ISSN: 2688-9323

Forecasting Impact on Tourism during the COVID-19

Yi Huang[1], Wei Long[1], Yanwei Wang[2,*]

[1]College of Economics and Management, Yunnan Agricultural University, Kunming 650201, China

[2]College of Architecture and Engineering, Yunnan Agricultural University, Kunming 650201, China

*2009030@ynau.edu.cn

Abstract

Tourism contributes to the sustainable development and booming of global economy, However, the pandemic caused by COVID-19 has hit hard on tourism with rapidly declining number of tourists. With the intention to investigate the uncertainty brought by this pandemic, this study thus develops a Shapley value combination forecasting model. In this study, an empirical case study was undertaken to analyze the data from Yunnan Province China. It was found that the measurement error of combined forecasting model based on Shapley value is smaller than that of SARIMA model,GM(1,1) model and Holt-winters respectively, and the accuracy of combined forecasting model is higher. The results also show that cumulative loss rate of domestic tourists in Yunnan Province from January to December in 2020 is 47.38%. Moreover, with the loss rate of each month slowing down and implementation of various measures, the tourism industry of Yunnan Province is gradually recovering.

Keywords

COVID-19; Tourism; Shapley Value; Combination Forecasting.

1. Introduction

Recently years, tourism has made an important contribution to the sustainable development of the global economy. However, impacted by pandemic, the number of tourists has dropped sharply, and tourism has been hit hard. There is a need to understand the current context and foresee the future of tourism by measuring the impact. However, the main focus of the existing qualitative studies has mainly been shifted on coping strategies of how to deal with the stress on tourism caused by COVID-19. For example, Shi et al. [1] found during the pandemic period, the tourism industry did not show retaliatory growth, and control strategies should be implemented into the protection of the industry capacity. Li [2] also found that the tourism industry has received the most serious damage by pandemic, as it requires high level of collaboration and involvement of mobility of people in different locations. Other researchers, such as Aburumman [3], investigated the COVID-19 impact on both the global and the UAE MICE markets and identified a competitive survival strategy for tourism companies based upon examples in the UAE. Xia et al. [4] studied the impact on domestic tourists and found the impact level of total tourism expenditure loss are directly related to the progress and pandemic coping context. Moreover, the current quantitative research mainly focuses on using time series regression model to analyze the impact on tourism industry. For instance, You et al. [5] established a time series regression model based on the domestic tourism income data from 1999 to 2018 to calculate the short-term and annual tourism economic losses. Uğur et al. [6] found that the tourism sector was easily affected by global crises. Mykhailo et al. [7] used a statistical analysis to predict the economic losses and recommended that city government authorities reduce local taxes and fees in the tourism sector as soon as possible to restore its potential normality. He et al. [8] used the ARIMA bivariate model to predict the impact on domestic and inbound tourism changes, and proposed relevant strategies.

In summary, existing literature has been focused on adopting the effective strategies to cope with the pandemic, and measuring the impact on tourism industry by using one single method. Moreover, it was found that there was limited research using various combinations of methods. Thus, in order to more accurately predict and overcome the limitation of "year-on-year", this study used the combined forecasting model based on Shapley value to quantitatively analyze the impact level of pandemic on the tourism industry.

2. Method

This study selects Yunnan Province of China as an example to study. SARIMA model, GM(1,1) model and Holt winters model were chosen to analyze the impact individually first. Then the Shapley value model was used to analyze the three models. Finally, a combined forecasting model was developed to predict the number of tourists impacted by the pandemic.

2.1. SARIMA Model

SARIMA model is the product of ARMA(P,Q) and ARMA(p,q) after d-order trend difference and d-order seasonal difference with period s as step[9-11]. It is developed on the basis of ARMA. It is not only a random time series prediction method, but also a time series short-term prediction method with high accuracy. If a series repeats a certain feature in a certain time interval, it is called a time series with seasonality, such as holiday tourism consumption and other time series with obvious seasonal changes. SARIMA model should be used for non-stationary time series with seasons and trends. Its basic idea is: first, test the stationarity of the time series. If the time series is not stable, it is necessary to eliminate the regularity of the series through phase by phase difference and seasonal difference to make the series stable; second, test the non white noise; the third step is to determine the order of the model; the fourth step is to estimate the model parameters; the fifth step is to test the model.

