

Prediction of Car License Auction Price in Guangzhou: Application of the Time-series Decomposition Model, the Holt-Winters Exponential Smoothing Model and the SARIMA Model

Yuliang Chen

The University of Sydney, Business School, Finance, Sydney, Australia

yche3212@uni.sydney.edu.au

Abstract. This paper aims to predict car license auction prices in Guangzhou and provide useful references on the car license auction prices for the potential participants. Nowadays, China is experiencing the development of urbanization and the economy. Conflict occurs between the increasing population and the limitation of urban spatial resources, which boosts the demand for the expansion of traffic infrastructures. As one of the solutions to relieve the congestion, the car license held by Guangzhou is known as a way to reduce the congestion and provides opportunities for the residents to attain car licenses at the level of prices that they are willing to offer. The Time-series Decomposition Model, the Holt-Winters Exponential Smoothing model (the Holt-Winters model), and the Seasonal Autoregressive Integrated Moving Average model (SARIMA) are employed as tools in data prediction based on the time series data of the car license prices in Guangzhou from January 2013 to December 2021, and predictions are made on the prices on a monthly basis in the year 2022. By describing the features of estimated prices and comparing the MAPE (Mean Absolute Percent Error) from the three models, SARIMA shall be taken as a better model than the Holt-Winters model and the time-series decomposition model with a smaller value of MAPE and a more accurate prediction in the trend of price changing.

Keywords: License auction, time series analysis, time-series decomposition model, Holt-Winters exponential smoothing model, SARIMA model.

1. Introduction

Currently, China is experiencing the development of both urbanization and the economy, revealed by the rapid increase in the urban population and the expanding scale of land use in cities. Meanwhile, the demand for urban traffic by residents has increased profoundly, as the number of car holdings has continuously mounted, from 6,250,000 in 2000 to 130,000,000 with an average increase of approximately 9,500,000 per year [1]. The decline of urban spatial resources will inevitably lead to traffic congestion, when the transportation infrastructure cannot meet the sharp increase in the number of cars without any restrictions or mitigation policies imposed against the increasing number of cars. At present, there are two popular solutions for urban congestion: the charge of congestion tax and direct restriction on the growth of car holdings [2]. Nevertheless, China has not yet established any relevant regulations on congestion tax from the perspective of legislation. Major cities in China, such as Shenzhen, Guangzhou, and Shanghai attempt to shrink the potential number of cars by imposing restrictions and conducting car license auctions and lotteries. The nature of auction centers on the bidding prices, based on which can build a rational interaction mechanism for effective communication between the sellers and the bidders, as well as to maintain the balance of a market economy [3]. Moreover, fairness should be taken into account in license auctions, and the price-time priority principle utilized by Guangzhou in the auction complies with the rule of fairness [4]. The license auction can effectively prevent the irrational rise in prices of car licenses, and gains on the license auctions will be applied to the construction of traffic infrastructures [5].

Currently, Guangzhou city offers two approaches to attain a car license, the car license lottery, and the license auction. The characteristic of fairness in the two approaches implemented by Guangzhou [5]. By using the combination of the license lottery approach and the license auction approach, it ensures fairness for participants. On the one hand, the license lottery approach gives opportunities to

the general public to attain a car license free of charge. On the other hand, the license auction enables those participants who are eager to attain the car license as quickly as possible and are of strong financial ability to pay at the level they are willing to. However, fairness is not established for benefiting the poor [6]. Guangzhou's license lottery and auction approach can prevent the concentration of social resources as well as comply with the concept of environmental protection and promoting the development of new energy cars [7, 8]. Simultaneously, strict restrictions on fuel cars continue to be set, which provides precious opportunities for the growth of clean energy car holdings and the development of new energy technology by car manufacturers. In the economy of China, issues related to automobiles have become the forefront problem in a debate on contemporary policies. Though the increasing penetration of cars has improved the living standards of Chinese residents, simultaneously it has led to the exacerbation of air pollution problems and severe traffic congestion and has increased the demand for clean energy [7, 9]. Major cities in China, such as Guangzhou, Tianjin, Shenzhen, and Hangzhou do not impose restrictions on the usage of clean energy cars and give access to clean energy cars to drive on the bus and fast traffic lanes [9]. Under the loose restriction policies on clean energy cars, the market share of clean energy cars is steadily gaining in China. In 2018, six major cities that are the earliest to impose restrictions on fuel cars have made a contribution of 77% of the sales of clean energy cars in China. However, the restriction imposed on car licenses has only a limited impact on residents' willingness of car purchases. In a survey on customers' attitudes towards new energy cars, two main reasons affect the purchasing of new energy cars, including the immature technology of new energy cars, and the short travel mileage [10].

