Research on New Energy Vehicles Based on Multiple Linear Regression and Gray Prediction

Nurjamal Sayit [a, *], Zhengrong Xiao [b], Haowei Yu [c], Zihan Zhang [d], Dengxun Sun [e]

Xinjiang Institute of Technology, Xinjiang Aksu, Xinjiang, 843100, China

[a] 1067224402@qq.com, [b] 2787572608@qq.com, [c] 2118913767@qq.com,

[d] sj45434649164@163.com, [e] sundnegxun@163.com

Abstract. The development of new energy vehicles has gradually become the focus of the global transportation field, and China has included the new energy vehicle industry in the seven strategic emerging industries, but at this stage there are still many factors affecting the development of China's new energy automobile industry. This paper collects the main index data of the old and emerging enterprises, and establishes a gray prediction model based on the analysis of the development level of the two kinds of enterprises in each year, predicts the indexes in turn, and trains the BP neural network model with the original index data. Collect the sales of new energy electric vehicles and traditional energy vehicles, draw the sales comparison chart and market share chart, so as to conclude that the traditional vehicles currently dominate the market, and secondly draw the scatter plot of the two found that the relationship between the two is similar to a hyperbola, and establish a multivariate linear regression model.

Keywords: New energy vehicles, conventional energy vehicles, gray prediction model, BP neural network, multiple linear regression Regression model.

1. Introduction

In the face of increasingly serious environmental problems and energy problems, the development of new energy automobile industry has become a general trend [1]. The development of new energy vehicles has gradually become the focus of the global transportation field, and China is the new energy automobile industry is included in the seven strategic emerging industries, but at this stage there are still many factors affecting the development of China's new energy automobile industry. This paper will comprehensively analyze the factors affecting the development of new energy electric vehicles in China in terms of policy, economy, society and other aspects, and describe and predict the future development of new energy electric vehicles in China.

(1) Analyzing data. Find out the relationship between China's new energy electric vehicles and their influencing factors and the degree of influence; select some representative cities in China, collect and analyze the data of their environment, and derive the impact of new energy vehicles on the urban ecological environment.

(2) Collect data. Analyze the impact of the targeted policies introduced by some countries on China's new energy vehicles on the development of China's new energy vehicles; pay attention to the development trend of China's new energy electric vehicles in the next ten years, and further analyze the impact of new energy electric vehicles on the global traditional energy industry.

2. Problem Description and Analysis

This question is to comprehensively analyze the factors affecting China's new energy electric vehicles and predict the future development of China's new energy electric vehicles. First of all, it needs to collect a large amount of data; secondly, the information is initially analyzed and classified to facilitate a more intuitive grasp of the characteristics of the data; finally, the outliers and irrelevant data in a large amount of data are dealt with to ensure the effectiveness of the model in the latter part of the paper as well as the accuracy of the prediction on the development of China's future new energy electric vehicles. In this paper, BYD, Tesla, Ningde Times, Ideal Automobile, Azure Automobile, Xiaopeng Automobile and other six new energy electric vehicles are selected as the representative brands of new energy automobile industry to carry out exploration.

3. Assumptions of the Model

(1) It is assumed that in the next 30 years, governments will not introduce mandatory policies to block traditional energy vehicles.

(2) Assuming that the historical sales data of new energy vehicles can represent the future sales trend, this paper uses historical sales data to predict future sales.

(3) It is assumed that in the next 50 years, there will be many breakthroughs in the technology of new energy vehicles in China.

(4) Assuming that the consumer population of new energy vehicles is basically stable and consumer behavior can be predicted in the short term;

4. Modeling and Solving the Problem

4.1. Solution Ideas

In the first step, we collect the development data of the new energy automobile industry as a way to predict the development of new energy automobiles in China in the next ten years. We selected a total of six representative brands of new and old brands in the new energy industry as research objects, analyzed the development of these representative industries of new energy vehicles, and analyzed the future development of the representative industries of new energy vehicles through data.

The representative old brands of new energy automobiles in the current era are BYD, Tesla, and Ningde Times, and the representative new energy automobile industries are Azure, Ideal, and Xiaopeng [2]. Collect the industry development data of these two types of brands, carry out the weighted processing of each development data, and use the gray prediction model and BP neural network model to calculate the development level of the new and old brands, respectively, and then predict the development level of China's new energy vehicles in the next ten years [3].

