NRIAG Journal of Astronomy and Geophysics 7 (2018) 175–179

Contents lists available at ScienceDirect

NRIAG Journal of Astronomy and Geophysics

journal homepage: www.elsevier.com/locate/nrjag

Review article

Statistical method to predict the sunspots number

H.I. Abdel-Rahman[a,b], B.A. Marzouk[a,*]

[a] National Research Institute of Astronomy and Geophysics (NRIAG), 11421, Helwan, Cairo, Egypt
[b] Department of Mathematics, Faculty of Science and Arts, Shaqra University, Shaqra, Saudi Arabia

ARTICLE INFO

Keywords:
Statistics
Time series analysis
Sun: sunspots number forecasting

ABSTRACT

Sunspots are dark regions, it is appearing in the deepest layer of solar atmosphere "photosphere", and the average number of sunspots has cycle every 11 year approximately. Sunspots cycles were numbered with solar cycle 1 beginning in 1755 and the most recent solar cycle, is cycle 24, began in December, 2008. Many authors used different techniques and precursor methods to forecast solar sunspots cycles. In this work we applied a new statistical method "autoregressive integrated moving average Models (ARIMA)" on sunspots number data observed by National Oceanic and Atmospheric Administration (NOAA) during the period 1991 – 2017 (27 years), we predict the sunspot number for the ending of second brunch of current solar sunspots cycle 24. Our prediction of sunspot numbers was compared with the international sunspot number predicted by National Oceanic and Atmospheric Administration (NOAA). Our predictions of monthly mean sunspot numbers along ending years of present solar cycle 24 are in a good agreement with international sunspots number predictions published by NOAA.

1. Introduction

As known the solar activity rises and fallsover an 11 year sunspots activity cycle, moreover it can be shorter or longer 11 years, i.e. the solar cycles can be short as 9 years and long as 14 years; where the sun goes through cycles of high and low activity that repeat every 11 yearsapproximately. The sunspot cycle is a useful way to mark the changes in the sun.

Many authors used different methods and techniques to forecast sunspots numbers for the next solar cycles.

Schatten and Myers (1996) predicted the solar cycle 23, and improved the timing of solar predictions; they found the peak of next solar cycle (no. 24) will reach a mean sunspots number R_z of 138 ± 30.

Sello (2003) predicted the solar cycle 23, and showed that the complexity of the solar processes related to the evolution of magnetic fields prevents any accurate forecasting of the solar cycle curve shape and in particular an accurate estimation for both amplitude and phase of future cycles.

Cher et al (2004) used an iterative regression analysis to the reconstruction and prediction of sunspot number series, they found the periodicities significant at the 95% confidence level that have been selected in order to reconstruct the sunspot series during 1933–1996 with correlation coefficient of r = 0.91.

Svalgaard et al. (2005) predicted the approaching solar cycle 24, and found a peak of smoothed monthly mean sunspot number of 75 ± 8, making it potentially the smallest cycle in last 100 years.

Hamid and Galal (2006) predicted solar activity cycle no. 24; they found the preliminary prediction of cycle 24 indicated that the maximum amplitude is 90.7 ± 9.2 and time of rise 4.6 ± 1.2 years.

Baranovski et al. (2008) developed a modern approach to solar cycle forecasting, and applied a nonlinear least squares fitting (NLF) to the annual sunspot time series, then they applied two methods to the NLF output parameters, then they derived a prediction of the upcoming solar cycle 24. They derived some global statistical time series, which contains many more cycles than in the available post observations.

Forecasting of the peak of the sunspot cycle is highly importantfor space weather applications. At the present time, precursormethods are the most favored for the prediction of the strengthof the next solar cycle (Kane, 2008; Hathaway, 2009; Podladchikova and Van der Linden, 2011).

