Environment, Resource and Ecology Journal (2022)
Clausius Scientific Press, Canada

DOI: 10.23977/erej.2022.060304
ISSN 2616-3756 Vol. 6 Num. 3

Research on Prediction Method of Photovoltaic Building Integration Plate Index Based on ARIMA Model

Wenbo Zhang, Zhenyang Li

Nanjing Institute of Technology, School of Electric Power Engineering, Nanjing, Jiangsu, 211167, China

Keywords: Time series analysis, ARIMA model, Photovoltaic building integration plate

Abstract: As an emerging sector, the photovoltaic building integration sector can understand the future development trend of corresponding industries in the sector by predicting the development trend of the sector index. When establishing the time series ARIMA model, this paper first extracts the historical stock closing price as the series, then substitutes the data into the model, and finally trains the model, and applies the trained model to the future stock price prediction.

1. Introduction

Photovoltaic building integration (BIPV) is a distributed photovoltaic power generation project that makes full use of resources such as industrial buildings and public building roofs. It plays an important role in realizing "carbon peak" and "carbon neutralization" in China. At present, 31 places have issued relevant policies for photovoltaic building integration in the next three to five years, which will have a certain impact on the development of listed enterprises related to photovoltaic building integration.

There is a plate index in the stock market, which is an overall reflection of the trend of the plate. Now, the stock data of 37 photovoltaic building integration related enterprises in Shanghai and Shenzhen stock markets are counted as a whole, which is called photovoltaic building integration plate. As an emerging sector, the photovoltaic building integration sector can understand the future development trend of corresponding industries in the sector by predicting the development trend of the sector index.

This paper uses the existing data to analyze the error of the established model and modify the model, forecasts the future development trend of the plate according to the modified model, and draws the daily moving average of 20 trading days, the weekly moving average of 3 weeks and the monthly moving average of 2 months after May 28.

2. Model Establishment and Solution

2.1 ARIMA model

Model establishment: ARIMA, differential autoregressive moving average model, also known as autoregressive moving average model, is one of the prediction and analysis methods of time series.

ARIMA (p, d, q) model can be expressed as:

$$(1-\sum_{i=1}^{p}\phi_i L^i)(1-L)^d X_t = (1+\sum_{i=1}^{q}\theta_i L^i) \tag{1}$$

Where L is the lag operator, $d \in Z, d > 0$

The specific formula of autoregressive model is as follows:

$$y_t = \mu + \sum_{i=1}^{p}\gamma_i y_{t-i} + \theta_t \tag{2}$$

The structural formula of moving average model MA is as follows:

$$y_t = \mu + \sum_{i=1}^{q} \varepsilon_i \theta_{t-i} + \tag{3}$$

If we write ∇ as a difference operator, then we have

$$\nabla^2 y_t = \nabla(y_t - y_{t-1}) = y_t - 2y_{t-1} + y_{t-2} \tag{4}$$

For delay operator B, there are

$$y_{t-p} = B^p y_t, \forall p \geq 1 \tag{5}$$

$$\nabla^k = (1-B)^k \tag{6}$$

If d-order homogeneous nonstationary time series y_t is set, then $\nabla^d y_t$ is a stationary time series, it can be set as ARMA (p, q) model, i e

$$\lambda(B)(\nabla^d y_t) = \theta(B)\varepsilon_t \tag{7}$$

Where $\lambda(B) = 1 - \lambda_1 B - \lambda_2 B^2 - \cdots - \lambda_p B^p$, $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_p B^p$ are autoregressive coefficient polynomial and moving average coefficient polynomial respectively.

2.2 SPSS data analysis and solution of the model

2.2.1 Plate index stationarity analysis

The stationarity test of data is the premise of establishing ARIMA conditions. If it is not met, the stationarity of time series data needs to be adjusted and corrected.

Figure 1: Timing chart of plate index

The time series of plate index fluctuates up and down, so it is preliminarily judged that the series is a non-stationary time series. Through the unit root test, the calculated p value of ADF test is 0.045, which is greater than 0.01 of significance level. Therefore, it can be judged that the sequence is in a non-stationary state.

