

# Time series analysis of Shanghai license plate price

Wangyi Mei 20307110042

June 24, 2023

## Abstract

As an international city, Shanghai has become the center of attention of East Asia, along with the expedited development, traffic became one of the most cumbersome issues. In this report, we will do some exploratory data analysis for overall car license by year and by month, perform the ARIMA time series model and ARIMA GARCH time series model. The performance of the two models were compared.

## 1 Introduction

Shanghai uses an auction system to sell a limited number of license plates every month and the average price of it was ¥92412 in April 2023. So it is often jokingly called 'the most expensive piece of metal in the world'. Since the auction system requires you to offer one of the highest price in all bidders with one chance but you must not want to waste too much money, a precise prediction based on time series model may be useful.

## 2 Dataset

The license plate price data comes from the website of Shanghai International Commodity Auction Co., Ltd. It included five dimensions including Date, Total number of license issued, lowest price, average price and total number of applicants. The date ranged from January, 2002 to May 2023 with a total number of 256 time intervals. By Augmented Dickey-Fuller test, it was stationary after first difference. There was no missing data.

## 3 Exploratory Data Analysis

In this section, we will carry out a comprehensive EDA on the dataset. Firstly, we will check the overall characteristics of the lowest price data and the average price data.

### 3.1 Overall Analysis

We can find out that the lowest price and the average price are very close to each other, especially after 2014 when the largest difference was within 300 yuan.

Then, we will check the relationship between the total number of applicants and the total number of license issued. Thereafter, the winning rate.

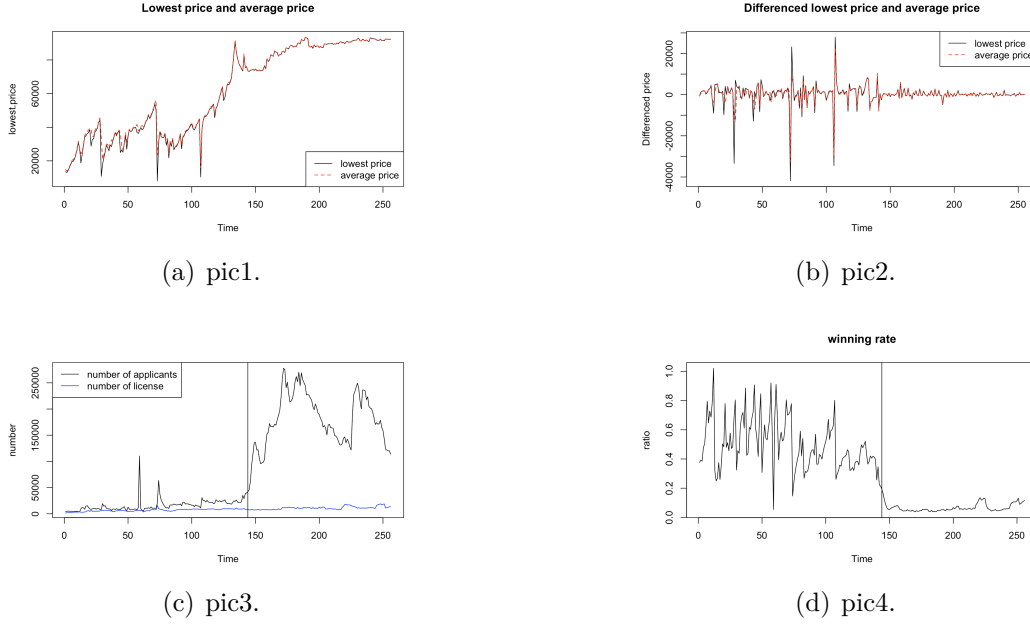


Figure 1: Overall Analysis

At around 2014, the number of applicants dramatically increased. However, the number of license was significantly lower than the number of applicants. Thus, the winning rate dropped significantly. Under this supply and demand relationship, the average price upsurged was natural.

## 3.2 Yearly and Monthly Analysis

### Yearly Analysis

Then, we will explore the yearly pattern and the monthly pattern of the data. Firstly, we did some data preprocessing during which we separated the year and month data.

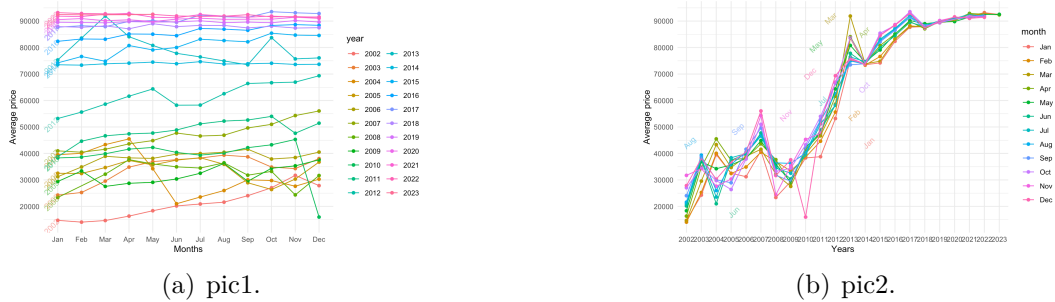


Figure 2: Yearly Price

The general trends within the year were relatively stable. Besides, we can see that the yearly pattern of average prices in different months are around the same as the general trend.

