ISSN 2522-3488 Vol. 9, Issue 2: 51-56, DOI: 10.25236/IJNDES.2025.090209

Prediction Study of Wind Power Generation Power Based on Arima Model

Jipeng Han[1,#], Siyuan Hao[2,#], Shijie Chen[2]

[1]College of Mechanical and Electrical Engineering, Sichuan Agricultural University, Ya'an, China, 625099
[2]Department of Physics, Xinzhou Normal University, Xinzhou, China, 034000
[#]These Authors Contributed Equally.

Abstract: *With the rapid development of renewable energy, the share of wind and solar energy in the energy sector is gradually increasing. However, the volatility of their generating power poses challenges to the stable operation of power grids. Wind and solar farms pose new challenges to the stable operation of the grid due to their volatility. The aim of this study is to predict the volatility of wind (wind) power generation using the Arima model so as to enhance the grid's ability to accommodate renewable resources. In this study, historical power generation data from wind farms and solar power plants in a certain region are used, and the triple standard deviation method is applied to deal with outliers, combined with the Lagrangian method to fill in the missing values, and the Arima model is built to make predictions. The results show that the model can better capture the fluctuation trend of power generation. Predictions are made 300 seconds before a significant increase or 120 seconds before a significant decrease in the PV power generation rate value, and the prediction accuracy is obtained by comparing the predicted value with the true value. The predictions show that the model has a high prediction accuracy. The study shows that the Arima model is suitable for short-term prediction of wind power generation, which can provide effective data support for grid scheduling and help to improve the stability and reliability of new energy power generation systems.*

Keywords: *Wind Power Generation, Power Prediction, Arima Model*

1. Introduction

Against the backdrop of global advocacy for sustainable development, wind and solar power, as clean and renewable sources of energy, have become increasingly important in the electricity supply system and delivery system, and are rapidly becoming an important part of the renewable energy sector. However, while wind power and photovoltaic power generation are developing rapidly, they are also facing challenges. For example, the intermittent and fluctuating problems of wind and solar power generation have seriously affected the stable operation of the power grid and the reliability of power supply. Many scholars at home and abroad have carried out research on this issue, and some of them have used physical models to predict the power of wind power generation. For example, Li Tian used a fluid dynamics (CFD) model to predict the short-term wind power in wind farms [1], M. Lydia used the power curve method to establish the wind turbine power curve, and accurately predicted the wind power [2], and there are also studies that use the traditional statistical methods, such as regression analysis, for example, Mantas Marčiukaitis combined regression method with the power curve, which can not depend on the initial data set to predict the power of wind power. The power generation can be predicted without relying on the initial data set [3], and Sylwester Borowski predicts wind power generation and balanced energy systems based on linear regression models [4]. In addition, P. Sivaprakash built a decision tree model to analyse a large amount of complex environmental data and accurately predicted the hourly wind power generation for the next 24 hours [5], and Zahraa Tarek optimized the LSTM using the SFS-PSO model to predict the wind power values in the R-mode equal to 99.99% [6]. Wu, Linhan et al. adopted the TCN-Transformer model to verify that this model is more effective in processing time series data and large-scale datasets, and demonstrated high accuracy and generalization ability in the photovoltaic power prediction task [7].

In view of this, this study introduces the Arima model to predict the power fluctuation of wind power generation. This model is more convenient for prediction compared with traditional methods, has lower requirements for data synthesis, a shorter model response time, higher prediction accuracy

ISSN 2522-3488 Vol. 9, Issue 2: 51-56, DOI: 10.25236/IJNDES.2025.090209

(RMSE=0.0236, MAPE=0.79%), and superior performance. It is expected to provide an efficient and accurate research scheme for the power prediction of wind power generation, thereby enhancing the power grid's capacity to consume new energy and promoting the wide application of renewable energy.

2. Data Processing

2.1 Source of data

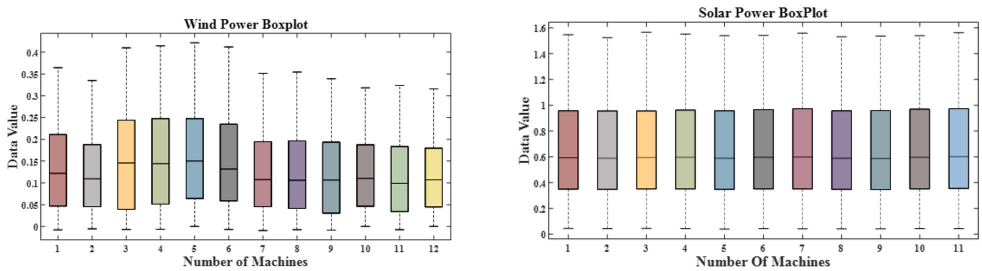
This wind farm data set was provided by wpd windmanager GmbH, Bremen, Germany. The sampling frequency was 1Hz, the time span was one month (29.7.2009-30.8.2009), the size of the wind farm was 4x4 km^2, it contained twelve turbines and the rated power was 2.xMW. The rated power is 2.xMW.This solar data set was recorded on a platform on the roof of the university of Oldenburg, Germany.Eleven sensors were used, the data were collected for a period of one month (1993.6.1-1993.6.30 This ensures that the data adequately reflect the fluctuating characteristics of wind power generation.

