|  |  |  |
| --- | --- | --- |
| **Problem Chosen** C | **2023 ShuWei Cup Summary Sheet** | **Team Control Number** 202300000001 |

Summary

**Keybord:**

**Content**

[Content 2](#_Toc18138)

[**1. Introduction** 3](#_Toc9307)

[1.1 Background 3](#_Toc19661)

[1.2 Work 3](#_Toc9355)

[**2. Problem analysis** 3](#_Toc11846)

[2.1 Data analysis 3](#_Toc4513)

[2.2 Analysis of question one 3](#_Toc2967)

[2.3 Analysis of question two 3](#_Toc144)

[2.4 3](#_Toc19524)

[2.5 3](#_Toc12220)

[**3. Symbol and Assumptions** 3](#_Toc7754)

[3. 1 Symbol Description 3](#_Toc431)

[3.2 Fundamental assumptions 4](#_Toc7300)

[**4. Model 4**](#_Toc18026)

[**5.Test the Models 4**](#_Toc15673)

[**6. Sensitivity Analysis 4**](#_Toc4154)

[**7.Strengths and Weakness 4**](#_Toc29211)

[**8.Conclusion 4**](#_Toc30780)

[**References 4**](#_Toc27966)

[**Appendix 4**](#_Toc6131)

**1. Introduction**

1.1 Background

New energy vehicles refer to cars that use unconventional vehicle fuels as their power source. These vehicles combine advanced power control and driving technologies, featuring advanced technical principles, new technologies, and new structures, and are primarily categorized into four types. The advantages of new energy vehicles include being environmentally friendly and providing a smooth driving experience. In the field of new energy vehicles, Western countries began paying attention and exploring this area in the mid-19th century. Since 2011, China has implemented a series of preferential policies, leading to significant development in the new energy electric vehicle industry. China has achieved breakthroughs in enterprises, technology, and market areas. Under the pressure of energy and environmental concerns, new energy vehicles will undoubtedly become the main direction of future automotive development.

1.2问题的数据

1.3 Work

**Question 1**: Consider the main factors influencing the development of China's new energy electric vehicles, describe and analyze the impact of these factors on the development of China's new energy electric vehicles.

**Question 2**: Collect data on the development of China's new energy electric vehicle industry, describe and predict the development of China's new energy electric vehicles in the next 10 years.

**Question 3**: Collect relevant data to analyze the impact of global new energy electric vehicles on the traditional energy automotive industry.

**Question 4**: Research policies in some countries that resist the development of China's new energy electric vehicles, and analyze the impact of these policies on the development of China's new energy electric vehicles.

**Question 5**: Assuming a city population of 1 million, analyze the impact of urban new energy electric vehicle electrification on the ecological environment.

**Question 6**: Based on the conclusions of Question 5, write an open letter to the citizens, including the benefits of new energy electric vehicles and the contributions of the electric vehicle industry in various countries around the world.

# 2. Problem analysis

## 2.1 Analysis of question one

## 2.2 Analysis of question two

## 2.3 Analysis of question three

## 2.4 Analysis of question four

# Symbol and Assumptions

## 3.1 Symbol Description

对应“符号说明”，写别填，没弄好

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

## 3.2 Fundamental assumptions

对应“问题假设”

# 模型的建立与求解

## 4.1 Question one

4.1.1Model Establishment

(One) Data Preprocessing

For Question 1, this paper primarily explores the relationship between the sales volume of China's new energy electric vehicles and other variables over time, that is, investigating the impact of various factors on sales volume in a time series. Here, the sales volume is used to reflect the development of China's new energy electric vehicles.

To this end, this paper establishes the following time series model.

(Two) Establishment of Time Series Model

(1) ADF Test

Before establishing a time series model, it is necessary to perform a stationarity test on the data. In this paper, the Augmented Dickey-Fuller (ADF) test is used. The model is as follows:

Assume the null hypothesis that there is a unit root, indicating the time series is non-stationary; assume the alternative hypothesis that there is no unit root, indicating the time series is stationary.

