Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: https://doi.org/10.61784/jcsee3049

THE FUTURE DEVELOPMENT OF NEW ENERGY VEHICLES BASED ON ARIMA TIME SERIES PREDICTION MODEL

YiTong Liu[1*], YiLin Wang[2]
[1]Department of business, Accounting, Xi'an International Studies University, Xi'an 710128, Shaanxi, China.
[2]Department of business, Business English, Xi'an International Studies University, Xi'an 710128, Shaanxi, China.
Corresponding Author: YiTong Liu, Email: 18991002666@163.com

Abstract: This paper mainly adopts the evaluation model based on gray correlation method, multiple linear regression model, the secondary polynomial regression model, studied the influence of various factors of new energy vehicles in China and the influence of traditional fuel vehicles, and then choose the prediction of the next ten years, collect new energy vehicles in the past seven years, using the Pearson correlation coefficient test the development of new energy electric vehicles and predictor, found that the strong correlation, and derived the corresponding index of the correlation coefficient. And the ARIMA time series model is used to predict the trend of new energy electric vehicles in the next decade. Research shows that the research and development of new energy electric vehicles is very important for environmental protection. This paper calls on people to buy and ride in new energy electric vehicles to reduce greenhouse gas emissions and promote green development.
Keywords: New-energy electric vehicles; Multiple linear regression model; Pearson correlation analysis; ARIMA time series analysis

1 INTRODUCTION

These years has witnessed rapid economic growth and globalization process, coupled with environmental crises, posing a great threat to the world [1]. For China the world second large economy, it is essential to implement environmental protection measures, addressing the rapidly growing global energy demand [2]. Therefore, the New Energy Vehicles represent a choice dictated by history.

New energy vehicles (NEVs) are vehicles which feature advanced technical principles, novel technologies, and new structures, encompassing four main categories: hybrid electric vehicles (HEVs), pure electric vehicles (PEVs), fuel cell electric vehicles, and other types. NEVs utilize unconventional fuels (excluding gasoline and diesel) as their power source and incorporate advanced vehicle power control and drive technologies, noted for their low energy consumption and reduced pollution [3,4].

Previously, Huang, X., Zhang, D., and Zhang, X. proposed an intelligent building energy management strategy using deep reinforcement learning [5]. However, they overlooked external factors' impact on new energy vehicle development within this context. Zeng B ,Yin F ,Wang J , et al proposed an IBO and a three - parameter gray prediction model for Chinese NEV sales, tackling data and prediction trap issues.Yet, their research under-emphasizes product-related factors like performance and after-sales on sales [6]. Wang, Z. et al, based on deep neural networks, explored sustainable design factors in NEVs for promoting sustainable consumption [7-8], but didn't quantify economic, technological, and policy impacts or predict long - term market trends.

This study addresses these gaps through a combination of methods. By applying gray correlation analysis, the complex factors can be comprehensively analyzed, which influence the new energy vehicle industry, including economic, technological, and policy factors. This helps in understanding the relative importance of different factors and their impact on the industry's development. In addition, the use of ARIMA time-series prediction allows us to forecast new energy vehicle trends accurately. It helps us to not only predict sales but also project long-term market trends. Moreover, by integrating two methods to conduct a comprehensive analysis, the relationships between various factors and the development of new energy vehicles can be qualified, providing a more in-depth analysis of the quantitative impacts of economic, technological, and policy factors that the former studies failed to achieve.The results of this study not only provide a theoretical basis for the development of the new energy vehicle industry, but also provide an important reference for government decision-making and enterprise strategic planning.

2 FACTORS AFFECTING NEW ENERGY VEHICLE DEVELOPMENT

2.1 Variable Declaration

When analyzing the main factors affecting the development of new energy electric vehicles in China,consider the following factors: economy, science and technology, policy and infrastructure and many other aspects, which are interrelated and affect each other, so in order to clarify the impact of these factors on the development of China's new energy electric vehicles more clearly, establish a multivariate linear regression model.