Hamid and Galal (2013) investigated the spotless sun and its impact to various aspects of our life on the Earth and space environment. They found that the distribution of spotless events across successive solar cycles are responsible for the variety of their strengths. The more frequently and longer are the spotless events at a given cycle, the smaller the values of the maximum Wolf number reached.

Peer review under responsibility of National Research Institute of Astronomy and Geophysics.
* Corresponding author.
E-mail address: bmarzoke@yahoo.com (B.A. Marzouk).

https://doi.org/10.1016/j.nrjag.2018.08.001
Received 29 April 2018; Received in revised form 10 July 2018; Accepted 7 August 2018
Available online 14 August 2018

Table 1
Sample of Sunspot number data.

Year	Month	SSN	Year	Month	SSN
1991	1	213.5	1993	1	92.1
1991	2	270.2	1993	2	126.1
1991	3	227.9	1993	3	107.4
1991	4	215.9	1993	4	98.6
1991	5	182.5	1993	5	79.1
1991	6	231.8	1993	6	68.5
1991	7	245.7	1993	7	81.6
1991	8	251.5	1993	8	59.4
1991	9	185.8	1993	9	33.5
1991	10	220.1	1993	10	73.5
1991	11	169	1993	11	51
1991	12	217.7	1993	12	75.9
1992	1	217.9	1994	1	86.4
1992	2	238.2	1994	2	60.5
1992	3	160.5	1994	3	52.4
1992	4	144	1994	4	29.3
1992	5	106.3	1994	5	35.4
1992	6	104.7	1994	6	42.6
1992	7	121.4	1994	7	52.7
1992	8	99.5	1994	8	38.4
1992	9	93.8	1994	9	40.5
1992	10	136.2	1994	10	67.1
1992	11	124.3	1994	11	33
1992	12	127.4	1994	12	38.7

Hamid and Marzouk (2018) used precursor techniques in particular spotless event, geomagnetic index, and solar flux F10.7 to predict the maximum peak of solar cycle 24, moreover predict the exact date of the maximum, they found that the solar cycle 24 achieved a very low amplitude, hence they said it is weaker than solar cycles 21–32.

In this paper we applieda new statistical method “an autoregressive integrated moving average Models (ARIMA)”onthe previous sunspots data observed by NOAAduring the period 1991–2017to predict sunspot number for the last two years of the descending brunch of current solar cycle 24. Our results of sunspot number prediction werecompared with the international prediction published of NOAA.

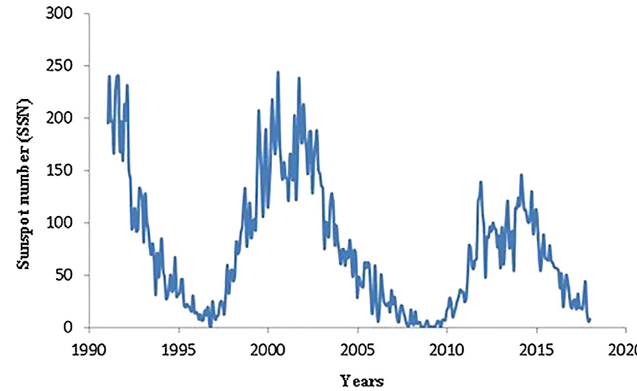

Fig. 1. The original time series of mean monthly sunspot number (SSN) during 1991–2017.

2. Data used

We selected amonthly mean of sunspot number (SSN) data whichobservedby the National Oceanic and Atmospheric Administration (NOAA) during 27 years (from 1991 till2017). Table 1. illustrated the sample of monthly mean data of sunspots number observed by NOAA.

