

Forecasting the trend of tourism industry in the United States: using ARIMA model and ETS model

Siyang Lin *

School of Macau University of Science and Technology, Macau, China

* Corresponding Author Email: 2009853fb111006@student.must.edu.mo

Abstract. As a result of the 2019 epidemic, all industries around the world have been hit hard, especially the tourism industry. Governments have introduced many measures to maintain the tourism industry, but the results are not satisfactory, and the industry has even experienced negative growth. In order to predict the future trend of the tourism industry and to make decisions, this paper has selected five different data on airlines, hotels, car rentals and travel agencies in the US tourism industry as the basis, and will use two models, ETS and ARIMA, to forecast the data from 2000 to 2020 respectively, to obtain the data without the impact of the epidemic, and then compare the forecast results with the historical real data. The results will then be compared with the historical data. Finally, a model suitable for each data was obtained.

Keywords: ARIMA model; ETS model; Tourism industry; Forecasting.

1. Introduction

Since the outbreak of the epidemic in 2019, Covid-19 has become a hot issue of concern to the whole society, and various countries have introduced different epidemic prevention policies to combat the epidemic. However, under the restrictions of these epidemic prevention policies, many industries have been affected to a certain extent, especially the tourism industry. In order to predict the impact of the epidemic on the tourism industry, this paper selects five data, including the passenger volume data of the US aviation industry and the reservation price index of travel agencies and cruise companies, as the research objects for further prediction. This prediction can not only get the degree of impact of the epidemic on the future tourism industry, but also can help the US tourism industry to better determine future actions.

This paper will collect historical data of five data, such as the passenger volume data of the US airline industry, the booking price index of travel agencies and cruise companies, etc., and because the passenger volume data of US airlines included these five data Long-term trends and seasonal change patterns, these data will suddenly rise or fall due to epidemics and policy changes, we choose to use ARIMA model and ETS model to predict these time series, and then compare with historical data to choose the most suitable 's model. Finally, It is concluded that Load Factor for Domestic and International Air Carrier in the United States, as well as Exports of Services: Travel, are suitable for prediction by the ARIMA model, and Producer Price Index by Industry: Travel Companies: Producer Price Index by Industry: Cruise and Tour Bookings Passenger Car Rental: Rental of Passenger Cars for Business Travel and Revenue Passenger Miles for Domestic and International US Air Carrier are suitable for ETS prediction.

A stationary stochastic model called the ARIMA combines the moving average and autoregressive models [1]. The basis of this group of approaches is the ARIMA model, which is typically tailored to longer travel movements. In recent years, the use of seasonal ARIMA models has increased [2]. The univariate and multivariate ARIMA models are used to forecast air transportation demand [3]. An ARMA-based methodology is used to forecast tourism demand [4]. To determine the number of tourists in Croatia, ARIMA is used [5]. The model which is based on historical data, also takes into account the intervention of war activities in Croatia between 1991 and 1995. International visitor arrivals in Malaysia are forecast using RIMA models [6]. ARIMA model are used to predict tourism demand in Montenegro [7].

Meanwhile, exponential smoothing (ETS) models are used less frequently in tourism demand studies, although there is ample evidence that ES often makes reasonable predictions of shifts in the direction and trend of tourism demand. Similarly, in tourism demand forecasting, the performance of singular spectrum analysis was compared to ARIMA, exponential smoothing (ETS), and neural networks (NN) [8]. After using Nave 1, simple exponential smoothing (SES), double exponential smoothing (DES), ARIMA, and other methods to estimate tourist arrivals in Hong Kong [9], Annual international tourist arrivals in Zambia were studied using ARIMA and Holt-Winters exponential smoothing [10]. The findings indicate that ETS is a suitable forecasting model with a high degree of forecasting accuracy.

2. Data analysis

In order to accurately predict future trends in the US tourism industry, this paper collects data on different aspects of the industry to forecast.

2.1. Load Factor for U.S. Air Carrier Domestic and International (Load)

Data on the total number of passengers carried by the US airline industry, with the y-axis showing the total number of passengers carried, increasing slowly and incrementally until 2020, falling off a cliff by 2020, and then slowly increasing again. See Fig. 1.