Assuming that the development level of new energy electric vehicles is Y, and the factors of economy, science and technology, policy and infrastructure are X1, X2, X3 and X4 respectively, the multiple linear regression model can be expressed as follows:

$$Y = \beta 0 + \beta 1X1 + \beta 2X2 + \beta 3X3 + \beta 4X4 + \varepsilon \quad (1)$$

Where $\beta 0$ is the constant term, β_1, β_2, β_3 and β_4 are the regression coefficients of each factor, and ε is the random error term.
Specifically, the impact of each factor on the development of new energy electric vehicles in China can be described as follows:
(1) The sales volume of new energy vehicles in China affects the market share occupied by new energy vehicle producers, which in turn affects consumer acceptance and producer production motivation. When the sales volume of new energy vehicles increases, the market share it accounts for will increase, which means that the influence and status of the enterprise in the industry will be enhanced.
(2) The energy density of Chinese new energy vehicles will greatly affect the range of new energy vehicles, which in turn affects the user satisfaction of new energy vehicles. The higher the energy density, the longer the range, the higher the consumer satisfaction and user experience, and the more inclined to buy new energy vehicles. Therefore, the regression coefficient β_2 of X2 should be positive.
(3) The cumulative amount of subsidies provided by the Chinese government for new energy vehicles will reduce the cost of purchasing new energy vehicles, promote consumers' motivation to buy new energy vehicles, stimulate market demand. The stronger the government subsidy, the faster the new energy industry will develop. Therefore, the regression coefficient β_3 of X3 should be positive.
(4) The degree of perfection of infrastructure related to new energy vehicles is inextricably linked to the user experience, which will greatly affect the market demand for new energy vehicles and other factors. For example, the more charging piles in the country, the more convenient it is to use new energy vehicles, and the greater the market demand, the regression coefficient β_4 of X4 should be positive.

2.2 Model Establishment

Determine the parent sequence and feature sequence, and prepare the data format. Discuss each major step in the comparative sequence analysis process as follows:

$$\left[X_1', X_2', \ldots, X_n'\right] = \begin{bmatrix} x_1'(1) & x_2'(1) & \cdots & x_n'(1) \\ x_1''(2) & x_2'(2) & \cdots & x_n'(2) \\ \vdots & \vdots & & \vdots \\ x_1''(m) & x_2'(m) & \cdots & x_n'(m) \end{bmatrix} \quad (2)$$

The parent sequence is as follows:

$$X_0' = \left(x_0'(1), x_0'(2), \ldots, x_0'(m)\right)^T \quad (3)$$

Carryout dimensional normalization of the index (usually required), in order to truly reflect the actual situation, eliminate the influence caused by the difference of each index unit and the disparity between the value orders, and avoid the occurrence of unreasonable phenomena, it is necessary to carry out dimensional normalization of the index.
Calculate the gray relational coefficient between the master sequence and the characteristic sequence. Calculate the correlation coefficient between each element in the comparison sequence and the corresponding element in the reference sequence according to the formula below:

$$\gamma\left(x_0(k), x_i(k)\right) = \frac{\Delta \min + \rho \Delta \max}{\Delta_{ik} + \rho \Delta \max} \quad (4)$$

$$\Delta \min = \min_i \min_k |x_0(k) - x_i(k)| \quad (5)$$

$$\Delta \max = \max_i \max_k |x_0(k) - x_i(k)| \quad (6)$$

$$\Delta_{ik} = |x_0(k) - x_i(k)| \quad (7)$$

The resolution coefficient takes values within [0.5], where the smaller the resolution coefficient, the greater the difference between the correlation coefficients, indicating stronger discriminative ability. Usually, a resolution coefficient of 0.5 is chosen.
Solve the correlation degree value:

$$\gamma_{0i} = \frac{1}{m}\sum_{k=1}^{m} W_k \zeta_i(k) \tag{8}$$

Sort the correlation values to draw conclusions. Establish the correlation sequence of each evaluation object based on the size of the gray weighted correlation. The greater the correlation degree, the greater the importance of the evaluation object to the evaluation criteria.