2.2 Methodology

In order to study the fluctuation patterns of wind and solar farms by analysing the generation data found above. Firstly, the outliers can be detected by the triple standard deviation method and combined with the Lagrangian interpolation method to fill in the missing data to ensure the completeness and accuracy of the data. Secondly, the power change trend can be analysed through visualisation to provide a basis for quantitative study of the fluctuation pattern.

2.3 Triple standard deviation method for outliers

In this paper, the triple standard deviation method [8] is used to detect and process the data. The standard deviation of the sample is calculated first, and three times of the standard deviation is used as the allowable range of the sample, and then the extreme values are eliminated to make the sample variance more accurate. The results of processing are shown in Figure 1, and no outliers exist in both data sets after processing.

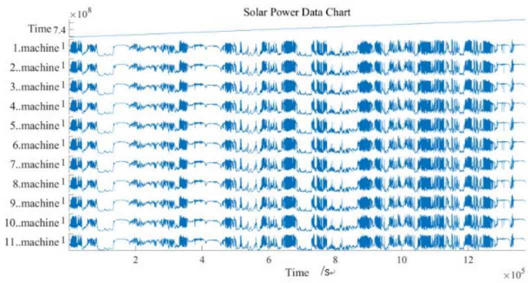


(a) Solor Power Boxplot *(b) Wind Power Boxplot*

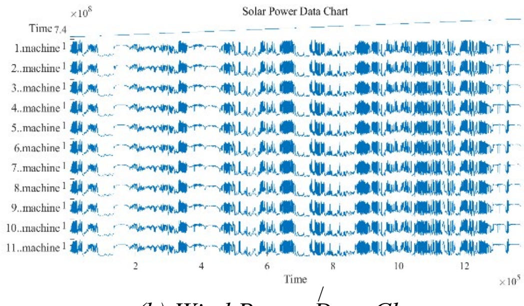
Figure 1 Boxplot

2.4 Lagrange interpolation for missing values

For missing values in the data, Lagrange interpolation [9] is used in this paper. A polynomial function is constructed by taking the given data points so that the function takes the given value at all the given data points. The processing results are shown in Figure 2.



(a) Wind Power Data Chart



(b) Wind Power Data Chart

Figure 2 Data Chart

The output is processed to output the image of a polynomial function that passes exactly through a number of known points in the two-dimensional plane, thus taking the observed values at each observation point.

3. Models

3.1 Forecasting of photovoltaic power generation by the ARIMA model

The value of K in the wind power curve [10] is an important parameter to measure the performance of wind speed, which can also be used to measure the performance of different WTGs, and is calculated as the formula (1) , the higher the value of K, the better the performance of WTGs in generating electricity, and the more electricity is generated under the same situation.

$$k = |p - q| / q \tag{1}$$

Firstly, the value is calculated using the formula. According to the calculation part of the value is too large. In order to facilitate the subsequent prediction and improve the prediction accuracy, this paper assigns a value of 20 instead of the infinity value for calculation. At the same time, due to the large amount of sensor data, the missing values are eliminated for subsequent calculations. The resulting output image is shown in Figure 3.

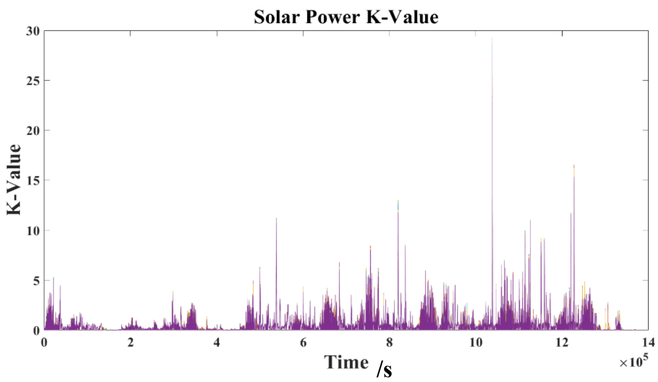


Figure 3 Solar Power K-Value

Table 1 Forecast data

Number of data rows	Predicted value	Actual value	Prediction Error
316362	1.7136	1.7022	0.006697
499688	5.4279	5.4548	0.005502
545637	1.6528	1.6745	0.012959
683762	5.9037	5.9473	0.007331
819073	2.3726	2.3640	0.003638
911439	1.0737	1.0928	0.017478
1046812	4.2103	4.2026	0.001832