3. Statistical method

General models of predictable numbers was discovered by (Box-Jenkins) in 1970 and is called autoregressive integrated moving average Models (ARIMA), Montgomeryet al. (2008). One of the most important techniques was used to build different models for time series analysis, these models are linear. There are three broad categories of

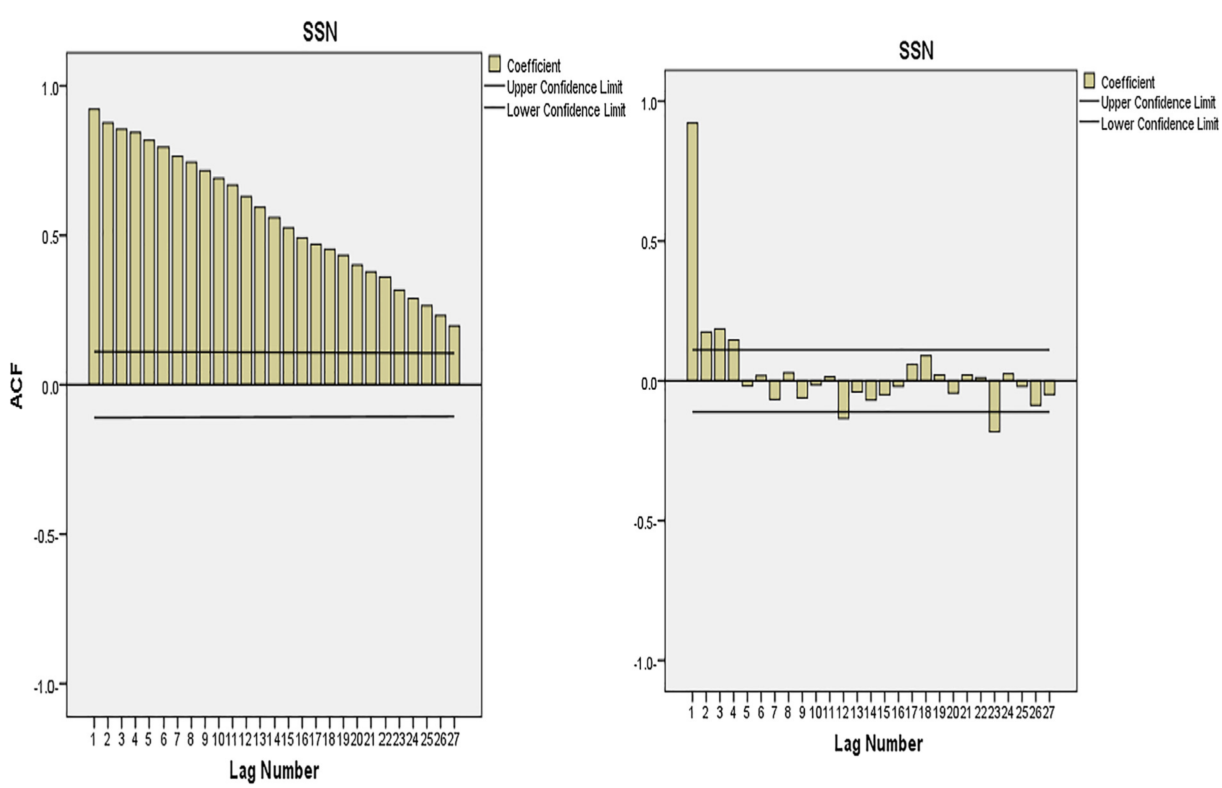

Fig. 2. The autocorrelation and partial autocorrelation functions for the original series.

