Team Number:	apmcm2306529
Problem Chosen:	C

2023 APMCM summary sheet

In the context of global energy transition and climate change, new energy vehicles, as clean and efficient modes of transportation, have attracted widespread attention globally. This paper constructs various mathematical models such as **regression**, **prediction**, **difference**, and **evaluation** to model and analyze the **factors influencing** the development of new energy vehicles in China and the world, **trend predictions**, and **benefit assessments**.

For the first question, we conducted **grey relational analysis** on collected index data to evaluate the eight major influencing factors on the development of China's new energy vehicles. Then, we established a **Gradient Boosting Decision Tree** model to regress the market share of China's new energy vehicles, using the **Shapley Additive Explanation** model to interpret the specific impact mechanisms of each factor.

For the second question, we analyzed the characteristics of the indices related to the development of China's new energy vehicles over the past twelve years. Based on the cyclical nature of the data, we established a **seasonal ARIMA model** to forecast the development status for the next ten years.

For the third question, we constructed a **Vector Error Correction Model (VECM)** and analyzed the long-term impact of new energy vehicles on the traditional energy vehicle industry through **cointegration tests**, discussing the short-term effects of new energy vehicles on traditional vehicles through impulse analysis. The results indicate that in the long run, the development of new energy vehicles has a certain inhibitory and substitutive effect on the traditional energy vehicle industry; in the short term, the impact of new energy vehicles increases the uncertainty in the traditional automobile industry.

For the fourth question, we established a **DID (Difference in Differences) model** to analyze the actual export volume under the impact of foreign boycott policies compared with the predicted export volume without policy influence. The results show that foreign boycott policies will reduce the export volume of China's new energy vehicles in the short term but cannot stop the long-term trend of rapid development of China's new energy vehicles.

For the fifth question, we introduced the **Life Cycle Assessment (LCA) model**, quantifying the impact of traditional fuel vehicles and new energy electric vehicles on the ecological environment throughout their **life cycles** into specific environmental costs. Based on this, we calculated that the annual reduction in greenhouse gas emissions from the complete electrification of vehicles in a city of one million people is equivalent to the CO2 absorption of about 1.5 million trees in a year.

Finally, we conducted **sensitivity analysis** and **model evaluation**, and wrote a letter to citizens to promote this concept.

Keywords: New Energy Vehicles, Grey Relational Analysis, SHAP Model, Vector Error Correction Model, DID Double-Difference Model, LifeCycle Assessment

Contents

I. Introduction

1.1 Background

At the beginning of the 21st century, with the rapid development of science and technology, especially the proposal of the sustainable development strategy, the new energy automobile industry, represented by pure electric vehicles, hybrid vehicles and fuel cell vehicles, has realized rapid development under the strong support of governments around the world[1].Under the double urgent situation of energy crisis and environmental pollution, the transformation of automobile industry to new energy has become the common strategic choice of all countries in the world, and its popularization and application in the world has also become the general trend.

Against this background, it is particularly important to measure the development prospects of new energy vehicles, especially electric vehicles. How to accelerate the development of China's new energy automobile industry, in order to cope with the fierce competition of the world's new energy automobile industry, seize the new heights of industrial technology development, and win the competitive advantage, is a problem that needs to be solved urgentl.

1.2 Problem Restatement

1 Review the data and establish a mathematical model to analyze the main factors affecting the development of new energy electric vehicles in China, and specifically analyze the specific impact of these factors on the development of new energy electric vehicles in China.
2 To build a mathematical model with the collected data on the development of China's new energy electric vehicle industry to describe and predict the development trend and scale of China's new energy electric vehicles in the next 10 years.
3 Collect data to analyze what impact new energy electric vehicles are having on the global conventional energy automotive industry.
4 Collect information and analyze what impact the boycott policies introduced by some countries against China's new energy electric vehicles have had on the development of China's new energy electric vehicles. Use a mathematical model to analyze the extent and characteristics of this impact.
5 Modeling and analyzing the ecological impacts of new energy vehicle electrification in a city with a million population.
6 Based on the results of the analysis in the previous question, write an open letter to educate citizens about the benefits of new energy electric vehicles and the contributions of the electric vehicle industry in countries around the world.

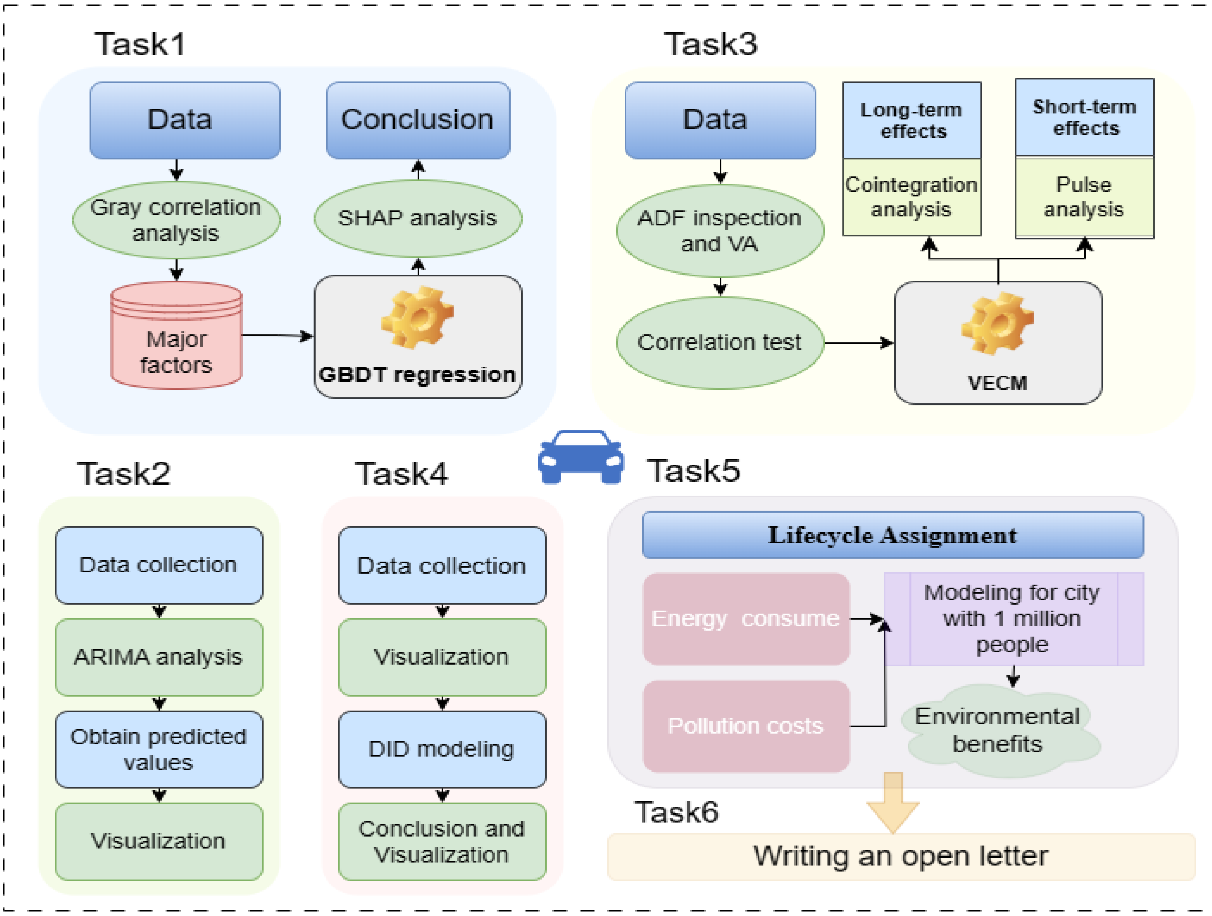


Figure 1 flow chart

1.3 Our Work

1 In the first question, first we use the collected data to assess the eight main influencing factors of China's new energy vehicle development through gray correlation analysis, and then we build a GBDT model to regress the market share of China's new energy vehicles, and use the Shapley Additive Explanation (SHAP) model to explain the specific influence mechanism of each factor.

2 In the second question, we collect data related to the development of new energy vehicles in China in the past twelve years and build a seasonal ARIMA model to forecast the development in the next ten years.

3 In the third question, we build a vector error correction model (VECM) to understand and analyze the long-term relationship between the data of new energy vehicle and fuel vehicle development as well as the short-term dynamics.

4 In the fourth question, we build a DID double difference model to analyze the impact of foreign boycott policies on China's new energy vehicle development by comparing the actual export volume under the imposed policy with the predicted export volume without the policy.

5 In Q5, we use the LifeCycleAssessment (LCA) approach to model and assess the environmental benefits of electrifying a city of one million people with new energy vehicles.

6 Finally, in conjunction with Q5, we wrote an open letter to educate citizens about the benefits of electric vehicles and the contributions of electric vehicle industry in countries around the world.

II. Assumption and Justification

1 **It is assumed that the development and sales of new energy vehicles are basically only related to the ten indicators selected in this paper, and other factors will not have a significant impact on them.** This paper has rigor and science in the process of indicator selection, has taken into account the various aspects of new energy vehicles affecting factors,so it can be considered that the indicators selected can basically explain the impact of the development of new energy vehicles.

