
y[w+1]<-(1/2)*(x1[1]*exp(a)-u/a*exp(a)+x1[f]*exp(a*f)-u/a*exp(a*f))*exp(-a*w-a)+u/a
}
Cat ("analog value of X (1): ', n', y ',' \ n ')
xy<-numeric(length(y))
xy[1]<-y[1]
For (o in 2: T) {#, the model input sequence x0 prediction sequence is obtained by using the post subtraction operation
xy[o]<-y[o]-y[o-1]
}
Cat ("analog value of X (0):", ', XY,' \ n ',' \ n ')
#Calculate the residual error E
e<-numeric(length(x0))
for(l in 1:length(x0)){
Residual error of E [l] < - x0 [l] - XY [l]
}
Cat ("residuals: ', n', e, '\ n')
#Calculate the relative error
e2<-numeric(length(x0))
for(s in 1:length(x0)){
Relative error of E2 [S] < - (E [S]) / x0 [S] ◌
}
Cat ("relative error: ', n', E2 ',' \ n ')
Cat ("sum of squares of residuals =", sum (e ^ 2), '\ n')
Cat ("average relative error =", sum (E2) / (length (E2) - 1) * 100, "%", '\ n')
Cat ("relative precision =", (1 - (sum (E2) / (length (E2) - 1)) * 100, "%", '\ n', '\ n')
#Posterior difference ratio test
```

```
Avge < - mean (ABS (E)); esum < - sum ((ABS (E) - avge) ^ 2); EVAR = esum / (length (E) - 1); SE = sqrt (EVAR) ໑ calculate the variance of residuals
Avgx0 < - mean (x0); x0sum < - sum ((x0-avgx0) ^ 2); x0var = x0sum / (length (x0)); SX = sqrt (x0) # calculate the variance of the original sequence x0
CV < - Se / SX ◌ɔ
Cat ("posterior difference ratio test: ', n'," C value = ", CV,'n ')). The posterior difference ratio is tested and compared with the general standard to judge whether the prediction result is good or not
if(cv < 0.35){
Cat ("C value < 0.35, GM (1,1) prediction accuracy grade: good", '\ n')
}else{
if(cv<0.5){
Cat ("C value belongs to [0.35,0.5), GM (1,1) model prediction accuracy grade: qualified", ', n', '\ n')
}else{
if(cv<0.65){
Cat ("C value belongs to [0.5,0.65), and the prediction accuracy grade of GM (1,1) model is: barely qualified", '\ n'
}else{
Cat ("C value > = 0.65, prediction accuracy grade of GM (1,1) model: unqualified", '\ n')
}
}
}
#Draw the prediction sequence of input sequence x0 and the comparison image of x0
Plot (XY, col = blue ', type = B', PCH = 16, xlab ='time series', ylab ='value ')
points(x0,col='red',type='b',pch=4)
Legend ('topleft ', C ('forecast price', 'original price'), CEX = 0.6,text.width=1.4,pch=c(16,4),lty=l,col=c('blue','red'))
}
a<-
c(4.9192,5.7135,6.0567,6.6559,6.4037,6.2892,8.5112,8.1757,8.004,7.574,7.913,8.0016,8.5437)
gm11(a,length(a)+1)
```

(2). Calculate F

```
clear all;
X = [- 0.06900973 -0.04268335 0.02075474 -0.05045404 -0.1038813 0.158146 0.09549206 0.04645758 -0.03999574 -0.02736665 -0.04857514 -0.013538331];% original value
[n m]=size(X);
%C = [100 100 100 100 100 100 100 100 100 100 100 100 100];% weight
f=[-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-0.12 -0.09 -0.03 0.03 0.09 0.15 0.18;
-12-0.09 -0.03 0.03 0.09 0.15 0.18;];% interval and turning point
[m s]=size(f);
F=zeros(m,s-2,n)
```