2. METHODOLOGY

2.1. Explanation of auction rules and procedures

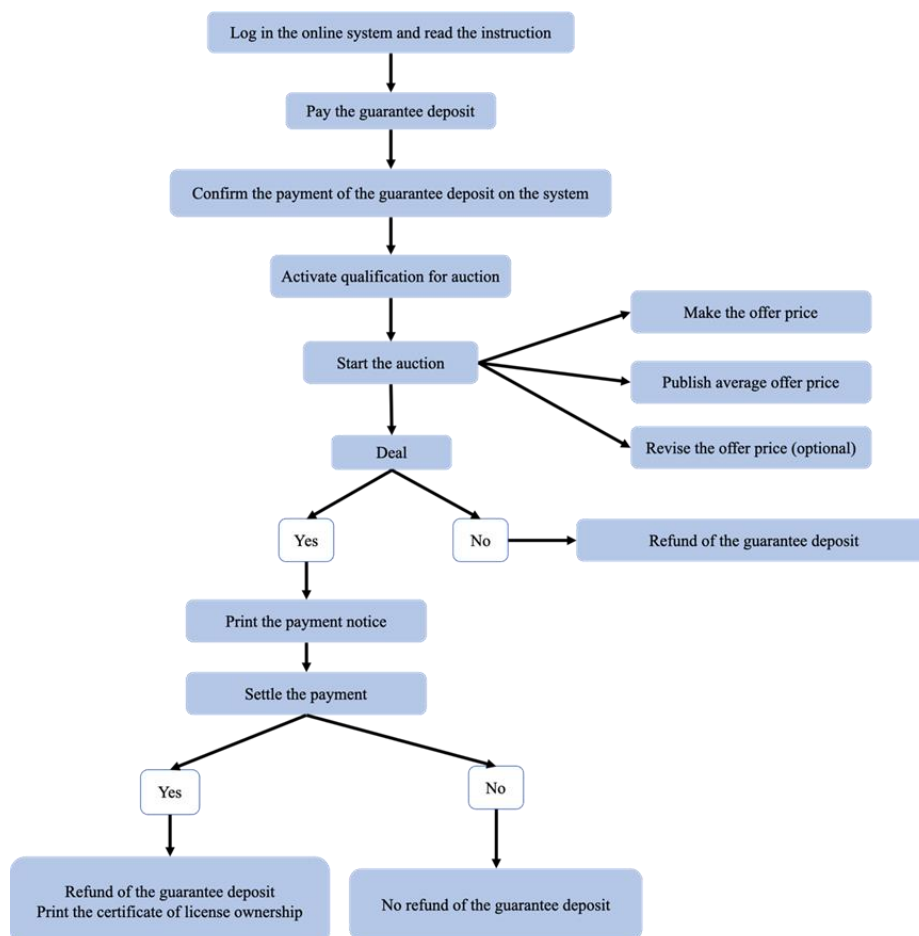


Figure 1. Online auction procedures

Figure 1 depicts the procedures of a car license auction. Before the auction proceeds, the participants should log in to the online auction system and carefully read the instruction for the auction. Besides, the participants must pay the guarantee deposit, and only when the guarantee deposit is confirmed by the system can the participants activate their qualification for the auction. During the auction, the participants can make their price offers two times for each individual. After the first price offers have been made by all the participants within the fixed time, an average offer price will be published as a reference. The participants can choose whether to revise their offer price according to the published average offer price, and they will not be informed of the auction results by the system until the auction is over. For each offer made, the offer price shall not be lower than the previous offer and shall be increased by the multiplier of 100 yuan and the minimum offer price is 10,000 yuan. If the participants do not win the auction, a refund will be made based on guarantee deposit. For the participants who win the auction, they are required to settle the payment for the car license according to the offer prices, after which they obtain the right to make a refund of the guarantee deposit and can print the certificate of license ownership. However, those who win the auction but do not settle the payment will not be granted a refund of the guarantee deposit.

2.2. Exploratory data analysis

Data of car license auctions in Guangzhou spans from August 2012 to July 2022 on a monthly basis. Each data contains 11 values, including the certain month of the auction occurred, the number of quotas given, the number of participants, the first offer average price (in CNY), the second offer average price (revised price, in CNY), the lowest offer price by the winning participants (in CNY), the average offer price published after the auctions (in CNY), the number of lowest price offers made by the participants, the number of lowest price offers accepted by the system, the number of quotas sold out but not yet paid, and the number of quotas sold out and paid by the winning participants. There are a total number of 240 data of the car license auction and contains no missing data after checking.

Figure 2 shows the number of quotas given and the number of participants in the auction for each month from August 2012 to July 2022 on a monthly basis, and the overall trend for the monthly quotas had been fluctuated in two time periods, August 2012 to May 2013 and May 2019 to February 2021. From August 2012 to May 2013, the quotas given for the auctions had a maximum value of 9121 (November 2012) and a minimum value of 3840 (August 2012). Another apparent fluctuation occurred from May 2019 to February 2021, in which the maximum value is 10649 (August 2020) and the minimum value is 3616 (February 2021). In the remaining period of time, the number of quotas remained stable. Since the auction held been held in August 2012, the number of quotas had a maximum value of 10649 (August 2012) and minimum values of 3520 (July of 2013 - 2017). As for the number of participants, it fluctuated more fiercely than that of quotas. The overall trend for the number of participants is generally upward from August 2012 to April 2020, reaching its peak at 30768 in June 2020, and then experiencing a sharp decrease to 14601 in October 2020. Followed by a rapid increase from November 2020 to December 2020, the number of participants increased to 26627, however, it continuously dropped to 4719 in February 2022.

To make a clear understanding of the level of engagement in the auction, the term “competitiveness index” can be introduced to evaluate the level of engagement by participants, illustrated by the ratio of the number of quotas and participants. Figure 2 presents the change in the competitiveness index in the auctions from August 2012 to July 2022. It can be apparently viewed that the competitiveness index in the first 8 months of the auctions exceeded 100%, meaning the number of quotas is much higher than that of participants. Except for September to October 2014, the demand for the license quota had constantly surpassed the number of quotas given.