2 **Not considering extremely exceptional factors for the impact on model.** Due to the difficulty in quantifying and the lack of regularity in extreme factors such as war, natural disasters, etc.This paper does not take them into account in the establishment and analysis of the model.

3 **It is assumed that the data collected from the Internet are real and reliable.** We choose data from the most typical regions in the world, so the breadth of the data is guaranteed; all the data come from the official website, so the reliability of the data is guaranteed.

4 **It is assumed that in the process of analyzing the factors affecting the development of new energy vehicles, it is reasonable to keep only the factors with strong correlation for analysis and remove the factors with small correlation.**The factors affecting the development of new energy are complex. If factors with low correlation are also considered, the results of the study will be less significant and may even produce misleading conclusions.

5 **It is assumed that there is a cointegration relationship between new energy trams and traditional energy vehicles, which is the premise of our VCEM analysis.** Since new energy vehicles and traditional energy vehicles are alternative commodities, it has become a consensus in the field of economics that there is a long-term stable relationship between the two variables[2]; and we have proved this assumption in our VCEM analysis.

III. Symbol Decriptions

Symbol	Description	Symbol	Description
ρ	Grey correlation resolution cod=efficient	**Cov**	Covariance
r	Correlation	D_i	Policy grouping dummy variables
μ	Data loss ratio	**C**	Total cost of polluting gases
∇	Differential operators	ε	Noise

IV. Model building and solution of question1

4.1 Indicator determination

Factors affecting the development of new energy vehicles can be broadly categorized into policy aspects, technology aspects, supporting services and other aspects. In this paper, we take these four aspects as primary indicators. Under each level 1 indicator, we consider 33 official indicators in the literature, and finally retain the 10 most representative level 2 indicators to construct our model, as follows:

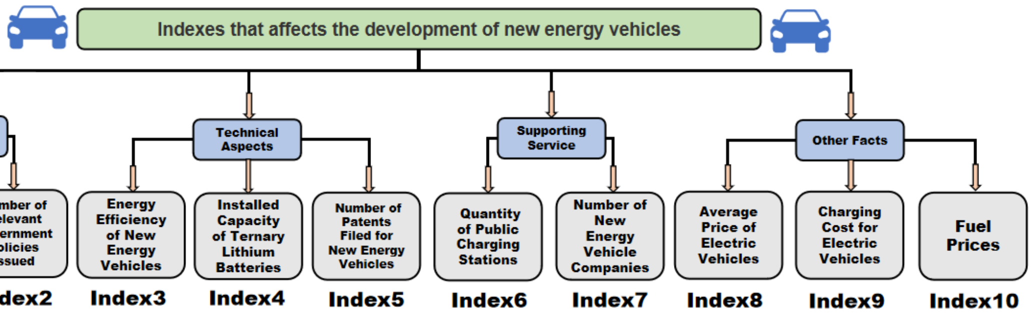


Figure 2 indexes that affects the development of new energy vehicles

4.2 Overview of data

4.2.1 *Data Source*

Since the question does not give data directly, we collected data through official documents, annual reports of enterprises, authoritative survey organizations and other channels. Through the analysis of the question, we mainly organized the data in Table 2. Due to the large amount of data, it is impossible to list them all, and only some of them are shown here.

Data Description	Data Source	Data Description	Data Source
Sales-related metrics (Sales, Retention Rate, Market Share)	https://www.qianzhan.com/ https://www.mmsonline.com.cn/ http://www.stats.gov.cn/ http://www.wnatces.com/	**Policy-related Metrics**	http://www.ndanev.com https://wap.miit.gov.cn/
Technology Innovation -related Metrics	https://www.ne-time.cn/	**Infrastructure -related Metrics**	https://www.gov.cn/

Figure 3 Table1 Data Sources Related to Electric Vehicles(EVs)

4.2.2 *Data Preprocessing*

First, we used the isnull() function in python to find the missing values. Considering analyzing the results, we decided to use the year (2013 2022) with more perfect data for each item as the final data for modeling.

Then, different evaluation indicators often have different magnitudes and units of magnitude, in order to eliminate the effect of magnitude between indicators, this paper standardizes the data: $x' = \frac{x-\min(x)}{\max(x)-\min(x)}$ Where X is the raw data for each indicator and the linear normalization function converts the raw data linearization method to the range of [0 1].

4.3 Determination of main influencing factors based on gray correlation analysis

There are many factors affecting new energy vehicles in China, and this paper uses gray correlation analysis to determine the main influencing factors. Gray correlation analysis is a method of evaluating the degree of correlation between sub-factors and parent factors, and the basic idea is to find out the degree of correlation between different factors by comparing their development trends. Among them, the parent factor here refers to the market share of China's new energy vehicles, and the sub-factors here specifically refer to the ten secondary indicators selected above.

First, we build the data matrix in the order of increasing years:

The sub-factor matrix (comparison matrix) is:

$$\begin{bmatrix} \boldsymbol{X}_1' & \boldsymbol{X}_2' & \cdots & \boldsymbol{X}_n' \end{bmatrix} = \begin{bmatrix} x_1'(1) & x_2'(1) & \cdots & x_n'(1) \\ x_1'(2) & x_2'(2) & \cdots & x_n'(2) \\ \vdots & \vdots & & \vdots \\ x_1'(m) & x_2'(m) & \cdots & x_n'(m) \end{bmatrix} \tag{1}$$

The parent factor sequence (reference sequence) is: $\boldsymbol{X}_0' = \left(x_0'(1), x_0'(2), \cdots, x_0'(m)\right)^{\mathrm{T}}$ Secondly the association coefficients of each comparison sequence with the corresponding elements of the reference sequence are calculated separately by the following equation:

$$\begin{gathered} \gamma\left(x_0(k), x_i(k)\right) = \frac{\Delta\min + \rho\Delta\max}{\Delta_{ik} + \rho\Delta\max} \\ \Delta\min = \operatorname{minmin}_i \left|x_0(k) - x_i(k)\right| \\ \Delta\max = \max_i \max_k \left|x_0(k) - x_i(k)\right| \\ \Delta_{ik} = \left|x_0(k) - x_i(k)\right| \end{gathered} \tag{2}$$

ρ is the resolution coefficient within $(0, 1)$. The smaller the resolution coefficient is, the greater the difference between the correlation coefficients and the stronger the differentiation ability, and we take 0.5 here.

Finally, the weighted average of the correlation coefficients between each indicator and the corresponding element of the reference series is calculated to reflect the correlation between each indicator and the reference series, which is called the correlation degree.$r_{0i} = \frac{1}{m}\sum_{k=1}^{m} W_k \zeta_i(k)$ The results are shown in the table below.

Relevance			
Index1	0.921	Index3	0.745
Index6	0.915	Index5	0.712
Index8	0.895	Index4	0.687
Index9	0.879	Index7	0.454
Index10	0.864	Index2	0.351

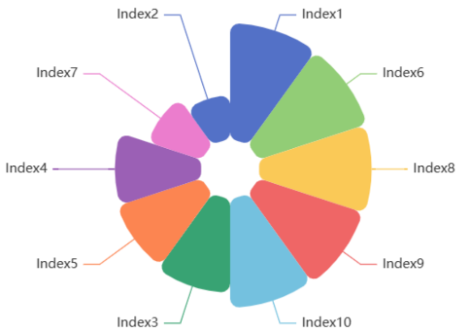


Figure 4 Grey correlation degree result

According to the calculation of gray correlation, we find that the number of public charging piles, subsidy policy and the average selling price of new energy vehicles are the factors with the greatest correlation with the development of new energy electric vehicles in China, which can be judged as the most important influencing factors. In the future, this paper will further analyze the eight major influencing indicators other than Index2 and Index7, and dig deeper into their specific impact on the development of new energy vehicles in China.

4.4 Impact analysis based on GBDT and SHAP models

4.4.1 *Constructing the GBDT model*

The GBDT model is an additive model that serially trains a set of CART regression trees, and eventually sums the predictions of all the regression trees, which results in a strong learner where each new tree fits the negative gradient direction of the current loss function. The final output is the sum of this set of regression trees, thus obtaining the regression results [3]. We use python's sklearn library to construct the GBDT regression model, in which the eight influence indicators are the independent variables, China's new energy vehicle occupancy rate is the dependent variable, and the ratio of the divided training set and test set is 7:3, and the regression results are as follows:

	MSE	MAE	R^2
Train Data	0	0	1
Test Data	0.079	0.25	0.972

[1] MSE (mean square error)
[2] MAE (mean absolute error)

Figure 5 Regression result assess

4.4.2 *Constructing the SHAP model*

Although integrated machine learning methods such as GBDT can complete complex fitting tasks through complex computational transformations, they are basically in a "black box" state internally, making it difficult to analyze the specific impact of each indicator on the dependent variable. Therefore, this paper introduces the SHapley Additive exPlanation (SHAP) model to deeply analyze the influence mechanism of each indicator through the method of ex-post explanation.