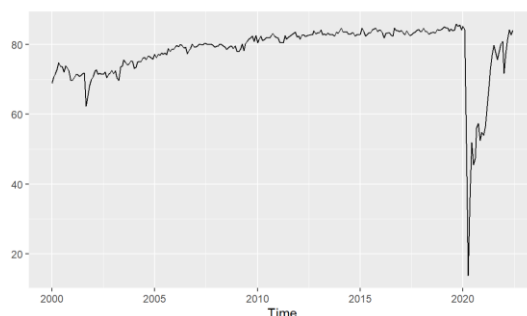


Fig 1. Load Factor for Domestic and International Air Carrier Scheduled Passenger Flights in the United States (Load)

Source: Load Factor for U.S. Air Carrier Domestic and International, Scheduled Passenger Flights (LOADFACTOR11) | FRED | St. Louis Fed [11]

2.2. Revenue Passenger Miles for U.S. Air Carrier Domestic and International (Rpm)

Data on total US aviation revenues, with the y-axis in US dollars, show seasonal growth until 2020, declining precipitously to 2020 and then slowly increasing again. See Fig. 2.

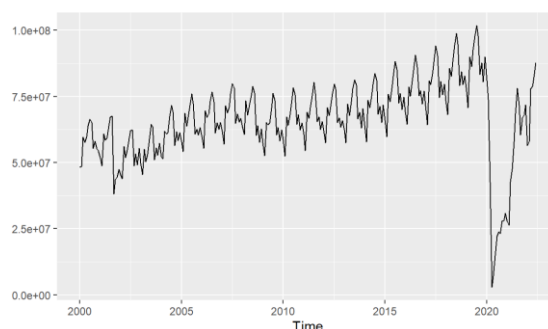


Fig 2. Revenue Passenger Miles for Domestic and International US Air Carrier Scheduled Passenger Flights (Rpm)

Source: Revenue Passenger Miles for U.S. Air Carrier Domestic and International, Scheduled Passenger Flights (RPM) | FRED | St. Louis Fed [12]

2.3. Producer Price Index by Industry: Travel Agencies: Cruise and Tour Bookings (Pcu)

Travel agents. Price index data for cruise and tour bookings, with the y-axis showing the price index, are flat until 2012, increasing slightly after 2012 and falling off a cliff to 2020, before slowly increasing again. See Fig. 3.

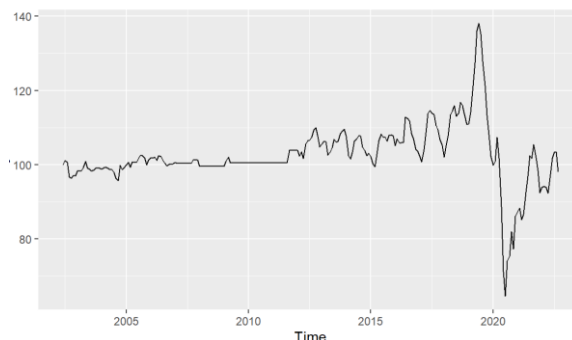


Fig 3. Producer Price Index by Industry: Travel Agencies: Cruise and Tour Bookings (Pcu)

Source: [13] <https://fred.stlouisfed.org/series/PCU561510561510222>

2.4. Exports of Services: Travel (for all purposes including education) (Iea)

Data for exports of tourism services, with the y-axis in millions of dollars, growing slowly until 2020, falling off a cliff in 2020 and then recovering sharply. See Fig. 4.

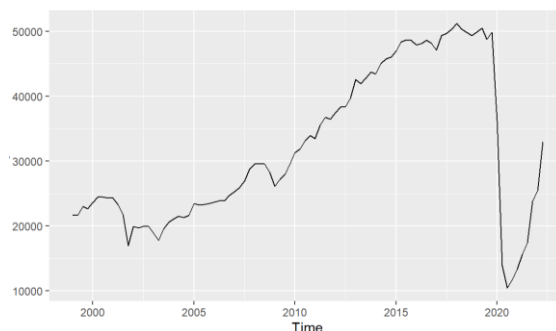


Fig 4. Travel Services Exports (for all purposes including education) (Iea)

Source: Exports of Services: Travel (for all purposes including education) (IEAXSTV) | FRED | St. Louis Fed [14]

2.5. Producer Price Index by Industry: Passenger Car Rental: Passenger Car Rental for Business Travel (Pcuc)

Data from the rental car industry price index, with the y-axis showing the price index, fell off a cliff in 2020 and then slowly rebounded. See Fig. 5.

