

Using Oil and Lithium Carbonate to Predict New Energy Vehicle Sales in China by Linear Regression Analysis

Yilin Wang*

Department of mathematics, Wuhan University, Hubei, China

* Corresponding Author Email: 2020302191890@whu.edu.cn

Abstract. Recently, a heated discussion has been centered on new energy vehicles, with derivative topics such as the violent rise and new height of oil price and also lithium price, the technical development, and sales volume of Lithium batteries, the rising price of new energy vehicles, and so on. As the new energy vehicle has been expected as one of the most important measures to protect the environment and save traditional fuel energy, it is crucial to accurately understand and explain how the price and import volume of oil and Lithium carbonate impact the sales of new energy vehicle in China and the relationships between each factor, based on linear regression analysis. In this paper, the output of Lithium batteries represents the sales of new energy vehicles in China. The result reveals that the price and import volume of oil and Lithium carbonate all have significant linear relationships with the output of Lithium batteries in China, while the most influential factor is Lithium carbonate price and the weakest factor is Oil import volume. Moreover, these factors interact with each other obviously when they affect the output of Lithium batteries together. The above results of the research provide a theoretical reference for relevant policies to slightly adjust the price of oil and lithium, for companies to modify sales strategies and for consumers to take action to accommodate the general trend of prosperous development of the new energy vehicle industry.

Keywords: New energy vehicle; oil; Lithium carbonate; Lithium battery; linear regression analysis.

1. Introduction

1.1. Background and related research

As one of the most important strategic energy resources internationally, oil is intensely related to the political and economic development and stability of the whole world, therefore it has always attracted attention. Since 2000, the international oil price has fluctuated frequently due to the influence of economic changes, political conflicts, military actions, terrorist attacks, and other factors. Especially in recent two years, international events emerged in an endless stream, oil prices are high, and it runs into the spring of the development of new energy industry in Europe, America, and Asia, so to get rid of the constraints of traditional energy, the pursuit of large quantity, high quality, cheap and alternative new energy has become the trend of the times. Given that the automobile industry is the pillar industry of the national economy and is one of the key industries to support the high-quality development of the national trade [1], there is no doubt that in the new energy industry, new energy vehicle is a very important aspect. Since the early 21st century, with the breakthrough of battery technology, countries began to large-scale apply new energy vehicles, but the market in Europe and the United States is more than earlier and faster in the Chinese market [2]. Until recent years, Chinese people have gradually accepted and bought new energy vehicles, letting the Chinese market fully open. The study of this research is about how the development of new energy vehicles is affected by certain factors from 2021 to the middle of 2022 in China.

At the same time, new energy vehicles use lithium batteries whose main raw material is lithium carbonate as the power support [3], so what impact will the rapid development of the new energy vehicle industry in China bring to the price and sales of lithium carbonate? Is the interaction between them linear? What does the combined effect of the prices of oil and Lithium carbonate and the import volume of both on the sales of lithium carbonate batteries in China? This research is carried out under such doubts and background.

From existing research, the new energy vehicle industry is affected by high oil prices in recent years [4], but whether the relationship between them is linear? And how to use the relationship to predict the future variation tendency of the development of the new energy vehicles? Also, a detailed analysis of the relevance between the Lithium carbonate price and its import volume and the output of Lithium batteries in China is hard to see, though there are some studies showing the rough relevance [5]. Therefore, this research does something about it. In addition, researches about the mechanism of action of the co-effect that the two raw materials bring to the output of the new energy battery are relatively empty, and this study will interpret it at an angle of linear regression analysis, including which is the most significant factor and the degree of factors' interaction. As a consequence, this study appears to explain the above unknown questions.

1.2. Objection

The motivation of this research is to figure out the general rule of how oil and Lithium carbonate affect the new energy vehicle industry in China, thus giving reasonable suggestions to different groups of people to profit and avoid loss by adjusting their ideas and behavior.

As for the framework, the first step is to analyze whether there are linear regression relationships between variables with the measure of simple and multivariate linear regression analysis. Based on that, some conclusions and enlightenments are supposed to be summed up. At last, this paper puts forward several suggestions for reference.

