Predicting of Weekly Solar Energy Generation Using ARIMA Approach

OMAR . M .ELMABROUK
Faculty of Engineering, University of Benghazi.
omar.elmabrouk@uob.edu.ly

NAJI. S .ABDELWANIS
Faculty of Engineering, University of Benghazi.
naji.abdwlwanis@uob.edu.ly

ABDELAZIZ M. BADI
Faculty of Engineering, University of Benghazi.
abdelaziz.badi@uob.edu.ly

Abstract

In this study, a statistical modeling approach named ARIMA has been utilized to predict the mean solar radiation for 10 weeks in Benghazi city. Since the collected solar radiation data described as non-stationary, it was firstly changed to stationary time series to be suitable for the application of the ARIMA model. The model is validated by means of determining the mean absolute error (MAE). The mean absolute error (MAE) was found to be 7 % which demonstrate the accuracy of the developed model and the efficiency of the used approach.

Keywords: solar radiation; predicting; ARIMA; modeling; time series.

الملخص

في هذه الدراسة ، تم استخدام أسلوب النمذجة الإحصائية المسمى ARIMA للتنبؤ بمتوسط الإشعاع الشمسي لمدة 10 أسابيع في مدينة بنغازي. نظرًا لأن بيانات الإشعاع الشمسي التي تم جمعها والتي تم وصفها على أنها غير ثابتة ، فقد تم تغييرها أولاً إلى سلاسل زمنية ثابتة لتكون مناسبة لتطبيق نموذج ARIMA. تم التحقق من صحة النموذج من خلال تحديد متوسط الخطأ المطلق (MAE). تم العثور على متوسط الخطأ المطلق (MAE) ليكون 7 ٪ مما يدل على دقة النموذج المطور وكفاءة النهج المستخدم.

1. Introduction

Since the traditional source of fossil fuel having significant carbon footprint, it is limited source of energy and the word demand for the electricity is still increasing. Based on those factors and due to the advanced of the technologies, the utilization of the renewable energy is increasing (Khan. et al., 2017).

The smarter girds motivate and increase the penetration of the renewable energy. Photovoltaic (PV) based solar energy is the most promising renewable energy (Majumder. et al., 2017).

Practically, the connected grid PV power is increasing fast and can interfere with the stability of the network. Therefore, an efficient predicting method will improve the grid operators to effectively manage the electrical balance between demand and power generation (Diagne. et al., 2013).

PV based solar energy depends mainly on the solar irradiance and other environmental factors such as humidity, temperature and geographic location (Hassan. et al., 2017).

The solar radiation data could be obtained from several sources such as Numerical Weather Prediction (NWP) models, Satellite-based forecast, All-sky imagers, Ground measurements and also there are many prediction approaches such as empirical regression, neural networks (Mellit. et al., 2006), and time-series models (e.g., ARMA, ARIMA)(Perdomo. et al., 2010, Raji and Boyo, 2012).

In this paper, ARIMA, referred to Autoregressive Integrated Moving Average, characterized by simplicity of implementation and used of the famous Box-Jenkins methodology was utilized for prediction of weekly solar radiation intensity in Benghazi city. The time series data is transformed form non stationary to a stationary one, the data then analyzed to determine the model parameters (p d q) and validated using mean absolute error (MAE).

The paper is structured as follows. In Section 2, previous literatures are reviewed. In Section 3, time series analysis and ARIMA prediction models are presented. In Section 4, application of ARIMA model with case study is introduced. Finally Section 5 is dedicated to conclusion.

2. Literatures

Forecasting energy consumption has an important role in planning energy strategies for both policy makers and related organizations in any country. Coal, oil, natural gas, renewable and total energy consumption data for 1970-2015 were used to forecast energy consumption of Turkey for the following 25 years, using (ARIMA) model. The ARIMA models used were ARIMA (1, 1, 1) for coal consumption, ARIMA (0, 1, 0) for oil consumption, ARIMA (0, 0, 0) for natural gas consumption, ARIMA (1, 1, 0) for renewable energy consumption and ARIMA (0, 1, 2) for total energy consumption. The results indicated that Turkey's energy consumption will continue to increase to the end of 2040. Consumption of coal, oil, natural gas, renewable energy and total energy will continue to increase at an annual average rate of 4.87 %, 3.92 %, 4.39 %, 1.64 % and 4.20 %, respectively in 25 years of the forecast (S. Ozturk and F. Ozturk, 2018)

Based on the daily average of solar irradiance, solar irradiance data from a 4.0 kW PV panel in the city of Awali, Kingdom of Bahrain, was utilized for solar energy forecast, using ARIMA models. These models were proved to effectively capture the autocorrelative structure of the solar irradiance (Shams. et al., 2016).