Figure 6 The role of the SHAP model

The core idea of the SHAP model is to calculate the marginal contribution of features to the model output, and then interpret the "black-box model" from both global and local levels, which is calculated by the formula:$g\left(z'\right) = \emptyset_0 + \sum_{j=1}^{M} \emptyset_j$.where g is the explanatory model, M is the number of input features, ∅ is the imputation value (Shapley value) for each feature, and ∅_0 is a constant[4].It is easy to see from the formula that solving the explanatory model for each impact indicator focuses on calculating the Shapley value for each indicator. For integrated tree models such as GBDT, for a certain indicator j, the shapley value needs to be calculated for all possible combinations of features (including different orders), and then weighted and summed, i.e:

$$\emptyset_j(val) = \sum_{S \subseteq \{x_1,\ldots x_p\} \setminus \{x_j\}} \frac{|S|!(p-|S|-1)!}{p!} \left(\mathrm{val}\left(S \cup \{x_j\}\right) - \mathrm{val}(S)\right) \tag{3}$$

Where S is a subset of the features used in the model, x is a vector of feature values to be interpreted, p is the number of features , and val(S) refers to the output value of the model under the feature combination S.

We use python's shap library to build Explainer to analyze and interpret the GBDT model constructed above, and present the results through visualization. Due to space limitation, we only do detailed analysis for some years.

Figure 7 The results of 2021 SHAP

Figure 6 presents a specific analysis of the impact of each indicator on the development of new energy vehicles in China in 2021. In particular, the red line segments indicate additional facilitating effects at certain benchmarks, and the blue line segments indicate additional inhibiting effects at certain benchmarks.

As can be seen, the factor that contributes most to the sales of new energy vehicles in China in this year is the average price of new energy vehicles. Meanwhile, China's new energy vehicle patents in 2021 showed a blowout development, the number of related patents increased by more than 2,000.The number of charging piles has a certain inhibitory effect on the development of new energy vehicles, due to the number of domestic charging piles in 21 years is relatively small, can not fully meet the needs of users, resulting in some consumers refused to buy new energy vehicles due to the inconvenience of charging. From the analyzed data in 2022, it can be seen that the problem of insufficient charging piles will be fully improved in 22 years.

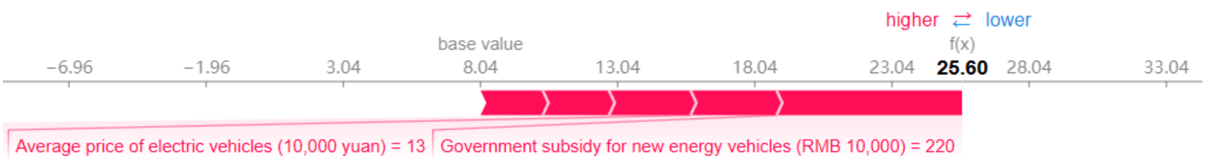


Figure 8 The results of 2022 SHAP

Specifically, many indicators other than government subsidies, in the early stage of development due to imperfections, the development of new energy vehicles for the promotion of limited, lower than the benchmark value; but with the increasing improvement of the indicators, the promotion of the development of new energy vehicles is increasingly obvious. Taking the number of patents for new energy vehicles as an example, it can be seen from Figure 8 (below) that in the early stage of development, the number of patents is insufficient, which inhibits the development of new energy vehicles in China to a certain extent; however, in the later stage, the number of patents for new energy vehicles in China has achieved a qualitative leap, and has become an important factor in the promotion of China's development of new energy vehicles.

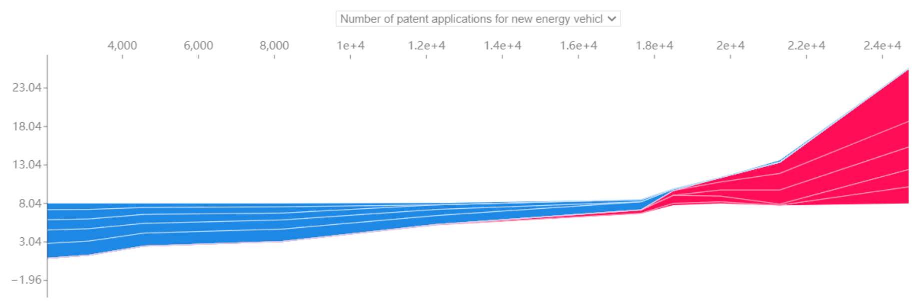


Figure 9 Analysis of the Impact of Patent Numbers on New Energy Vehicles

V. Model building and solution of question2

In the second question, we collect data related to the development of China's new energy electric vehicle industry and forecast the development of China's new energy vehicle industry in the next 10 years by building a seasonal ARIMA model using the collected data.

5.1 Data preprocessing

Consulting the information, we can see[5],The indicators to measure the development of the industry mainly include three categories: growth rate indicators, technology indicators and scale indicators, the growth rate indicators are reflected in the increase of the market size of new energy vehicles over the years, the technology indicators are reflected in the number of patent applications of new energy vehicles over the years, and the scale indicators can be measured by the total amount of sales, so we plan to collect data from the above three aspects for analysis and take the sales of China's new energy vehicles as a representative. Specifically analyze the development of the industry.

In the process of data collection, we found that occasionally, data values were missing for a small period of time and could not be collected, therefore, it was necessary to replace the data with missing values. The calculation of the missing rate was carried out by the following equation:μ = d/ sum(μ is the missing proportion, d is the number of missing cases, sum is the total number of cases)

Taking the 2011.12 2022.10 China new energy vehicle sales volume dataset as an example, the calculated missing proportion of this data is 4.8%, which is lower than 10%, and statistically it is considered that the interpolation of different missing values may have no significant effect, so we use the Lagrangian interpolation method in the Spss software to fill in the missing values.

5.2 Visualization

We not only want to predict the development of the industry for the whole year, but also want to explore the pattern of change in each year, so the next sales data and sales data in units of months and years, as well as the visualization of the other two indicators, the results are as follows:

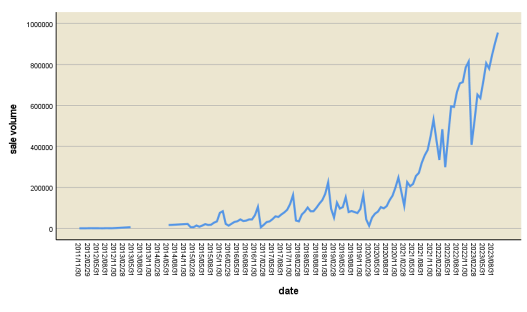


Figure 10 sale volume (months)

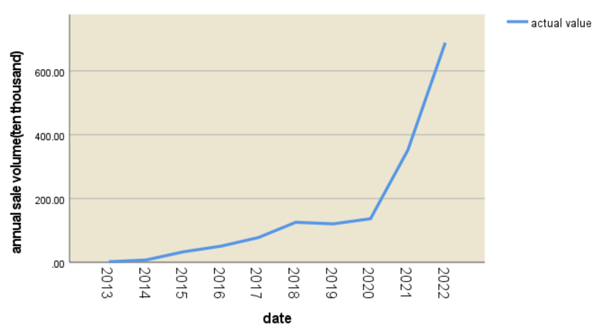


Figure 11 sale volume (years)

Figure 9 (the first graph) and Figure 10 (the second graph) show the sales data in terms of months and years, respectively, and it can be seen that China's sales of new energy electric vehicles are generally on the rise with the advancement of the year, and there are obvious cyclical fluctuations.

Specifically, from 2011 to 2016, the sales of new energy vehicles experienced slow growth. However, starting from 2017, the growth rate of new energy vehicles significantly increased. By the end of 2022, there was a remarkable 1200% increase from the sales of 109,000 units in 2017.

Simultaneously, an analysis of monthly data reveals certain patterns and periodicity. Sales volume tends to steadily rise in the first and second quarters of each year, with a significant increase in growth rate in the mid to late second quarter or early third quarter, reaching a peak rapidly. Subsequently, there is a rapid decline at a faster pace. The occurrence of this periodicity is associated with factors such as the characteristics of the batteries in new energy vehicles and consumer purchasing desires.

Figures 11 and 12 (below) show the market size and number of patents for new energy vehicles in China. It can be seen that China's new energy vehicle market size and the number of patent applications have been increasing steadily year by year, of which the market size had a low growth rate before 2016, with an average of RMB 6.5 billion per year, and after 2016, with the accumulation of resources and base, the growth rate has increased significantly and tended to stabilize. And the number of patent applications increased year by year at a relatively stable rate.