The second step, this paper takes with linear regression method, first collected the global traditional energy vehicle sales and global new energy vehicle sales in the past ten years, analyzed: in the past ten years, the global new energy vehicle market share is the largest year by year; and collected in the past ten years, the global demand for oil (days) and energy battery (GWh). In this way, to comprehensively analyze the impact of new energy electric vehicles on the global traditional energy automobile industry.

4.2. Data Pre-processing

4.2.1. Key Indicators for Forecasting New and Old Branded Vehicles

According to the comprehensive analysis of the factors affecting China's new energy electric vehicles [4], for the key indicators of the old and new brands, the data collected in this question are shown in the table below:

Table 1. Key indicators for forecasting older branded vehicles

Old brand						
Date	R&d investment	Number of charging piles	Average oil price	Policy quantity	Marketing expenses	Average level of development
2017.1	11.51	156058	7278	264	23.53	52.8
2017.2	13.21	170930	7083	269	18.83	55.12
2017.3	11.22	192336	6849	287	23.05	54.5
2017.4	12.92	213899	7379	300	22.11	54.92
2018.1	12.1	253070	7826	320	22.75	51.52
2018.2	13.14	271747	8104	329	24.19	53.19
2018.3	13.54	284648	8426	343	25.28	59.77
2018.4	15.34	299748	8503	360	26.02	60.35
2019.1	13.99	383567	7578	374	25.03	53.92
2019.2	13.07	411615	8076	380	24.38	56.48
2019.3	14.35	466097	7548	390	23.87	58.69
2019.4	15.71	516392	7584	404	24.35	60.33
2020.1	11.5	568669	7519	424	21.25	56
2020.2	13.99	596717	6088	447	24.78	60.04
2020.3	18.36	651199	6303	469	31.24	64.73
2020.4	18.36	651199	6303	469	31.24	64.73
2021.1	22.09	752744	7673	499	35.1	64.95
2021.2	23.49	798718	8048	514	35.56	68.41
2021.3	26.39	860378	7998	527	37.12	69.72

Table 2. Emerging Brand Vehicle Forecast Key Metrics

New brand						
Date	R&d investment	Number of charging piles	Average oil price	Policy quantity	Marketing expenses	Average level of development
2019.3	5.87	466101	7550	392	5.58	31.69
2019.4	6.71	516396	7586	406	7.28	33.64
2020.1	3.4	568673	7521	426	4.29	34.71
2020.2	3.45	596721	6090	449	5.49	39.78
2020.3	5.16	651203	6305	471	8.33	46.33
2020.4	5.48	691498	6213	492	8.58	50.15
2021.1	5.76	752748	7675	501	8.13	45.55
2021.2	7.87	798722	8050	516	11.35	48.27
2021.3	11.11	860382	8000	529	14.66	49.77

4.2.2 Global Sales of Traditional Energy Vehicles vs. Global Sales of New Energy Vehicles

Comparison of global sales of conventional energy vehicles and global sales of new energy vehicles in the last decade Table 3.

Table 3. Global sales of conventional energy vehicles and global sales of new energy vehicles

Year	Global Automobile Sales (10,000 units)	Global Sales of Traditional Energy Vehicles (10,000 units)	Global Sales of New Energy Vehicles (million units)	Global New Energy Vehicle Market Share	Global Traditional Energy Vehicle Market Share
2011	6,438	6,433	4.5	0.07%	99.92%
2012	6,573	6,561	11.8	0.18%	99.82%
2013	6,921	6,899	21.8	0.31%	99.68%
2014	7296	7,264	32.1	0.44%	99.56%
2015	7,429	7,372	56.8	0.76%	99.23%
2016	7,672	7,595	77.4	1.01%	99.00%
2017	8,005	7,882	122.8	1.53%	98.46%
2018	8,201	7,992	208.8	2.55%	97.45%
2019	7,754	7,513	240.9	3.11%	96.89%
2020	5,794	5,489	305.2	5.27%	94.74%
2021	6,939	6,387	551.8	7.95%	92.04%
2022	7,529	6,498	1,065	14.14%	86.31%