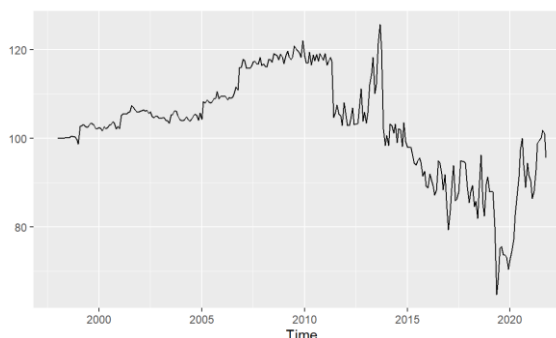


Fig 5. Industry Producer Price Index: Passenger Car Rental: Business Travel Passenger Car Rental (Pcuc)

Source: Producer Price Index by Industry: Travel Agencies: Hotel and Car Rental Bookings (PCU561510561510211) | FRED | St. Louis Fed [15]

3. Model

This paper will use ETS, ARIMA to forecast the historical data (2000-2020) and then compare it with the real data to select a more accurate forecast method.

3.1. ETS

The ETS algorithm is especially useful for data sets with a priori seasonal and other assumptions. ETS computes the weighted average of all observations in the input time series data set as its forecast. In contrast to the simple moving average method, the weights decrease exponentially over time. ETS is used as one of the forecasting methods in this paper due to its simplicity and ability to forecast in the short term.

The (A, N, N) model, which has additive error, no trend, and no seasonality, corresponds to the simplest ETS model, also known as simple exponential smoothing. The Holt method's state space equation is

$$\begin{aligned} y_t &= y_{t-1} + e_t \\ l_t &= l_{t-1} + \alpha e_t \end{aligned} \quad (1)$$

This state space formulation can be used to generate various formulas, predictions, and smoothing equations.

$$\begin{aligned} \hat{y}_{t|t-1} &= l_{t-1} \\ l_t &= \alpha y_{t-1} + (1 - \alpha)l_{t-1} \end{aligned} \quad (2)$$

In the simple exponential smoothing model, the forecast corresponds to the previous level. A weighted average of the prior level and observation is used in the second equation (the smoothing equation) to calculate the subsequent level.

3.2. ARIMA

The ARIMA model (autoregressive integrated moving average model) is a simple time series forecasting model that forecasts by combining its own past values and errors in a linear fashion.

ARIMA (p, d, q) represents the autoregressive, with p indicating the number of autoregressive terms, MA indicating the moving average, q indicating the number of sliding average terms, and d indicating the number of differences (orders) used to smooth the series.

The term "difference" does not appear in the English name of ARIMA, but it is an important step in smoothing the time series.

The ARIMA (p, d, q) model extends the ARMA (p, q) model, and it is denoted by the following formula:

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) (1 - L)^d X_t = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t \quad (3)$$

$$\text{where } L \text{ represents the lag operator: } d \in \mathbb{Z}, d \neq 0 \quad (4)$$

The data selected in this paper are all non-stationary time series, so this paper will first perform a difference to convert the data into a stationary time series, and then use ARIMA to make predictions and compare the predicted data with the real data.

4. Result

4.1. Forecast in ETS model

4.1.1 Load

According to the calculation, this paper will use ETS (A, N, N) to analyze the load data.

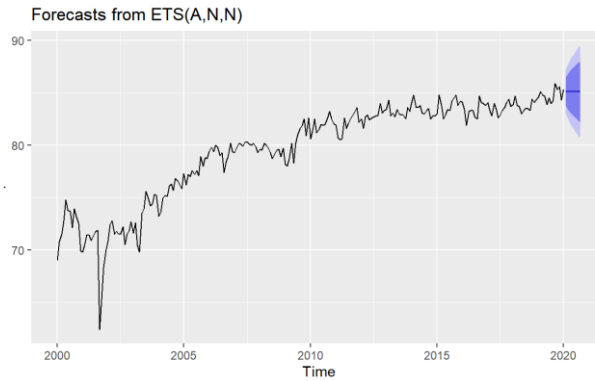


Fig 6. Load data forecasts from ETS model

The EST model's predictions are smooth straight lines. See Fig. 6.

4.1.2 Rpm

According to the calculation, this paper will use ETS (A, N, N) to analyze the rpm data.