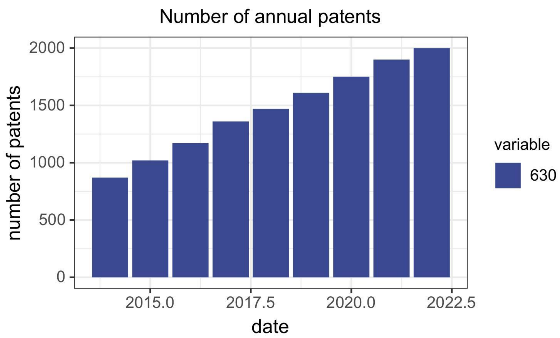


Figure 12 Number of annual patents

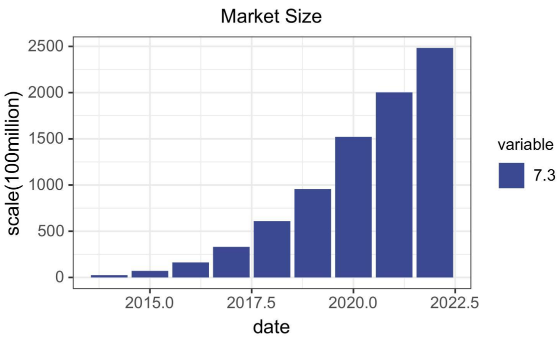


Figure 13 Market Size

5.3 Seasonal ARIMA modeling

From the visualization analysis, we found that the sales of China's new energy vehicles may have a certain cyclicality. Therefore, when forecasting the development of China's new energy vehicle industry in the next ten years, we decided to use the seasonal ARIMA model for analysis.

5.3.1 *Autocorrelation function validation*

Autocorrelation Functions ACF,Correlation may exist between the values of a time series at different points in time. If there is a significant correlation at a certain lag, then the time series is considered to be autocorrelated at that point in time. The ACF can be calculated and the autocorrelation coefficient is given by$\rho_k = \frac{\mathrm{Cov}(X_t, X_{t-k})}{\mathrm{Var}(X_t)}$ where X_t is the value of the time series at the time point t, X_{t-k} is the value of the time lag k period at the time point t, Cov denotes covariance, and Var denotes variance. We used Spss to calculate the ACF of China's new energy vehicle sales data from 2011.12 to 2023.10,the results are as follows:

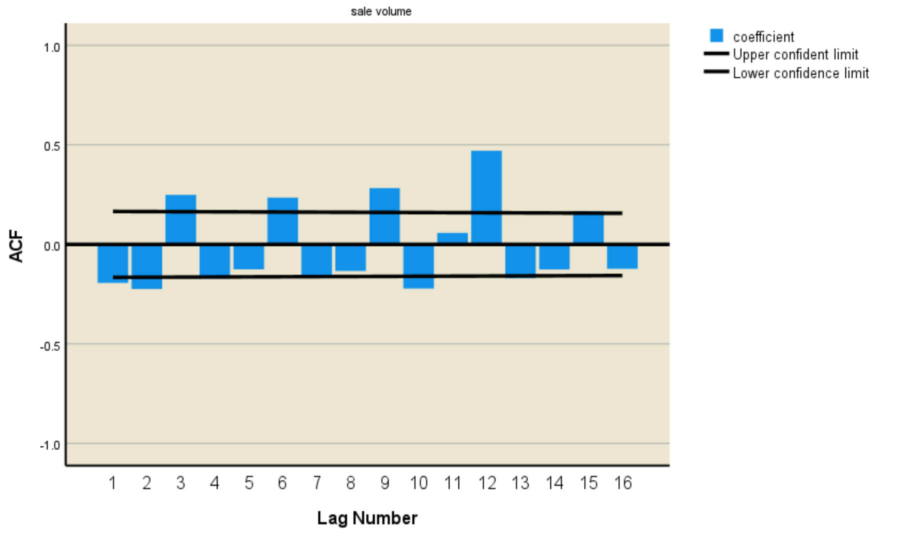


Figure 14 ACF result

Analyzing the ACF chart, it can be found that there are the following points: after the lag phase 0

1 There is a positive correlation and negative correlation, which means that the sales volume of new energy vehicles in a certain period of time will appear first growth and then decrease or first decrease after the growth of the change, this change indicates that the total sales volume is not simply increasing or decreasing year by year, but there is a repeated cyclical pattern.

2 After reaching the peak value at lag number 12, it gradually converges to zero. It indicates that the sales volume data has autocorrelation in different time periods, and the value of this period is 12 months. 3. After reaching the peak value at lag number 12, it gradually converges to zero. It indicates that the sales volume data has autocorrelation in different time periods, and the value of this period is 12 months.

5.3.2 *Forecast for the next decade*

We not only want to analyze the annual sales volume of new energy vehicles in the next ten years, but also want to further analyze the trend of the sales volume in the predicted one year, so we set the change period S to 12 months. Due to the influence of temperature, holidays and other environmental factors in a year, the data in a time period may be unstable, so the first need to smooth the data first, we carry out 1 seasonal difference processing of the data, on the basis of which we construct the 3rd-order autoregressive and 1st-order shifted panning seasonal time series model with the data period of 12 months:

$$y'_t = c + \phi_1 y'_{t-1} + \ldots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{4}$$

Where θ_t is the white noise, representing the smoothed external influences, and the y-value in the right-hand equation is the delayed value with the delayed error, indicating the adjustment made by the y-value to the influences.

We use the method of signal decomposition to visualize the analysis results of the seasonal ARIMA model. The observed values of new energy vehicle sales are decomposed into Trend, Seasonal, and Residual, where Trend shows the growth trend of new energy vehicles, which can be seen that the growth rate of new energy vehicle sales has accelerated significantly since 2020; Seasonal shows the cyclical characteristics of new energy vehicle sales: the sales volume grows in spring, summer, and fall, and the growth rate is the fastest in fall, while the sales volume drops significantly in winter; and Residual shows the cyclical characteristics of new energy vehicle sales: sales volume grows in spring, summer, and fall, and the growth rate is the fastest in fall. Seasonal shows the cyclical characteristics of new energy vehicle sales: the sales volume grows in spring, summer and fall, with the fastest growth rate in fall, while the sales volume drops sharply in winter; and Residual shows irregular noise.

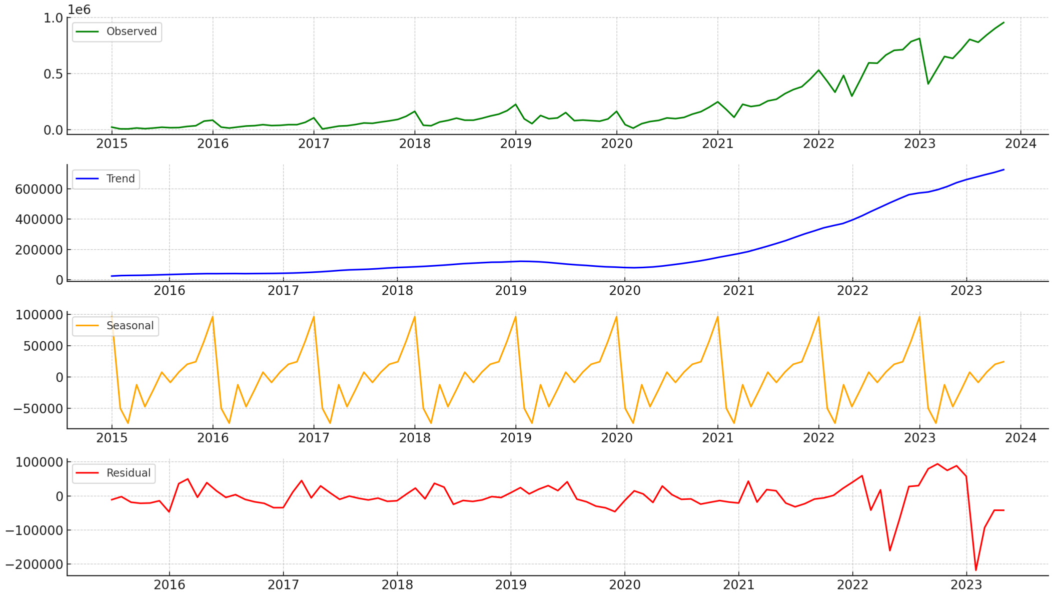


Figure 15 Result of ARIMA

After the model was constructed, we first predicted the annual sales volume of new energy vehicles in China in the next ten years. The results of the data prediction are as follows: Forecast results show that in the next decade,in 2023 China's new energy vehicle sales are expected to reach 8.5 million units, 2024 is expected to exceed 10 million units. From the point of view of growth rate, the growth rate from 2023 to 2028 is relatively stable, with an average annual growth rate of more than 12% in 6 years; after 2028, the growth rate has declined with a total growth rate of only 30% in 4 years, which may be related to the new energy car market is close to saturation.Considering the overall demand in China's auto market, new energy vehicles will replace fuel vehicles as the mainstream models in the market in ten years.