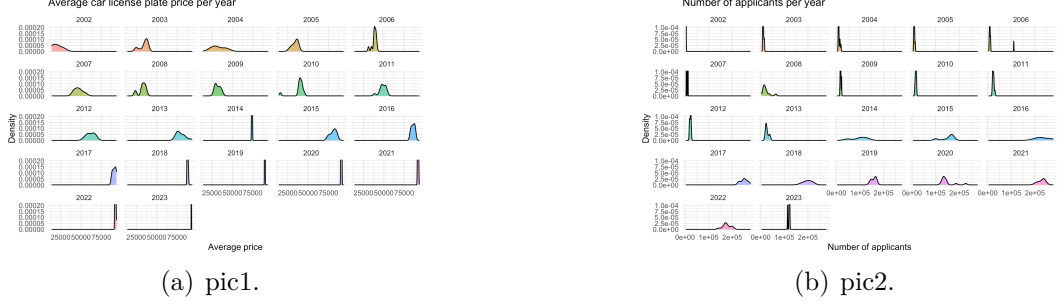


Figure 3: Distribution Yearly Price and Number of Applicants

From 2002 to 2023, the average license plate price per year had a similar increasing trend to the number of applicants per year.

### Monthly Analysis

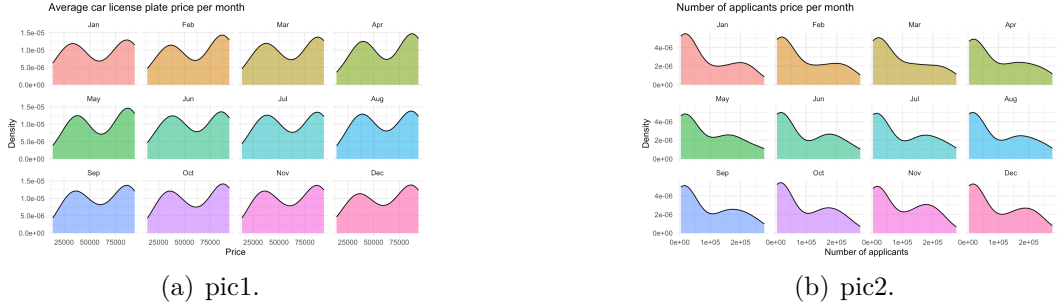


Figure 4: Monthly Analysis

Not surprisingly, the monthly price has nearly the same distributions. This is corresponding to yearly change of monthly price.

## 4 Models

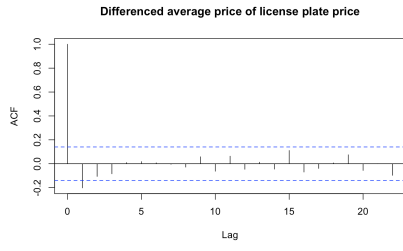
In this section, we will perform ARIMA and ARIMA GARCH models. Actually we also performed a linear model, you can see it in the code of the project.

### 4.1 ARIMA

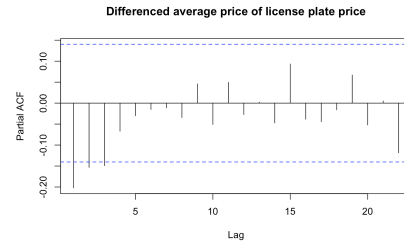
Firstly, we will show the ACF and PACF of differenced average price in Figure 5.

We can see that it has first order significant nonzero point in the ACF and second or third order significant nonzero point in the PACF. We excluded the last 60 and 7 data as two validation set and used other data as training set to fit an ARIMA(0, 1, 2) model and used it to predict the time series. Coefficients estimation is shown in the Table 1 and we can find out that it is not invertible. Its residual were also plotted in Figure 6.

We can see that the prediction is relatively good and the residuals approximately follows normal distribution and it passed the Ljung-Box test.



(a) pic1.

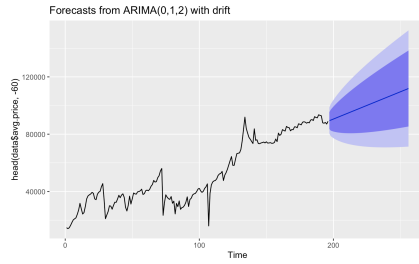


(b) pic2.