```
%The clustering coefficient of triangle whitening of center point is calculated
for i=1:n;
for k=2:s-3;
tmp=0;
for j=1:m;
if X(i,j)<f(j,k)||X(i,j)>f(j,k+2)
F(j,k,i)=0;
elseif X(i,j)>=f(j,k)&&X(i,j)<=f(j,k+1)
F(j,k,i)=(X(i,j)-f(j,k))/(f(j,k+1)-f(j,k));
elseif X(i,j)>=f(j,k+1)&&X(i,j)<=f(j,k+2)
F(j,k,i)=(f(j,k+2)-X(i,j))/(f(j,k+2)-f(j,k+1));
else;
end;
tmp=tmp+F(j,k,i);
end;
result(i,k)=tmp;
end;
for k=1;
tmp=0;
for j=1:m;
if X(i,j)<f(j,k)||X(i,j)>f(j,k+2)
F(j,k,i)=0;
elseif X(i,j)>=f(j,k)&&X(i,j)<=f(j,k+1)
F(j,k,i)=1;
elseif X(i,j)>=f(j,k+1)&&X(i,j)<=f(j,k+2)
F(j,k,i)=(f(j,k+2)-X(i,j))/(f(j,k+2)-f(j,k+1));
else;
end;
tmp=tmp+F(j,k,i);
end;
result(i,k)=tmp;
end;
for k=s-2;
tmp=0;
for j=1:m;
if X(i,j)<f(j,k)||X(i,j)>f(j,k+2)
F(j,k,i)=0;
elseif X(i,j)>=f(j,k)&&X(i,j)<=f(j,k+1)
F(j,k,i)=(X(i,j)-f(j,k))/(f(j,k+1)-f(j,k));
elseif X(i,j)>=f(j,k+1)&&X(i,j)<=f(j,k+2)
F(j,k,i)=1;
else;
end;
tmp=tmp+F(j,k,i);
end;
result(i,k)=tmp;
end;
end;
clc;
Disp ('the input object observation matrix is (where the row is the object and the column is the index value): ';
disp(X);
Disp ('the input index value matrix is (where the row is the indicator item and the column is the value range): ';
```

```
disp(f);
Disp ('the comprehensive clustering coefficient matrix of each object is (where the row is the object and the column is the cluster coefficient of the subclass): ';
disp(result);
disp(F);
```

(3). Calculate P1, P2, P3, P4, P5, R, Q

```
F
= [801/1232 431/1232 0 0 0 ;
245/1159 914/1159 0 0 0 ;
0 1276/8281 269/318 0 0;
916/2687 1771/2687 0 0 0 ;
1 0 0 0 0 ;
0 0 0 0 1 ;
0 0 0 1717/1890 173/1890;
0 0 590/813 223/813 0;
391/2347 1956/2347 0 0 0 ;
0 1416/1481 274/6243 0 0 ;
239/772 533/772 0 0 0 ;
0 767/1057 290/1057 0 0 ;
];
[n s]=size(F);
for i=1:s;
for j=1:s;
p1(i,j)=(F(1:n-1,i))'*(F(2:n,j))/sum(F(1:n-1,i));
End
End
disp('p1:');
disp(vpa(p1,3));
for i=1:s;
for j=1:s;
p2(i,j)=(F(1:n-2,i))'*(F(3:n,j))/sum(F(1:n-2,i));
End
End
disp('p2:');
disp(vpa(p2,3));
for i=1:s;
for j=1:s;
p3(i,j)=(F(1:n-3,i))'*(F(4:n,j))/sum(F(1:n-3,i));
End
End
disp('p3:');
disp(vpa(p3,3));
for i=1:s;
for j=1:s;
p4(i,j)=(F(1:n-4,i))'*(F(5:n,j))/sum(F(1:n-4,i));
End
End
disp('p4:');
disp(vpa(p4,3));
for i=1:s;
for j=1:s;
p5(i,j)=(F(1:n-5,i))'*(F(6:n,j))/sum(F(1:n-5,i));
```

```
End
End
disp('p5:');
disp(vpa(p5,3));
R=[F(n-s+1,1)*p1(1,:)+F(n-s+1,2)*p1(2,:)+F(n-s+1,3)*p1(3,:)+F(n-s+1,4)*p1(4,:)+F(n-
s+1,5)*p1(5,:);
F(n-s+2,1)*p2(1,:)+F(n-s+2,2)*p2(2,:)+F(n-s+2,3)*p2(3,:)+F(n-s+2,4)*p2(4,:)+F(n-
s+2,5)*p2(5,:);
F(n-s+3,1)*p3(1,:)+F(n-s+3,2)*p3(2,:)+F(n-s+3,3)*p3(3,:)+F(n-s+3,4)*p3(4,:)+F(n-
s+3,5)*p3(5,:);
F(n-s+4,1)*p4(1,:)+F(n-s+4,2)*p4(2,:)+F(n-s+4,3)*p4(3,:)+F(n-s+4,4)*p4(4,:)+F(n-
s+4,5)*p4(5,:);
F(n,1)*p5(1,:)+F(n,2)*p5(2,:)+F(n,3)*p5(3,:)+F(n,4)*p5(4,:)+F(n,5)*p5(5,:);]
disp('R:');
disp(vpa(R,3));
q=sum(R)/s;
disp('q:');
disp(vpa(q,3));
```

3. Linear Regression Prediction