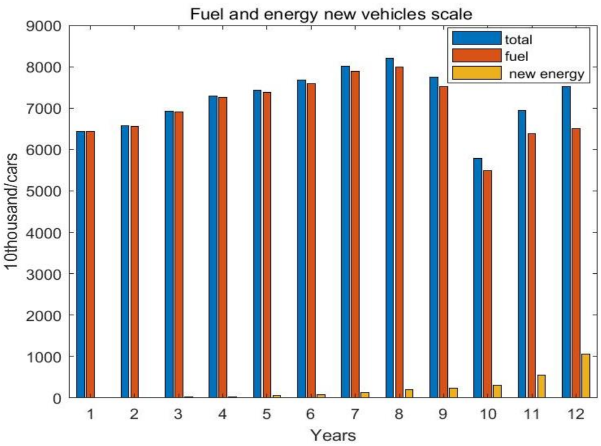

Figure 1. Sales share of conventional and new energy vehicles

In the figure, the horizontal coordinate starting value 0 represents 2011 and the end value 12 represents 2023. Since 2010 until 2018, the market sales of traditional energy electric vehicles have shown an increasing trend, but due to the competition of new energy electric vehicles and the impact of national epidemics in 2018, which led to a decline in the market sales of traditional energy vehicles in 2021, but at the same time the marketing amount of new energy electric vehicles has continued to grow, and gradually occupy a larger marketing market in the world.

4.3. Gray Prediction Model and BP Neural Network Model Calculation

With the indicators in the can be seen, the average development level of new and old brands of new energy vehicles are showing a growth trend Figure 1, so the prediction of China's future development level of new energy vehicles shows a growth trend.

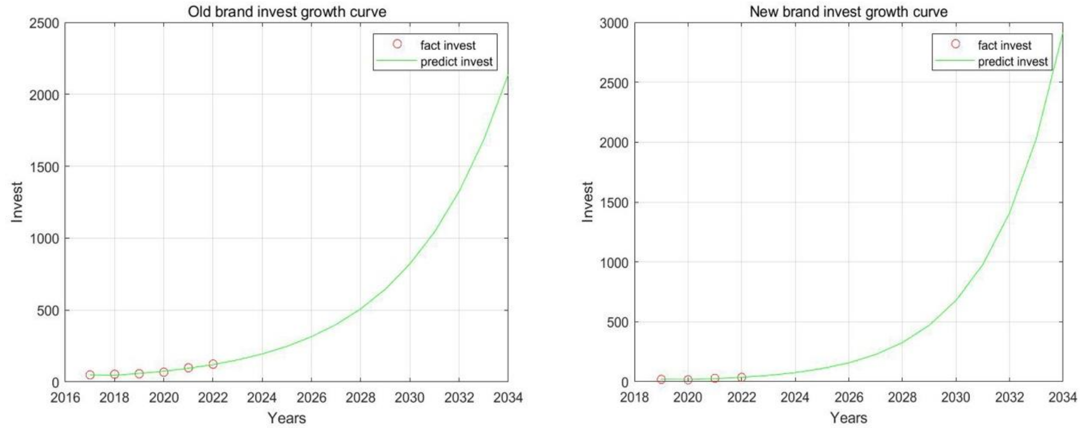

Figure 2. Comparison of R&D of new and old brands of automobiles

4.3.1. BP Neural Network Modeling

This question, need to be based on BP neural network time series prediction to solve, through the neural network model to calculate the next ten years of China's new energy vehicles inherent patterns and trends, and finally used in the future prediction, the next 10 years of China's new energy development trend.

Step 1: Modeling: the original formula:

$E^{(0)}(k) + \delta Z^{(1)}(k) = \nu$, In order to calculate the δ, Value of ν;

Step 2: Estimate the values of δ, ν using regression linear equation analysis, The corresponding whitening differential equations are:

$$\frac{d\tau^{(1)}}{dt} + \delta\tau^{(1)}(t) = \nu \tag{2.1}$$

The solution to the equation is:

$$\tau^{(1)}(t) = \left(\tau^{(0)}(1) - \frac{\nu}{\delta}\right) e^{-\delta(t-1)} + \frac{\nu}{\delta} \tag{2.2}$$

Therefore, it is possible to derive the predicted value:

$$\tau\hat{}^{(1)}(k+1) = (\tau^{(0)}(1) - \frac{\nu}{\delta})e^{-\delta k} + \frac{\nu}{\delta}, k = 1,2,\cdots,n-1 \tag{2.3}$$

