Mathematical Modeling and Algorithm Application
ISSN: 3006-0842 | Vol. 2, No. 1, 2024

Forecasting Development Trends and Analysing Influencing Factors of New Energy Electric Vehicles in China: Based on Multi-level Analysis and Time Series Modelling

Zihan Wei*

College of Mathematics and Physics, Chengdu university of Technology, Chengdu, 610059, China.
* **Corresponding author**: Zihan Wei (Email: 202220010110@stu.cdut.edu.cn.)

Abstract: This study comprehensively analyses the current development status and future trends of China's new energy electric vehicle industry, and constructs a multilevel analysis model by integrating data from multiple sources, systematically assessing the influence weights of factors such as policy orientation, economic development, technological innovation, infrastructure construction and environmental protection on the development of new energy vehicles. On this basis, ARIMA model, Holt's index smoothing model and multiple linear regression method are used to quantitatively forecast the development of China's new energy electric vehicle market in the next ten years. Meanwhile, this study explores the potential impact of the popularity of new energy vehicles on the global traditional energy vehicle industry pattern, and analyses the effectiveness of domestic and international government support policies. The results of the study show that policy promotion and technological advancement are the key drivers of the rapid growth of China's new energy vehicle market, and it is expected that the penetration of new energy vehicles will continue to profoundly change the ecology of the automotive industry in the next ten years, and exert an important thrust on the transformation of the global energy structure.

Keywords: Multi-level Analysis; Time Series Model; New Energy Electric Vehicles.

1. Introduction

With the increasingly severe global energy crisis and environmental pollution, new energy electric vehicles, as a representative of strategic emerging industries, are gradually becoming the focus of the transformation and upgrading of the global automotive industry [1]. As the world's largest automobile market, the development of new energy vehicles in China not only concerns national energy security and environmental sustainability, but also has a far-reaching impact on the global automobile industry [2]. In recent years, the Chinese government has introduced a series of supportive policies aimed at promoting the rapid development of the new energy vehicle industry, while accelerated technology iteration and infrastructure development have also provided strong support for the industry. However, the development of new energy vehicles still faces many uncertainties, including challenges such as fluctuations in market demand, breakthroughs in technological bottlenecks, and intensified international competition. Therefore, accurately predicting the development trend of new energy electric vehicles and identifying and evaluating the key influencing factors have important practical and theoretical values for policy makers, industry participants and researchers. Through scientific modelling and empirical analysis, this study aims to provide insights and strategic recommendations for the development path of new energy vehicles in China and the world.

2. Data processing

2.1. Methodological note on the establishment of impact factor indicators

The analysis of the influencing factors of China's new energy electric vehicle industry from a macro perspective is diverse, so this paper is based on the PEST analysis model[3] and thinking about it, mainly from the macro factors such as policy, economy, technology and other macro factors to examine the indicators of influencing factors in each issue.

V (Sales volume): represents the annual sales volume of new energy electric vehicles, and reacts to the development trend of electric vehicles through the change of sales volume.

P (Policy government support): refers to the government's policy support for new energy electric vehicles, including purchase subsidies and tax reduction policies.

E (Economic growth): the impact of social-economic growth on the development of the automobile market, including the growth rate of national per capital income.

S (Society): mainly includes the degree of public attention to environmental issues and environmental awareness, etc.

As new energy electric vehicles have problems such as short range, long charging time and difficulty in finding charging piles, we also need to consider the following factors:

T (Technology): Consider the impact of technological innovations in battery range, charging modes, etc. on the new energy vehicle market.

I (Infrastructure) : examines factors such as the number and distribution of charging piles.

2.2. Data cleaning

2.2.1. Treatment of missing values

Given the large span of years examined and the difficulty in quantifying the selected impact factors, there are some gaps in the data collected. In order to ensure that there is available data for each impact factor indicator, this paper uses time series regression interpolation to deal with the missing data [4].

2.2.2. Standardized processing

Multiple linear regression equations have been developed to address a number of issues raised in the paper. In order to facilitate the model construction, the collected data have been preprocessed by standardization of standard deviation method [5]. The specific formula is as follows:

$$Z_{ij} = \frac{x_{ij}-\mu_i}{\sigma_i} \quad (1)$$

Where represents the specific value of the particular impact factor indicator, represents the mean, and represents the standard deviation.