```
x=c(1,2,3,4,5,6,7,8,9,10,11,12,13)
y=c(4.9192,5.7135,6.0567,6.6559,6.4037,6.2892,8.5112,8.1757,8.004,7.574,7.913,8.0016,8.543
7)
plot(x,y)
model=lm(y~x)
summary(model)
```

4. Time Series Prediction

```
A=[4.9192 5.7135 6.0567 6.6559 6.4037 6.2892 8.5112 8.1757 8.004 7.574 7.913 8.0016
8.5437]
h = adftest(A)
%B = dtrend(A)
B = diff(A)
H = adftest(B)
figure(1)
autocorr(B)
figure(2)
parcorr(B)
x = A;
w = B;
n = 2;
s = 1;
M1 = length (a);% number of original data
for i = s+1:m1
Y (I-S) = x (I) - x (I-S);%
End
Toestmd = ARIMA ('arlags', 1, 'malags', 1:1,' constant ', 0);% specifies the structure of the model
[estmd, estparamcov, logl, info] = estimate (toestmd, W ');% model fitting
W_Forecast = forecast(EstMd,n,'Y0',w');
yhat = y(end) + cumsum(w_Forecast);% restored value of first order difference
for j = 1:n
X (M1 + J) = yhat (J) + X (M1 + J-s);% X's predicted value
End
```

```
x(1:end)
```

5. Drawing r Program

```
sample<-seq(1,2,by=1)
truevalue<-c(8.7301,8.9371)
grey<-c(8.9376,9.1753)
impgrey<-c(8.9371,9.135)
greymarkov<-c(8.8035,9.0377)
impgreymarkov<-c(8.6944,8.9135)
huigui<-c(8.9991,9.2653)
timeseires<-c(9.1005,9.6697)
value<data.
frame(sample,truevalue,grey,impgrey,greymarkov,impgreymarkov,huigui,timeseires);value
Plot (true value ~ sample, PCH = 15, col = darkturquoise ", ylim = C (8.5,10), xlab = year, ylab = number of inbound passengers in Shanghai (million)")
#PCH represents the shape of scatter points, col represents color, ylim represents Y-axis range, ylab represents Y-axis title, and main represents image title
Points (sample, grey, PCH = 16, col = deeppink ", CEX = 1) ○ CEX represents the size of scatter points
points(sample,impgrey,pch=17,col="RosyBrown",cex=1)
points(sample,greymarkov,pch=18,col="green",cex=1)
points(sample,impgreymarkov,pch=11,col="red",cex=1)
points(sample,huigui,pch=12,col="blue",cex=1)
points(sample,timeseires,pch=13,col="black",cex=1)
Lines (true value, col = darkturquoise ", lty = 1) ○ lty = 1 means to connect with a solid line
Lines (grey, col = deeppink ", lty = 2) ○ lty = 2 means to connect with a dotted line
Lines (impgrey, col = rosybrown ", lty = 3) ○ lty = 3 means to connect with a dot line
lines(greymarkov,col="green",lty=4)
lines(impgreymarkov,col="red",lty=5)
lines(huigui,col="blue",lty=6)
lines(timeseires,col="black",lty=7)
Legend ("topleft", CEX = 0.5, C ("true value", "GM (1,1) model", "improved GM (1,1) model", "Grey Markov model", "Improved Grey Markov model", "linear regression", "time series"), col = C ("darkturquoise", "deeppink", "rosebrown", "green", "red", "blue", "black"),text.col=c("DarkTurquoise","DeepPink","RosyBrown","green","Red","blue","black"),pch =c(15,16,17,18,11,12,13),lty=c(1,2,3,4,5,6,7))
#12 means the x-axis coordinate is 12400, and the y-axis coordinate is 400, which means the left and upper boundaries of the legend,text.colRepresents the legend text color
```

6. EEMD Program

```
function allmode=eemd(Y,Nstd,NE)
%function allmode=eemd(Y,Nstd,NE)
%Y=xlsread('D:/Program Files/bin/data.xlsx).
Y=importdata('new zealand.txt).
%part1.read data, find out standard deviation ,devide all data by std
xsize=length(Y);
dd=1:1:xsize;
Ystd=std(Y);
Y=Y/Ystd;
NE=100;
Nstd=0.2;
%part2.evaluate TNM as total IMF number,ssign 0 to N*TNM2 matrix
TNM=fix(log2(xsize))-1; % TNM=m
```