ARIMA time series prediction method was used based on solar irradiance and surface air temperature data collected from 10 meteorological stations in Europe and 4 stations in Asia. It was found that various climate time series are dependent on long-range variability of solar irradiance(ARIMA, 2009). The hourly solar irradiance observations from National Solar Radiation Data Base (NSRDB) site between 2008 and 2009 were investigated and analyzed for prediction of solar irradiance. It was concluded that the hybrid ARIMA-BP does not outperform ARIMA (Nazaripouya. et al., 2016).

ARIMA time series approach was implemented for prediction of solar radiation for 1–5 h time horizon. 14 years of hourly solar radiation data from solar were collected and it was found that ARMA outperforms the persistence model for short and medium term solar predictions (Huang, 2012).

3. ARIMA Models

ARIMA models are, in theory, the most commonly used to forecast future values of times series data. ARIMA model was first popularized by Box and Jenkins (Box and Jenkins, 1970). It forecasts future values of a time series as a linear combination of its own past values and/or lags of the forecast errors (also called random shocks or innovations). Box and Jenkins, (1976) stated that these models do not involve independent variables, but rather make use of the information in the series itself to generate forecasts. Therefore, ARIMA models depend on autocorrelation patterns in the series.

An ARIMA (p, d, q) model has three parameters. AR parameter ‘p’ represents the order of autoregressive process, I parameter ‘d’ represents the order of difference to obtain stationary series if the series are non-stationary, and MA parameter ‘q’ represents the order of moving average process. Autoregressive revolves around regression the variable on its prior terms. The I parameter of the model is generally applied when the data in the sample are non-stationary. If the series are stationary, then d=0, and if the series are first-difference stationary then d=1 and so forth.

The moving-average parameter states that the variable linearly depends on the present and past values of a stochastic term. The generalized univariate ARIMA model with p, d, q process has the following specification:

$$Yt = \mu + \alpha_1 Y_{t-1} + \ldots + \alpha\, p\, Y_{t-p} - \theta_{1et-1} - \ldots - \theta_{qet-q} \qquad (1)$$

where Y_t is the differenced time series value, α and θ are unknown parameters α and θ are independent identically distributed error terms with zero mean. The lagged autoregressive (AR) process are symbolized by p and that of a moving average (MA) process are symbolized by q.

4. Data Preparation and Analysis

A. *Development of ARIMA Model*

The original data for solar radiation intensity taken from National Meteorological Center at Benina city in 2011 are shown table 1. Thirtyeight (38) weeks data were used for developing the model and 10 weeks for its validation the model. The data was read and plotted as a time series using Minitab 17 statistical software as shown in Figure1

Table.1 Actual data for solar radiation intensity

Mean Solar Radiation MSR kW/m^2					
Week	***MSR***	***Week***	***MSR***	***Week***	***MSR***
1	3.03	17	7.14	33	6.54
2	3.45	18	7.72	34	6.17
3	3	19	7.71	35	5.39
4	3.57	20	7.93	36	5.09
5	3.44	21	7.22	37	4.47
6	4.27	22	7.61	38	5.02
7	4.06	23	7.05	39	4.35
8	3.58	24	8.26	40	4.43
9	4.18	25	8.43	41	3.99
10	4.76	26	7.93	42	3.3
11	4.87	27	7.77	43	3.82
12	5.84	28	7.68	44	3.45
13	6.11	29	7.48	45	3.27
14	6.34	30	7.33	46	3.02
15	7.35	31	7.01	47	2.88
16	6.03	32	6.68	48	2.78

