Multidimensional Analysis of Development Indicators for New Energy Vehicles: A Study Based on Principal Component Analysis and Grey Correlation Analysis Model

Jianya Zhao *, #, Libing Chen #, Xuhan Chen

Jinan University-University of Birmingham Joint Institute, Guangzhou, China, 511443

* Corresponding Author Email: 1223477371@qq.com

#These authors contributed equally.

Abstract. The recent rapid growth of the new energy vehicle(NEV)industry has established it as a pivotal force in sustainable development. Yet, challenges in technology, cost, market, and policy hinder its high-quality development. Understanding these factors is essential for policy-making, fostering innovation, and guiding investment. This paper study employed principal component analysis to examine primary indicators' impact on NEV development and a grey relational model to assess various secondary indicators' significance and correlation with industry progress. The study found that while policy measures in China have persistently hampered the industry, market factors have begun positively influencing it since 2019. Furthermore, NEV indicators increasingly shape their own development, with economic factors exerting relatively limited influence. Market and economic indicators like Per Capita GDP and Carbon Trading Price exhibit strong correlations, while indicators linked to NEV characteristics such as range, fuel prices, and vehicle prices closely drive industry advancement, particularly in China's NEV sector.

Keywords: New Energy Vehicles, Development Indicators, Principal Component Analysis, Grey Correlation Analysis.

1. Introduction

Global warming is increasing and threatening the sustainability of the earth through increasing frequency and intensity of natural disasters causing loss of lives, and capital and initiating economic instability[1]. In the context of energy structure transformation and environmental protection requirements, the transportation industry, as the third largest contributor to global greenhouse gas emissions and primary consumers of energy resources[2, 3], urgently needs to be transformed and upgraded. Therefore, developing new energy vehicles is gradually becoming the direction of the automobile industry[4], attracting world widespread attention.

In 2022, 14% of vehicles sold worldwide were electric vehicles in comparison to 9% in 2021 and 5% in 2020[5]. To some extent, the widespread application of electric vehicles has reduced carbon emissions and improved urban air quality[6, 7]. Due to the significant potential for carbon emission reduction of new energy vehicles, countries around the world are paying increasing attention to the development of the new energy vehicle industry and taking a series of measures to promote its growth. For instance, China emphasizes the vigorous development of green and low-carbon industries, with the new energy vehicle industry being one of the key focuses to achieve the goal of "peak carbon emissions and carbon neutrality".

Accordingly, despite the increasing sales volume of new energy vehicles, it is still imperative to explore the main factors affecting the development of new energy vehicles to ensure precise and effective support for the progress of new energy vehicles. Prior studies have already proposed methods to analyze partial influencing factors of the development of new energy vehicles. Through the evolutionary game model, Liu et al. analyzed the development of new energy vehicles based on subsidies and carbon taxes, exploring the evolutionary stable strategies between local governments and automobile manufacturers[8]. They found that constant subsidies and carbon tax policies are not conducive to the popularization of NEVs, and then subsidies and carbon tax policies applicable in the long term and short term are different. Li et al. presented that the sales of new energy vehicles and the construction of charging infrastructure have a mutually reinforcing and constraining relationship, forecasting that the ratio of sales of new energy vehicles to charging stations is expected to continue decreasing in the future, with an estimated ratio of 2:1 by 2025[9]. Ren identified that technology maturity, new energy vehicle technical standards, and research and development funding for new energy vehicles are the most important driving factors for promoting the sustainable development of the Chinese new energy vehicle industry[10].

The above discussion indicates that while there have been numerous studies on the influencing factors of new energy vehicle development, there is a limited amount of research that systematically and comprehensively analyzes the influencing factors of new energy vehicle development. Furthermore, existing researches predominantly focus on investigating the impact of individual factors on new energy vehicle development, without considering the interactions among these influencing factors. Therefore, this study aims to address this research gap by constructing a scientifically robust index system for the development of the new energy vehicle industry and employing principal component analysis and grey correlation analysis model for multidimensional analysis of new energy vehicle development to fill this literature void.

Principal Component Analysis serves as reducing data dimensionality and eliminating noise to reveal patterns and correlations among data. It effectively prevents information redundancy, thus enhancing the accuracy and reliability of model evaluation analysis. Furthermore, Grey Correlation Analysis is applied in diverse fields to deal with complex systems and exploring interrelationships across different domains.