2.2.2 Differential processing

It can be seen from the above that the initial plate index time series is a non-stationary series. Next, the data of the series are processed by difference. After the first-order difference, the plate index time series will be in a stable state.

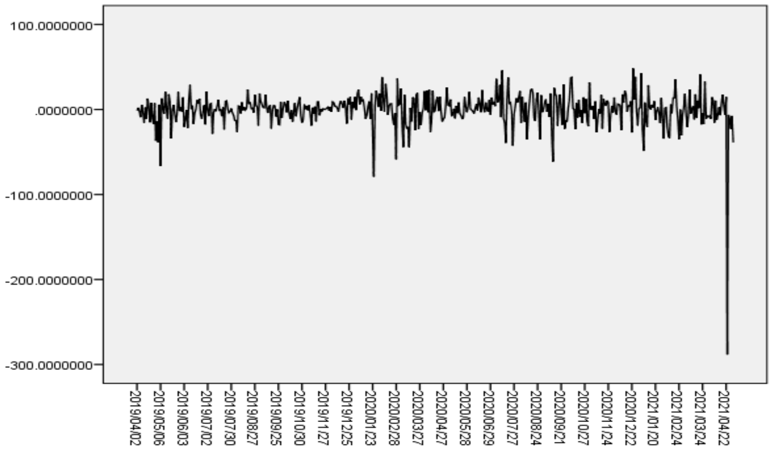

Figure 2: First order difference sequence diagram of plate index

2.2.3 Model order determination and parameter estimation

After the difference, the plate index time series reaches a stable state. Because the first-order difference has been made for the initial series, the value of d is 1. By observing Fig. 3, it is considered to establish ARIMA (2,1, 2), ARIMA (3,1,2) and ARIMA (4,1,2) models. The most appropriate orders p and q are determined through the AIC minimum criterion. When p = 2 and q = 2, the AIC value is the smallest, so the model ARIMA (2,1,2) is determined as the final model. Thus, the data from May 6 to May 28, 2021 can be predicted.

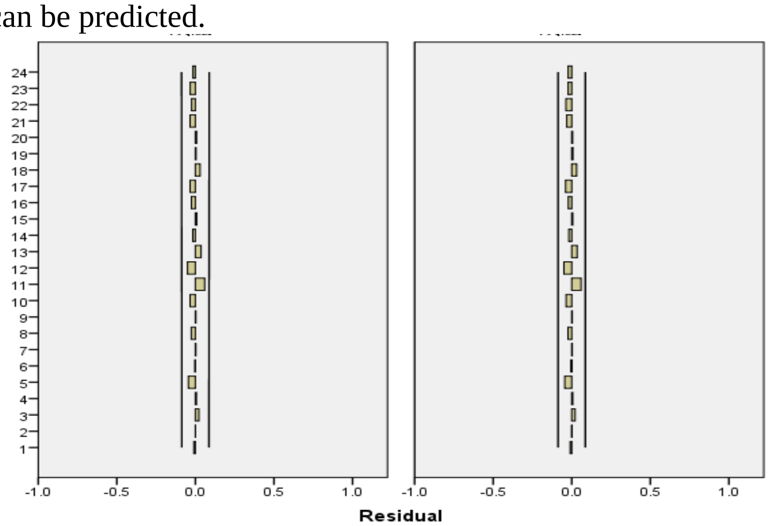

Figure 3: Residual diagram

2.3 Error Analysis and Test of Model

Use the data from May 6 to May 28 to calculate the real plate index. At this time, analyze the error between the predicted value and the real value and correct the prediction model.

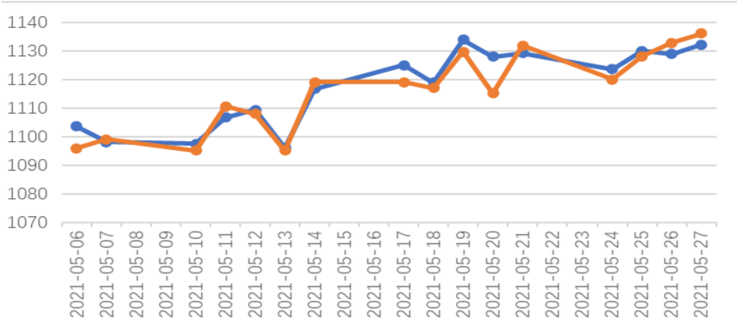

Figure 4: Comparison between actual plate parameters and predicted plate parameters

Through comparison, it is found that the established model performs well, and the predicted value is basically consistent with the actual value without correction.