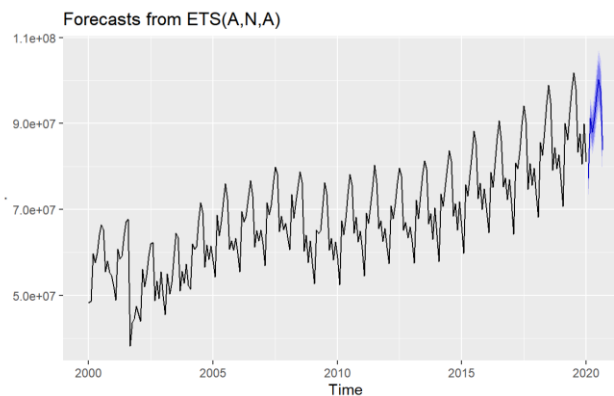


Fig 7. Rpm data forecasts from ETS model

EST model forecasts continued seasonal growth. See Fig. 7.

4.1.3 Pcu

According to the calculation, this paper will use ETS (A, N, N) to analyze the pcu data.

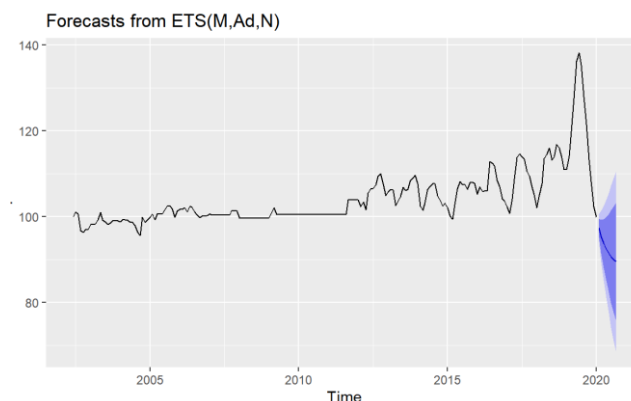


Fig 8. Pcu data forecasts from ETS model

The EST model predicts a continuous decline. See Fig. 8.

4.1.4 Iea

According to the calculation, this paper will use ETS (A, N, N) to analyze the Iea data.

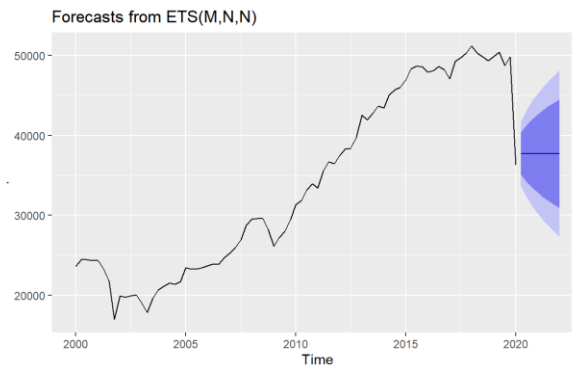


Fig 9. Iea data forecasts from ETS model

The EST model's predictions are smooth straight lines. See Fig. 9.

4.1.5 Pcuc

According to the calculation, this paper will use ETS (A, N, N) to analyze the pcuc data.

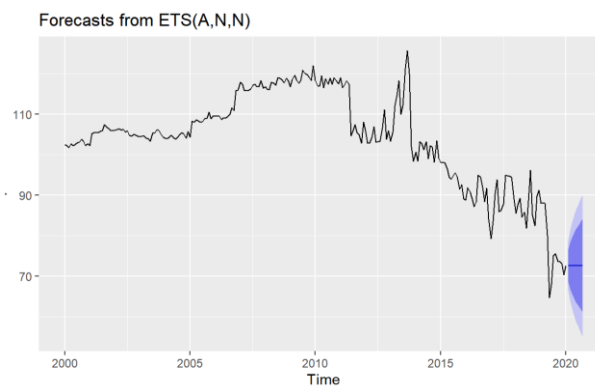


Fig 10. Pcuc data forecasts from ETS model

The EST model's predictions are smooth straight lines. See Fig. 10.

4.2. Forecast in ARIMA model

4.2.1 Load

This paper determine that the load data should have a first-order difference. After first-order differencing, we use ARIMA (1,1,1) (2,0,0) [12] to predict the data.