2.2. GM(1,1)Model

The grey theory was put forward by Professor Deng Julong in the 1980s. It is a method to predict the uncertainty when some information is known and some information is unknown. By analyzing the certain and confirmed information, what will happen in reality can be predicted and described. Its basic idea is: first, preprocess the original sequence by one-time accumulation to generate the sequence of the nearest mean value; second, construct the coefficient matrix and data vector, and calculate the parameter column; third, construct GM(1,1) and solve the grey differential equation; fourth, regressive reduction to get the predicted value, and test the accuracy[12-13].

2.3. Holt-winters Exponential Smoothing Model

The Holt-winters model is an exponential smoothing forecasting method. It is fit for time series with decreasing or increasing tendency, seasonality and described by additive model. The Holt-winters method depends on three parameters to estimate the level, slope, and seasonality of the current time point [14].

$$F_{t+m} = (S_t + b_t m) I_{t-L+m} \tag{1}$$

Where L is the length of the season, and I is the seasonal correction index. The steady, trending, and seasonal equations are:

$$S_t = \alpha \frac{x_t}{I_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1}), 0 < \alpha < 1 \tag{2}$$

$$b_t = \beta(S_t - S_{t-1}) + (1-\beta)b_{t-1}, 0 < \beta < 1 \tag{3}$$

$$I_t = \gamma \frac{x_t}{S_t} + (1-\gamma)I_{t-L}, 0 < \gamma < 1 \tag{4}$$

In Eqs. (2)-(4), α is used to estimate the current time level; β is used to estimate the slope of the trend portion of the current time; and γ is used to estimate the seasonal portion of the current time. The values of the three parameters, α, β, and γ are the same. In the interval between 0 and 1, the closer the parameter value is to 0, the smaller the weight of the recent observation value for the future prediction value. The closer the parameter value is to 1, the more recent is the observation, and the observation would have greater weight for future predictions.

2.4. Establishment of Shapley Value Combination Model

In this study, the weights of SARIMA, GM (1,1) and Holt-winters methods were determined by Shapley value method. Shapley value method is often used to calculate the contribution of the team members, which can reflect the different roles that each member play in the cooperative team. All parties or members are treated fairly and the method makes it easy to be recognized by all [15-16].

Using m methods to predict the same problem, m models are established respectively. Thus, the combined forecasting model is as follows:

$$A = \sum B_i a_i \quad i = 1, 2 \cdots, m \tag{5}$$

In Formula (5), A is the final value of the combined forecasting model; a_i is the predicted value of model i, and B_i represents the weight of the predicted model i, and

$$\sum B_i = 1 \tag{6}$$

3. Case Study

3.1. Prediction Results of SARIMA Model

On the basis of SARIMA model, the data from January to December in 2019 were predicted by Eviews[17]. Table 2 shows the prediction results. In 2019, the actual number of domestic tourists in Yunnan Province is 80.05 million, and the number predicted by SARIMA model is 78.47 million. The error between the sum of the predicted values and the actual values in each month in 2019 was 1.9%, which was within the acceptable range.

3.2. Prediction Results of GM (1,1) Model

This paper used GM(1,1) to predict the number of domestic tourists in Yunnan Province from January to December 2019, and used MATLAB 7.11 to test the model. Since the tourism data involved has obvious seasonal trend, the data from January 2013 to December 2018 were divided into 12 groups according to the monthly data (see Table 1). The GM(1,1) method was then used to predict the data from January to December in 2019.

ISSN: 2688-9323

Table 1. Data grouping results from January 2013 to December 2018 (Unit:ten thousand people)

date	2013	2014	2015	2016	2017	2018
January	1557.23	1771.07	1984.79	2547.98	4545.67	4507.53
February	2228.56	2894.99	3141.69	4033.16	4650.43	6916.59
March	1793.86	2051.49	2390.41	3068.70	4306.40	5776.43
April	2045.25	2287.06	2637.86	3386.36	3934.41	5202.57
May	2047.99	2362.19	2764.19	3548.54	4652.63	5612.86
June	2109.73	2247.39	2661.24	3416.38	4326.52	5586.99
July	2195.55	2648.06	3022.42	3880.048	5234.94	6162.39
August	2466.91	2884.54	3459.29	4440.88	5977.20	6942.44
September	1861.08	2283.17	2702.65	3469.54	4918.95	5594.02
October	2371.99	2750.42	3069.00	3939.84	6088.96	7186.30
November	1667.76	1927.78	2282.54	2930.22	4129.68	4758.33
December	1626.44	2021.91	2226.04	2857.69	3906.34	3901.26