Meng et al. proposed a method combining laser-induced fluorescence and principal component analysis to effectively identify microplastics in the ocean, utilizing Principal Component Analysis to differentiate different types and proportions of microplastic samples[11]. Zhu et al. conducted a study using Principal Component Analysis on the mineral and elemental composition of the Baiyun Obo deposit in northern China [12]. The ore samples were classified into three categories: bastnaesite REE ore, Fe-REE-F ore, and Nb-rich ore. This study revealed information on mineralization and alteration, providing an effective method for classifying the Baiyun Obo ores.

In terms of construction industry, research on permeable concrete has shown that pore structure is a key factor influencing the mechanical properties of pervious concrete[13]. Moreover, in the study of aeolian sand concrete, through Grey Correlation Analysis, Zhou and Dong suggested that by appropriate aggregate substitution and optimization of structural parameters, the pore structure of aeolian sand concrete can be optimized, thereby enhancing its mechanical properties and frost resistance[14]. As for energy sector, Zhang et al. utilized orthogonal experimental design and grey correlation analysis to investigate the effects of different conditions on the reactivity and yield of pyrolytic carbon and then identified that pyrolysis temperature is the most significant factor influencing carbon quality, while the pyrolysis atmosphere has a significant impact on the yield and reactivity of carbon[15]. In the medical field, research in the area of traditional Chinese medicine constitution has used Grey Relational Analysis to identify key factors affecting the Yang deficiency constitution and conduct quantitative analysis, offering recommendations for the prevention and treatment of diseases related to the Yang deficiency constitution in traditional Chinese medicine[16].

Based on previous research, it is known that the combined model of principal component analysis (PCA) and grey relational analysis (GRA) can comprehensively analyze various influencing factors, improve the model's interpretability, accuracy, and robustness. Therefore, this study is based on the development data of China's new energy vehicle industry from 2013 to 2022. By employing a relatively systematic and comprehensive index system for the development of the new energy vehicle industry, the study utilized a combined model of PCA and GRA to conduct multidimensional analysis of the development indicators of new energy vehicles.

Conducting a thorough analysis of the key factors influencing the development of new energy vehicles holds significant theoretical and practical significance for guiding industrial policy formulation, providing data support for policy-making, driving the transformation and upgrading of the transportation industry, and guiding capital investment.

2. Data Processing and Preparation

This paper searches extensively on journal databases like National Bureau of Statistics, China Automobile Association, etc for literature on new energy vehicles, collecting the most probable indicators influencing the development of new energy vehicles, establishing an indicator system, identifying primary indicators (policy indicators, market indicators, etc.), and determining secondary indicators such as fuel prices, number of charging stations, range of new energy vehicles, prices, per capita GDP, and more.The indicator system is shown in Table1.

Table. 1 Indicator System

S（Primary Indicators）	P（ Secondary Indicators）
Policy Indicators S1	New Energy Vehicle Subsidies P1
	Vehicle Restriction Policies P2
	Carbon Trading Prices P3
Market Indicators S2	Crude Oil Production P4
	Number of Charging Stations for New Energy Vehicles P5
	Private Car Ownership P6
	Fuel Prices P7
	Industrial Investment Amounts P8
Performance Metrics S3	Vehicle Resale Value P9
	New Energy Vehicle Prices P10
	New Energy Vehicle Range/Endurance P11
Economic Indicators S4	Per Capita GDP P12
	GDP P13
	CPI P14
	Per Capita Disposable Income P15

Outlier and Missing Value Treatment: This study conducts outlier detection and missing value imputation on the collected data. Subsequently, missing values are supplemented, and outliers are removed. Detailed examination results are depicted in the following figure1 :