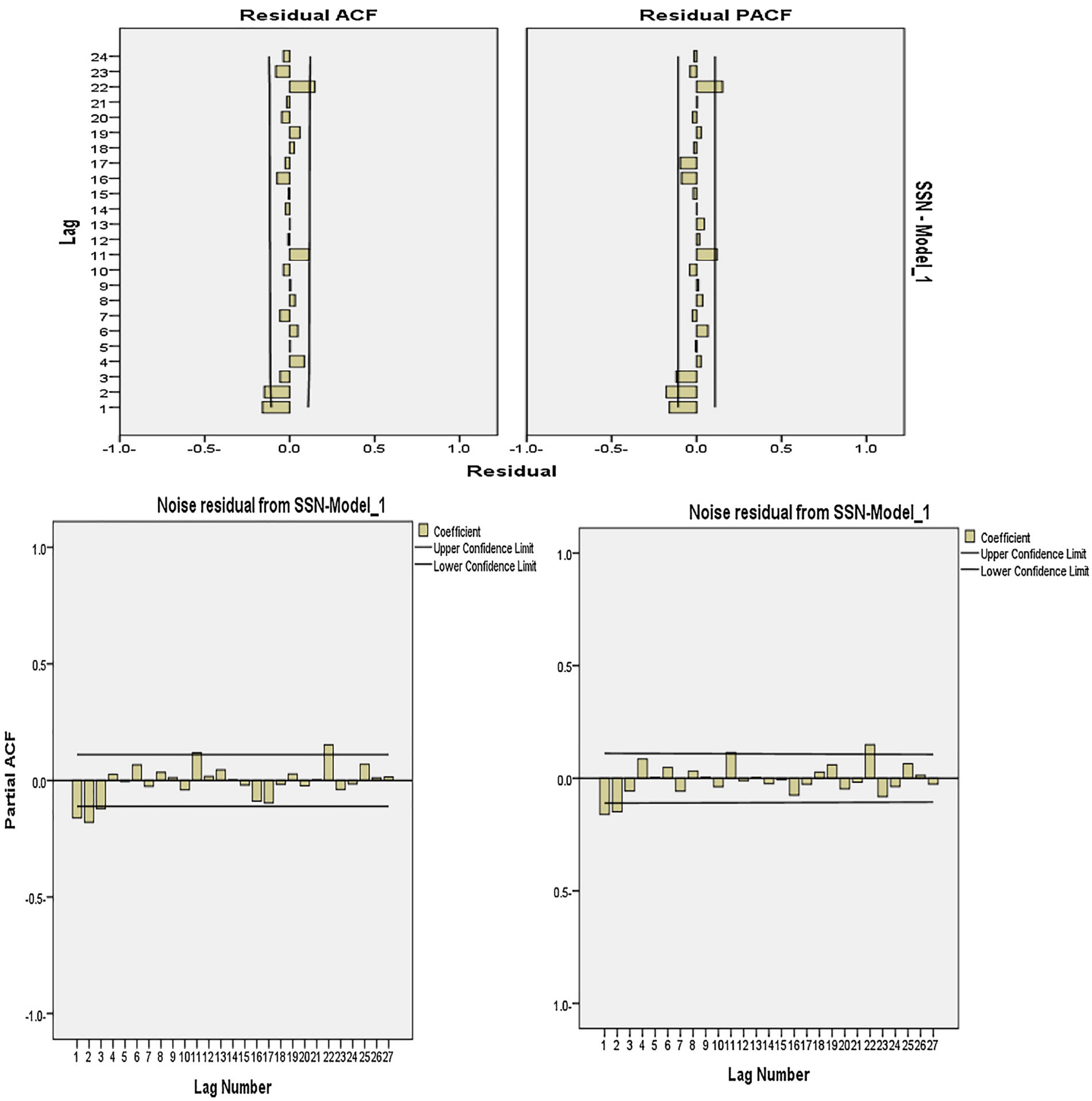

Fig. 3. The autocorrelation and partial autocorrelation functions for the residuals and the noise residual.

Table 2
Comparison between our prediction of sunspots number and that of NOAA along two years.