Year	Sale volume(ten thousand)	Year	Sale volume(ten thousand)
2023	849.67	2028	1790.80
2024	1061.87	2029	1878.89
2025	1361.87	2030	1898.45
2026	1403.23	2031	1992.34
2027	1498.45	2032	2035.03

Figure 16 Forecast results

Meanwhile, our seasonal ARIMA model can reasonably predict the sales volume of each month in the future. Due to the limited space, we only show the forecast results of China's new energy vehicle sales volume in the next 24 months (with a 95% confidence interval). The forecast results show that

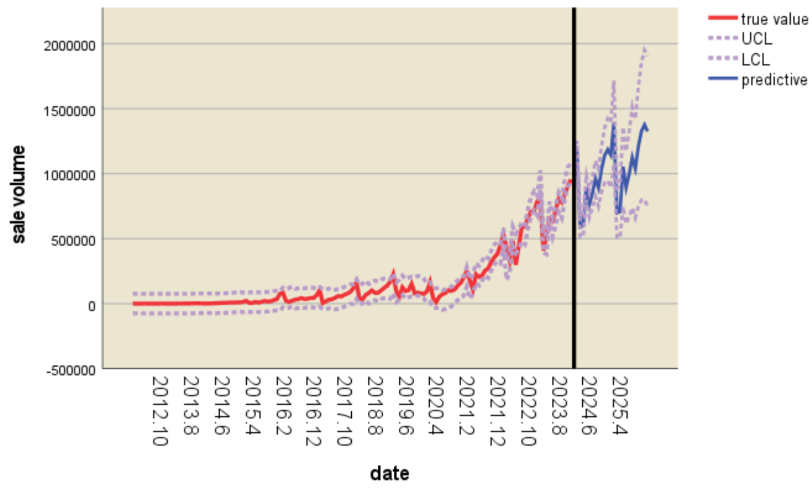
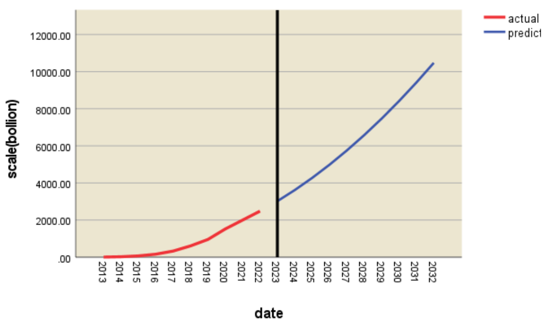


Figure 17 Forecast results

in the next 24 months, China's new energy vehicles will still maintain the original growth trend with seasonal fluctuation growth.Similarly, the forecast results of the market size of new energy vehicles and the number of patents filed are shown in the following chart, and will not be analyzed further here.



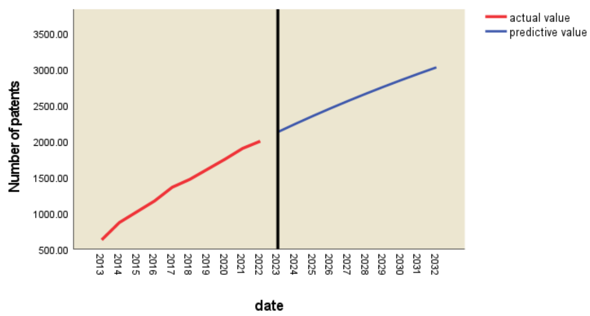


Figure 18 Scale **Figure 19 Number of patents**

VI. Model building and solution of question3

In order to study the impact of new energy trams on conventional energy vehicles, data analysis of sales of both models over the last decade is required. Here, the Vector Error Correction Model (VECM) is able to analyze and understand the long-run relationship as well as the short-run dynamics between multiple non-stationary time-series data, especially when there is reason to believe that there may be a cointegration relationship between new energy trams and traditional energy vehicles, with some kind of smooth linear combination that reflects the long-run equilibrium relationship between them. It takes the form of the equation:

$$\Delta \boldsymbol{x}_t = \boldsymbol{\alpha}\boldsymbol{\beta}^T \boldsymbol{x}_{t-1} + \sum_{j=1}^{p-1} \boldsymbol{\Phi}_j^* \Delta \boldsymbol{x}_{t-j} + \boldsymbol{a}_t + \sum_{j=1}^{q} \boldsymbol{\Theta}_j \boldsymbol{a}_{t-j} \tag{5}$$

In order to construct a VECM model to analyze the relationship between the global sales volume of new energy trams and traditional energy vehicles from 2014 to 2023, first we need to carry out a unit root test (ADF test) on the data in order to determine whether the time series is stationary or not, and the p-value of new energy trams and traditional energy vehicles is calculated to be greater than 0.05, which indicates that both of them are not stationary.So we conducted two differentials to obtaining stationary data.

After that, we performed Vector Autoregressive to determine the optimal cointegration rank and lag order for the cointegration test.The Johansen cointegration test returned the largest p-value and the corresponding cointegration rank of 14.642 and 0.832, respectively, indicating that suggests that sales of new energy electric vehicles and sales of traditional fuel vehicles show some long-term linear dependence relationship. Based on the results of Johansen's cointegration test, using the information criterion, the minimum AIC was obtained as: 4842.716, and the lag order was determined to be 12, which suggests that the effect of data within the past year should be considered in the model.

Subsequently, we built the VECM model and estimated its parameters, including the cointegration vector and heteroskedasticity test.The results show that there is a certain nonlinear relationship between the two. In the long run, the development of new energy vehicles has an inhibitory effect on the traditional automobile industry; in the short run, their short-term sales fluctuations show a more complex interconnection. This short-term linkage may result from the immediate impact of market demand changes, policy changes, oil price fluctuations, technological innovations or consumer preferences.

Table 1 Diagnostic Results

Variable	**AS**	**AP**	**HS**	**HP**
Electric Cars	0.7851	0.3756	21.6897	0.0168
Oil Cars	0.1114	0.7385	14.2529	0.1618

Notes:

- AS: Autocorrelation Statistic
- AP: Autocorrelation P-value
- HS: Heteroskedasticity Statistic
- HP: Heteroskedasticity P-value

In order to quantify the long-term impact of new energy vehicles on fuel vehicles, we conducted a normality test on the two sets of data and chose to use the Spareman coefficient to measure the correlation between the two based on the results of the normality test. The Spareman coefficient indicates that the development of new energy vehicles has a certain inhibitory effect on the traditional energy automobile industry, but the correlation of the long-term data is not obvious.

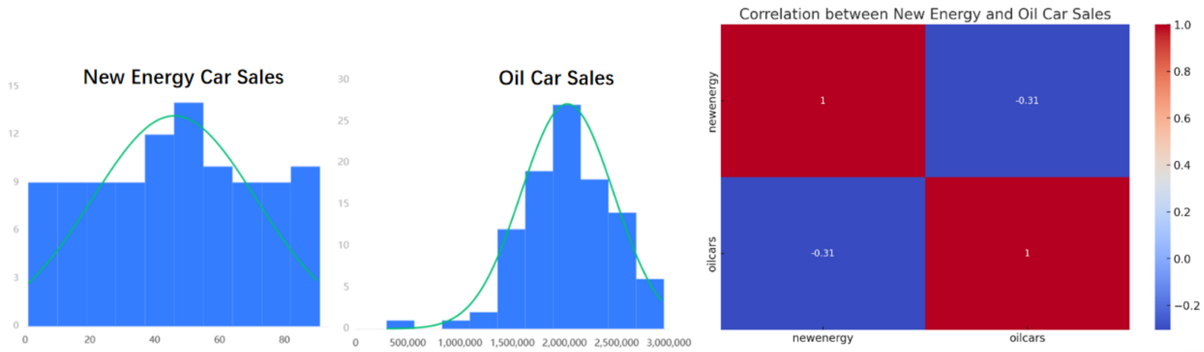


Figure 20 Correlations test

6.1 Impulse Response Analysis

The VECM model is not only able to show that there is a long-term cointegration relationship between new energy electric vehicles and traditional energy vehicles, but also provides us with the possibility to analyze in depth the interaction between these two markets in the short term. We calculated the 12-period impulse response function of traditional energy vehicles to specifically analyze what will happen to the traditional automobile industry in the short term under the impact of new energy vehicles.

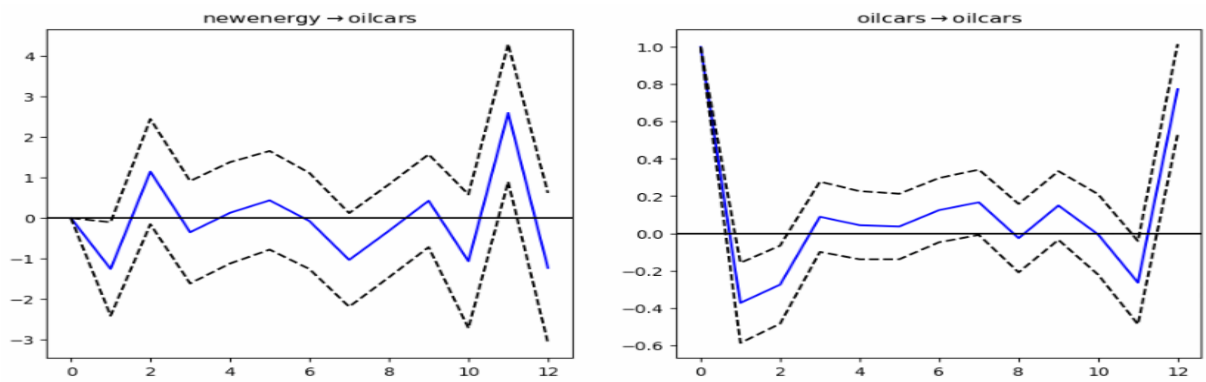


Figure 21 The mutual influence results of different vehicles

Impact of New Energy Vehicles on Conventional Oil Vehicles (left chart): As can be seen from the chart, a shock to New Energy Vehicles can lead to significant positive and negative responses to sales of conventional oil vehicles, which can increase market uncertainty in the short term.