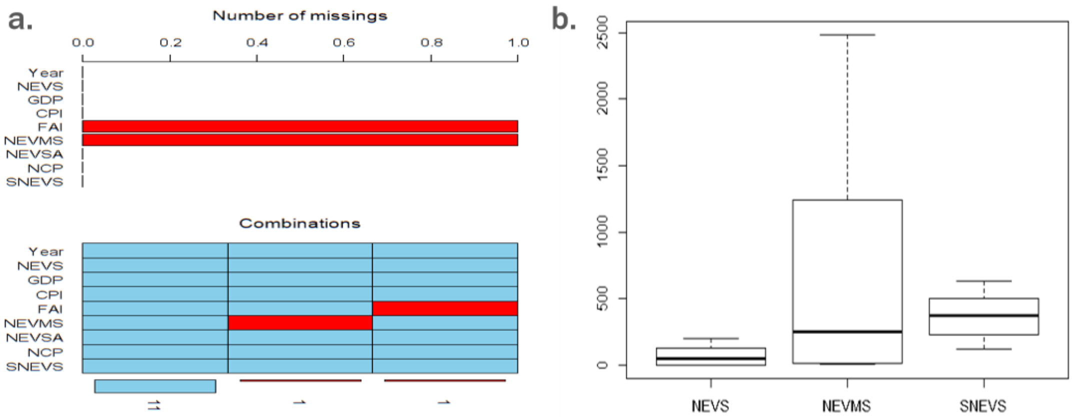

Figure. 1 Missing Values and Outlier Detection

score standardization for data processing: To eliminate the influence of different measurement scales, the data undergoes Z-score standardization. The calculation method is as follows: $X_{ij} = \frac{X_{ij}-\overline{X}_j}{S_j}$ $(i = 1,2,\cdots,n; j = 1,2,\cdots,p)$， In the equation,signifies the $\overline{X}_j = \frac{1}{n}\sum_{i=1}^{n} X_{ij}$ represents the sample mean of the j-th indicator， $S_j = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_{ij}-\overline{X}_j)^2}$ sample standard deviation of the j-th indicator.

3. Dimensionality Reduction of Influencing Factors in New Energy Vehicles

To explore the impact of two layers of different indicators on new energy vehicles, we conducted dimensionality reduction using principal component analysis. Initially standardizing the data, we applied principal component analysis, resulting in reduced dimensions. This yielded the impact scores and rankings of four primary indicators on the development of new energy vehicles.

Visual representation in the form of a flowchart illustrating the procedural steps of Principal Component Analysis (PCA) is shown in the figure2.

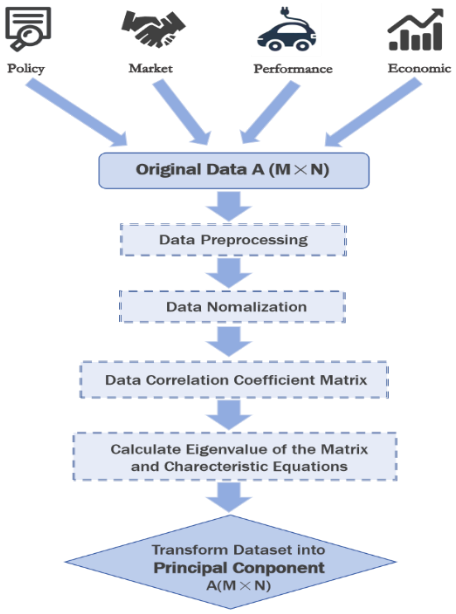

Figure. 2 Flow chart of PCA

The results of the principal component analysis model solution are shown in the Table2 :

Table. 2 Results of the principal component analysis

year	S1	S2	S3	S4
2013	1.34	-1.11	4.2	-1.60
2014	1.25	-1.01	4.5	-1.11
2015	0.63	-0.79	4.8	-0.7
2016	-0.07	-0.60	5.1	0.00
2017	-0.16	-0.32	5.4	0.28
2018	-0.16	-0.01	5.7	0.44
2019	-0.22	0.34	6	0.53
2020	-0.87	1.07	6.3	0.53
2021	-0.87	0.93	6.6	0.93
2022	-0.86	1.51	6.9	0.70

(Where S1 represents policy indicators, S2 represents market indicators, S3 represents new energy vehicle-specific indicators, and S4 represents economic indicators.)

The results from the principal component analysis reveal distinctive trends among these factors. In terms of policy indicators, from 2013 to 2022, the values consistently remained negative, suggesting that national policies related to new energy vehicles, including subsidies and carbon trading prices, have had an adverse impact on the sales of new energy vehicles.

Regarding market indicators (such as the number of charging stations and industrial investment), before 2019, these indicators exhibited negative values, indicating an unfavorable effect of market factors on the development of new energy vehicles. After 2019, these indicators turned positive, signifying a shift in market factors toward a positive impact on new energy vehicles.

Analyzing the indicators specific to new energy vehicles themselves, these indicators have consistently shown positive values, displaying an increasing trend over the years. This suggests significant improvements in aspects like pricing and range, which have significantly boosted the sales of new energy vehicles.

Regarding economic factors, the transition from negative to positive started in 2016. However, in recent years, the positive values for economic indicators have been relatively small, indicating that factors such as GDP and per capita disposable income have had limited influence on the development of new energy vehicles.