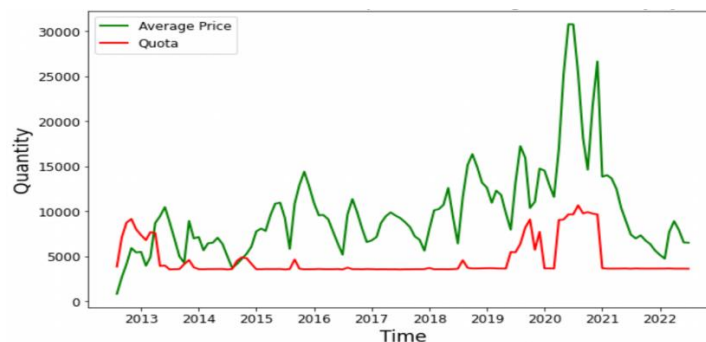


Figure 2. Quotas and number of participants from August 2012 to July 2022

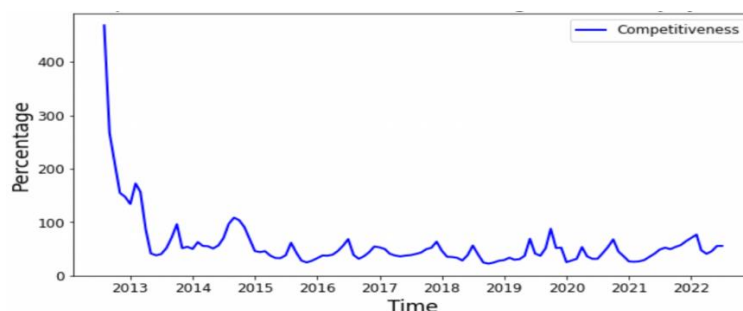


Figure 3. Competitiveness of the auctions from August 2012 to July 2022

Figure 4 displays the correlations between variables and the depth of color implies the level of correlation between variables. The deeper color means a high correlation. The heatmap shows a highly correlated relationship between the average offer price with the first and the second offer prices made by the participants, respectively at 0.963 and 0.969, meaning that the final average price is highly dependable on the first and the second offer prices. By contrast, the average price of the car license has a low correlation with the number of participants and even a negatively low correlation with the number of quotas. Besides, the number of quotas and the number of participants does not present a high correlation, which indicates that the number of quotas issued in each auction has a limited effect on participants' willingness to engage in the auctions.

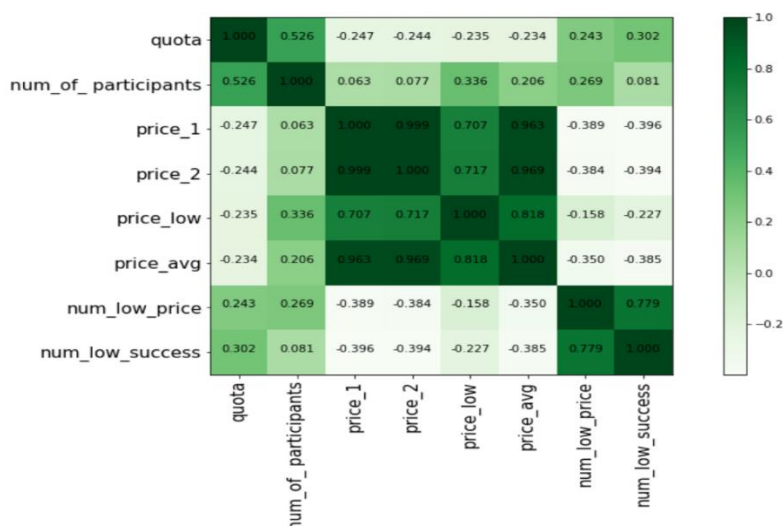


Figure 4. Correlation heatmap

2.3. Time-series decomposition model

The time-series decomposition model has two main forms: additive decomposition and multiplicative decomposition. Character m is used to describe a seasonal period, for example in this paper, $m=12$ represents the monthly data in a year.

In both methods, the first step is the computation of the trend-cycle component $\hat{T}_t$, which refers to the moving direction of the time series ($2 \times m - \text{MA}$ when m is an even number; $m\text{-MA}$ when m is an odd number). Then, the de-trend series can be realized by subtracting the trend estimates from the series in the additive method ($y_t - \hat{T}_t$) or dividing the series by the trend values ($y_t/\hat{T}_t$) in the multiplicative method. Besides, the seasonal component for each season can be estimated by averaging the de-trend series, which measures the effect for each month within a year. The result will be $\hat{S}_t$ as the seasonal component. Finally, the component for reminder can be computed by subtracting the trend and seasonal components from the series in the additive method ($\hat{R}_t = y_t - \hat{T}_t - \hat{S}_t$) or dividing the series by the product of trend and seasonal components ($\hat{R}_t = y_t/(\hat{T}_t \times \hat{S}_t)$).

2.4. The Holt-Winters exponential smoothing model

The Holt-Winters Exponential Smoothing Model has two main forms: the additive Holt-Winters model and the multiplicative Holt-Winters model. With an aim to capture the seasonality of the time series, the model includes both the forecast equation and three smoothing equations, in which ℓ_t refers to the level component, b_t refers to the trend component, and s_t refers to the seasonal component with smoothing parameters α , β^* and γ for each component. The character m is used to reflect the frequency of seasonality.