Conventional Tanker Impact on Ownership (Right): The response of the conventional tanker market to a shock to its own ownership is initially positive, then fluctuates and gradually stabilizes. This indicates a tendency for the market to recover from internal shocks, suggesting a degree of stability.

The analysis of the impulse response functions shows that there is interaction between the new energy vehicle and traditional oil vehicle markets, but the strength and duration of this interaction varies. The new energy vehicle market appears to have a strong short-term impact on the traditional oil vehicle market, which may be related to the reaction of market participants to the development of traditional energy technologies and potential market shifts.

These findings are important for understanding how new energy vehicles are reshaping the automotive industry landscape. They also emphasize the importance for automakers, investors, and policymakers of monitoring market dynamics and adapting strategies in a timely manner.

6.2 Predictive analysis

In order to further understand the future trend of the impact of new energy vehicles on hot oil starter cars, we forecasted the sales data of new energy vehicles and fuel car vehicles for the next 24 months, and the visualization results are as follows: The forecast shows that in the next 24 months, with the sales

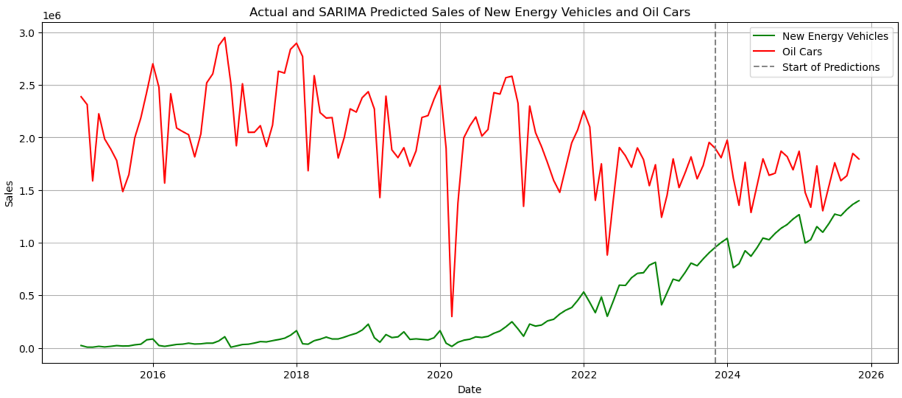


Figure 22 Predicted sales data

of new energy vehicles continuing to rise, the sales of fuel vehicles will show a fluctuating and slight decline. New energy vehicles are showing some replacement of traditional fuel vehicles, but will still coexist and influence each other in the short term.

VII. Model building and solution of question4

In this question, we analyze the impact of foreign boycott policies on the development of China's new energy vehicle industry. Since foreign policies are complex and diverse, and most of the data are not logical, continuous, or regular, we select typical policies to impose nodes, and establish a DID double-difference model to quantitatively analyze the typical boycott policy as a representative.

7.1 Data Collection

Since China is at the world's leading level in the field of new energy vehicles and basically has no technological dependence, foreign targeted boycott policies generally refer to setting up higher trade barriers externally to prevent their imports. Considering at the same time that China's domestic new energy market is not significantly affected by foreign targeted policies, we decided to model and analyze China's foreign exports of new energy vehicles over the years as an indicator for evaluating the development under the influence of foreign policies.

By reviewing relevant information, the initial scale of China's new energy vehicle exports occurred in 2011, and the effectiveness of foreign boycott policies on China's new energy vehicles can be considered to begin in 2019[7], so we collected data on China's calendar year new energy vehicle exports from 2011 to 2019 and from 2019 to 2022 as the object of the study, respectively. A time series model is built using the data before the boycott policy began (2011 2019) to forecast China's new energy vehicle exports from 2019 to 2022 without the impact of the boycott policy. The impact of foreign boycott policies on China's new energy vehicles is quantitatively analyzed by comparing with the actual values.

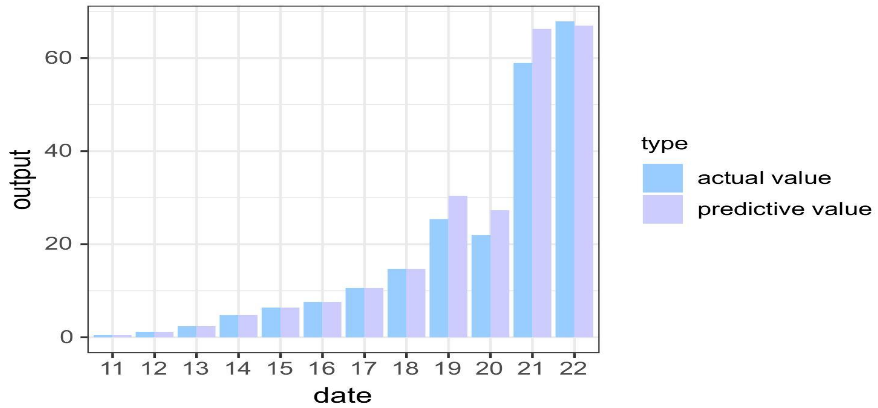


Figure 23 Data on the export volume of new energy vehicles

7.2 Difference-in-Difference Model

In order to better characterize the impact of China's new energy vehicles receiving foreign policies, we introduce the DID double difference model.

The principle of Difference in Difference is the use of observational data to simulate the design of an experimental study. The basic idea is to divide the survey sample into two groups: an experimental group affected by the policy, and a control group that is not affected by the policy. First, the amount of change in an indicator before and after the policy is calculated for the experimental group, then the amount of change in the same indicator before and after the policy is calculated for the control group, and then the difference between the two variables is calculated, thus reflecting the net impact of the policy.

According to the above principle, let Y_{it} be the real export quantity, D_i be the policy grouping dummy variable, affected by the policy is recorded as 1, not subject to it, that is, 0, T_t be the time of the study, and$D_i * D_t$ be the interaction term of the two, and construct the following model:

$$Y_{it} = \alpha + \delta D_i + \lambda T_t + \beta \left(D_i \times T_t\right) + \varepsilon_{it} \tag{6}$$

β is the causal effect of interest in the question, i.e., the impact of the boycott policy on China's exports of new energy vehicles. Then, by constructing an interaction term that identifies the average treatment effect of the policy shock on the affected individuals (the treatment group), we can approximate the "change" of the control group, which has not been affected by the boycott policy, in the period of 2011 to 2022 as follows The "change" in the control group not affected by the boycott between 2011 and 2022 can be approximated by constructing an interaction term that identifies the average treatment effect on affected individuals (the treatment group), as follows:

$$\beta = [\mathrm{E}(Y \mid D=1, T=1) - \mathrm{E}(Y \mid D=1, T=0)] - [\mathrm{E}(Y \mid D=0, T=1) - \mathrm{E}(Y \mid D=0, T=0)] \tag{7}$$

E denotes to find the mean.

We used spsspro software to build the DID model and analyze it, and the results are shown in the table below: From the table, it can be seen that after the policy imposed by foreign countries, the P-value

Conclusion	Standard error	coefficient	P
Before	5.005	0	1
After	5.005	-18.65	0.01***
DID	8.669	-18.65	0.045**

Figure 24 DID analysis results 1

decreases significantly and is less than 0.05, which indicates that the policy has a more significant impact on the export of China's new energy vehicles, and in general, the P-value is less than 0.05, which rejects the original hypothesis of the ineffectiveness of the boycott policy, and the coefficient value of -18.65, which is less than 0, which indicates that the boycott policy is effective and negative. We then visualized

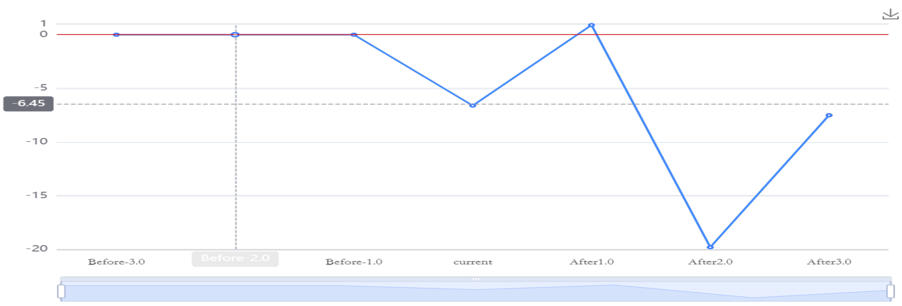


Figure 25 DID analysis results 2

the impact, with the red line in the figure representing the baseline and the blue line representing the impact suffered. As can be seen from the figure, the impact of the boycott policy on China's exports in the three years following its imposition is generally negative, with an average impact of -6.45%.