Table 2 shows the predicted number of domestic tourists from January to December 2019. The actual number of domestic tourists in Yunnan Province in 2019 is 80.05 million, and the number predicted by GM (1,1) model is 817.87 million. The error between the predicted value and the actual value in 2019 was 2.2%, within the acceptable range.

3.3. Prediction Results of Holt Winters Model

Based on Holt-winters model, the data from January to December in 2019 were predicted by Eviews. Table 2 shows the prediction results. In 2019, the actual number of domestic tourists in Yunnan Province is 80.05 million, and the number predicted by Holt winters model is 79.36 million. The error between the sum of the predicted values and the actual values in each month in 2019 was 0.8%, which was within the acceptable range.

3.4. Weight Analysis of Each Method based on Shapley Value Combination Predicting Model

According to the results of SARIMA, GM(1,1) and Holt-winters forecasting model, the average value of absolute error was calculated respectively. The average value of absolute error of grouped GM(1,1) forecasting model was found the smallest, while the SARIMA model is the second smallest, Holt-winters is the third smallest. Using the average value of the absolute error, the average value of the total error of the combined model was calculated.

$$Q=(209.06+194.43+212.18)/3=205.22$$

According to Shapley value principle, the single forecasting model in the combination model was G={1,2,3}, The combination errors of all subsets were Q({1}), Q({2}), Q({3}), Q({1,2}), Q({2,3}), Q({1,2,3}), The average values of vector errors were 209.06, 194.43, 212.18, 201.75, 210.62, 203.31, 205.22. Shapley value of each member was calculated.

$$\begin{aligned} Q_1 &= \frac{2!\times 0!}{3!}[A\{1\}-A(\{1\}-\{1\})]+\frac{1!\times 1!}{3!}[A\{1,2\}-A(\{1,2\}-\{1\})]+\frac{1!\times 1!}{3!}[A\{1,3\}-A(\{1,3\}-\{1\})]+\frac{0!\times 2!}{3!}[A\{1,2,3\}-A(\{1,2,3\}-\{1\})] \\ &= \frac{1}{3}(209.06)+\frac{1}{6}(201.75-209.06)+\frac{1}{6}(203.31-212.18)+\frac{1}{3}(205.22-210.62) \\ &= 71.28 \end{aligned}$$

ISSN: 2688-9323

In the same way, the error amount allocated by group GM(1,1) was calculated Q2=60.31, the error amount allocated by group Holt-winters was calculated Q3=73.63, and Q1+Q2+Q3= 205.22. the total error was found equal to the sum of the errors of three single methods. Therefore, the weights of each method in the combination model were as follows:

$$\omega_1 = \frac{205.22 - 71.28}{205.22(3-1)} = 0.326 \quad \omega_2 = \frac{205.22 - 60.31}{205.22(3-1)} = 0.353 \quad \omega_3 = \frac{205.22 - 73.62}{205.22(3-1)} = 0.321$$

The combination forecasting model was: A =0.326×(the predicting results of SARIMA) +0.353 ×(the predicting results of GM (1,1)) +0.321×(the predicting results of Holt-winters).

3.5. Implementation of Combination Model

The data from January 2013 to December 2018 are used for training, and the data from January to December 2019 are calculated. Based on the combined forecasting method, the empirical simulation is carried out. The simulation result is shown in Table 2.