Component form of the additive Holt-Winters model:

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)} \quad (1)$$

$$\ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad (2)$$

$$b_t = \beta^* (\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \quad (3)$$

$$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m} \quad (4)$$

Component form of the multiplicative Holt-Winters model:

$$\hat{y}_{t+h|t} = (\ell_t + hb_t)s_{t+h-m(k+1)} \quad (5)$$

$$\ell_t = \alpha \left(\frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad (6)$$

$$b_t = \beta^* (\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \quad (7)$$

$$s_t = \gamma \left(\frac{y_t}{\ell_{t-1} + b_{t-1}} \right) + (1 - \gamma)s_{t-m} \quad (8)$$

2.5. SARIMA model

The SARIMA model is considered as an extension of the ARIMA model with additional seasonal terms, which can support the model of time series with a seasonal component. The character m indicates the number of observations per year. SARIMA model includes a non-seasonal part called a trend part, as well as a seasonal part. Uppercased notation represents the seasonal parts of the model while lowercased notation is for the non-seasonal parts of the model. All the components are listed with descriptions in Table 1. The general form for the model is $SARIMA(p, d, q)(P, D, Q)_m$.

Table 1. Components and descriptions

Component	Description
p	Non-seasonal autoregressive order
d	Non-seasonal difference order
q	Non-seasonal moving average order
P	Seasonal autoregressive order
D	Seasonal difference order
Q	Seasonal moving average order
m	Number of observations per year

2.6. Data descriptions

Data on car license prices in Guangzhou spans from January 2013 to December 2021 on a monthly basis and relevant statistics are collected and extracted from the official document issued by Guangzhou government. Each data contains the specific month of the auction that occurred, and the average offer price published after the auctions (in CNY). There are a total number of 108 car license auctions that occurred during this period and contain no missing data. Before applying the model to the dataset, the validation set is separated from January 2021 to December 2021 (12 data in total), and the train set is the remaining data from Jan 2013 to December 2020. The reason for removing the data from Aug 2013 to December 2013 is that at the initial stage of the auction, participants might not be familiar with the rules and regulations of the auction, which makes the auction price in the early months high and fluctuated. Under such consideration, this part of the data shall be removed because it might lead to bias and cannot represent the true situations.

3. Results and discussion

3.1. Time-series decomposition model results

After finding the seasonality, another step is to make a trend estimate using the centered moving average (CMA) method. In the previous steps, we have proved with the line chart that there might be a seasonality component of 12 months within a year. Therefore, CMA-12 is used to smooth the line as well as to remove the seasonality. Compared to the original plot (see Figure 6), the shape of the line in Figure 7 becomes less volatile than before and does not fluctuate dramatically.

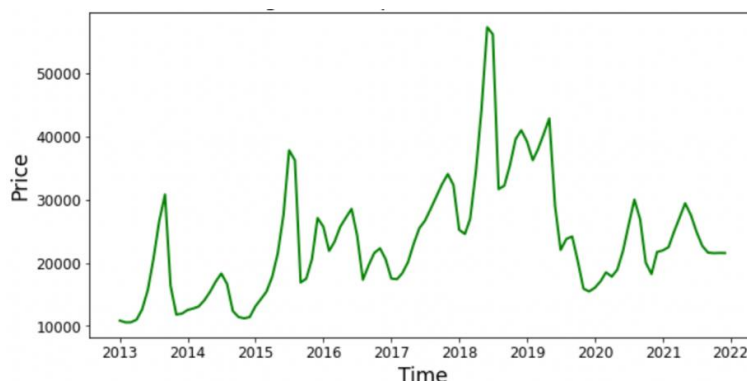


Figure 5. Original plot of the time series

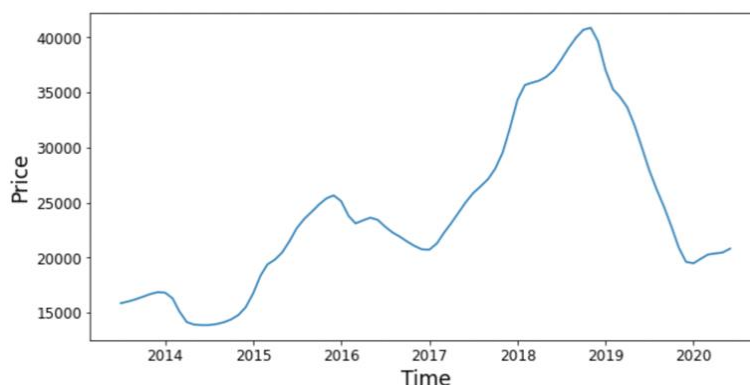


Figure 6. License price from Jan 2013 to Dec 2021 after CMA=12 smoothing

Two plots in Figure 8 representing the multiplicative method and the additive method respectively, show high similarity in the shape of the lines so more metrics should be taken into consideration in the choice of model. As Table 1 shown, the range (which derives from the maximum value minus the minimum value) of prices is quite different in the two methods, and apparently, the multiplicative

method has a much smaller range than the additive method does. With a considerably smaller standard deviation, the multiplicative method is considered to be the proper one.