This section illustrates the key steps in the data preparation process, focusing on data cleaning. Dealing with missing values and standardization is fundamental to ensuring the accuracy of analytical results when faced with large datasets, especially when they span many years and contain variables that are difficult to quantify. Dealing with missing values through time series regression interpolation, as well as standardizing the data, can effectively improve the plausibility of model building and the accuracy of predictions.

3. Modeling and solving

3.1. Hierarchical analysis method AHP modeling

Establishment of hierarchical model: the problem is decomposed into three hierarchical structures, the top layer is the target layer A, selecting the appropriate important indicators affecting the development of new energy electric vehicles; the middle layer is the criterion layer, including policy (B1), economy (B2), technology (B3), infrastructure (B4), and the environment (B5); and the bottom layer is the scenario layer, which means that the ten influencing factors of the vehicle purchase subsidy (C1), the tax exemption policy (C2), GDP growth rate (C3), per capita income level (C4), battery technology (C5), charging technology (C6), charging station distribution (C7), the degree of charging facilities (C8), carbon emissions (C9), air quality (C10).

3.2. Model solution

3.2.1. Constructing the judgment matrix M-C

The process of constructing the judgement matrix M-C involves comparing factors from adjacent tiers two by two and using a scale of 1-9 to reflect the relative importance between them. This scale is used for quantitative comparisons, where a larger number indicates that one factor is more important relative to another. Table 1 demonstrates the importance level scale used to construct the judgement matrix, which is a key step in the overall hierarchical analysis method (AHP) that ensures that the decision maker's subjective judgements are systematic and consistent. By constructing the judgement matrix in this way, we can transform a complex multi-criteria decision-making problem into a mathematical model, which further analyses and determines the weights of the factors to guide the decision-making process.

Table 1. Judgment Matrix A-B

A	B1	B2	B3	B4	B5	W1 Arithmetic mean	W2 geometric mean	W3 eigenvalue method	W average
B1	1	3	5	7	9	0.5028	0.51	0.5128	0.51
B2	0.33	1	3	5	7	0.2602	0.02638	0.2615	0.26
B3	0.20	0.33	1	3	5	0.1344	0.1296	0.1290	0.13
B4	0.14	0.20	0.33	1	3	0.0678	0.0636	0.0632	0.06
B5	0.11	0.14	0.20	0.33	1	0.0348	0.0329	0.0333	0.03
λ_{max}=5.2375; CR=0.0530<0.1 (consistency test passed)									

The following tables are all consistency matrices, so there is no need to perform consistency tests.

Table 2. Judgment Matrix B1-C

B1	C1	C2	W
C1	1	3	0.75
C2	0.33	1	0.25

Table 3. Judgment Matrix B2-C

B2	C3	C4	W
C3	1	0.25	0.2
C4	4	1	0.8

Table 4. Judgment Matrix B3-C

B3	C5	C6	W
C5	1	4	0.8
C6	0.25	1	0.2

Table 5. Judgment Matrix B4-C

B4	C5	C6	W
C7	1	4	0.75
C8	0.25	1	0.25

Table 6. Judgment Matrix B5-C

B5	C9	C10	W
C9	1	0.2	0.17
C10	5	1	0.83

3.2.2. Reach a verdict

Using MATLAB software, the combined weights of the program layer on the target layer were further calculated and ranked, i.e., the influence weights of the 10 indicators on the development of new energy electric vehicles in China were, in descending order, as follows: C1(0.38), C4(0.21), C2(0.13), C5(0.10), C3(0.05), C7(0.05), C6(0.03), C8(0.02), C10(0.02), and C9(0.01). Therefore, we conclude that the top three factors that have the greatest impact on the development of new energy electric vehicles in China are: C1 government subsidies, C4 per capita income level, C2 tax exemption policy, and C3 tax exemption policy.

3.3. Predictive modelling

3.3.1. Selection of key indicators

Sales volume (V): indicates the annual sales volume of new energy electric vehicles, which is a key indicator.

Government support (P): Measures the government's policy support for new energy electric vehicles, including purchase subsidies, tax cuts and so on.