2.3 Results

The related curves and bar charts derived from these data are as follows Figure 1 and Figure 2:

Figure 1 The Related Curves

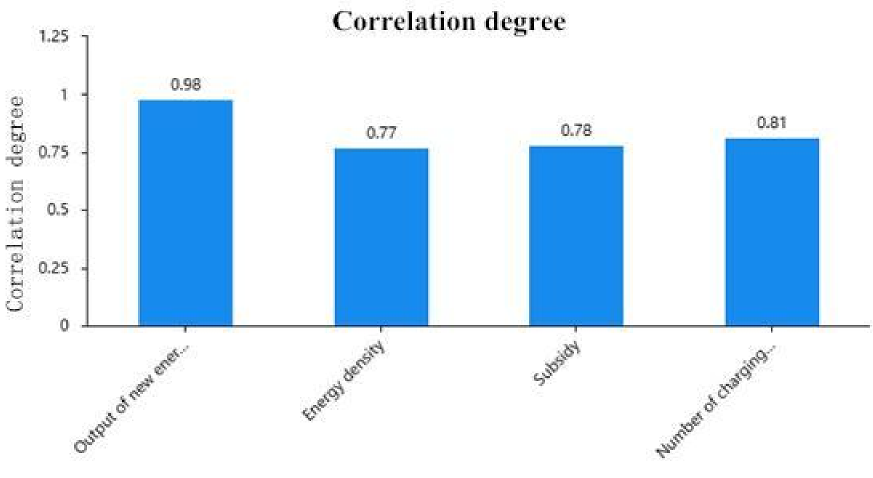

Figure 2 The Correlation Degree

According to the gray correlation algorithm,it is finally concluded that the production of new energy electric vehicles and the number of charging piles in China are the most important factors affecting the development of new energy electric vehicles in China, with the greatest degree of correlation. Therefore, according to the obtained data, in order to promote the development of the new energy electric vehicle industry, the enterprises should be incentivized to increase the cost of production investment and scientific and technological R&D costs of new energy vehicles, and continue to increase the new energy vehicle's market share, accelerate technology iteration, and develop new energy vehicles with longer range and higher performance as soon as possible. At the same time, more charging piles should be arranged to make the use of new energy vehicles more convenient.

3 PREDICTION ON THE FUTURE DEVELOPMENT OF NEW ENERGY VEHICLES

3.1 Basic Data Analysis

(1) x_1 is the utilization efficiency of each unit of charging fee for new energy electric vehicles.

(2) x_2 is the percentage of government subsidies for new energy EVs in the total subsidy amount, and a 1:100 ratio is adopted to standardize the order of magnitude;

(3) x_3 is the vehicle-to-pile ratio, representing the ratio of new energy vehicles per charging pile;

(4) x_4 is the ratio of new energy vehicle production to China's total automobile production.

3.2 ARIMA Forecasting in the Development Trajectory of Nevs

3.2.1 Pearson correlation analysis

Through Pearson correlation analysis, x_1, x_2, x_3, x_4 have significant correlation on China's new energy vehicle market share.

(1) Market share(%) and Utilization efficiency of charging charges(%): correlation coefficient 0.981, significant at 0.01 level, significant positive relationship.

(2) Market share(%) and Subsidy ratio (1:100): correlation coefficient 0.778, significant at 0.05 level, significant positive correlation.

(3) Market share(%) and Vehicles per charging pile ratio (%): correlation coefficient -0.776, significant at 0.05 level, significant negative correlation.

(4) Market share(%) and proportion of new energy vehicle production: correlation coefficient 0.997, significant at 0.01 level, significant positive correlation.

3.2.2 Model establishment

Based on the foregoing, then formulate the subsequent equation by employing the correlation coefficient as the coefficient thereof, and subsequently derive the corresponding function F(x) for the period from 2016 to 2022.