Secondly, the selection of t-value is carried out. In this paper, the 80th percentile of the data is selected as the value, and the values of 11 generators are calculated and averaged to obtain the average fluctuation amplitude. After calculation, the value is 0.5955. Again, each point exceeding the t-value is predicted one by one. Finally, the prediction accuracy of each point is averaged to obtain the prediction accuracy of the model for wind power generation. The ARIMA model [11] was also used to predict the wind power generation for the next 20 seconds. After the prediction, the predicted mean value of 0.00792 was obtained, and the predicted data are shown in Table 1.

Some of the predictions are shown in Figure 4.

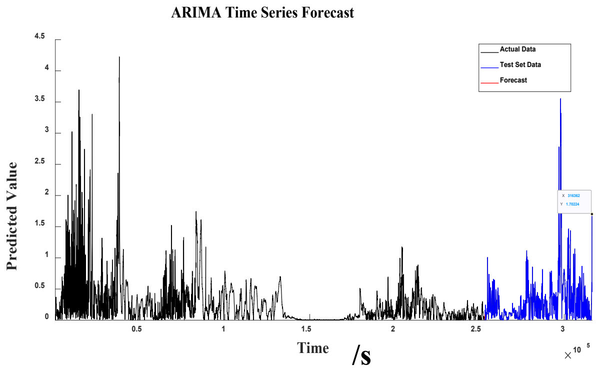


Figure 4 Prognostic Chart1

3.2 Forecasting of wind power generation by the ARIMA model

Similar to the above text. First, use the formula to calculate the values, remove the missing values and assign a value of 20 to replace the infinite values. The output image of the calculated values is shown in Figure 5.

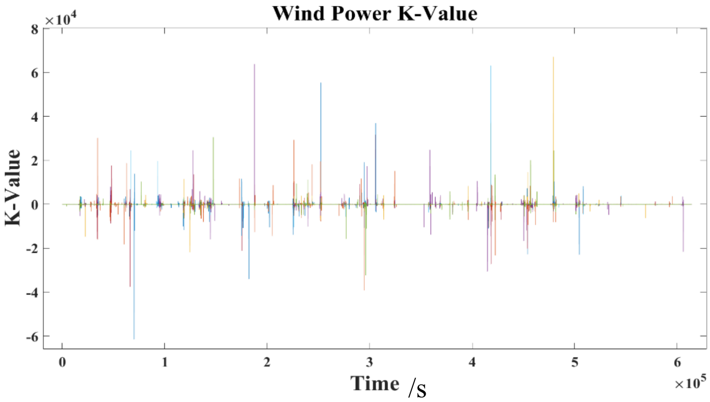


Figure 5 Wind Power K-Value

Secondly, the selection of value is carried out. In this paper, the 95th percentile of the data is selected as the value, and after calculation. The values of 11 generators are calculated respectively and averaged to get the average fluctuation amplitude. After calculation, the value is 0.9686, which is a significant increase. Therefore, the prediction is made after 120 seconds for this point. The value of 0.9686 is predicted to be a significant increase in wind power generation, and the value of -48.5785 is predicted to be a significant decrease in wind power generation. 7843 points are calculated to be a significant increase in wind power generation, and 7 points are predicted by selecting 1 point for every 1000 points, and comparing the predicted value with the actual value to get the prediction accuracy of the point. Finally, the prediction accuracy of each point was averaged to get the prediction accuracy of the model for k-value prediction of wind power generation. And the ARIMA model is used to predict the wind power generation for the next 20 seconds. Some of the predicted images are obtained as shown in Figure 6. The prediction accuracy was obtained by comparing the predicted values with the true values. The RMSE value was 0.0236, the R-squared value was 0.9998, and the MAPE value was 0.0079. Among them, the median absolute error was 0.0101 and the median relative error was 0.35%, both of which were significantly lower than the mean values (RMSE=0.0236, MAPE=0.79%). The errors of most predicted values were extremely low, but there were a few high error points that raised the average value. The error significantly increased after 75% quantile (for example, the relative error at 95% quantile reached 1.65%), indicating that there were a few high-error outliers in the tail. The maximum relative error is 3.2%, which may be caused by abnormal

ISSN 2522-3488 Vol. 9, Issue 2: 51-56, DOI: 10.25236/IJNDES.2025.090209

power generation at the end of the note or data noise.