Economic growth (E): Considers the impact of the overall growth of a country's economy on the automobile market.

Technological innovation (T): Consider battery technology, charging technology and other aspects of the promotion of new energy vehicles.

Infrastructure development (I): measures the extent to which charging infrastructure has been developed, including the number and distribution of charging stations.

3.3.2. Indicator data collection

By examining government support using the amount of subsidies (billion yuan), economic growth using the GDP growth rate, technological innovation using the number of patents published each year, and infrastructure construction using the number of charging stations several sets of data, the historical data of various indicators of new energy electric vehicles (which have been standardized and processed) were collected and made into the following table:

Table 7. Historical Data for Indicators

Vintages	Level of government support	Economic growth	Technological innovation	Infrastructure development	Sales volume
2010	2.7	0.021806641	1098	5154	48854
2011	6.6	0.037280669	1307	4915	52086
2012	9.1	0.026719076	1463	6606	54642
2013	6.11	0.031800453	1165	6216	52922
2014	8.9	0.016446319	1723	6804	55704
2015	7.8	0.035331792	1751	6447	57608
2016	9.8	0.033759582	1771	7735	57239
2017	12.8	0.048445346	1938	7887	61610
2018	11.4	0.06217894	2487	7499	61893
2019	9.8	0.076771154	2223	8481	65758
2020	11.0	0.084138791	2291	8621	64994
2021	13.1	0.069600388	2735	9285	67332
2022	13.8	0.072182654	3248	9490	67319
2023	13.3	0.088175163	3005	9905	69375

3.3.3. Development of time-series pre-future impact factor indicators

(1) ARIMA model indicator factor determination

For this question we use ARIMA model to predict the indicator factors affecting the sales volume of automobiles.

Policy Support Forecast: Forecasting future government support for new energy vehicles.

Economic forecasting: forecasting future national economic trends.

3. Technological innovation prediction: predict the future technological trends of new energy electric vehicles, taking into account battery technology innovation, etc.

4. Forecast of infrastructure construction: Forecast of future development of charging infrastructure construction.

(2) ARIMA model description

ARIMA forecasting model through time series analysis, combined with the specific data of previous years can predict the future data of a specific variable, and its model is simple, only need endogenous variables to participate in the prediction without the need for external factor variables, can be used to describe the past and predict the future [6]. ARIMA model has a high adaptability in approximating many smooth processes, and therefore is commonly used in linear forecasting.

The mathematical expression of ARIMA model can be expressed as:

$$Y_t = \mathrm{c} + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_P Y_{t-P} + \theta_1 \epsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2)$$

(3)Model building

We used the SPSS expert modeler for modeling. SPSS Expert Modeler: The Expert Modeler automatically finds the best-fitting model for each dependent series. If dependent (predictor) variables are specified, the Expert Modeler selects those models for what is in the ARIMA model that have a statistically significant relationship with that dependent series. When appropriate, the model variables are transformed using difference and/or square root or natural logarithm transformations. By default, the expert modeler considers both exponential smoothing method models and ARIMA models. Automatic detection of outliers can also be specified. By analyzing smoothness, periodicity, autocorrelation, and partial autocorrelation, the government support model is best modeled using the ARIMA (0,0) model, economic growth is modeled using ARIMA (0, 1,0), and infrastructure development and technological innovation are modeled best when fitted and forecasted using the Holt exponential smoothing approach.

The sample autocorrelation coefficients (ACF) and partial autocorrelation coefficients (PACF) of the four series graphically show a trailing nature and none of them are significantly different from 0. The time series can be fully recognized by the model. The smoothed R-square is close to 1, indicating that the linear model is well fitted. If outliers are found: SPSS will use the mean of the series as a substitute for the outliers found.

3.3.4. Multiple linear regression model fitting for electric vehicle sales volume

The mathematical model between the development of new energy electric vehicles and key indicators was established using a multiple linear regression model.

Model Representation:

$$\mathrm{V} = \beta_0 + \beta_0 \mathrm{P} + \beta_0\ \mathrm{E} + \beta_3\ \mathrm{T} + \beta_4\ \mathrm{I} + \varepsilon \quad (3)$$

V denotes the annual sales volume of new energy electric vehicles, denotes the measure of government policy support for new energy electric vehicles, E denotes economic growth, T denotes the value of technological innovation, I denotes the degree of infrastructure development, and The coefficients of the coefficients are indicated. is the error.