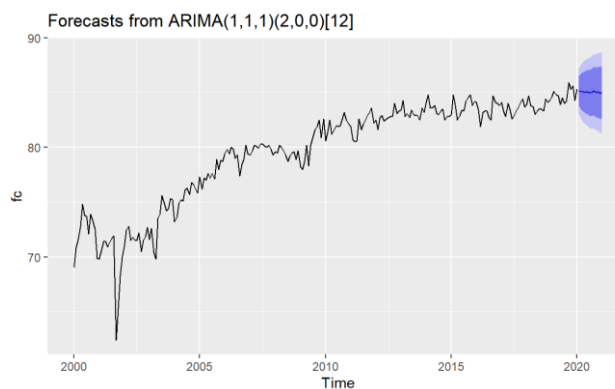


Fig 11. Load data forecasts from ARIMA model

The ARIMA model yields a prediction of a slow decline. See Fig. 11.

4.2.2 Rpm

This paper determine that the rpm data should have one first-order differential and one seasonal differential. After two differentials, we used ARIMA (1,1,2) (2,1,1) [12] to predict the data.

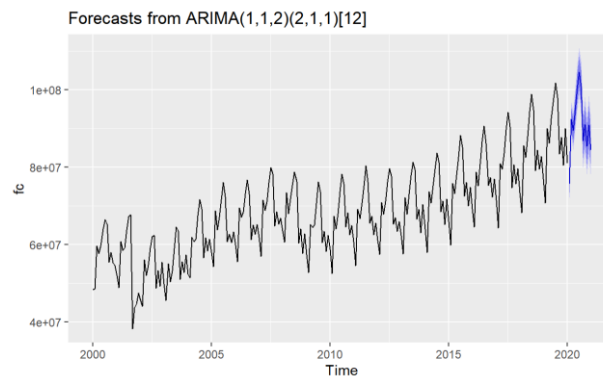


Fig 12. Rpm data forecasts from ARIMA model

The ARIMA model yields a forecast of sustained seasonal growth. See Fig. 12.

4.2.3 Pcu

This paper determine that the pcu data should have a first-order difference. After one difference, we used ARIMA (1,1,3) (1,0,0) [12] to predict the data.

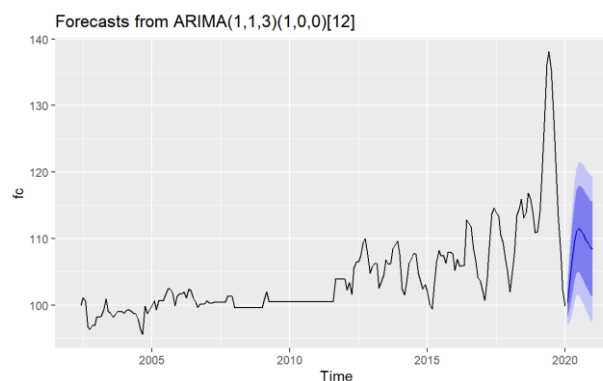


Fig 13. Pcu data forecasts from ARIMA model

The ARIMA model yields a forecast of a small increase followed by a decrease. See Fig. 13.

4.2.4 Iea

This paper determine that the iea data should have a first-order difference. After one difference, we used ARIMA (0,1,0) to predict the data.

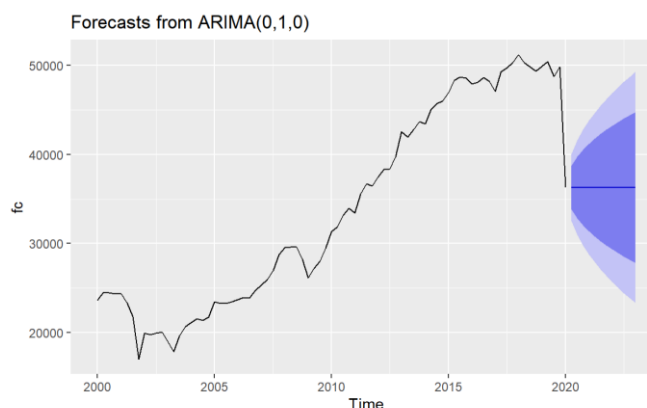


Fig 14 Iea data forecasts from ARIMA model

The ARIMA model yields a forecast of a continuous straight line. See Fig. 14.

4.2.5 Pcuc

This paper determine that the Pcuc data should have a first-order difference. After one difference, we used ARIMA (0,1,2) (0,0,1) [4] to predict the data.