Table 2. Results of each prediction model (Unit:Ten thousand people)

date	Actual value	SARIMA	GM(1,1)	Holt-winters	The combination forecasting model
2019.01	5381	5195	5455	5369	5343
2019.02	8519	7957	8516	8204	8234
2019.03	6526	6644	6688	6518	6619
2019.04	6081	5995	6107	6249	6116
2019.05	6541	6465	6657	6655	6594
2019.06	6523	6437	6692	6538	6559
2019.07	7097	7096	7286	7249	7212
2019.08	7790	7997	8032	8152	8059
2019.09	6315	6440	6685	6454	6531
2019.1	8231	8267	8759	7910	8326
2019.11	5680	5480	5862	5354	5574
2019.12	5321	4497	5048	4706	4759
Total	80005	78470	81787	79358	79926
the error compared with the actual value (%)	—	1.9	2.2	0.8	0.1

From Table 3, GM (1,1) is better than the other two models in three single models, and the prediction error of linear regression model is the largest under all error indicators. the various error indices of the combined model are smaller than the other three single models, and the advantage of the combination forecasting model is very obvious.

Table 3. Errors of each prediction model

	MAE	SSE	MSE	MAPE	MSPE	RMSE
SARIMA model	209	96973	90	0.033	0.015	311
GM(1,1) model	194	57670	69	0.029	0.010	240
Holt-winters model	212	75125	79	0.032	0.013	274
Combination model	159	47138	63	0.025	0.010	217

3.6. Prediction Results

According to the above combined forecasting model, the number of people not affected by the epidemic from January to December in 2020 was measured, the predicted baseline value and loss rate were obtained. The difference between the actual value and the predicted value is the decreased number of tourists due to the impact. The percentage of the difference between the two values and the predicted value is the loss rate. Table 4 shows the predicted value and loss rate.

Table 4. Predicted value and loss rate of domestic tourists from January to December 2020

date	Predicted number of tourists (ten thousand people)	Actual number of tourists (ten thousand people)	The loss rate (%)
2020.01	6639.17	4122.94	38.71
2020.02	10467.69	309.45	97.10
2020.03	7942.87	2147.91	73.38
2020.04	7523.66	3396.78	55.45
2020.05	7957.46	4330.26	46.52
2020.06	7935.13	4735.21	41.57
2020.07	8630.58	5261.96	40.28
2020.08	9530.22	6101.23	37.26
2020.09	7706.91	5362.27	32.16
2020.10	10021.31	6903.12	33.32
2020.11	6780.05	5288.47	24.60
2020.12	6081.38	5040.40	17.12
Total	91135.04	47959.60	47.38

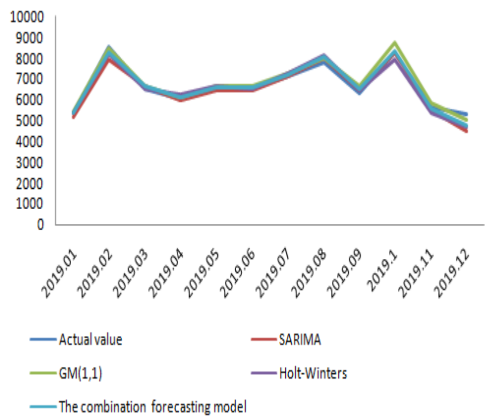

Figure 1. Trend fitting of each prediction model results and actual values in 2019

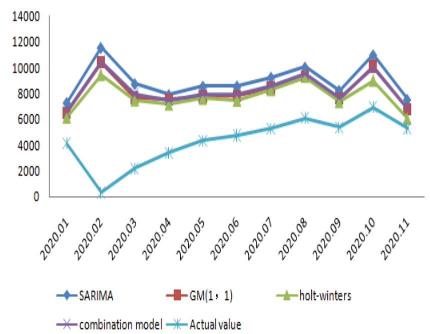

Figure 2. Trend fitting of each prediction model results and actual values in 2020

ISSN: 2688-9323

As can be seen from the above, the cumulative loss rate from January to December in 2020 is 47.38%, and the highest loss rate is 97.1% in February.

4. Conclusion

This study used the Shapley value combination forecasting model to investigate the impact on Yunnan's tourism industry, and it was found the error was smaller than that of SARIMA model, GM(1,1) model and Holt-winters model. Thus, the combination model has high prediction accuracy.

Yunnan's tourism industry has been hit hard since January 2020, and the number of domestic visitors has dropped sharply. The total loss rate was 47.38% from January to December in 2020. In particularly, during February 2020, the number of visitors lost in Yunnan was 97.1%. This made February the highest loss month in 2020.Since March, the loss rate of the number of domestic tourists has slowed down, which shows that with various prevention and control measures of the pandemic context, the tourism industry in Yunnan Province is gradually recovering.