2. Methodology

2.1. Source of data

A data set is collected from several websites (<http://huaon.com/>, <http://www.caam.org.cn/>, <https://yte1.com/datas/idx-yhtl>, <http://www.customs.gov.cn/>, <http://www.niuche.com/>), is used to figure out whether the development of new energy vehicle industry in China is affected by certain factors. The data covering from January 2021 to June 2022 contains 7 numerical variables and they are all recorded monthly. In total, the data set consists of 152 items with 8 columns. One of the columns is the date, which is not relevant to this study.

2.1.1 Dependent variable

In this data set, the main dependent variable is the monthly output of Lithium batteries in China, and another one is the monthly sales of new energy vehicles in China. Since the new energy vehicle is driven by a Lithium battery and the former variable is the upstream industry compared to the latter one, meaning that it is greater impacted by chemical raw materials, the output of the Lithium battery is set as the main dependent variable.

2.1.2 Independent variable

There are many elements that influence the new energy vehicle industry. However, in this project, it is only 5 independent variables are chosen to be discussed. They are Brent oil price (1), China's 95 # gasoline price (2), oil import volume (3), Lithium carbonate price (4), and Lithium carbonate import volume (5). Brent oil price is representative of the international crude oil price and China's 95 # gasoline is representative of the oil for fuel vehicles in China. The reason for this replacement is that international crude oil price is always observed by WTI and Brent oil, and these two indexes' change trends are similar [6]. What's more, in China, fuel vehicles mainly use 92 # gasoline and 95 # gasoline, and they are close in sales volume.

2.2. Machine learning models

This project is studied by means of a linear regression analysis model with the help of SPSS software. The general theory of linear regression is a statistical analysis method used to determine the

quantitative dependence of two or more variables. Its purpose is to predict trends and find out the laws found in the model, or to find a suitable expression for the data to express a certain trend.

First, the monthly output of Lithium batteries in China is studied with the 5 independent variables respectively, which is called single-variable linear regression analysis. Its formula is

$$y = \beta_0 + \beta_{1x1} + \varepsilon \quad (1)$$

Based on the outcomes, it is studied with the 5 variables together, which is called multivariate linear regression analysis. Its formula is

$$y = \beta_0 + \beta_{1x1} + \beta_{2x2} + \dots + \beta_{pxp} + \varepsilon \quad (2)$$

It's the general form of the linear regression model. Where y is the dependent variable, $x_1, x_2, \dots, x_p$ is p independent variables, ε is the random factor, $(0, \sigma^2)$, and $\beta_0, \beta_1, \dots, \beta_p$ are parameters to be estimated [7]. When $p=1$, Equation (2) is a simple linear regression model (Equation (1)).

2.3. Evaluation metric

For evaluating indicators, through this study, β value, R square, p-value, t value, and F value are concerned about.

β value represents the standardized regression coefficient, it shows the correlation between the independent variable, also known as the predictor variable, and the dependent variable. The reason for standardization is that only when the units of each independent variable and the dependent variable are unified can the results more accurate and reduce the error caused by different units. Usually, when $\beta > 0.8$, the relationship is strong, and when $0.3 < \beta < 0.8$, it is weak.

R square and adjusted R square are the descriptions of the fitting effect of the model. The closer the R square is to 0, the lower the goodness of fit of the regression equation. And the closer it is to 1, the higher the goodness of fit.

the p-value indicates whether the regression model is significant or not. The influence is significant when $p < 0.01$, and $p < 0.05$ means that there is a linear relationship between variables.

the t value is the linear regression test statistic and the F value is the variance test of the regression model as a whole.

3. Results And Discussion

Grounding on the data set, a scatter matrix is firstly made to roughly but visually display the linear relationship between the independent variables and the dependent variable.

In order to verify the conclusion drawn from the scatter matrix and have more accurate cognition, the single-variable linear regression analysis experiment is done. Some evaluation indexes such as p-value and β value are tested and recorded in Table.1.