When the basic trend of a series exhibits irregular increases or decreases repeatedly, it is categorized as an autoregressive process without a drift term. The corresponding regression formula is:

When the basic trend of a series shows a clear increase or decrease over time and the trend is not too steep, it is categorized as an autoregressive process with a drift term. The corresponding regression formula is:

When the basic trend of a series shows a rapid increase over time, it is categorized as a regression process with a trend term. The corresponding test regression formula is:

In this model, is the constant term, is the time trend term, and is the random disturbance term.

If the p-value is greater than 0.05, the null hypothesis cannot be rejected, indicating that the time series is non-stationary. If the p-value is less than 0.05, the null hypothesis is rejected, indicating that the time series is stationary.

(2) Differencing Adjustment

If the time series is non-stationary, it is necessary to use first-order differencing to adjust the series. The formula for this is:

If the time series remains non-stationary after first-order differencing, it is necessary to use second-order differencing for further adjustment. The formula for this is:

In the same way, the d-order integral is

In addition, the seasonal difference is

Among them, m is the period.

(Two) Testing of time series models

(1) Granger Causality Test []

The Granger causality test is used to determine if one time series variable is a cause of changes in another time series variable. The null hypothesis posits that the time series variable X does not cause changes in the time series variable Y. The alternative hypothesis is that the time series variable X does cause changes in Y. The formula is represented as:

If "time series variable X is the cause of changes in time series variable Y", according to the p-order autoregressive model of Y

Among them, is a constant term, is the maximum number of lag periods of Y, and is white noise. At the same time, considering the influence of X on Y, the infinite regression model of Y can be obtained

Among them, is the maximum number of lag periods for X.

Then use the residual sum of squares and of these two regression models to obtain the statistics

Among them, n is the maximum sample size.

According to the above model, the previously proposed hypothesis "time series variable X is the cause of the change in time series variable Y" is equivalent to testing that β\_i is significantly not equal to 0, that is

If , then the time series variable X is not the cause of the change in the time series variable Y.

(2) Cross-Check []

For the time series and with the same sample length N, let the cross-correlation function be

Then the correlation test statistic is

The correlation statistic obeys the chi-square distribution with m as the degree of freedom (the degree of freedom m is between 1-1000). If there is no cross-correlation between the two time series, the value of the correlation statistic is the same as the chi-square The distribution is consistent; on the contrary, as m increases, the correlation statistics exceed the standard chi-square distribution value, indicating that the cross-correlation between the two time series is significant.

4.1.2 Model Solution and Results

First, the data is processed to obtain the results of descriptive statistics, as shown in the table below.

Table Descriptive Statistics of Question 1 Data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Sales volume** | **Number of public charging piles\_Nationwide**  **(cumulative)** | **…** | **Battery capacity (km)** |
| **mean** | 103501.2072 | 490761.8874 | … | 494.5045045 |
| **min** | 750 | 518 | … | 320 |
| **25%** | 16772.5 | 74965 | … | 480 |
| **50%** | 50211 | 278736 | … | 480 |
| **75%** | 139121 | 751294.5 | … | 545 |
| **max** | 474475 | 1900386 | … | 545 |
| **std** | 127141.6221 | 523393.5695 | … | 64.42104031 |

Note: See Appendix 1 for complete data.

As can be seen from the table, the sales volume ranges from 750 to 474475, and the standard deviation is 127141.6221, which reflects the large difference between sales volumes; the number of charging piles ranges from 518 to 1900386, and the standard deviation is 523,393.5695, indicating that the number of charging piles varies in different There are large differences between time points; the range of battery capacity is 320~545km, and the standard deviation is 64.42104031, showing the relative stability of battery capacity between different data points.

In order to study the relationship between sales volume and other factors, this article successively conducted the ADF data stationarity test, Granger causality test, and cross-examination test. The specific solution results are as follows. The following takes the relationship between sales volume and infrastructure construction (number of charging piles) as an example. See Appendix 2 for other processes.