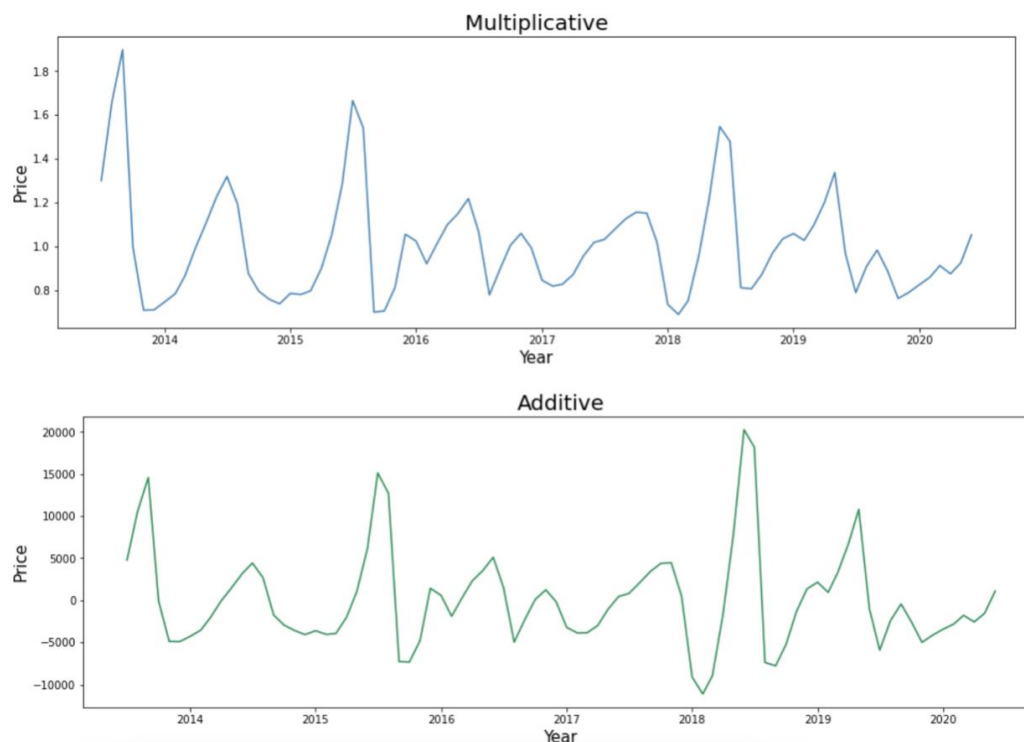


Figure 7. Remove the trend pattern by multiplicative and additive methods

Table 2. Descriptions of data by multiplicative and additive methods

	Max	Min	Range	Standard deviation
Multiplicative	1.90	0.69	1.21	0.24
Additive	20253.54	-11129.71	31383.25	5857.44

Table 3. Estimated prices by decomposition model

Time	Decomposition Model
	Estimated price
2021/1/1	18828.58
2021/2/1	18472.83
2021/3/1	19625.97
2021/4/1	21643.59
2021/5/1	24281.96
2021/6/1	26111.83
2021/7/1	27161.95
2021/8/1	25015.39
2021/9/1	22867.19
2021/10/1	20142.70
2021/11/1	19521.12
2021/12/1	19883.69

As Figure 9 shown, the line that represents the estimated price in 2021 has a similar trend as the validation set, but there seems to be a lag in each prediction and the general values for each estimated price remain smaller than that for the validation set.

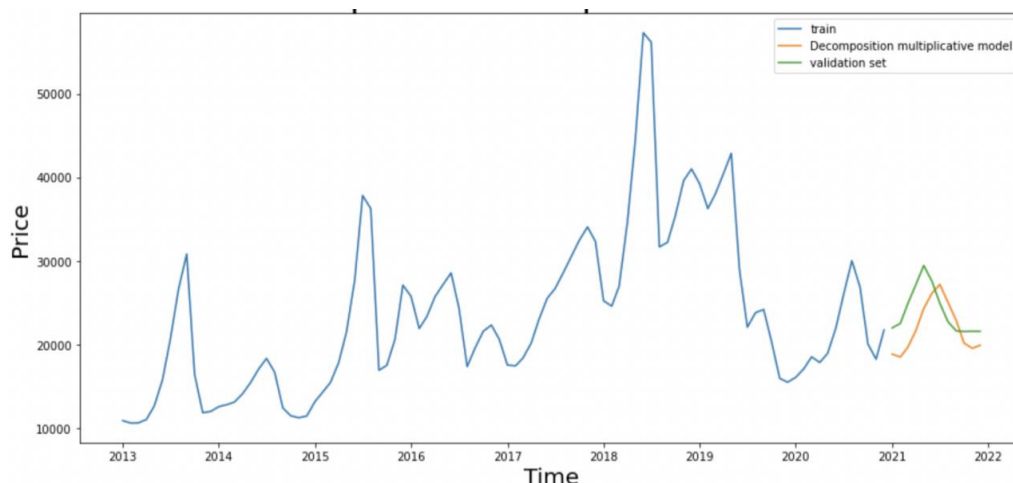


Figure 8. Decomposition multiplicative forecast

3.2. Holt-Winters exponential smoothing method

Table 4 shows the estimated price under both the Holt-Winters additive model and multiplicative model, which are plotted and make comparisons with the validation set in Figure 10. The line in red and green respectively represent the estimated prices for the additive method and the multiplicative highly deviates from the aspect of the trend when compared to the validation set. In addition, the two methods display an opposite trend. The result of additive methods shows a general declining trend while that of the multiplicative method shows a general growing trend.