7.3 Summary

From the above analysis, it can be seen that the boycott policy formulated by some countries can play a certain role in suppressing China's new energy automobile industry at the beginning of the implementation of the new energy automobile industry in three years, making the new energy automobile export volume reduced by 15% compared with the forecast value. However, in the fourth year, China's new energy vehicle exports rebounded significantly, exceeding the predicted value by 3%, which indicates that China's new energy vehicle industry has a strong resilience and acceptance of the challenge of the ability of foreign boycott policy in a short period of time to achieve a certain degree of suppression of the effect, but can not effectively block the rapid development of the industry in China.

VIII. Model building and solution of question5

8.1 LCA-based environmental impact assessment

The emergence of new energy electric vehicles is considered to be a more effective alternative to traditional internal combustion engine vehicles due to their environmental friendliness.

Although electric vehicles can be driven with virtually zero emissions and consume less energy, it must be recognized that a large amount of pollution is generated during the battery production and electrical energy manufacturing processes.

LifeCycleAssessment (LCA) is a method to evaluate the environment and resources in the whole life cycle of a product from "cradle to grave". In recent years, many scholars at home and abroad have comparatively analyzed the energy consumption and environmental effects of the whole life cycle of traditional internal combustion engine vehicles and new energy vehicles. According to the theory of life

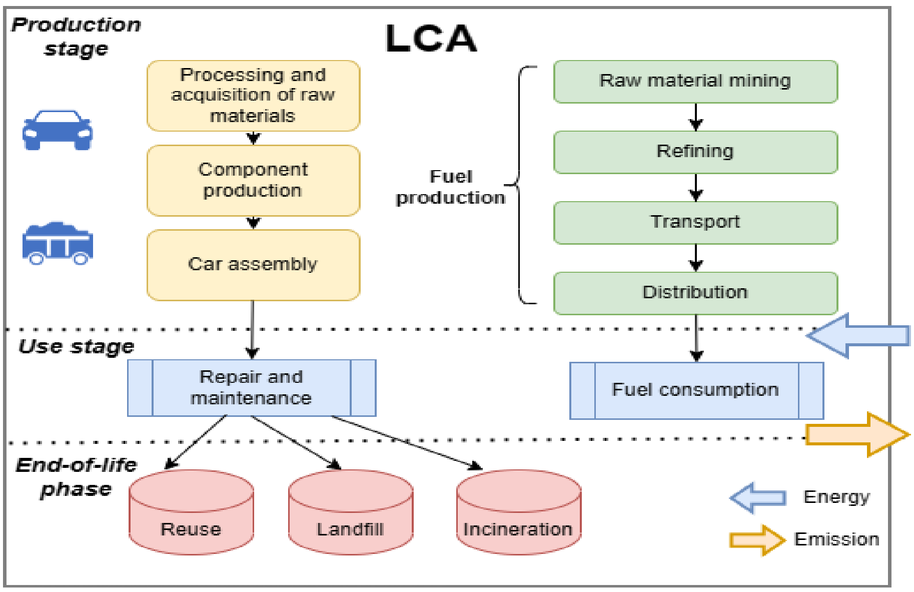


Figure 26 LCA model

cycle assessment, this paper adopts the LCA life cycle system and analyzes the energy consumption and emission per 100 km of conventional internal combustion engine vehicles and pure electric vehicles in the whole life cycle based on GREET software.

We plan to describe the ecological impact of the electrification of new energy vehicles in two ways: 1. Compare the degree of energy consumption of new energy vehicles with that of conventional vehicles under the same energy structure and technological routes 2. Quantify the emissions of pollutant gases from the electrification of new energy electric vehicles, and quantitatively compare the quantified values with those of conventional vehicles as a means of describing the impact on the ecological environment.

1 **Analysis at the level of energy consumption**

We have gathered energy consumption data throughout the lifecycle for traditional internal combustion engine vehicles, battery electric vehicles, hybrid electric vehicles, and hydrogen fuel cell vehicles, respectively. Subsequently, we conducted visualized analyses on these datasets.

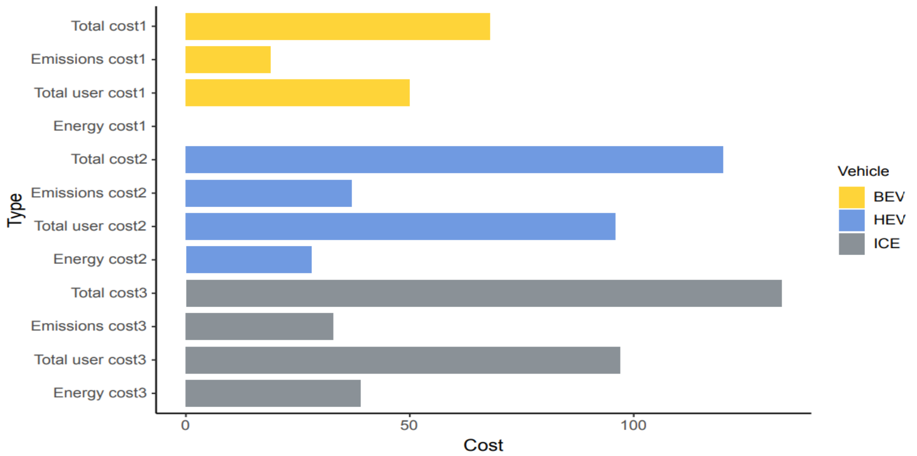


Figure 27 Energy consumption of different cars

From the visualized results, under the same energy structure and technological pathway, new energy electric vehicles demonstrate a greater ability to conserve energy compared to traditional internal combustion engine vehicles. Battery electric vehicles (BEVs) exhibit the lowest lifecycle energy consumption, representing a reduction of 42% compared to traditional internal combustion engine vehicles.

2 **Analysis at the level of emission of pollutant gases**

This paper takes China as an example to analyze the impact of new energy vehicles in terms of pollutant gas emissions. Checking the information, at present, pure electric vehicles account for the largest proportion of new energy electric vehicles in China, about 73%[8]; followed by hybrid vehicles. We use this as a weight, set the impact of new energy vehicles on the environment as X, pure electric vehicles as X_1, hybrid vehicles as X_2, then there are:$X = 0.73X_1 + 0.27X_2$

In order to quantify the impact on the environment, we refer to the text to calculate the total social cost of pollutants by corresponding to the unit cost of the pollutant and the emission of the pollutant

(Meng Xianchun, 2007), the specific formula is as follows:$C_{\text{pollution}} = \sum e_i d_i$
C is the total cost of pollutants, e_i is the pollutant emissions, d_i is the corresponding pollution-free unit emissions.Firstly, the following nine pollutants are selected as the research standard of emission in combination with the main causes of air pollution in China: Among them, CO_2, CH_4 and N_2O

Table 2 Categories And Main Emission

Type	Emissions	Type	Emissions
Greenhouse Gases	CO_2, CH_4, N_2O	**Particle Pollutants**	$PM2.5, PM_{10}$
Acid Rain Pollutants	NO_x, SO_x	**Other Atmospheric Pollutants**	CO, VOC

are all greenhouse gases, and since there are obvious differences in the magnitude of the impacts of different greenhouse gases on the greenhouse effect, the internationally recognized Global Warming Potential (GWP) index is used to measure them. GWP is an index parameterized by CO2, and the correlation coefficients of CO_2, CH_4 and N_2O on the greenhouse effect are 1, 21, 310.
In this paper, the total social cost of pollutants is calculated by corresponding to the unit cost of pollutants and the emission of pollutants, and the specific results are as follows: For a single new

Emissions	Unit cost(yuan/kg)	Emissions	Unit cost(yuan/kg)
CO_2	779.4	CO	864.5
CH_4	823.3	VOC	939.1
N_2O	834.6	$PM_{2.5}$	1042.68
NO_x	840.5	PM_{10}	1171
SO_x	846.6		

Figure 28 Cost of pollutants

energy vehicles and traditional fuel vehicles, the difference between the environmental pollution costs of the two respectively for the above nine kinds of emissions of gas pollution cost difference and here set the new energy vehicle emissions of each kind of pollutant gases relative to the traditional fuel vehicles for the decline in the rate of μ, there is a single new energy vehicles compared to the traditional fuel vehicles per 100 kilometers to save the environmental pollution costs for:

$$dx = 0.73 \times \sum_{1}^{9} (1 - \mu_{ci}) * \text{cost}_{ci} + 0.27 \times \sum_{1}^{9} (1 - \mu_{hi}) * \text{cost}_{hi} \tag{8}$$

The subscript c denotes a pure electric vehicle, h denotes a hybrid vehicle, and collecting the corresponding data above we can calculate dx = 4.617 yuan That is, compared with the traditional fuel car, the new energy car can save 4.617 yuan of environmental pollution costs per 100 kilometers. In terms of air pollutant gases, new energy electric vehicles have obvious emission reduction effects on greenhouse gases (CO_2, CH_4, N_2O), and the greenhouse gas emission reduction effect of pure electric vehicles is 21.54% compared with that of traditional internal combustion engine vehicles. For acid rain pollutants (NO_x, SO_x) and solid particulate matter ($PM_{2.5}$,PM_{10}), new energy vehicles increase the emissions, and pure electric vehicles increase the emissions of acid rain pollutants by 35.36% compared with conventional internal combustion engine vehicles, which is closely related to the battery production process and the preparation of electric energy.