The simple linear regression analysis leads to the multivariate linear regression analysis, which goes deeper in the research. The multivariate analysis discusses how the independent variables interact with each other and act together on the output of Lithium batteries in China. Additionally, it also interprets the degree of the influential significance of every independent variable to the dependent variable.

3.1. Data Visualization and interpretation

Fig. 1 first illustrates the independent variables generally have a strong single-variable linear relationship with the main dependent variable.

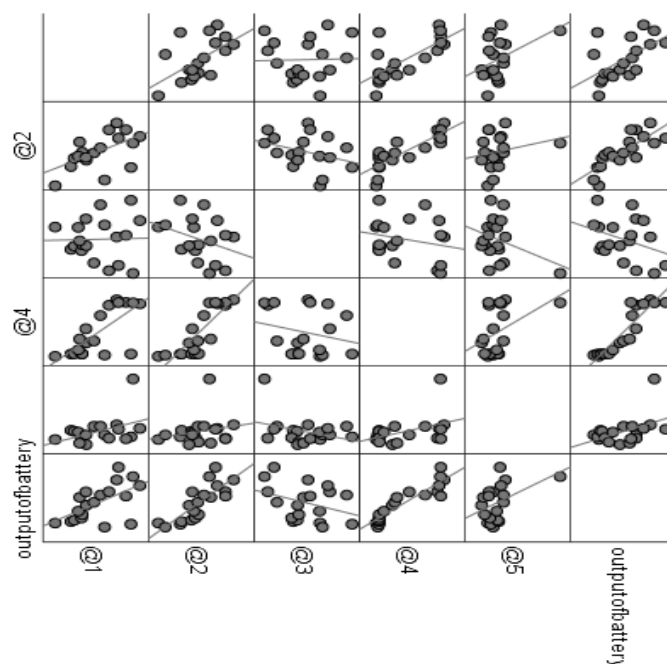


Fig. 1 Scatter Matrix about 5 Independent Variables and the Output of Lithium Battery in China
(Photo credit: Original)

A Scatter matrix is a tool that arranges all of the scatter diagrams showing the linear relationship between any two variables into the form of a matrix. This project has 6 rows and 6 columns, and every variable matches one row and one column. Every square represents the scatter of the row variable and the column variable. If the spots in a scatter distribute irregularly and scattered, the linear dependence is weak. And the more intense and making up of a line the spots are, the stronger the linear relationship the scatter diagram presents. There are also linear fitting lines shown in the graph to help observe the distribution of spots.

When unscrambling, it only needs to observe the upper triangle or lower triangle. Considering the research object is the linear regression analysis of 5 independent variables and the dependent variable, the last row of the graph is where to pay attention to.

The first five pictures in the last row of the scatter matrix in Fig. 1 show that variable 1.2.4.5 have a strong linear relationship with the output of the Lithium battery, while variable 3 has a weak relationship with the dependent variable for the spots are comparatively dispersed.

Fig. 1 is gained by graph—chart builder—scatter diagram in SPSS.

Table 1. Single-variable Linear Regression Analysis of 1.2.3.4.5 Variables Forecasting the Output of Lithium Battery in China Respectively (Merge Form)

variable	β	R^2	p	t	F
Brent oil price	0.574	0.330	0.010	2.892	8.364
China's 95 # gasoline price	0.797	0.636	<0.001	5.446	29.655
Oil import volume	-0.342	0.117	0.152	-1.498	2.245
Lithium carbonate price	0.914	0.836	<0.001	9.297	86.436
Lithium carbonate import volume	0.461	0.212	0.047	2.140	4.580

Table. 1 is a collection of five separate tables to save space. It shows that Brent oil price, China's 95 # gasoline price, and Lithium carbonate price all can significantly predict the output of Lithium battery in China (for $p \leq 0.01$, β is generally big) and they respectively explain 33.0%, 63.6%, 83.6% (R^2 square) of the variance of the output of Lithium battery in China. Though Lithium carbonate import volume can predict the output, it is not remarkable (for $0.01 < p < 0.05$, $0.3 < \beta < 0.8$). Oil import volume cannot predict the output of the battery (for $p = 0.152$, $|\beta| \approx 0.3$). Table. 1 verifies what figure. 1 finds out.