$F_{(x)}$ is the comprehensive future development trend and make ARIMA forecast on it, which can lead to the development trend of new energy vehicle industry in the next 10 years.

$$F(x) = ax_1 + bx_2 + cx_3 + dx_4 \tag{9}$$

The ARIMA model comprises three components: AR, I, and MA. AR stands for Auto Regression (autoregressive model), I for Integration (single integer). As a time series model, ARIMA requires the time series to be smooth for establishing an econometric model. Hence, a unit root test is initially performed on the time series. If non-smooth, differing is needed to convert it into a smooth series, and the number of times differing is applied is called the order of single-integer. MA represents the Moving Average (moving average model). Thus, the ARIMA model is essentially a combination of the AR and MA models.

AR model is used to describe the relationship between the current value and the historical value, and use the historical time data of the variable itself to forecast itself, and its formula is as follows:

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i y_t - i + \varepsilon_t \tag{10}$$

y_t denotes the current value, μ denotes the constant term, p denotes the order, γ_i denotes the autocorrelation coefficient, and ε_t denotes the error.

The MA model is concerned with the accumulation of error terms in the autoregressive model, which can effectively eliminate the random fluctuations in the prediction, and its formula is as follows.

$$y_t = \mu + \sum_{i=1}^{p} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{11}$$

Generally speaking, the MA(q) model must be smooth when q is constant, so for the ARIMA model, it is only necessary to verify the smoothness of the AR model.

The ARIMA autoregressive moving average model combines the advantages of both AR and MA models. In the ARMA model, the autoregressive process is responsible for quantifying the relationship between the current data and the previous data, and the moving average process is responsible for solving the problem of solving the term of the immediate change, and its formula is as follows.

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t \sum_{i=1}^{p} \theta_i \varepsilon_{t-i} \tag{12}$$

3.2.3 Solutions

Since the smoothness of the original data is poor, the original data are processed by first-order differing: the purpose of differing is to eliminate the instability of the series and make its fluctuation curve smoother. Through the time series of first-order difference and second-order difference, to carry out the ADF unit root test for the ARIMA model, this paper utilizes SPSS PRO software to carry out first-order and second-order difference processing on the data. The results are shown in the Table 1: The result of ADF test below:

Table 1 The Result of ADF Test

f(x)-ADF test					
Difference in order	t	p	Threshold value		
			1%	5%	10%
0	0.079	0.965	-5.354	-3.646	-2.901
1	-2.964	0.038	-6.045	-3.929	-2.987
2	-3.235	0.018	-7.355	-4.474	-3.127

As it can be seen from the above table, the p=0.965>0.1, the original hypothesis cannot be rejected and the series is not smooth. The first-order difference is performed on the series and then the ADF testis performed.
The ADF test result of the data after the first-order difference shows p=0.038<0.05, there is a higher than 95% certainty that the original hypothesis is rejected, and the series is smooth at this time.
Therefore, the 1st difference order in ARIMA model is chosen.
Then, perform the white noise residual test, as the Table 2: Model Q statistic result shown:

Table 2 Model Q Statistic Result

Model Q statistic table		
Items	Statistic	*p*
Q1	2.314	0.128
Q2	2.336	0.311
Q3	2.377	0.498
Q4	2.377	0.667
Q5	2.377	0.795

From the results of Q statistics, the p-value of Q5 is greater than 0.1, then the original hypothesis cannot be rejected at the significance level of 0.1, the residuals of the model are white noise, and the model basically meets the requirements.
According to the ARIMA prediction model, this paper predicts the development trend of new energy electric vehicles in the global decade as shown in Figure 3:

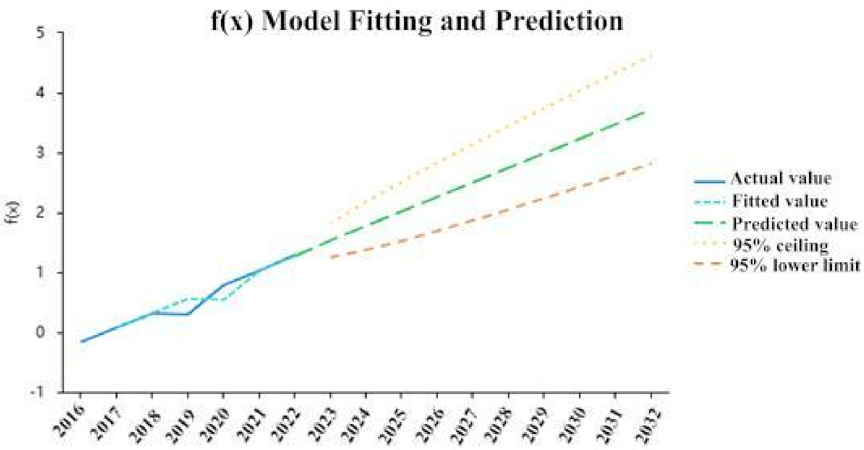

Figure 3 f(x) Fitting and Prediction

The LM test fails in the ARIMA model testing, and the test model residual series do not have serial correlation, which may be caused by the data volume is too small. However, in the partial autocorrelation plot, the p-value greater than 0.05 does not directly determine whether ARIMA forecasting is possible.The selection of ARIMA model needs to consider the characteristics of autocorrelation plot (ACF) and partial autocorrelation plot (PACF). As the result above, combined the ADF test and partial autocorrelation analysis to finally decide to use the ARIMA (0, 1, 0) model to predict f(x) and arrive at the above results.

4 CONCLUSION

The results show that among the factors affecting the development of new energy vehicles, policy subsidies, production of new energy vehicles play a key role in the number of new energy vehicles. The output of new energy vehicles and the number of charging piles are the most closely related to the development of new energy vehicles in China, indicating that these factors are the main driving force for the development of the new energy vehicle industry. Secondly, new energy vehicles will show a growth trend in the next decade. With the enhancement of global awareness of environmental protection and the transformation of energy structure, new energy vehicles will gradually become the mainstream of the automobile market with their characteristics of low carbon, environmental protection, energy saving and high efficiency. Governments of various countries have introduced relevant policies to increase the support for new energy vehicles and promote their industrialization process. Technological innovation will continue to promote the development of new energy vehicles, and the popularization of new energy vehicles will promote the optimization of the energy structure. To sum up, China's new energy vehicle industry has broad development prospects and huge market potential. The results of this study not only provide a theoretical basis for the development of the new energy vehicle industry, but also provide an important reference for government decision-making and enterprise strategic planning. Research shows that new energy vehicles are of great significance to environmental protection, and China's new energy vehicle industry has broad development prospects and huge market potential in the future.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Zhang Q, Du M A, Lin B. Driving total factor productivity: The spillover effect of digitalization in the new energy supply chain. Research in International Business and Finance, 2025, 75, 102764-102764.

[2] Fan B, Wen Z, Qin Q. Competition and cooperation mechanism of new energy vehicle policies in China’s key regions. Humanities and Social Sciences Communications, 2024, 11(1): 1640-1640.

[3] Liu B, Song C, Wang Q, et al. Research on regional differences of China’s new energy vehicles promotion policies: A perspective of sales volume forecasting. Energy, 2022, 248: 123541.

[4] Liu Q, Wen X, Cao Q. Multi-objective development path evolution of new energy vehicle policy driven by big data: From the perspective of economic-ecological-social. Appl Energy, 2023, 341: 121065.

[5] Huang Xiaoqing, Zhang Dongliang, Zhang XiaoSong. Energy management of intelligent building based on deep reinforced learning. Alexandria Engineering Journal, 2021, 60(1): 1509-1517.

[6] Zeng B, Yin F, Wang J, et al. Prediction and rationality analysis of new energy vehicle sales in China with a novel intelligent buffer operator. Engineering Applications of Artificial Intelligence, 2025, 143, 110030-110030.

[7] Wang Z, Niu S, Hu S, et al. How to promote sustainable consumption and development of NEV? Decoding complex interrelationships in consumer requirements and design practice. Journal of Cleaner Production, 2025, 486, 144524-144524.

[8] Jiang Jiachen, Mei Yuxin, Pan Yaoyao, et al. Development evaluation and sales forecast of new energy vehicles. Journal of Taizhou University, 2024, 46(03): 9-15.