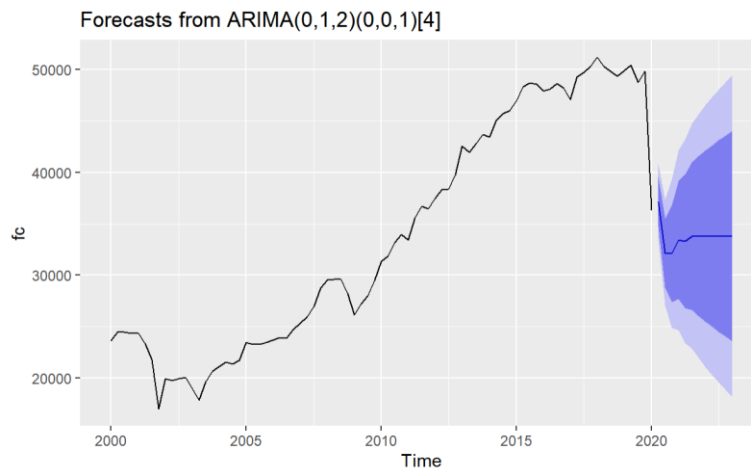


Fig 15. Pcuc data forecasts from ARIMA model

The ARIMA model yields a prediction of a decline followed by a slow rise, followed by a flat, straight line. See Fig. 15.

4.3. Comparison

4.3.1 Load

The figure shows that the ETS model's prediction is a straight line, whereas the ARIMA model's prediction is a downward curve. If there is no impact of the epidemic, the Load data will fluctuate smoothly and slightly. Due to the impact of the epidemic, the data after 2020 will plummet. Because the ARIMA model's predicted value is closer to the original value than the historical data of metadata, the ARIMA model is better suited for predicting Load data. See Fig. 16.

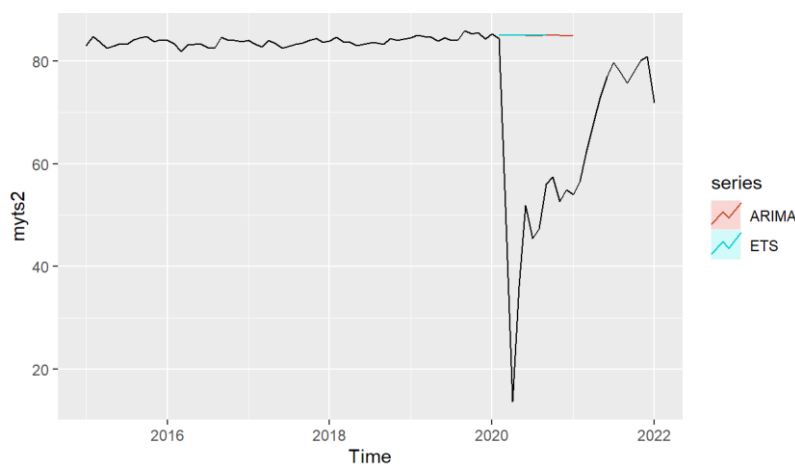


Fig 16. Comparison of Forecasting Between ARIMA Model and ETS Model of Load data

4.3.2 Rpm

The figure shows that the ETS model's prediction is a seasonal curve lower than the peak value of last year, while the ARIMA model's prediction is a seasonal curve higher than the peak value of last year, with a small fluctuation at the end. If there is no impact of the epidemic, the Rpm data will be Continuing to show a seasonal rise, due to the impact of the epidemic, the data after 2020 plummeted. Compared with the historical data of metadata, Because the ETS model's predicted values is closer to the original value, the ETS model is better suited to predicting Rpm data. See Fig. 17.