```
TNM2=TNM+2;
for kk=1:1:TNM2,
for ii=1:1:xsize,
allmode(ii,kk)=0.0;
End
End
%part3 Do EEMD -----EEMD loop start
for iii=1:1:NE, %EEMD loop NE times EMD sum together
%part4 --Add noise to original data,we have X1
for i=1:xsize,
temp=randn(1,1)*Nstd; % add a random noise to Y
X1(i)=Y(i)+temp;
End
%part4 --assign original data in the first column
for jj=1:1:xsize,
mode(jj,1) = Y(jj); % assign Y to column 1of mode
End
%part5–give initial 0 to xorigin and xend
xorigin = X1;
xend = xorigin;
%part6–start to find an IMF-----IMF loop start
nmode = 1;
while nmode <= TNM,
xstart = xend; %last loop value assign to new iteration loop
%xstart -loop start data
iter = 1; %loop index initial value
%part7–sift 10 times to get IMF—sift loop start
while iter<=10,
[spmax, spmin, flag]=extrema(xstart); %call function extrema
%the usage of spline ,please see part11.
upper= spline(spmax(:,1),spmax(:,2),dd); %upper spline bound of this sift
lower= spline(spmin(:,1),spmin(:,2),dd); %lower spline bound of this sift
mean_ul = (upper + lower)/2; %spline mean of upper and lower
xstart = xstart - mean_ul; %extract spline mean from Xstart
iter = iter +1;
End
%part8–subtract IMF from data ,then let the residual xend to start to find next IMF
xend = xend - xstart;
nmode=nmode+1;
%part9–after sift 10 times,that xstart is this time IMF
for jj=1:1:xsize,
mode(jj,nmode) = xstart(jj);
End
End
%part10–after gotten all(TNM) IMFs ,the residual xend is over all trend
% put them in the last column
for jj=1:1:xsize,
mode(jj,nmode+1)=xend(jj);
End
%after part 10 ,original + TNM IMFs+overall trend —those are all in mode
allmode=allmode+mode;
end %part3 Do EEMD -----EEMD loop end
%part11–devide EEMD summation by NE,std be multiply back to data
allmode=allmode/NE;
```

```
allmode=allmode*Ystd;
figure(1)
m = size(allmode,2);
for i = 2:m
subplot(m,1,i)
plot(allmode(:,i))
End
%Calculate the average period
figure(2)
for n=2:m
subplot(m,1,n)
maximum= findpeaks(allmode(:,n))
period=length(maximum)
aperiod=xsize/period
findpeaks(allmode(:,n))
End
imf1=allmode(:,2)
maximum1= findpeaks(imf1)
period1=length(maximum1)
aperiod1=xsize/period1
imf2=allmode(:,3)
maximum2= findpeaks(imf2)
period2=length(maximum2)
aperiod2=xsize/period2
imf3=allmode(:,4)
maximum3= findpeaks(imf3)
period3=length(maximum3)
aperiod3=xsize/period3
imf4=allmode(:,5)
maximum4= findpeaks(imf4)
period4=length(maximum4)
aperiod4=xsize/period4
imf5=allmode(:,6)
maximum5= findpeaks(imf5)
period5=length(maximum5)
aperiod5=xsize/period5
imf6=allmode(:,7)
maximum6= findpeaks(imf6)
period6=length(maximum6)
aperiod6=xsize/period6
res=allmode(:,8)
maximum= findpeaks(res)
period=length(maximum)
aperiod=xsize/period
%Calculate variance contribution rate
imf=allmode(:,2:m)
imf=imf'
[a,b]=size(imf);
for i=1:a
%Calculate the variance contribution rate of each IMF
%Definition: the mean square of variance minus the square of the mean
%The square of the mean
%imfp2=mean(c(i,:),2).^2
%Mean square
```

```
%imf2p=mean(c(i,:).^2,2)
%Variance of each IMF
mse(i)=mean(imf(i,:).^2,2)-mean(imf(i,:),2).^2;
end;
MMSE = sum (MSE);% total variance
for i=1:a
%Variance percentage, that is, variance contribution rate
mseb(i)=mse(i)/mmse*100;
end;
figure(3);
bar([1:a],mseb);
%Title ('variance contribution rate of each IMF component ');
Xlabel ('variance ') and ylabel ('percentage')
```

Citation: Bin Zhao, et al (2020) The Forecast of The Number of Inbound Tourists and The Analysis of The Source Market During the Epidemic of Coronavirus Disease. Int J Biol Genetics; 3(1): 1-53

DOI: 10.31829/2692-4242/biogen2020-3(1)-109