Table 2. Multivariate Linear Regression Analysis of 1.2.3.4.5 Variables Forecasting the Output of Lithium Battery in China together

variable	β	t	p	Adjusted R ²	F	p
Brent oil price	-0.160	-1.014	0.329	0.818	17.201	<0.001
China's 95 # gasoline price	0.100	0.475	0.643			
Oil import volume	-0.072	-0.577	0.574			
Lithium carbonate price	0.885	3.772	0.002			
Lithium carbonate import volume	0.104	0.784	0.447			

Table. 2 shows that there is only Lithium carbonate price significantly influences the output of Lithium battery (for $p < 0.01$, $\beta > 0.8$), indicating that these 5 independent variables interact intensely with each other, and the strongest factor in the output of the battery is Lithium carbonate price.

Moreover, $F = 17.201$, $p < 0.001$ also indicates that at least one independent variable explains part of the variation of the dependent variable (these variables account for 81.8% of the variation in the output of Lithium batteries in China as a whole), so that the regression variation is large, and the model is successfully established.

Table 3. Regression Analysis of the Output of Lithium Battery in China Forecasting the Sales of New Energy Vehicles in China

variable	β	R ²	p	t	F
The output of Lithium battery in China	0.852	0.726	<0.001	6.707	44.986

Table.3 shows that the output of Lithium batteries in China can significantly predict sales of new energy vehicles in China (for $p < 0.001$, $\beta > 0.8$). It illustrates how the 5 independent variables are supposed to act on the sales of new energy vehicles in China is similar to how they act on the output of Lithium batteries in China. The strategy of this research is practicable and scientific.

The data in Table. 1. 2. 3 are all worked out by analysis—regression—linear in SPSS.

3.2. Suggestion

Based on data analysis results, Lithium carbonate price's continuous rise has a significant relationship with the power battery manufacturing industry and the new energy vehicle market. It cannot contain the development of the new power vehicle industry; however, it has truly caused the price rise of power batteries and new energy vehicles [8].

The government supervision department may be responsible for cracking down on unfair competition, such as hoarding and speculation. For more policy suggestions to strengthen the development and exploration of lithium resources, continue to improve the investment in the research and development of lithium extraction technology in salt lakes to reduce production costs. At the same time, optimizes the lithium electricity industry chain to improve the recycling utilization rate of lithium-ion batteries [9]. And some related enterprises are advised to expand their production scale to balance Lithium carbonate's supply and demand, thus maintaining a reasonable price and being responsible for consumers. For example, SQM, the Chilean miner that is the world 's second-largest producer of lithium, raised \$1.1billion in capital earlier in 2021 to expand the production of lithium, iodine, and nitrate, with the aim of more than doubling Chile's annual capacity for lithium carbonate from 70,000 tonnes to 180,000 tonnes. The company expects to produce 140,000 tonnes of lithium in 2022 [10].

Meanwhile, the authority can probably encourage people to choose new energy vehicles by adjusting gasoline prices, since another strong independent variable—China's 95 # gasoline price represents gasoline price in China. There is research showing that compared with fuel vehicles, people's purchase intention of the pure electric vehicle is significantly and positively affected by oil prices [4].

The overall trend is that the power battery manufacture and new energy vehicle industry enjoy good momentum of growth [11]. As a result, besides the government and firms, consumers are also better to enhance awareness and broaden their horizons of the new energy industry as early as possible to go with the tide closely when deciding college major, finding a job, investing, and so on. It will bring them a number of benefits. And there is already a sign of consumers' increasingly obvious tendency toward new energy vehicles in part of China [12].