Table 4. Estimated prices by the Holt-Winters model

Time	HW-Multiplicative Estimated price	HW-Additive Estimated price
2021/1/1	23759.33	16127.23
2021/2/1	22519.33	13399.25
2021/3/1	20989.64	14626.36
2021/4/1	19615.96	16203.38
2021/5/1	21570.71	18702.75
2021/6/1	23053.59	18066.09
2021/7/1	20931.75	18580.12
2021/8/1	19943.31	17862.18
2021/9/1	16267.61	18611.10
2021/10/1	12962.46	21007.43
2021/11/1	14246.34	22789.06
2021/12/1	19030.58	21786.51

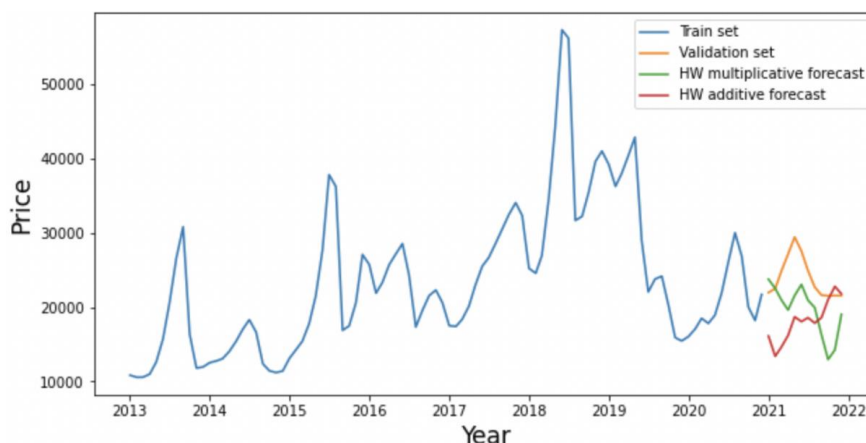


Figure 9. Prediction plot using multiplicative and additive forecast

3.3. SARIMA (Seasonal Autoregressive Integrated Moving Average)

Autocorrelation function plots (ACF plots) and Partial autocorrelation function (PACF plots) are used to determine the appropriate parameters in a SARIMA model, with the assistance of grid search. As a preliminary step, the ACF and PACF plots were utilized to narrow down some possible parameters, followed by the ideal parameters to be identified through a grid search. By checking the original plot, with some seasonality, the data is obviously non-stationary. And the spikes die down slowly to zero in the ACF of the original plot.

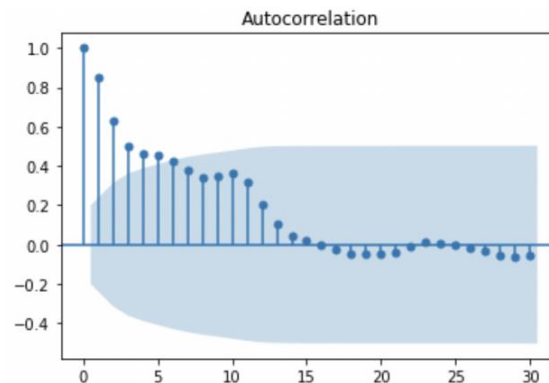


Figure 10. ACF of the original plot

Therefore, a seasonal difference ($D=1$) should be taken as shown in Figure 12, which indicates the data appears to be non-stationary, due to the slow decline of the lags.

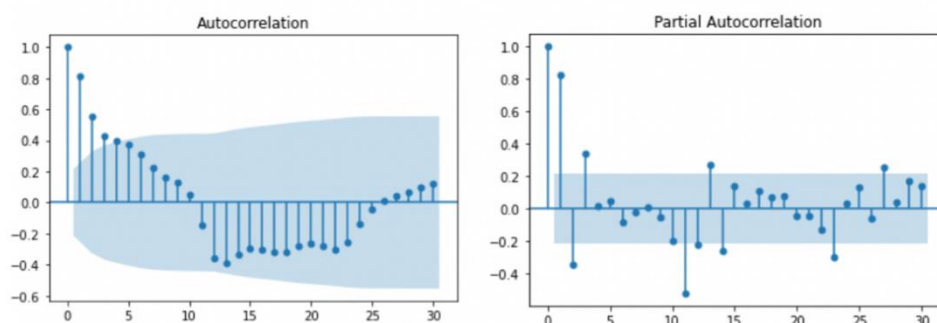


Figure 11. Plot of differenced price, ACF plot and PACF plot ($D=1$)

An additional normal difference process ($d=1$) is necessary to further make the plot stationary, as shown in Figure 13. Though the data is stationary, the significant spike appears at lag 2 in the ACF plot implying a non-seasonal MA (2) component, and the significant spike at lag 12 in the ACF implies a seasonal MA (1) component. Therefore, a possible model might be $SARIMA(0,1,2)(0,1,1)_{12}$.