4. Assessment of Key Factor Influences on New Energy Vehicles

To delve deeper into understanding how different factors within the secondary indicators influence the advancement of new energy vehicles, we employed correlation analysis and conflict analysis to assess their relationships. Following this, we constructed a grey relational analysis model that integrates these correlation coefficients into its formula, quantifying the significance of each factor concerning the development of new energy vehicles.

Next, utilizing correlation coefficients as the measure, calculated by the following formula

$$r_s = 1 - \frac{6\sum_{i=1}^{n} {d_i}^2}{n(n^2-1)} \tag{1}$$

Where d_i represents the rank difference between X_i and Y_i ,n stands for the sample size.

In the case of a large sample, the statistic follows a normalx distribution $r_s\sqrt{n-1} \sim N(0,1)$。 Therefore, the hypothesis test is as follows：

$H_0: r_s = 0$;

$$H_1: r_s \neq 0$$

At a 95% confidence level, if the p value is less than 0.05, the null hypothesis is rejected; otherwise, it cannot be rejected.

The conflict coefficient is represented by the correlation coefficient formula.

$$R_j = \sum_{i=1}^{p}(1 - r_{ij}) \tag{2}$$

The correlation coefficient R_{ij} between evaluation indicators i and j signifies that a larger coefficient indicates a higher level of redundancy between the evaluation content reflected by these indicators. This redundancy weakens the evaluation strength of the respective indicator, suggesting a need to reduce its assigned weight.

Ultimately, the p-values among the indicators are all less than 0.05, meeting the assumption and passing the correlation test. The part results of the correlation analysis for a randomly selected few indicators are shown as Table3:

Table. 3 Correlation Analysis

	P1	P5	P6	P10	P12	P15
P1	1.00	0.77	0.62	0.80	0.94	0.79
P5	0.77	1.00	0.94	0.69	0.84	0.52
P6	0.62	0.94	1.00	0.50	0.69	0.34
P10	0.80	0.69	0.50	1.00	0.80	0.82
P12	0.94	0.84	0.69	0.80	1.00	0.77
P15	079	0.52	0.34	0.82	0.77	1.00

The formulas for calculating the Grey Relational Coefficient and for computing the Grey Relational Coefficient between corresponding elements of each sub-sequence and the main sequence are as follows:

$$\begin{gathered}\gamma(x_0(k),x_i(k))=\frac{\Delta\min+\rho\Delta\max}{\Delta_{ik}+\rho\Delta\max}\\ \Delta\min=\min_i\min_k|x_0(k),x_i(k)|\\ \Delta\max=\max_i\max_k|x_0(k),x_i(k)|\\ \Delta_{ik}=|x_0(k),x_i(k)|\end{gathered}\tag{3}$$

The resolution coefficient ranges ρ between (0, 1), with smaller values indicating larger differences between Grey Relational Coefficients.

Here, the indicators are denoted by $(x_0,\ x_1,x_2,\ldots,x_m)$, the new energy vehicle sales, represented as x0

The relevance reflects the correlation relationship between each manipulated object and the reference sequence. Calculate the weighted average of the Grey Relational Coefficients between the secondary indicators and the new energy vehicle sales, as shown in the formula:

$$r_{oi}=\frac{1}{m}\sum_{k=1}^{m}W_k\zeta_i(k)\tag{4}$$

Based on the correlation analysis and conflict analysis of various indicators, incorporated into the Grey Relational Analysis formula, the part results for randomly selected indicators are shown as table4:

Table. 4 Grey Relation Degree Results

Variables	Relevance	Variables	Relevance
P12	0.87	P15	0.76
P5	0.86	P6	0.67
P11	0.77	P4	0.65
P1	0.76	P10	0.65
P3	0.76	P2	0.45
P7	0.76	P14	0.12

Analyzing the aforementioned relevance analysis reveals that among the various indicators, per capita GDP demonstrates the highest relevance, reaching as high as 0.87, while the Consumer Price Index exhibits the lowest relevance, as low as 0.12. Considering 0.5 as a threshold for indicator relevance, values below 0.5 indicate a relatively minor impact, whereas values above 0.5 suggest a significant impact. Thus, it can be inferred that both the vehicle restriction policy and the Consumer Price Index have relevance scores below 0.5, indicating a relatively weaker association with the development of new energy vehicles. Conversely, indicators such as the number of charging stations and the range of new energy vehicles exhibit a stronger association with the development of new energy vehicles.