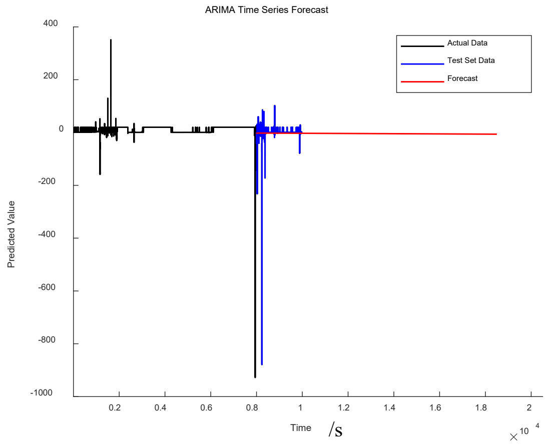


Figure 6 Prognostic Chart

4. Conclusions

In this study, the Arima model was successfully applied to predict the future power of wind power generation. The data quality is effectively improved through reasonable data methods, which lays the foundation for the model construction, and the Arima model shows certain advantages in the short-term prediction of wind power, which is able to capture the fluctuation trend of wind power. It provides data support for grid scheduling, which helps to optimise grid operation and improve the stability of the power generation system.

Although the Arima model has achieved certain results in this study, it is still defective in dealing with extreme weather and other factors that cause the power generation to fluctu-ate in a nonlinear manner. In the future, we can consider combining the Arima model with deep learning models such as LSTM model to give full play to the advantages of both and improve the prediction accuracy.

References

[1] S. Rong, L. Qiang, L. Jiankun, et al. Short-Term Wind Power Forecasting Based on Multi-Scale Fusing Deep Learning[J]. 2021 IEEE Sustainable Power and Energy Conference (iSPEC), 2021: 596-601.
[2] M. Lydia, S. Suresh Kumar, A. Immanuel Selvakumar, et al. A comprehensive review on wind turbine power curve modeling techniques[J]. Renewable and Sustainable Energy Reviews, 2014, 30: 452-460.
[3] Mantas Marčiukaitis, Inga Žutautaitė, Linas Martišauskas, et al. Non-linear regression model for wind turbine power curve[J]. Renewable Energy, 2017, 113: 732-741.
[4] Borowski, Sylwester & Migawa, Klaudiusz & Sołtysiak, et al. Regression model in the operation of wind turbines[J]. MATEC Web of Conferences, 2021, 351(01021).
[5] P. Sivaprakash, M. Sankar, R. Chithambaramani, et al. Regression Method-Based Prediction of Wind Power Generation[J]. 2023 International Conference on System, Computation, Automation and Networking (ICSCAN), 2023: 1-6.
[6] Zahraa Tarek, Mahmoud Y. Shams, Ahmed M. Elshewey, et al. El-dosuky,Wind Power Prediction Based on Machine Learning and Deep Learning Models[J]. Computers, Materials and Continua, 2022, 74(1): 715-732.
[7] Wu, Linhan and Yu, Jizhou and Dai, Yuxin, et al. Photovoltaic Power Generation Forecasting Based on TCN-Transformer Model[J]. 2024 5th International Conference on Artificial Intelligence and Electromechanical Automation (AIEA), 2024: 620-626.
[8] X. Gong, F. Zhang, T. Lu, W. You. Comparative Analysis of three outlier detection methods in univariate data sets[J]. 2022 3rd International Conference on Electronic Communication and Artificial Intelligence (IWECAI), 2022: 209-213.
[9] P. Manembu, A. Kewo, B. Welang. Missing data solution of electricity consumption based on Lagrange Interpolation case study: IntelligEnSia data monitoring[J]. 2015 International Conference on Electrical Engineering and Informatics (ICEEI), 2015: 511-516.

ISSN 2522-3488 Vol. 9, Issue 2: 51-56, DOI: 10.25236/IJNDES.2025.090209

[10] Tianhao Liu, Kunye Lv, Fengjie Chen, et al. Wind power curve model combining smoothed spline with first-order moments and density-adjusted wind speed strategy[J]. Energy, 2024(133628).
[11] F. Mahia, A. R. Dey, M. A. Masud, et al. Forecasting Electricity Consumption using ARIMA Model[J]. 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), 2019: 1-6.