Year	month	Our SSN prediction	NOAASSN prediction	Year	Month	Our SSN prediction	NOAA SSN prediction
2018	1	7.42	12.9	2019	1	5.86	8.9
2018	2	7.62	12.7	2019	2	5.72	8.3
2018	3	7.44	11.8	2019	3	5.59	7.8
2018	4	7.26	11.3	2019	4	5.46	7.3
2018	5	7.09	11.7	2019	5	5.33	6.8
2018	6	6.92	12.2	2019	6	5.2	6.4
2018	7	6.76	12.5	2019	7	5.08	5.9
2018	8	6.6	12.4	2019	8	4.96	5.5
2018	9	6.45	11.5	2019	9	4.84	5.1
2018	10	6.29	10.8	2019	10	4.73	4.8
2018	11	6.15	10.1	2019	11	4.62	4.4
2018	12	6	9.5	2019	12	4.51	4.1

these models using the following symbols:

X_t: refers to the time series in general and the value of the phenomenon in thedifferent time periods t = 0, 1, 2, …, m.

$\varnothing_i$: refers to the parameters of the Autoregressive model (AR).

θ_j:refers to the parameters of the Moving Average (MA).

e_t: refers to errors in the differenttime periods t = 0, 1, 2, …, m.

Then we could clarify the categories of the Box-Jenkins models for time chains as follows:

1- Autoregressive Models

The AR(1) of the first order takes the following form:

$$X_t = \varnothing X_{(t-1)} + e_t \quad (1)$$

In general, the model AR(P) in order (P) is taken by the following formula

$$X_t = \varnothing_1 X_{(t-1)} + \varnothing_2 X_{(t-2)} + \ldots + \varnothing_p X_{(t-p)} + e_t \quad (2)$$

2- Moving Average Models

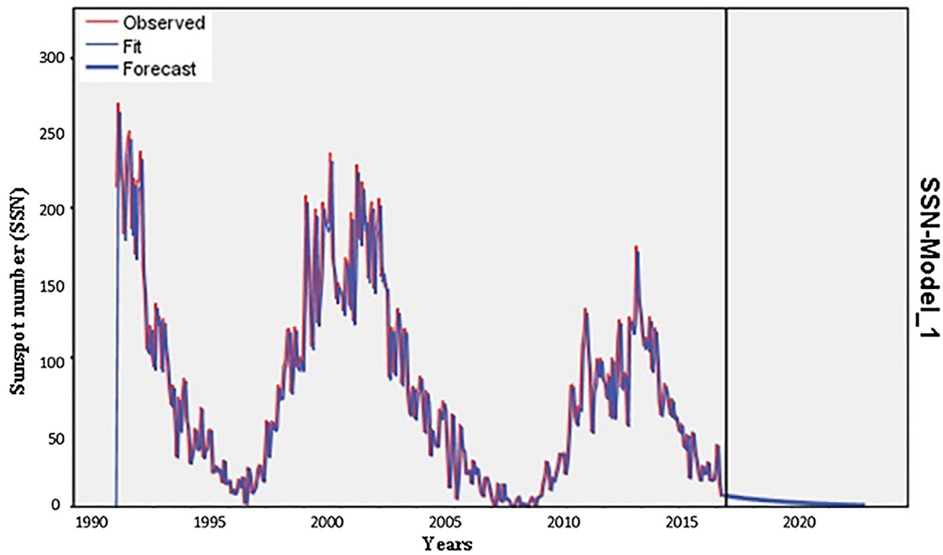

Fig. 4. Monthly observed and estimated time series of sunspots numbers during 1991–2017and our predicted sunspots numbers during 2018–2019 by using AR (1) model.

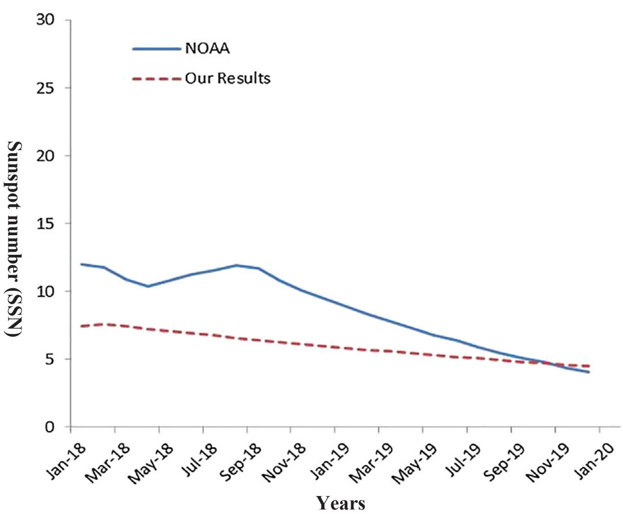

Fig. 5. Comparison between our predicted values of monthly mean sunspot numbers with that predicted by NOAA during 2018 & 2019.