5. Conclusions

This study utilizes principal component analysis to uncover the impact of various primary indicators on the development of new energy vehicles over the past decade. The findings reveal that policy measures in China have consistently exerted a negative influence on the new energy vehicle industry, while market factors have gradually shifted from hindering to promoting its development since 2019. Additionally, the positive impact of new energy vehicle indicators on their own development has become increasingly evident in recent years. Using the grey relational model, this study delves into the interrelation among various secondary indicators and the development of the new energy vehicle industry. The findings highlight a close correlation between certain self-factors of new energy vehicles, significantly impacting market acceptance and competitiveness. Notably, the industry faces new opportunities and challenges with the rise in per capita GDP levels and fluctuations in carbon trading prices.

In summary, this study proposes a novel perspective and methodology for evaluating the development of new energy vehicles. Through comprehensive employment of quantitative analysis methods such as principal component analysis and the grey relational model, it not only reveals the impact mechanisms of various primary indicators on new energy vehicle development but also conducts in-depth analyses of industry development from multiple dimensions including policy, market, economy, and technology. The practical application of this research framework has demonstrated its feasibility and effectiveness, providing valuable references and insights for related fields of research and practice. Moving forward, further refinement of this framework and conducting research in conjunction with real-world scenarios will advance the sustainable development of the new energy vehicle industry and promote the green transformation of the economy and society.

References

[1] Le, H., et al., *Dynamic Interlinkages between Carbon Risk and Volatility of Green and Renewable Energy: A Wavelet Analysis.* Available at SSRN 4359925, 2023.

[2] Sun, L., et al., *Reducing energy consumption and pollution in the urban transportation sector: A review of policies and regulations in Beijing.* Journal of cleaner production, 2021. **285**: p. 125339.

[3] Hassouna, F.M.A. and K. Al-Sahili, *Environmental impact assessment of the transportation sector and hybrid vehicle implications in Palestine.* Sustainability, 2020. **12**(19): p. 7878.

[4] Kendall, M., *Fuel cell development for New Energy Vehicles (NEVs) and clean air in China.* Progress in Natural Science: Materials International, 2018. **28**(2): p. 113-120.

[5] IEA, J., *Global EV Outlook 2022: Tech. Rep.* 2021.

[6] Solaymani, S., *CO2 emissions patterns in 7 top carbon emitter economies: The case of transport sector.* Energy, 2019. **168**: p. 989-1001.

[7] Dixon, J., et al., *Scheduling electric vehicle charging to minimise carbon emissions and wind curtailment.* Renewable Energy, 2020. **161**: p. 1072-1091.

[8] Liu, C., et al., *A time-delayed evolutionary game analysis of new energy vehicles development considering subsidy and carbon tax.* Heliyon, 2024.

[9] Li, B., X. Lv, and J. Chen, *Demand and supply gap analysis of Chinese new energy vehicle charging infrastructure: Based on CNN-LSTM prediction model.* Renewable Energy, 2024. **220**: p. 119618.

[10] Ren, J., *New energy vehicle in China for sustainable development: Analysis of success factors and strategic implications.* Transportation Research Part D: Transport and Environment, 2018. **59**: p. 268-288.

[11] Meng, X., et al., *Identification of marine microplastics based on laser-induced fluorescence and principal component analysis.* Journal of Hazardous Materials, 2024. **465**: p. 133352.

[12] Zhu, X.-F., et al., *Principal component analysis of mineral and element composition of ores from the Bayan Obo Nb-Fe-REE deposit: Implication for mineralization process and ore classification.* Ore Geology Reviews, 2024: p. 105972.

[13] Song, W., M. Zhang, and H. Wu, *Gray correlation analysis between mechanical performance and pore characteristics of permeable concrete.* Journal of Building Engineering, 2024: p. 108793.

[14] Zhou, M. and W. Dong. *Grey correlation analysis of macro-and micro-scale properties of aeolian sand concrete under the salt freezing effect*. in *Structures*. 2023. Elsevier.

[15] Zhang, B., et al., *Assessment of the impact of pyrolysis conditions on char reactivity through orthogonal experimental-based grey relational analysis.* Journal of Analytical and Applied Pyrolysis, 2024: p. 106426.

[16] Yue, L., et al., *grey correlation analysis on influencing factors of Yang deficiency constitution.* Digital Chinese Medicine, 2023. **6**(2): p. 151-159.