3.3. Limitation

Firstly, the data set is some small that each column only contains 19 figures matching a short period of 19 months, which means that the result of this study can only be true in recent days. Its universal significance is weak. Secondly, to analyze the change in the condition of the new energy vehicle industry in China, there are lots of factors and the relationship between them is complicated [13]. Merely choosing 5 variables cannot obtain great precise conclusions or the conclusions being expected. Furthermore, there is still vast extended space for this project, like analyzing the interaction of these independent variables, eliminating certain variables, and then making multivariate linear regression analysis again. In addition, the dimensions of data in this project are different, which may make difference to the final results. To improve it, it may be a good idea to normalize all the data before analyzing. Finally, the means linear regression analysis model can do nothing about the non-linear relationship. Barely using this model to predict variables is undoubtedly limited and not firm, consequently, future research can analyze whether there are non-linear relationships between the independent variables and the dependent variable.

4. Conclusion

This research studies the difference made by the price and import volume of oil and Lithium carbonate on the sales of new energy vehicles in China by linear regression analysis. The conclusion is that generally speaking, these independent variables have significant linear relationships with the output of Lithium batteries in the single-variable linear regression analysis, and among them, the most significant one is Lithium carbonate price. But Oil import volume cannot predict the dependent variable. From multivariate linear regression analysis, independent variables interact with each other to a large extent, yet they can explain largely the variation in the output of Lithium batteries in China. In the end, the output of Lithium batteries in China can significantly predict the sales of new energy vehicles in China, indicating that they have a consistent change trend.

Starting from existing data, after processing the data set by linear regression analysis, the change model is built and the future development trend is able to be forecast with the help of the linear regression model. Therefore, people and the authority can make use of favorable conditions and avoid unfavorable conditions according to the future trend gained before. And this is where the significance of the project lies.

5. References

- [1] 21st Century Business Herald. Release the multiplier effect of automobile consumption and help to stabilize the macroeconomic market, 2022-07-11(001), 2022-10-05. <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=SJBD202207110012&DbName=CCND2022>.
- [2] Wang Huifang, Shi shuling. Study on the enlightenment of foreign new energy vehicle policies on the development of China's automobile industry. Science Technology and Economy in Inner Mongolia, 2021, (24): 9-10.
- [3] Wang Ludan, Hebei Daily. Lithium battery: create a rechargeable world, 2022-09-08(011), 2022-10-05. <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=HBRB202209080110&DbName=CCNDTEMP>.
- [4] Jing Peng, Cai yunhao, Sun Huiqian, Wang Weichao, Wang Baihui, Ming Baixu. Can high oil prices promote consumers to buy new energy vehicles? Journal of Transportation Engineering and Information, 2022: 1-29.

- [5] Li Chunlian, Li Qiaoyu. Securities Daily, Performance of several lithium battery companies surged in the first two months, and Lithium salt enterprises are interviewed, maybe making lithium price calming down. 2022-03-21(B02), 2022-10-05, <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=CJRB20220321B022&DbName=CCND2022>.
- [6] Gou Zhili. The price relevance between international crude oil WTI and Brent. Global Market Information Guide, 2015, (43): 9+13.
- [7] Wei Xinxing, Lu Yugui. The application of several statistical software of linear regression analysis. Journal of Hechi University, 2017, 37(05): 64-69.
- [8] Luo Ahua, Zhang Xiaoyu. The rise in the price of lithium carbonate causes public anger. Observations on China's Petroleum and Chemical Industry, 2022, (04): 20-21.
- [9] Ru Cunyi. Evaluation of supply and demand situation of lithium mineral resources in China. China University of Geosciences (Beijing), 2021(04).
- [10] Electric vehicle market is booming, and the price of lithium carbonate exceeds one hundred thousand every ton. SQM estimates that global lithium demand will increase by about 40% this year. China Nonferrous Metallurgy. 2021, 50(04): 48.
- [11] Xu Yingjun, Ding Shaoheng, Luo yantuo. Long term trend prediction of China's new energy vehicle development and its impact on gasoline demand. International oil economy, 2022, 30(08): 32-40.
- [12] Bai Lin. Chongqing Daily, More and more Chongqing consumers favor new energy vehicle, 2022-10-04(001), 2022-10-05, <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=CQRB202210040011&DbName=CCNDTEMP>.
- [13] Gong Piming. International competitiveness of China's new energy vehicle industry: factors, characteristic and promotion path. Modern Management Science, 2022, (04): 63-72.