8.2 Environmental impact assessment for a city of one million people

According to question's requst, we select a typical city Kaifeng in He'nan provice with a population of 1.02 million, and use the number of new energy vehicles and the average daily kilometers traveled in the city as a reference number to quantify the cost of new energy vehicles on environmental pollution.Data show that the proportion of car ownership in Kaifeng is 46%, of which the proportion of new energy vehicle ownership is 25%, due to China's private passenger car average daily driving distance of 35.6 kilometers, substituting into the above formula can be calculated as follows:$\sum_0^n dx = 120432.96$

That is, these new energy vehicles can save an average of 12W yuan a day in environmental pollution costs. Assuming that these costs are used to deal with greenhouse gases, of which CO2 contributes 74%, CH4 17%, and N2O 6.2%, we find out that these 12W RMB environmental pollution costs can deal with about 17tCO2, which is equivalent to the CO2 absorbed by 944 trees in a year, and this figure will be doubled by 4 to 5 times if 46% of the cars are new energy cars.

Through the results calculated by the following model, we find that, for a large city with a population of 1 million, if all cars are electrified, the reduction of greenhouse gas emissions per year is equivalent to the CO2 absorbed by about 1.5 million trees per year, i.e., the electrification of new energy vehicles can significantly reduce the pollution of the environment.

To summarize, new energy vehicles can not only reduce the consumption of natural resources, but also reduce harmful gas emissions, which is of great significance for the protection of the ecological environment.

IX. Resolution of question six

In this question, we investigated national initiatives in the field of new energy vehicles, and combined the results of our analysis in Question 5 to write an open letter for the public to publicize the benefits of new energy electric vehicles and the contribution of the electric vehicle industry in various countries around the world.

This open letter was placed after our references.

X. Sensitivity analysis

In the seasonal ARIMA model, the parameter order can measure the size of the role of the influence exerted by the external environment in different time series segments, we take the ten years in the title requirements as the time measure, and get the new prediction result by changing the difference order in the model parameters, and compare it with the original prediction value, and analyze the sensitivity of the time series model to the influence of the external environment by comparing the difference between the two.

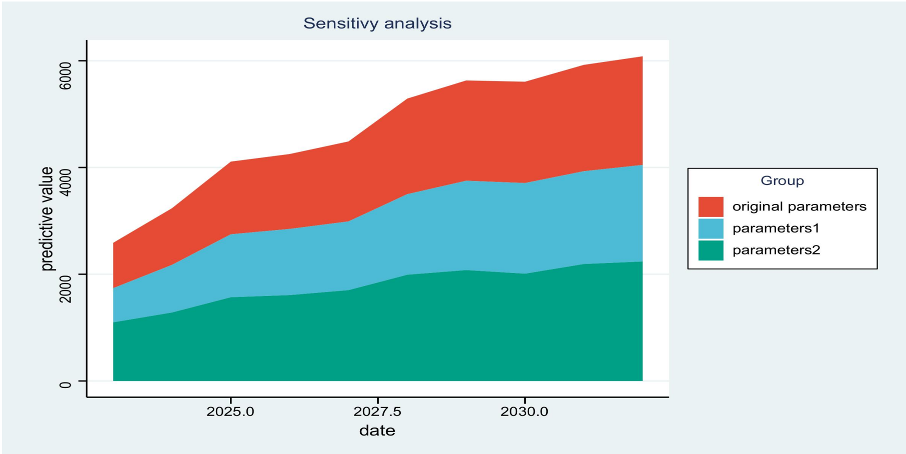


Figure 29 Sensitivy analysis

In the figure, the blue stacked graphs represent the original model (differential order = 2), the red and green stacked graphs represent the model when the differential order is 1 and 3, respectively, and the stacked area represents the predicted sales volume of new energy vehicles under different parameters. Compared with the original parameters, the sum of the predicted values after the first and second changes of the parameters decreased by 15.3% and increased by 24.1%, respectively, and the predicted values under the influence of different environments have obvious adaptive fluctuations, which demonstrates that the model has both excellent stability and reasonable sensitivity to the changes of the environment.

XI. Model Advantages and Disadvantages

11 . 1 Model Advantages

1 In Problem 1, not only the main influencing factors are identified through gray correlation analysis, but also the specific influence of each factor in the GBDT regression on the dependent variable is further analyzed in detail using the SHAP model, which improves the interpretability of the complex machine learning model.

2 The seasonal ARIMA model is established by capturing the cyclical pattern of data in problem two, which makes the model more specific in describing the future development of new energy vehicles in China;

3 In problem three, the vector error correction model is used to not only analyze the long-term impact of new energy vehicles on the traditional energy automobile industry through the cointegration test, but also discuss the short-term effect of the impact of new energy vehicles on traditional energy vehicles through impulse analysis.

4 In problem four, the DID model is established by skillfully setting the predicted value without policy impact as the control group and the real value with policy impact as the experimental group, which solves the problem of not being able to quantify the impact of foreign boycott policy on the development of this industry in China.

5 In question 5, through reviewing relevant literature, the LCA life cycle evaluation system is introduced to quantify the impact on the ecological environment during the entire life cycle of traditional fuel vehicles and new energy electric vehicles into specific environmental pollution costs, so as to realize the quantitative analysis of the impact of electrification of new energy electric vehicles on the ecological environment.

11.2 Disadvantages of the model

1 In this paper, it is difficult to consider the impact of unexpected factors in the modeling process, such as the new crown epidemic, the Russian-Ukrainian conflict, etc., which may have a significant impact on the development of new energy vehicles.

2 Due to data limitations, in the fourth question, this paper can't model and analyze the impact caused by all foreign boycott policies, and can't give the specific impact of specific policies of specific countries on the development of new energy vehicles in China.

XII. References

[1] https://www.iea.org/reports/global-ev-outlook-2021

[2] Gong H, Wang M Q, Wang H. New energy vehicles in China: policies, demonstration, and progress[J]. Mitigation and Adaptation Strategies for Global Change, 2013, 18: 207-228.

[3] https://en.wikipedia.org/wiki/Gradient_boosting

[4] Van den Broeck, G., Lykov, A., Schleich, M., Suciu, D. (2022). On the tractability of SHAP explanations. Journal of Artificial Intelligence Research, 74, 851-886.

[5] DOI: 10.3773/j.issn.1006-4885.2019.09.028 Chinese figure classification number: F062.9 Literature identification code: A Article number: 1002-9753(2019)09-0028-21

[6] Piccolo D. A distance measure for classifying ARIMA models[J]. Journal of time series analysis.

[7] Ma S C, Fan Y, Feng L. An evaluation of government incentives for new energy vehicles in China focusing on vehicle purchasing restrictions[J]. Energy Policy, 2017, 110: 609-618.

[8] Liu W, Tian J, Chen L, et al. Temporal and spatial characteristics of lead emissions from the lead-acid battery manufacturing industry in China[J]. Environmental Pollution, 2017, 220: 696-703.

Open Letter For Citizens

Dear Citizens,

With the rapid development of our cities, automobiles have become an indispensable part of our daily lives. However, traditional internal combustion engine vehicles not only consume a lot of natural resources, but also aggravate air pollution. In contrast, new energy electric vehicles offer an effective solution to environmental pollution and energy consumption due to their efficient energy utilization and ”zero emission” during exercise.

By comparing the energy consumption and pollutant emissions of new energy electric vehicles with those of conventional vehicles over their entire lifecycle, we find some exciting facts: new energy vehicles significantly outperform conventional vehicles in terms of energy utilization efficiency. For example, pure electric vehicles (BEVs) consume 42% less whole-life energy than conventional internal combustion engine vehicles. This means less energy consumption and lower operating costs, resulting in more effective energy savings than conventional combustion engine vehicles.

Second, our study found that pure electric vehicles have a significant reduction in greenhouse gas emissions, significantly reducing CO2, CH4 and N2O emissions by 21.54%. More specifically, in a typical Chinese city of 1 million people, new energy vehicles can save about 120,000 RMB per day in environmental pollution costs. This is equivalent to a reduction of about 17 tons of CO2 emissions, a figure equivalent to the amount of CO2 absorbed by 944 trees in a year. If more of our citizens and friends choose new energy electric vehicles and take new energy buses, this number will increase exponentially, which will be even more effective in reducing environmental pressure and greenhouse gas emissions.

Globally, many countries are actively promoting the popularization of electric vehicles, making significant contributions to reducing greenhouse gas emissions and mitigating climate change. For example, Europe, the United States, China and many other countries have set emission reduction targets and provided relevant policy support. Large investments in research and development have led to rapid innovation and advances in related technologies, further enhancing the performance and efficiency of new energy vehicles. These efforts have significantly reduced the number of conventional vehicles that rely on fossil fuels, thereby reducing environmental damage.

Therefore, when we consider choosing new energy electric vehicles, we are not only making a contribution to our city, but also participating in a worldwide environmental movement to promote a cleaner, greener and more sustainable future. Our choices are not just about transportation, they are responsible choices for the future and for future generations.

Together, let’s move towards a cleaner, greener and more sustainable future.

Sincerely,
Apmcm Team

November 26, 2023