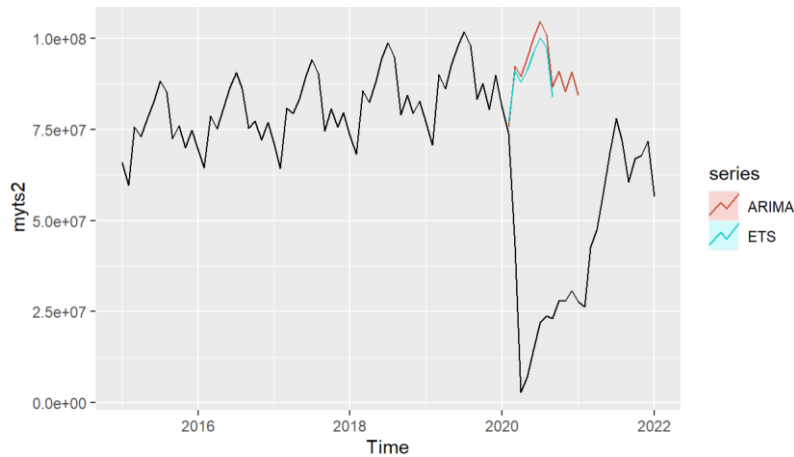


Fig 17. Comparison of Forecasting Between ARIMA Model and ETS Model of Rpm data

4.3.3 Pcu

The figure shows that the ETS model predicts a downward curve, whereas the ARIMA model predicts a curve that rises first and then gently declines. Due to the impact of the epidemic, the data after 2020 has plummeted. When compared to the historical metadata data, the ARIMA model’s predicted value is closer to the original value, indicating that the ARIMA model is better suited for predicting Pcu data. See Fig. 18.

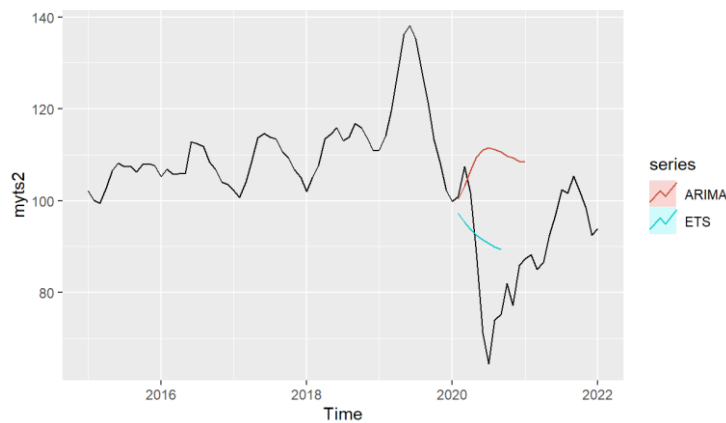


Fig 18. Comparison of Forecasting Between ARIMA Model and ETS Model of Pcu data

4.3.4 Iea

The figure shows that the ETS model's prediction is a straight line, whereas the ARIMA model's prediction is a straight line smaller than the ETS model's predicted value. If there is no impact of the epidemic, the IEA data will be relatively stable. Due to the impact of the epidemic, the data after 2020 will plummet. Compared with the historical data of metadata, the ETS model’s predicted value is closer to the original value, so the ETS model is more suitable for predicting Iea data. See Fig. 19.

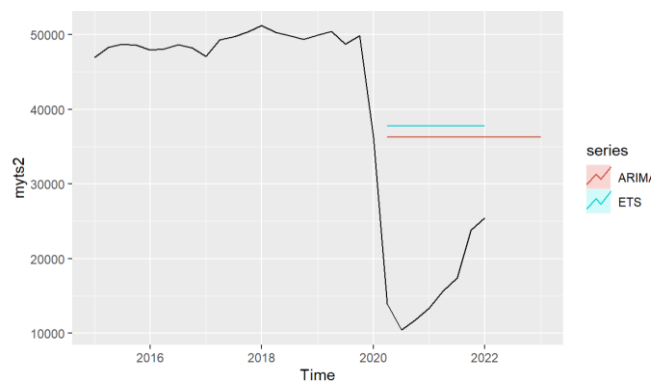


Fig 19. Comparison of Forecasting Between ARIMA Model and ETS Model of Iea data

4.3.5 Pcuc

The figure shows that the ETS model predicts a straight line, whereas the ARIMA model predicts a curve that first declines and then fluctuates slightly. Due to the impact of the epidemic, the data after 2020 will rise sharply. Compared with the historical data of metadata, the ETS's predicted value is closer to the original data, so the ETS model is more suitable for predicting Pcuc data. See Fig. 20.