(One) ADF inspection and differential adjustment

1. The relationship between sales volume and infrastructure construction (number of charging piles)

Table Sales Volume ADF Test and Difference Adjustment Results

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Before differential  Statistics | *p value* | First difference  Statistics | *p value* | Second order difference  Statistics | *p value* |
| -1.1846 | 0.680 | -1.3105 | 0.624 | -7.990 | 2.484 |

As can be seen from the table, the p-value of sales volume before the difference is 0.680, which is greater than 0.05, indicating that there is a unit root and the data is non-stationary; the p-value after the first-order difference is 0.624, which is still greater than 0.05, indicating that the data is still non-stationary; second The p value after the first-order difference is 2.484, which is less than 0.05, indicating that the data is a stationary sequence.

Table Infrastructure (number of charging piles) ADF inspection and differential adjustment results

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Before differential  Statistics | *p value* | First difference  Statistics | *p value* | Second order difference  Statistics | *p value* |
| 3.9663 | 1.000 | -0.2420 | 0.933 | -5.1425 |  |

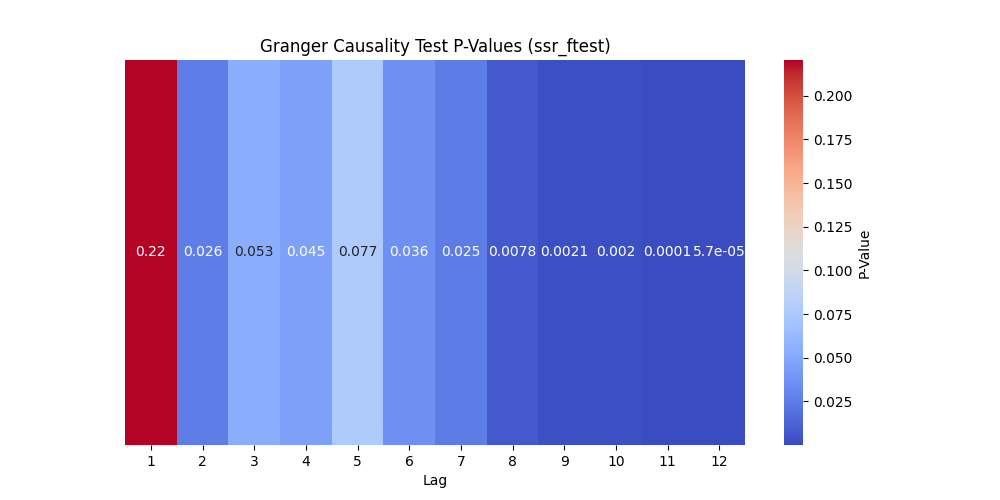
As can be seen from the table, the p-value of the sales volume before the difference is 1, which is greater than 0.05, indicating that there is a unit root and the data is non-stationary; the p-value after the first-order difference is 0.933, which is still greater than 0.05, indicating that the data is still non-stationary; second The p value after the first-order difference is , which is less than 0.05, indicating that the data is a stationary sequence.

(2) Differentially adjusted data

Table: Difference data between sales volume and infrastructure construction (number of charging piles)

|  |  |
| --- | --- |
| Date | Number of public charging piles\_Nationwide (cumulative) |
| 2014.3 | 0 |
| 2014.4 | 0.01 |
| 2014.5 | -0.01 |
| 2014.6 | 0 |
| 2014.7 | 0 |
| … | … |
| 2023.3 | 0 |

(Two) Results and analysis of Ganger causality test



Picture Geanger Causality Test P-Values

As can be seen from the figure, if the lag period is 4 and the subsequent p-value is less than 0.05, the null hypothesis is rejected, indicating that infrastructure construction (the number of charging piles) is the Granger cause of sales; and during these lag periods, infrastructure has statistically significant predictive ability for sales volume. Especially when the lag period is long, the p value is very small, indicating that the significance is stronger.