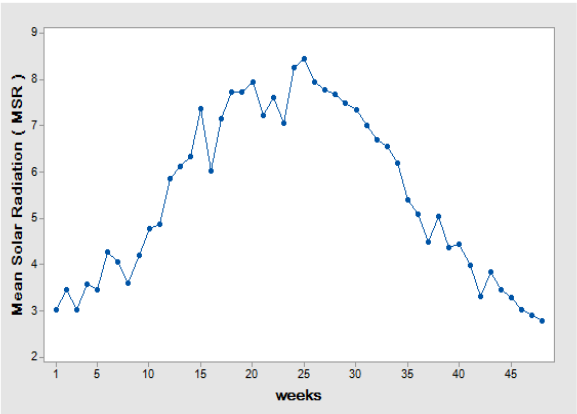

Figure 1. Collected solar radiation intensity

The next step is to plot the log base 10 for the data. If data are in saturation one can apply ARIMA model. However as given in shown Figure 2, the data do not have this characteristic.

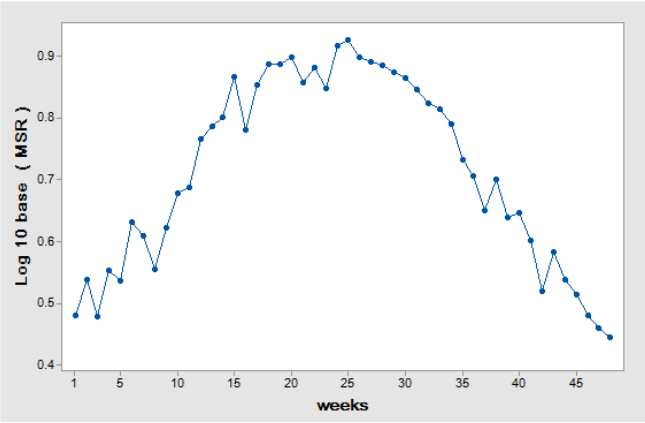

Figure 2. log base 10 solar radiation data

Thus, the next step is to calculate the differences of log base 10 data solar radiation intensity time series. Data are plotted again as shown in Figure 3. Figure 3 indicates that ARIMA model for forecasting can be applied where ARIMA parameter d = 1.

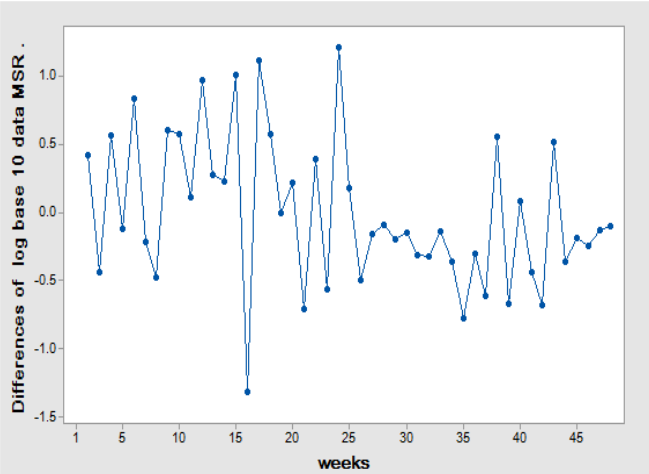

Figure 3. Solar radiation differences data

The parameters of ARIMA (p, q) were determined by plotting the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) as shown in Figure 4 and Figure 5 respectively. The values of ARIMA parameters p and q were found to be 2 and 2 respectively. The p is obtained as the Lag value was minus when the Lag value was 1 and the value raised to plus at Lag equal to 2. The q value is found to be 2, since the Lag value has direction towards zero only at Lag equal to 2.

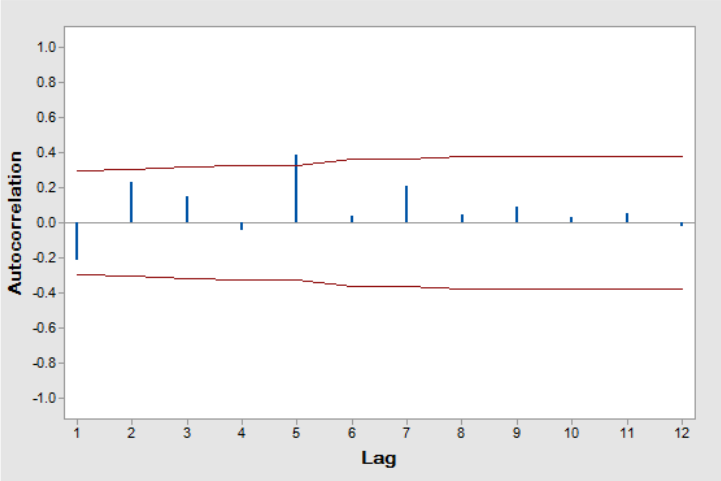

Figure 4. Autocorrelation Function (ACF)