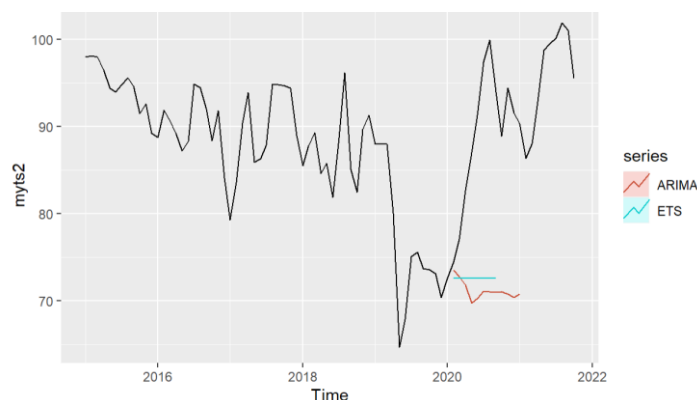


Fig 20. Comparison of Forecasting Between ARIMA Model and ETS Model of Pcuc data

5. Conclusion

This paper using ARIMA model and ETS model to predict the five data of the US tourism industry respectively, and the forecast data without the impact of the epidemic is obtained. At the same time, this paper using the historical real data to compare with the data obtained by the two models, it is concluded that the passenger capacity of the U.S. airline industry and export tourism price data are more suitable for forecasting using the ARIMA model. The price indices of travel agencies, car rentals and the total revenue of the U.S. airline industry are more suitable for forecasting using the ETS model.

References

- [1] Ghalekhondabi, I., Ardjmand, E., Young, W. A., & Weckman, G. R. (2019). A review of demand forecasting models and methodological developments within tourism and passenger transportation industry. *Journal of Tourism Futures*, 5(1), 75-93.
- [2] Coshall, J.T. Combining volatility and smoothing forecasts of UK demand for international tourism. *Tour. Manag.* 2009, 30, 495-511.
- [3] Andreoni, A., & Postorino, M. N. (2006). A multivariate ARIMA model to forecast air transport demand. *Proceedings of the Association for European Transport and Contributors*, 1-14.
- [4] Chu, F. L. (2009). Forecasting tourism demand with ARMA-based methods. *Tourism Management*, 30(5), 740-751.
- [5] Krsić and Gatti (2009). Forecasting methodology of maritime passenger demand in a tourist destination, *Promet/Traffic/Transportation*, Vol 21, 2009, 3 pp 183-190
- [6] Loganathan, N., Ibrahim, Y. (2010), "Forecasting international tourism demand in Malaysia using Box Jenkins Sarima application", *South Asian Journal of Tourism and Heritage*, Vol. 3 No. 2, 50-60
- [7] Bigovic, M. (2012). Demand forecasting within Montenegrin tourism using Box-Jenkins methodology for seasonal ARIMA models. *Tourism and Hospitality Management*, 18(1), 1-18.
- [8] Hassani, H.; Webster, A.; Silva, E.S.; Heravi, S. Forecasting US tourist arrivals using optimal singular spectrum analysis. *Tour. Manag.* 2015, 46, 322-335.
- [9] Xu, X.; Law, R.; Chen, W.; Tang, L. Forecasting tourism demand by extracting fuzzy Takagi-Sugeno rules from trained SVMs. *CAAI Trans. Intell. Technol.* 2016, 1, 30-42.
- [10] Jere, S.; Banda, A.; Kasense, B.; Siluyele, I.; Moyo, E. Forecasting annual international tourist arrivals in Zambia using Holt-Winters exponential smoothing. *Open J. Stat.* 2019, 9, 258-267.

- [11] Load Factor for U.S. Air Carrier Domestic and International, Scheduled Passenger Flights (LOADFACTOR11) | FRED | St. Louis Fed, <https://fred.stlouisfed.org/series/LOADFACTOR11>
- [12] Revenue Passenger Miles for U.S. Air Carrier Domestic and International, Scheduled Passenger Flights (RPM) | FRED | St. Louis Fed, <https://fred.stlouisfed.org/series/RPM>
- [13] Producer Price Index by Industry: Travel Agencies: Cruise and Tour Bookings (PCU561510561510222) | FRED | St. Louis Fed, <https://fred.stlouisfed.org/series/PCU561510561510222>
- [14] Exports of Services: Travel (for all purposes including education) (IEAXSTV) | FRED | St. Louis Fed, <https://fred.stlouisfed.org/series/IEAXSTV>
- [15] Producer Price Index by Industry: Travel Agencies: Hotel and Car Rental Bookings (PCU561510561510211) | FRED | St. Louis Fed, <https://fred.stlouisfed.org/series/PCU561510561510211>