The original prediction sequence:

$$\tau\hat{}^{(0)}(k+1) = \tau\hat{}^{(1)}(k+1) - \tau\hat{}^{(1)}(k), k = 1,2,\cdots,n-1 \tag{2.4}$$

Step 3: Construct the data matrix as well as the data vector Ψ:

$$B = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \vdots & \\ -Z^{(1)}(n) & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}(E^{(1)}(1) + E^{(1)}(2)) & 1 \\ -\frac{1}{2}(E^{(1)}(2) + E^{(1)}(3)) & 1 \\ \vdots & \\ -\frac{1}{2}(E^{(1)}(1)(n-1) + E^{(1)}(n) & 1 \end{bmatrix} \tag{2.5}$$

$$\Psi = \begin{bmatrix} E^{(0)}(2) \\ E^{(0)}(3) \\ \vdots \\ E^{(0)}(n) \end{bmatrix} \tag{2.6}$$

Let $u = \begin{bmatrix} \delta \\ \nu \end{bmatrix}$, Solving the above differential equation yields:

$$u=(B^T \cdot B)^{-1} B B^T \Psi \tag{2.7}$$

Then find the values of δ, ν, substituting the above equation, a predictive value function is fitted; Step 4: Residual Test:

$$\omega = \frac{\tau^{(0)}(k) - \tau^{\wedge(0)}(k)}{\tau^{(0)}(k)}, \text{ k=1,2,} \cdots, n \tag{2.8}$$

If the absolute value of all residuals is less than 0.001, it is considered to meet the higher demand; less than 0.02 meets the general demand;

Step 5: Grade deviation value test:

$$\Phi = 1 - \frac{1-0.5\delta}{1+0.5\delta} \lambda(k) \tag{2.9}$$

If the absolute value of all grade ratio deviations is less than 0.001, the higher demand is considered to be met; less than 0.02 meets the general demand;

This ultimately results in the most accurately fitted predictive function.

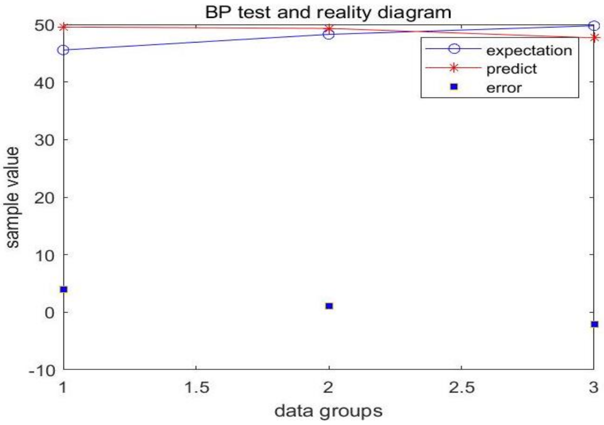

Figure 3. Comparison of predicted and actual BP neural model for new and old brands

4.4. Summary

Through the gray model prediction and BP neural network model [5], the key indicators of new and old brands of new energy vehicles are analyzed, and it is concluded that the development level of new and old brands of new energy enterprises all show a growth trend Figure 4 shows.

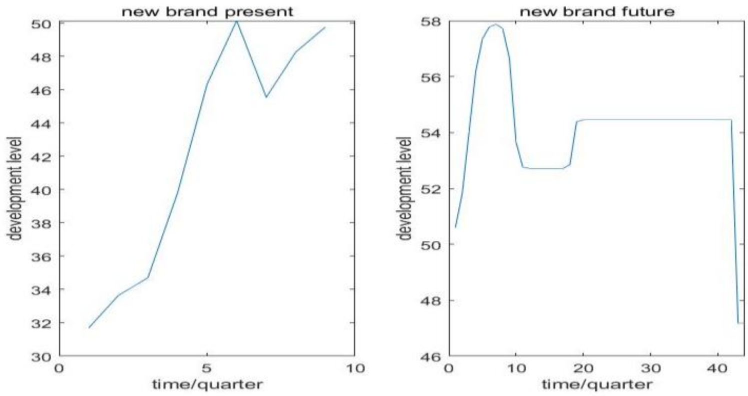

Figure 4. Comparison of current and future levels of development of new brands

Studying the development level of China's representative new energy enterprises, it is predicted that the future development trend of China's new energy is also showing a growth trend, as shown in Figure 4.