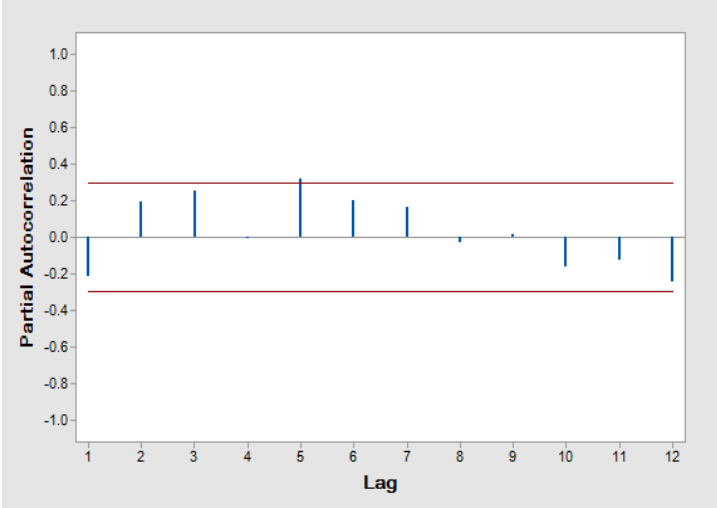

Figure 5. Partial Autocorrelation Function (PACF)

B. Validation of ARIMA Model.

The actual data of weeks 38 through 48 were used for purpose of model validation. The forecasted data are given in Table 4.2 and plotted in Figure. 6.

Table 2 The forecasting data for solar radiation.

week	**1**	**2**	**3**	**4**	**5**
Forecasting (SR)	4.36	4.03	3.78	3.56	3.36
week	**6**	**7**	**8**	**9**	**10**
Forecasting (SR)	3.18	3.02	2.87	2.73	2.60

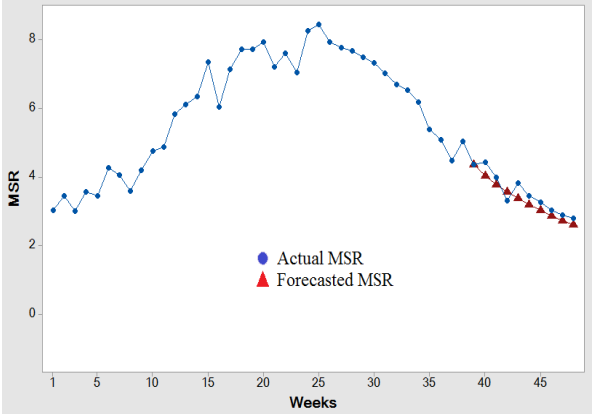

Figure 6. The actual and the forecasting solar radiation

The validation of the solar radiation model is made by comparing the actual with the predicted value of solar radiation intensity, as given in table 3. The absolute values of error and percentage error are given. The mean absolute percentage error was calculated and found to be 7 %.

Table 3. MODEL Validation.

Actual Data	Forecasted Data	Absolute Data	Error %
4.35	4.36	0.01	0.002
4.43	4.03	0.40	0.090
3.99	3.78	0.21	0.053
3.3	3.56	0.26	0.079
3.82	3.36	0.46	0.120
3.45	3.18	0.27	0.078
3.27	3.02	0.25	0.076
3.02	2.87	0.15	0.050
2.88	2.73	0.15	0.052
2.78	2.60	0.17	0.065
		MAPE	7%

C. Analysis of ARIMA Model.

The analysis of the developed ARIMA model parameters is shown in Table 4.

Table 4. The Analysis of ARIMA Model.