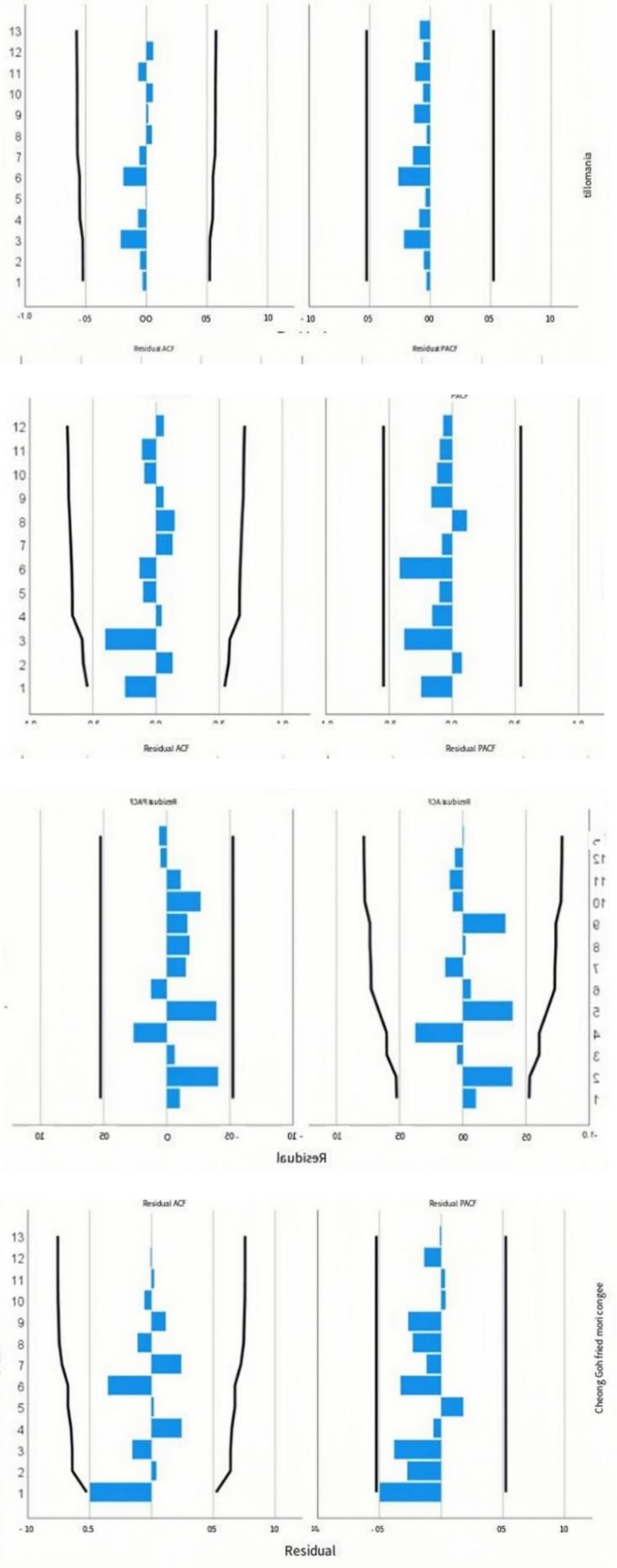


Figure 1. ACF and PACF images of the four metrics

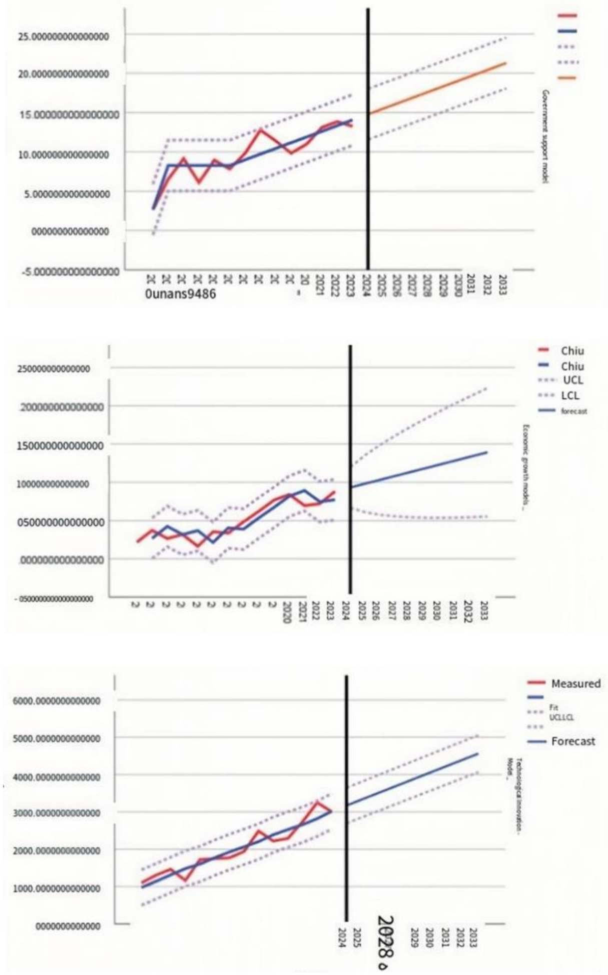


Figure 2. Predictive images of the four indicators

Model description

		Model type
Model ID	Technological innovation Model _1	Holt

Model statistics

Models	Number of predictor variables	Model fit statistics	Young-box Q(18) stationary R square Statistics	DF	Salience	Number of outliers
Technological innovation - Model_1	0	816.	.	0	.	0

Exponential smoothing method model parameters

			Estimation	Standard error	t	Significance