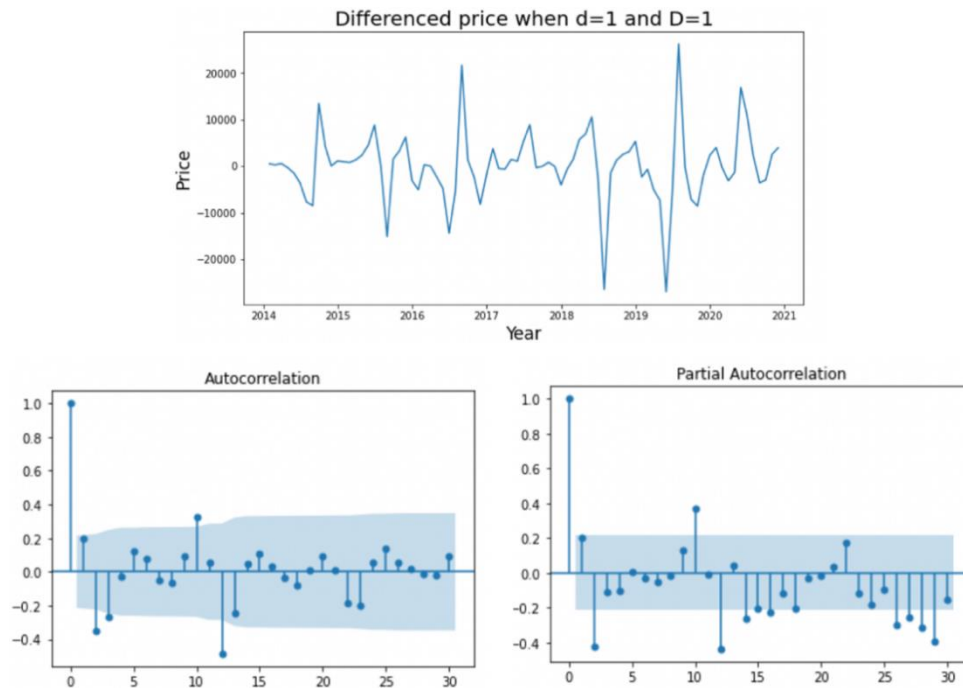


Figure 12. Plot of differenced price, ACF plot and PACF plot ($d=1$ and $D=1$)

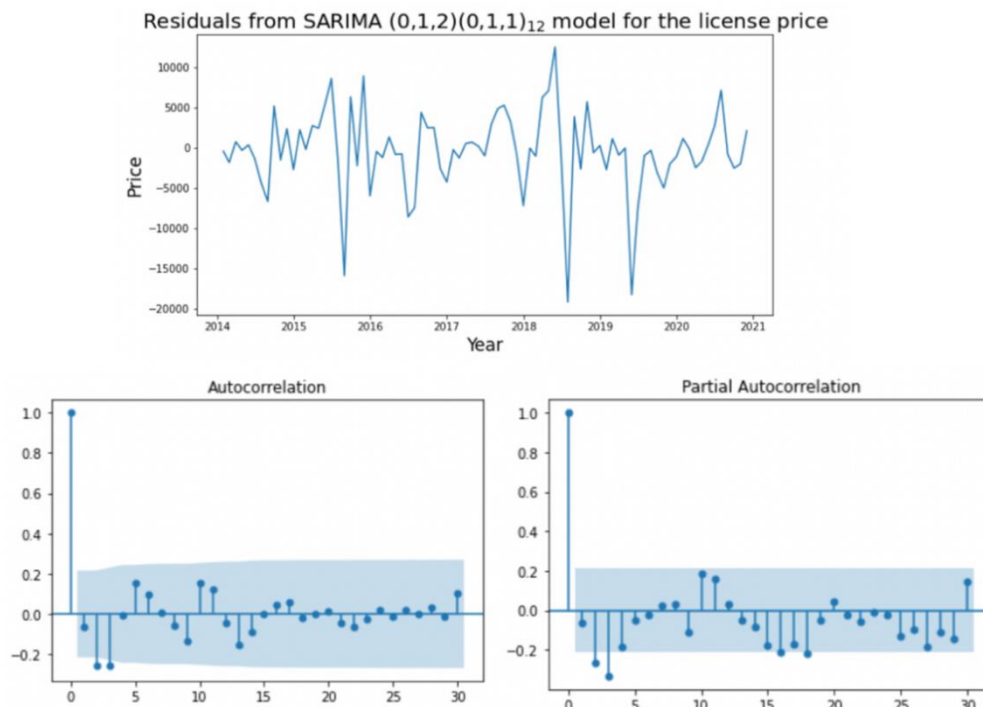


Figure 13. Plot of residuals, ACF plot and PACF plot in $SARIMA (0, 1, 2)(0, 1, 1)_{12}$

The residuals for the fitted model are shown in Figure 14. Both the ACF and PACF plots show a significant spike at lag 3, implying that some additional non-seasonal terms shall be covered in this model. The AIC of the $SARIMA (0, 1, 2)(0, 1, 1)_{12}$ model is 1662.023 while that of the $SARIMA (0, 1, 3)(0, 1, 1)_{12}$ model is 1658.277. Therefore, $SARIMA (0, 1, 3)(0, 1, 1)_{12}$ shall be chosen, with its residuals plotted in Figure 15. All the spikes are now contained within the significant level.

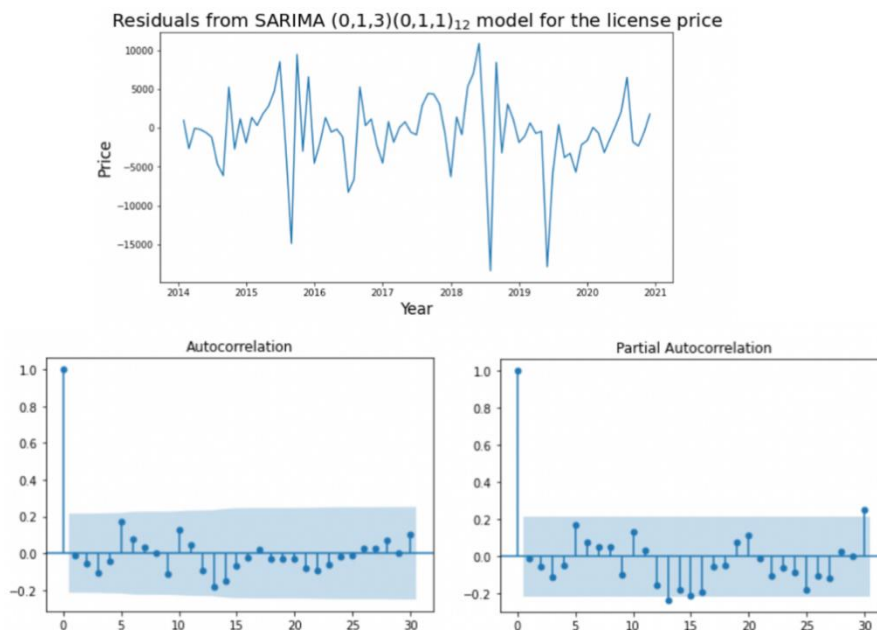


Figure 14. Plot of residuals, ACF plot and PACF plot in $SARIMA(0, 1, 3)(0, 1, 1)_{12}$

The estimated prices under the SARIMA model have a similar line shape when compared to the validation set (see Figure 16). However, the gap between the estimated price and the price in the validation set still exists.