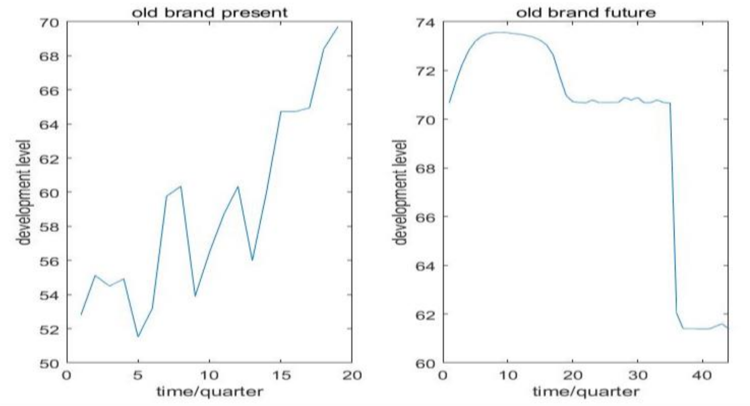

Figure 5. Comparison of current and future levels of development of older brands

As can be seen in Figures 4 and 5 above, established firms continue to increase their level of development over the next three years, but begin to decrease their level of development after 2027, stagnate from 2028 to 2030, and reach a low point in 2031. However, the emerging firms in comparison continue to increase over the next eight years, but begin to decline in their level of development after 2031.

5. Curve Transformations and Linear Regression Calculations

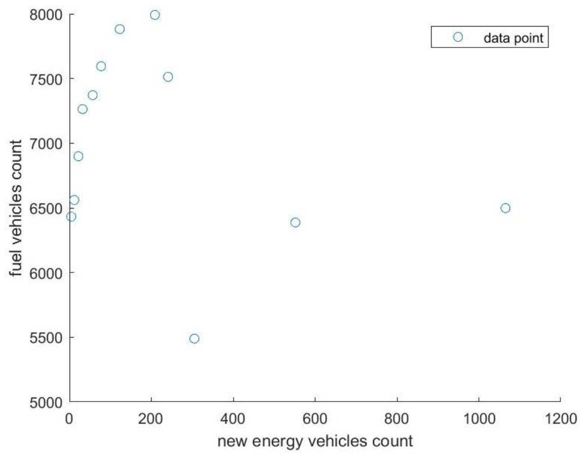

Figure 6. Scatterplot of sales of new and conventional energy vehicles

In order to more accurately analyze the impact of new-energy electric vehicles on conventional-energy electric vehicles, we drew their respective sales scatter plots based on the sales data of the two types of vehicles, and found that the sales scatter plots approximated the hyperbolic graphs that

We therefore set the relation curve as:

$$y = \frac{x}{a+bx} \tag{3.1}$$

Taking this a step further and incorporating the data we have collected, since the hyperbolic plot curve needs to pass through the origin, then we subtract 6433 from the intercept of the function and the modified formula is.

$$y = \frac{x}{a+bx}+6433 \tag{3.2}$$

After subtracting the intercept 6433, the adjusted data are shown in the table below:

Table 4. Market sales of conventional and new energy vehicles after modification

Year	Global sales of conventional energy vehicles (10,000 units)	Global sales of new energy vehicles (10,000 units)
2011	33.2	4.5
2012	161.8	11.9
2013	499.5	21.7
2014	863.9	31.9
2015	971.6	56.4
2016	1,194.5	76.9
2017	1,482.1	122.9
2018	1,593.2	208.5
2019	1,113.4	240.7
2020	1,186.5	305.3
2021	1,558.6	551.7
2022	1,771.2	1,032.1

Then use the curve transformation: Taking the inverse of the formula $y = \frac{x}{a+bx}$;

$$y' = \frac{x}{y} = \tag{3.3}$$

be tantamount to:

$$y' = ax+b \tag{3.4}$$

The linear regression equation REGRESS function was then used to calculate the linear programming regression for =ax+b. The coefficients a, b was determined, and the data from the above figure 7 was brought in, and observing this residual figure 8, it was found that the fit was respectable, and therefore it was possible to use: a=0.0792; b=0.0002, as the coefficients:

Bringing in the original hyperbolic formula yields:

$$\hat{y} = \frac{x}{0.0792+0.0002x} + 6433 \tag{3.5}$$

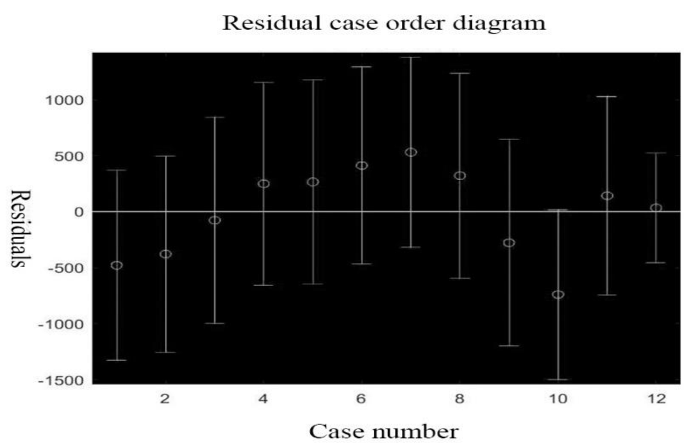

Figure 7. Residual plot of linear regression function for coefficients

Then make the biplot below to analyze the impact of new energy vehicles on conventional energy vehicles around the world.

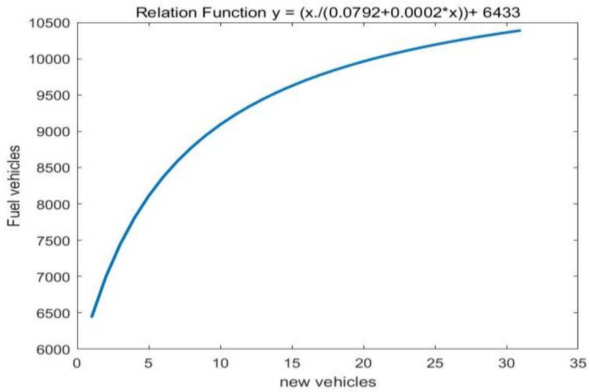

Figure 8. Industrial Relationship Curve between New Energy Vehicles and Conventional Vehicles

As can be seen from Figure 8, new energy electric vehicles and traditional energy electric vehicles will be analyzed. As the development of new energy electric vehicles continues to grow, the development of traditional energy electric vehicles will stabilize, and new energy electric vehicles will have a certain impact on the development of traditional energy vehicles.

6. Summary

In fitting the prediction of the number of China's new energy vehicle exports, the analysis of the number of China's new energy vehicle exports nearly ten years ago is not limited to the simple data collected, but combined with the impact of the epidemic that occurred in the first three years, objectively analyzed China's new energy vehicle exports in the last ten years as a way of fitting the level of China's new energy vehicle development in the next ten years.

The algorithm used in this paper has the advantages of strong optimization ability, strong parallelization ability, etc., which is very suitable for solving many nonlinear regression models. However, the model used still has some shortcomings, such as the gray prediction model to analyze the future is shallow, the data requirements are too high, and the data, if there are large outliers, may have a large impact on the results.

References

[1] Xie Zhao'an. Evaluation of the development level of new energy vehicle enterprises in China and analysis of influencing factors [D]. Jilin University of Finance and Economics, 2023. DOI: 10.26979/d.cnki.gccsc. 2022.000363.

[2] Zheng Longfei, Zhou Zonghong, Liu Jian et al. Research on rockburst grade prediction model based on entropy weight method-LGBM algorithm [J]. Chemical Minerals and Processing, 2023, 52(10):39-45. DOI: 10.16283/j.cnki.hgkwyjg. 2023. 10. 006.

[3] Zhang Xiyong. Ten future development trends of new energy automobile industry [J]. Innovation World Weekly, 2023, (08):42+6.

[4] Peng Hua. Research on the development and spatial layout of new energy automobile industry in China [D]. Jilin University, 2020. doi: 10.27162/d.cnki.gjlin.2019.000347.

[5] Xingshan Zhu, Xuesi Shen. From natural gas to hydrogen: the evolution of China's new energy system [J]. International Petroleum Economics, 2023, 31(08):1-15.