Model Technology Innovation - Model _1	No conversion	Alpha (horizontal)	099.	107.	921.	375.
		Gamma (Trend)	2.599 e-5	124.	000.	1.000

Figure 3. Model description, statistics and parameters for the four indicators

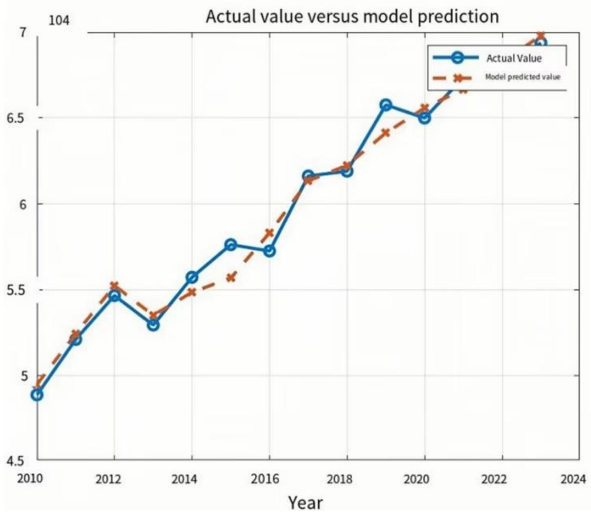
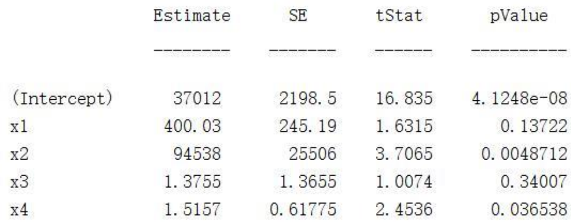


Figure 4. Comparison of actual values with model predictions

By visualizing the comparison of the curves of the predicted and actual values, we can intuitively understand the fit of the model and the accuracy of the prediction, and using the SPSS tool we conclude that the model is well fitted.