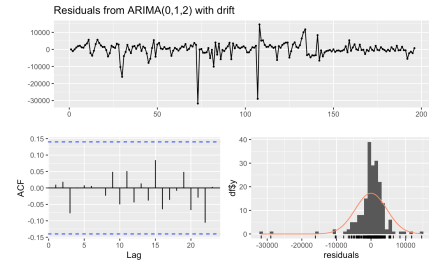
Figure 5: ACF and PACF

	ma1	ma2	drift
	-0.2800	-0.1525	380.8656
s.e.	0.0694	0.0679	186.6079

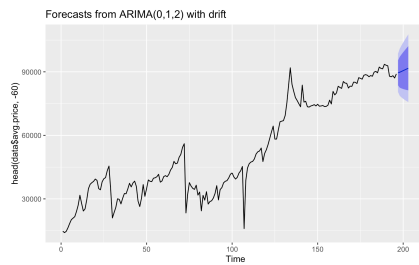
Table 1: Coefficients



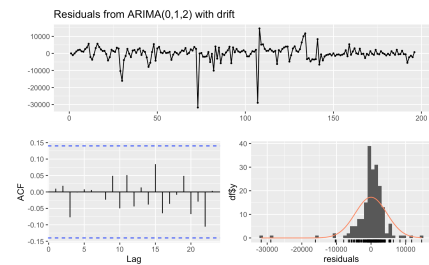
(a) pic1.



(b) pic2.



(c) pic3.



(d) pic4.

Figure 6: ARIMA

## 4.2 ARIMA GARCH

We found that the data has heteroscedasticity, we also found out that the ACF of the squared differenced average price data has a first order significant nonzero point. Then we added a GARCH model into the original ARIMA model.

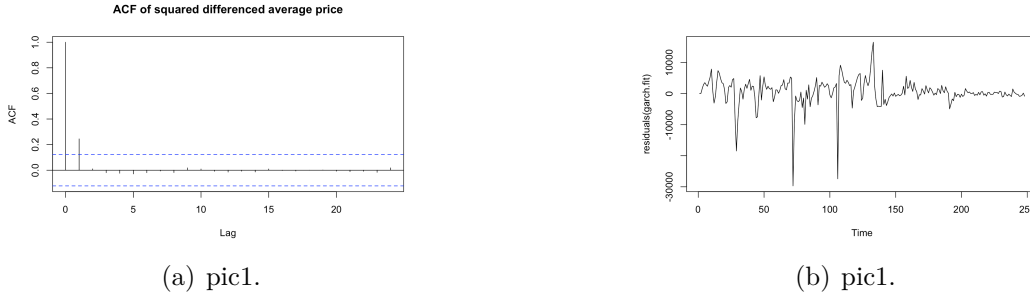


Figure 7: ARIMA GARCH

We fitted an ARIMA(0, 1, 2) + GARCH(1, 0) model to predict 60 and 7 future time intervals.

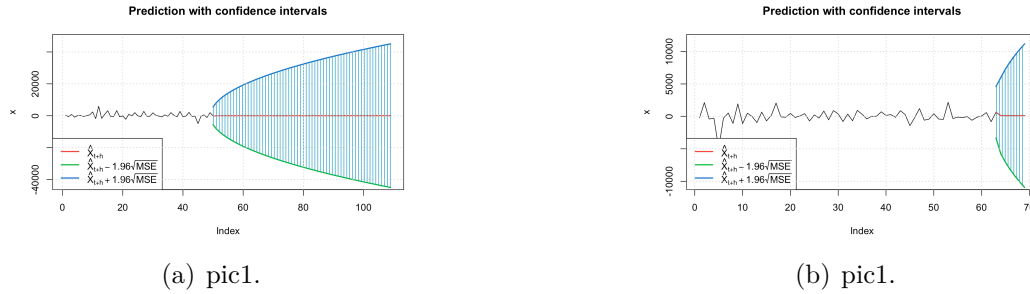


Figure 8: ARIMA GARCH

We plot the prediction of differenced series.

Then, we used the two models to predict the original series and here is the comparison.

Model	Step	ME	RMSE
ARIMA	7	1614.07	1688.47
ARIMA	60	-9810.86	11147.86
ARIMA GARCH	7	-398.21	482.58
ARIMA GARCH	60	513.26	995.70

Table 2: Comparison

## 5 Conclusion

In this project, we implemented two models to predict the Shanghai license plate price and we compared their prediction performance. By using ARIMA GARCH model, our mean prediction error is within 400 yuan in 7 months which is relatively instructive.