2.4 Draw Moving Average

According to the established model, we can predict the plate index in the two months after May 28, and then we can draw the daily moving average of 20 trading days, the weekly moving average of 3 weeks and the monthly moving average of 2 months after May 28, as shown in figures 5, 6 and 7.

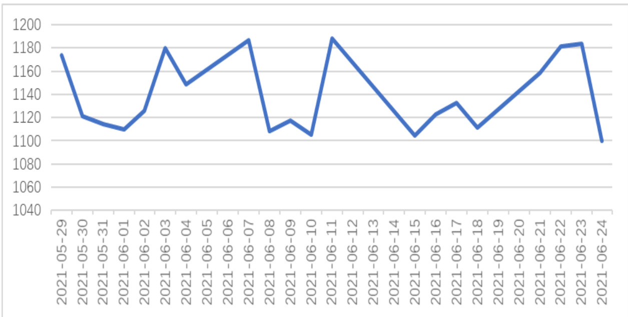

Figure 5: Daily moving average of 20 trading days

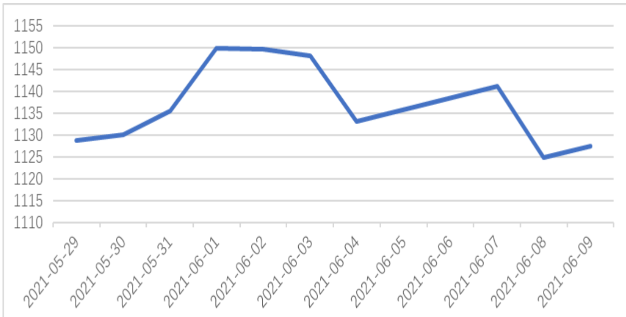

Figure 6: Weekly moving average of 3 weeks

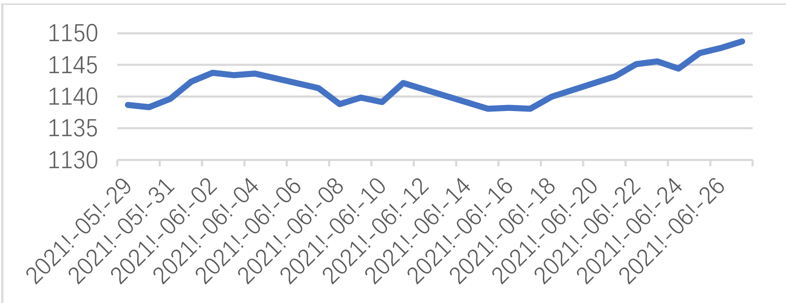

Figure 7: Monthly moving average of 2 months

3. Model Evaluation

In order to avoid the singleness of the model, we have adopted a variety of models, established different models from various aspects, and discussed various problems. Refer to various indicators, screen the data, ensure that the data is more in line with the model requirements, and provide a good data basis for establishing an accurate model. The time series of the obtained plate index is analyzed, and its historical stock data are fitted. Finally, ARIMA model which can make short-term prediction is established to better predict the future development trend. The establishment of multiple models avoids the single solution caused by a single model and reduces the generation of errors. The combination and extension of the models make the calculation more convenient and fast. The maximum withdrawal rate can better present the risk coefficient.

When the sample data changes, the parameter structure of ARIMA model will change, which shows that the model is very sensitive to the change of samples, and the prediction accuracy also changes due to the change of samples. Finally, due to the lack of relevant data, it is difficult to compare the influence of multiple factors on the weight in detail.

References

[1] Wang yanru. Application of portfolio model in stock price forecasting [D]. Chongqing: Chongqing University, 2018

[2] Yan Fang. Fitting and prediction of ARIMA model of stock price index based on wavelet denoising [D]. Nanchang: Jiangxi University of Finance and economics, 2016

[3] Sun Xianqiang. Application of time series model in stock price prediction [D]. Kunming: Yunnan University, 2016