Estimated coefficients:



	Estimate	SE	tStat	pValue
	--------	-------	------	----------
(Intercept)	37012	2198.5	16.835	4.1248e-08
x1	400.03	245.19	1.6315	0.13722
x2	94538	25506	3.7065	0.0048712
x3	1.3755	1.3655	1.0074	0.34007
x4	1.5157	0.61775	2.4536	0.036538

Figure 5. Model Estimated Coefficients

Number of observations: 14

Degree of freedom of error:9

Root Mean Square Error: $1.13e + 03$

R-square: 0.979

Adjustment R-square 0.97

F-statistic (constant model): 107

$p - \text{value} = 1.38e - 07$

By looking at the data, it can be seen that the is 0.979 and the adjusted is 0.97 , indicating that the model fits better; the significance level of the F-test is < 0.05, so the test passes; the root mean square error formula is $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$, where is the number of samples and is the observed value, and is the model predicted value. Multiple linear regression equation: $V = 37012 + 400.03P + 94538E + 1.3755\,T + 1.5157I + \varepsilon$

3.3.5. Multiple Linear Regression to Predict Electric Vehicle Sales

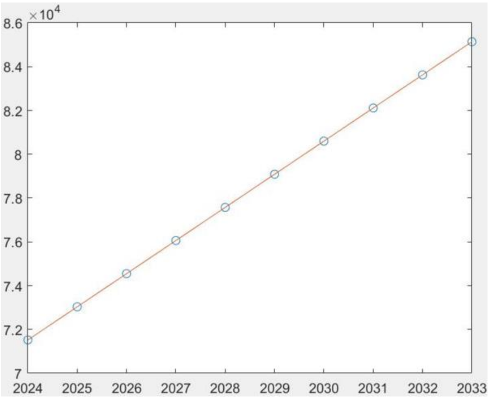


Figure 6. New Energy Electric Vehicle Sales Forecast

Each point represents the sales volume of new energy vehicles in each year of the next ten years, and this straight line represents the linear image derived from the multiple linear regression equation.

3.3.6. Reach a verdict

Through the establishment of time series model, respectively predicted the development of the four influencing factors indicators, and then use multiple linear regression prediction so as to arrive at the sales of new energy electric vehicles in the next 10 years as Table 8. Through the SPSS prediction image we can see that the sales of China's new energy electric vehicles in the next ten years is rising, and the future development of the future development continues to be good.

Table 8. Sales of New Energy Electric Vehicles in the Next 10 Years

VINTAGES	SALES VOLUME
2024	71521.08343
2025	73034.17584
2026	74547.26825
2027	76060.36067
2028	77573.45308
2029	79086.54549
2030	80599.63791
2031	82112.73032

4. Conclusion

This paper describes and predicts the development trend of new energy electric vehicles in China. By collecting and integrating data related to China's new energy electric vehicles, we constructed a hierarchical analytical model to analyse the impact of policy, economic, technological, infrastructural and environmental factors on the development of China's new energy electric vehicles. In order to depict and predict the development of new energy electric vehicles in China in the next ten years, we constructed an ARIMA model with a Holt exponential smoothing model, and predicted it by fitting a multiple linear regression model. By observing the car sales and ownership function model, we analysed the impact of new energy electric vehicles on the traditional energy automobile industry in China and globally. In addition, we evaluated the role of Chinese government support policies and foreign government support on the development of new energy electric vehicles in China by building a multiple linear regression model.

References

[1] Shor B, Bafumi J, Keele L, et al. A Bayesian multilevel modeling approach to time-series cross-sectional data[J]. Political Analysis, 2007, 15(2): 165-181.

[2] Kusano K, Kemmelmeier M. Multi-level modelling of time-series cross-sectional data reveals the dynamic interaction between ecological threats and democratic development[J]. Royal Society Open Science, 2020, 7(3): 191804.

[3] Ma J, Li T, Li G. Comparison of representative method for time series prediction[C]//2006 International Conference on Mechatronics and Automation. IEEE, 2006: 2448-2453.

[4] Ismail L, Materwala H, Dankar F. Machine Learning and Deep Learning Data-Driven Residential Load Multi-Level Forecasting with Univariate and Multivariate Time Series Models Towards Sustainable Smart Homes[J]. IEEE Access, 2024.

[5] Dang X H, Shah S Y, Zerfos P. seq2graph: discovering dynamic dependencies from multivariate time series with multi-level attention[J]. arXiv preprint arXiv:1812.04448, 2018.

[6] Al-Hajj R, Assi A, Fouad M M. Multi-level stacking of long short term memory recurrent models for time series forecasting of solar radiation[C]//2021 10th International Conference on Renewable Energy Research and Application (ICRERA). IEEE, 2021: 71-76.