Type	Coef	SE coef	T	P
AR 1	1.1106	0.3139	3.54	0.001
AR 2	-0.1850	0.2823	-0.66	0.516
MA 1	1.4768	0.2590	5.70	0.000
MA 2	-0.6706	0.2227	-3.01	0.004
Constant	-0.00141	0.01448	-0.10	0.923

Based on the data given in this table, The ARIMA model was developed with a significant level of 5%. As a result, the most significant parameter is AR (2) where the P value is 51.1 %. Moreover, the energy consumption function is as follows:

$$MSRt = -0.00141+1.4768MSRt-1 -0.6706MSRt-2 +1.1106\ \varepsilon t-1 -0.1850\ \varepsilon t-2 \qquad (2)$$

The ARIMA (2, 1, 2) is also found to fits the actual time-series data considerably. The results of the evaluation of the Box Pierce Chi Square Statistic, as given in Table 5, showed the independence of the errors, thus indicating that the model is dependable.

Table 5. Evaluation of Box Pierce

Lag	12	24	36	48
Chi-Square	6.4	26.3	31.9	*
DF	7	19	31	*
P-Value	0.499	0.122	0.421	*

5. Conclusions

After developing and implementing the ARIMA model for predicting of mean solar radiation for 10 weeks in Benghazi city and then comparing the actual data with the predicted solar radiations, the developed ARIMA (2 1 2) model showed that it could significantly predict the solar radiation values with mean absolute percentage error of 7 % . Thus model is a valuable tool to determine the requirements of solar panels for different applications using of solar energy.

References

- I. Khan, H. Zhu, J. Yao, and D. Khan, “Photovoltaic power forecasting based on elman neural network software engineering method,” in 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Nov 2017, pp. 747–750.
- I. Majumder, M. K. Behera, and N. Nayak, “Solar power forecasting using a hybrid emd-elm method,” in 2017 International Conference on Circuit Power and Computing Technologies (ICCPCT), April 2017, pp. 1–6.
- Hadja Maïmouna Diagne, Mathieu David, Philippe Lauret, John Boland, Nicolas Schmutz. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renewable and Sustainable Energy Reviews, Elsevier, 2013, 27, pp.65 - 76.
- M. Z. Hassan, M. E. K. Ali, A. B. M. S. Ali, and J. Kumar, “Forecasting day-ahead solar radiation using machine learning approach,” in 2017 4th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Dec 2017, pp. 252–258.
- A. Mellit, M. Benghanem, and S. Kalogirou, "An adaptive waveletnetwork model for forecasting daily total solar-radiation," Applied Energy, voJ. 83, no. 7, pp. 705-722, 2006.
- R. Perdomo, E. Banguero, and G. Gordillo, "Statistical modeling for global solar radiation forecasting in Bogota," in Photovoltaic Specialists Conference (PVSC), 2010 35th IEEE, Jun 2010, pp. 002 374-002 379.
- M. H. T.A. Raji, A.O. Boyo, "Analysis of global solar radiation data as time series data for some selected cities of western part of Nigeria," International Journal of Advanced Renewable Energy Research, voJ. 1, no. I, pp. 14-19,2012.
- Suat Ozturk, Feride Ozturk. forecasting energy consumption of turkey by ARIMA. Journal of Asian Scientific Research, 2018, 8(2): 52-60.
- Shams MB, Haji S, Salman A, Abdali H, Alsaffar A. Time series analysis of Bahrain's first hybrid renewable energy system. Energy. 2016;103:1-15.
- Kärner O. ARIMA representation for daily solar irradiance and surface air temperature time series. Journal of Atmospheric and Solar—Terrestrial Physics. 2009;71:841-847.
- Nazaripouya H, Wang B, Wang Y, Chu P, Pota HR, Gadh R. Univariate time series prediction of solar power using a hybrid wavelet-ARMA-NARX prediction method. In: Proceedings of the IEEE Transmission and Distribution Conference and Exposition (T&D); 2016.

- Huang R, Huang T, Gadh R. Solar generation prediction using the ARMA model in a laboratory-level micro-grid. In: Proceedings of the IEEE SmartGridComm 2012 Symposium– Support for Storage, Renewable Sources, and MicroGrid; 2012. pp. 528–533.
- G. P. E. Box and G. M. Jenkins, Time series analysis: Forecasting and control. San Francisco: Holden Day, 1970.
- G. E. P. Box and G. M. Jenkins, Time series analysis: Forecasting and control, Revised Edition ed. Oakland, CA: Holden-Day, 1976.