Table 3
the correlation matrix between our model and NOAA.

		WORK_PRED	NOAA_PRED
WORK_PRED	Pearson Correlation	1	.965**
	Sig. (2-tailed)		.000
	N	24	24
NOAA_PRED	Pearson Correlation	.965**	1
	Sig. (2-tailed)	.000	
	N	24	24

** Correlation is significant at the 0.01 level (2-tailed).

The formula of the 1st order model MA(1) as follows:

$$X_t = e_t + \theta e_{t-1} \tag{3}$$

Generally, the formula of (q) order MA(q) is:

$$X_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q} \tag{4}$$

3- Mixed model (ARMA)

It is a mixture of the previous two models is called Autoregressive and Moving average Models (ARMA), For example, the Model ARMA (p,q), where (p, q) are the orders of the model and it takes the following form:

$$X_t = \varnothing_1 X_{(t-1)} + \varnothing_2 X_{(t-2)} + \ldots + \varnothing_p X_{(t-p)} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q} \tag{5}$$

This model can be generalized taking the difference of the values of the series. If the series was not stationary, integrated mixed model is called autoregressive integrated moving average Models (ARIMA). ARIMA (p,d,q) where the letter (d) refers to the number of times differences of time series to stationary, for example, if d = 1, this means taking the differences in the series once only (in this state the differences is first order) and the difference continue up to the series stationary. i. e. d = 1, 2, 3, …

3.1. Identification of the model

To identify the appropriate model for time-series data we need to draw the time seriesof observed sunspot number (SSN) data, we found that it is regular as shown in Fig. 1.

Fig. 1 shows that the original time series of monthly mean sunspot number (SSN) data observed by NOAA during 1991–2017.

To determine the primary model with high accuracy, we are drawing the two functions of Autocorrelation function (ACF) and Partial Autocorrelation function (PACF) for the series; from the properties of previous functionswe can get the primary model. We noted that as in Fig. 2, ACF takes the exponential form and PACF finished after the first order. Moreover the primary model is the autoregressive from first order. i.e. the primary model is AR(1).

We can define the model AR(1) as follows:

$$\boldsymbol{X_t} = 0.976\boldsymbol{X_{t-1}} + \boldsymbol{e_t} \tag{6}$$

where X_t is the future sunspots numbers, X_{t-1} is the previous sunspots numbers and e_t is the errors in different interval, t = 1, 2, 3, …, m.

From the AR (1) model equation we can get a prediction of the future sunspots number(X_t), with knowledge of the previous number of sunspots (X_{t-1}).

3.2. Is the suggested model is correct?

To answer this question, Residuals must be checked for any defect in the conditions required for the validity of this model, and the most important that the ACF and PACF functions for residuals and the noise residual are not auto correlated and It does not have a certain patternas Fig. 3, hence, the suggested model is good to get the future sunspot number prediction.

4. Results and discussions

In the present work we appliedthe ARIMA (1, 0, 0) or AR (1) model to get the prediction of sunspot number during the near future times.

By using Eq. (6) we can get the prediction ofsunspots numbers for two years (2018 & 2019) as shown in Table 2, which indicates sunspots numbers during the descending phase of the current solar cycle 24.

Table 2 shows our prediction of sunspots number and NOAASunspots number prediction.