(Three) Cross-examination results and analysis

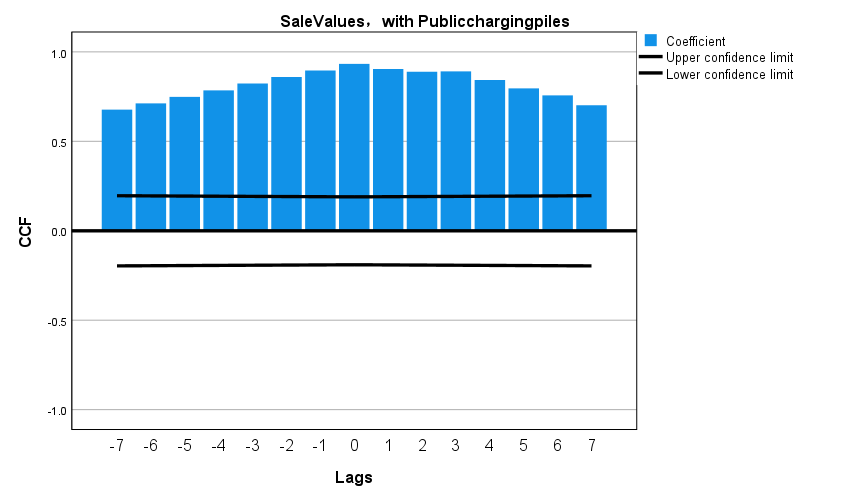


Figure Cross-validation results (before difference)

As can be seen from the figure, the correlation is high at each lag order. This may be due to the trend and seasonal components in the data, indicating that the time series of sales and infrastructure construction may share one or more unscheduled times. A common trend or cyclical structure that is eliminated by differencing.

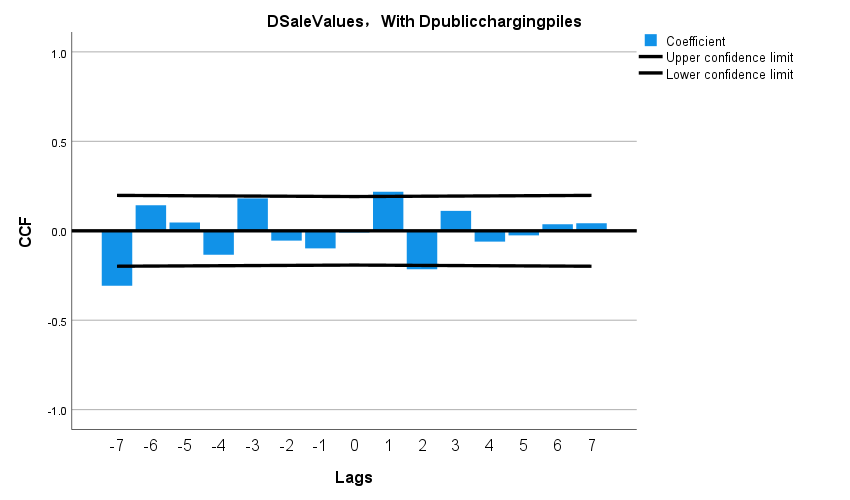


Figure Cross-validation results (after difference)

It can be seen from the figure that the correlation is close to zero at each lag order, indicating that there is no obvious linear relationship between the differentiated time series.

(Four) Comprehensive analysis and inference

The significant results of the Granger causality test indicate that changes in infrastructure precede changes in sales volume in time and may have an impact on them. However, cross-correlation analysis did not indicate a strong immediate relationship between the two. This indicates that the relationship between sales volume and infrastructure construction may be moderated by other variables, or that there is a non-linear relationship between sales volume and infrastructure construction.

That is to say, there is no obvious linear relationship between the development of new energy electric vehicles and infrastructure construction in China.

4.1.3 Summary

## 4.2 问题二

4.2.1 Model Establishment

(One)

(Two)

4.2.2 Model Solution and Results

4.2.3 Summary

## 4.3 问题三

4.3.1模型的建立

4.3.2模型的求解及结果

## 4.4 问题四

4.4.1模型的建立

4.4.2模型的求解及结果

# 5.模型的评价

## 5.1 模型的优点

## 5.2 模型的缺点

## 5.3 模型的推广

# 6.参考文献

# References

[1]

[2]

[3]

[4]

# 

# Appendix

## 1 支撑材料

## 2 附录一

## 3 附录二