This proposed model has some limitations. For example, the combination forecasting model has only combined three models; we shall add another forecasting method on the basis of the combination forecasting model in the future. Moreover, currently limited methods were used in the weight determination method. Other methods such as entropy method, and AHP can be used in the future studies. Finally, comparative studies can be conducted to compare and develop a further combination forecasting model.

Acknowledgments

This research is supported by Bureau Statistics of Yunnan, Province Republic of China and the data were supplied by the Tourism Development Commission of Yunnan Province, Republic of China.

Funding Resources: This research was funded by 16XJC630009, 2018FB111 and A2032021032.

References

[1] Shi,P.; Lu,M.M. Research on novel coronavirus pneumonia prevention and control and tourism health protection capability construction -- a new perspective and new change in the impact of new crown pneumonia on tourism industry and Countermeasures Research. Journal of Xinjiang Normal University (PHILOSOPHY AND SOCIAL SCIENCES). 2020,41 (06): 55-67.

[2] Li, J. Research on the impact of public health emergencies on tourism. Tourism overview (second half of the month). 2020 (10): 38-39.

[3] Aburumman Asad A.. COVID-19 impact and survival strategy in business tourism market: the example of the UAE MICE industry. Humanities and Social Sciences Communications.2020, 1.

[4] Xia,J.C.; Feng,X. X. Novel coronavirus pneumonia impact on tourism industry and Countermeasures. China circulation economy.2020, 03, 3-10.

[5] You, J.; Peng, G.C.; Li, Q. Impact of epidemic situation on tourism economy and countermeasures. Special Economic Zone. 2020, 07, 85-88.

[6] Uğur Naciye Güliz, Akbıyık Adem. Impacts of COVID-19 on global tourism industry: A cross-regional comparison. Tourism management perspectives. 2020, 36.

[7] Mykhailo Rutynskyi, Halyna Kushniruk. The impact of quarantine due to COVID-19 pandemic on the tourism industry in Lviv (Ukraine). Problems and Perspectives in Management. 2020, 2.

[8] He,X.R.;Peng, K.J. Impact of novel coronavirus pneumonia on Tourism in China: prediction and Countermeasures. Journal of Sichuan Tourism College.2020 ,04, 65-71.

ISSN: 2688-9323

[9] P. Kabbilawsh, D. Sathish Kumar,N. R. Chithra. Trend analysis and SARIMA forecasting of mean maximum and mean minimum monthly temperature for the state of Kerala, India. Acta Geophysica: Official Journal of The Institute of Geophysics, PAS and Polish Academy of Sciences. 2020,68(4).

[10] Paisit Khanarsa , Krung Sinapiromsaran. Automatic SARIMA Order Identification Convolutional Neural Network. International Journal of Machine Learning and Computing. 2020,10(5).

[11] Nwokike Chukwudike C.,Offorha Bright C., Obubu Maxwell, Ugoala Chukwuma B.,Ukomah Henry I.. Comparing SANN and SARIMA for forecasting frequency of monthly rainfall in Umuahia. Scientific African. 2020,10.

[12] Maolin Cheng,Guojun Shi, Mingyin Xiang. On the improvement of the parameter estimation of the grey model GM(1,1) and model application. Communications in Statistics - Simulation and Computation. 2020,49(5).

[13] Zeynep Ceylan,Serol Bulkan,Sermin Elevli. Prediction of medical waste generation using SVR, GM (1,1) and ARIMA models: a case study for megacity Istanbul. Journal of Environmental Health Science and Engineering. 2020.

[14] C. Chris and M. Yar, "Holt-winters Forecasting: Some Practical Issues," The Statistician, Vol.37, No.2, pp. 129-140, 1988.

[15] J.M. Gallardo, A. Jiménez-Losada. A characterization of the Shapley value for cooperative games with fuzzy characteristic function. Fuzzy Sets and Systems. 2020,398.

[16] Dhrubajit Choudhury, Surajit Borkotokey, Rajnish Kumar, Sudipta Sarangi. The Egalitarian Shapley value: a generalization based on coalition sizes. Annals of Operations Research. 2020.

[17] Yi Danhui. Data analysis and Eviews application. Beijing: China Renmin University Press, 2008.