We got sunspots number predicted by NOAA from the following website: https://www.swpc.noaa.gov/products/predicted-sunspot-number-and-radio-flux.

Fig. 4 shows the estimated curve of our monthly mean predicted sunspots numbers during 2018–2019 by using AR (1) model with the monthly observed sunspots number values by NOAA, and the fit line of curve. We can distinguish between the three chains easily, and we can note that there is a good converging in values, which indicates the accuracy of the used model. We found that the estimated sunspot numbers values in the future (our prediction) comes in consistent with the NOAA's observed sunspots number values curve (see Fig. 4).

Fig. 5 shows the comparison between our predictions of sunspot numbers during next two years (2018–2019) with predicted sunspots number published by National Oceanic and Atmospheric Administration (NOAA). We can notice that our predicted sunspots numbers are in a good agreement with published predictions by NOAA during two years.

In Table 3, we illustrate that the correlation coefficient between our work and NOAA sunspot number predictions, we found that the correlation coefficient value is 96.5% and it is very high correlation. The correlation is significant at the 0.01 level.

5. Conclusions

Autoregressive Model with first order AR(1) and observed sunspots number data by NOAA during 27 years (1991–2017) was used to get a prediction of sunspots number for the last two years of the descending phase of current solar cycle 24.

We were clarifying the values predicted by our model in one figure with the previous monthly sunspots numbers values observed by NOAA. We can note that there is a good converging in our predicted sunspots values, which indicates the accuracy of the used model, and we found that the estimated sunspot numbers values in the future (our prediction) is consistent with the NOAA's observed sunspots number values curve.

We compared our predictions of future sunspots number during last two years (2018–2019) of current solar cycle with predicted sunspots number published by NOAA. We found that our predicted sunspots numbers are in a good agreement with published predictions by NOAA during (2018–2019).

Finally our result of predicted sunspots number illustrates that there is a very strong correlation confection of about 96.5% between our results with these proposed results.

References

Baranovski, A.L., Clette, F., Nollau, V., 2008. Nonlinear solar cycle forecasting: theory and perspectives. Ann. Gephys. J. 26, 231–324.

Cher, E., Rigozo, N.R., Nordemann, D.J.R., Vieira, L.E.A., 2004. Prediction of solar activity on the basis of spectral characteristics of sunspot number. Ann. Gephys. J. 22, 2239–2243.

Montgomery, Douglas C., Jennings, Cheryl L., Kulahci, Murat, 2008. Introduction to Time Series Analysis and Forecasting. John Wiley & Sons. Inc., Hoboken New Jersey.

Hamid, R.H., Galal, A.A., 2006. Preliminary prediction of the strength of 24th 11-year solar cycle, Solar activity and its magnetic origin. In: Proceeding IAU, symposium No. 233. vol. 233, pp. 413–16, 2006.

Hamid, R.H., Galal, A.A., 2013. An early prediction of the maximum amplitude of the solar cycle 25. J. Adv. Res. 4 (3), 275–278.

Hamid, R.H., Marzouk, B.A., Elsevier Press, NRIAG Journal of Astronomy and Geophysics (Accepted February 2018).

Hathaway, D.H., 2009. Solar cycle forecasting. Space Sci. Rev. 144, 401–412.

Kane, R.P., 2008. Prediction of solar cycle maximum using solar cycle lengths. Sol. Phys. 248 (1), 203–209.

Schatten, Kenneth, Myers, Daniel J., 1996. Solar activity forecast for cycle 23. Geophys. Res. Lett. 23 (6), 605–608.

Podladchikova, T., Van der Linden, R., 2011. J. Space Weather Space Climate 1, A01.

Sello, S., 2003. Solar cycle Activity: A preliminary predication for cycle #24. A&A Journal 410, 691–693.

Svalgaard, L., Cliver, E.W., Kamide, Yohsuke, 2005. Sunspot cycle 24: smallest cycle in 100 years. Geophys. Res. Lett. 23, L01104.