Table 5. Estimated prices by the SARIMA model

Time	SARIMA Estimated price
2021/1/1	22730.88
2021/2/1	21649.70
2021/3/1	22269.55
2021/4/1	24189.27
2021/5/1	26897.31
2021/6/1	28791.48
2021/7/1	29830.78
2021/8/1	27464.55
2021/9/1	25527.13
2021/10/1	23492.38
2021/11/1	23348.45
2021/12/1	24249.34

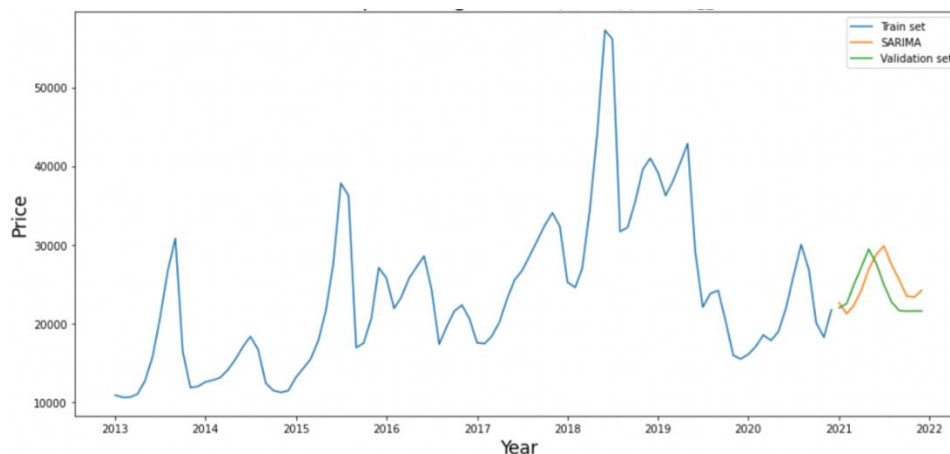


Figure 15. Prediction plot using $SARIMA(0, 1, 3)(0, 1, 1)_{12}$ model

3.4. Valuation by MAPE

Besides evaluating the model in a visualization way, mean absolute percent error (MAPE) is introduced to measure the accuracy of the forecast by a specific model. The rule is that the higher MAPE value indicates a less accurate forecast. Among all the models, the SARIMA model performs best with the smallest value of MAPE, indicating the estimated prices under the SARIMA are much closer to the validation set value than the prices predicted by other models.

Table 6. M MAPE of different models

Decomposition	HW-Multi	HW-Add	SARIMA
12.05%	19.39%	24.09%	10.87%

4. Conclusion

From the results derived in the last section, the Time-series Decomposition model and the SARIMA model provide a better estimation of the trend of the price changes. However, the Holt-Winters Exponential Smoothing model performs worst in both its multiplicative and additive methods as shown in estimating the trend of prices with the highest value of MAPE. Besides, the Time-series Decomposition model and the SARIMA model have smaller MAPE values, and the SARIMA performs best under MAPE valuation. Since a different set of data could have different adaptability to different models, this paper only focuses on analyzing the feasibility of different models in predicting future car license prices in Guangzhou, based on which can conclude that the SARIMA can predict the future prices more accurate when compared to the other two models. By using these models, although predictions on future prices can be made based on the historical time-series data, these models do not consider the effects of other variables, such as the number of participants and the number of quotas given in each auction. In the future, efforts will be taken to make a more comprehensive analysis and prediction of the license auction prices in Guangzhou with more variables to be included.

References

- [1] Xu, M.M., "Analysis and measures of urban traffic fairness," Urban and rural governance and planning reform, 18-27 (2014).
- [2] Wang, W.J. "Analysis of the incremental regulation policy of cars," (2019).
- [3] Wang, S. and Zhao, J., "The distributional effects of lotteries and auctions—License plate regulations in Guangzhou," Elsevier, (2017).
- [4] Chen, L. and Zhang, B., "Supply policy of the license plate and its 'making up the short board': Viewing from Eight Cities," Reform, 112-124 (2017).
- [5] Bu, Y.T., "Analysis on the rationality of Guangzhou automobile license restriction policy," ECONOMIC RESEARCH GUIDE, 51-53 (2013).
- [6] Wang, S. and Zhao, J., "Transportation research part A: policy and practice," Elsevier, (2017).
- [7] Xing, L.Z., et al., "First study on vehicle registrations via hybrid distribution of lottery and auction in Beijing," JOURNAL OF BEIJING UNIVERSITY OF TECHNOLOGY (SOCIAL SCIENCES EDITION), 17 (2017).
- [8] Li, X.M., et al., "Research on the impact of support policies on the promotion of new energy vehicles in China," Management Review, 34(03):55-65 (2022).
- [9] Zhou, M.X., et al., "New energy vehicle development and battery technology analysis," NEW ENERGY AUTOMOBILE, (2021).
- [10] Chow, Vinci., "Predicting auction price of vehicle license plate with deep recurrent neural network," Elsevier, (2019).