Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Prediction of Sunspot Number During Solar Cycle 25: Deducing a New Model by Box-Jenkins Technique

H.I. Abdel Rahman, M. M. Farghaly and M.A. Semeida
**Solar and Space Research Department, National Research Institute of astronomy and geophysics, Cairo, Egypt*
E-mail: helal_ismaeil@yahoo.com , mohebmagda@live.com, helal.ismaeil@nriag.sci.eg

ARTICLE INFO

Received: 31 Dec 2024

Revised: 20 Feb 2025

Accepted: 28 Feb 2025

ABSTRACT

Solar cycle is known as the solar magnetic activity cycle, is a nearly periodic 11 years change in the activity of the sun measured by observing the number of sunspots, its solar maximum and minimum refer to the periods of highest and lowest sunspot counts, respectively. Our recently studied cycle 25, which commenced in December 2019 with a minimum smoothed sunspot number of 1.8. It is expected to continue until around 2030. In our study, we developed a new statistical prediction model by using observed sunspot data from 1749 to August 2024 (approximately 276 years) to forecast sunspot numbers at the end of Solar Cycle 25. The correlation between output of the suggested model and the observed sunspot numbers (SSN) from 1749 to August 2024 is a strong at 93.3% while the correlation coefficient between our predicted results and the published predicted data of NOAA demonstrates exceptionally very strong correlation at 98.7%.

Keywords: Solar cycle-Time series analyses-sunspots-statistical model

INTRODUCTION

The Sun has entered its 25th solar cycle and is on the verge of awakening. Over the past few years, our star has been relatively quiet, showing few sunspots, bright flares, or massive ejections of magnetized plasma from its surface. This period of reduced activity is known as the solar minimum; however, signs indicate that activity begins to intensify once more. Many authors have used different methods and techniques to forecast sunspot numbers for the next solar cycles (Kane, 2008; Hathaway, 2009; Podladchikova and Van der Linden 2011). Modern approach for solar cycle forecasting, they applied two methods of nonlinear least squares fitting to annual sunspot series (Baranovski; Clette; and Nollau, 2008), they derived a prediction of the upcoming solar cycle. (Abdel Rahman and Marzouk, 2018) used another model from Box-Jenkins techniques to predict the sunspots number for cycle 24.

Diverse predictions on the strength of Cycle 25 have been proposed, ranging from forecasts of a very weak cycle, hinting at a gradual descent into a Maunder Minimum-like state (Upton and Hathaway, 2018) to projections of a weak cycle similar to the previous Cycle 24 (see., Bhowmik, and Nandy, 2018), and even anticipations of a strong cycle.

(Upton and Hathaway, 2018) predicted that the weakness of cycle 25 would make it part of the Modern Gleissberg Minimum. The Solar Cycle 25 Prediction Panel predicted in December 2019 (Sarp, V. et al 2018) that solar cycle 25 will be similar to solar cycle 24, with the preceding solar cycle minimum in April 2020 (± 6 months), and the number of sunspots reaching a (smoothed) maximum of 115 in July 2025 (± 8 months). This prediction is in line with the current general agreement in the scientific literature

(see., https://www.esa.int/Space_Safety/Solar_cycle_25_the_Sun_wakes_up) which holds that solar cycle 25 will be weaker than average (i.e. weaker than during the exceptionally strong Modern Maximum). However, observations from 2020 to 2022, the first three years of the cycle, significantly exceed predicted values (Sarp, V. et al 2018). (Sabry M. A. et al. 2012) used some statistical methods to determine the members of the Hickson compact group with high accuracy, so it can be used to know the beginning and end of the solar cycle.

Time series encompasses sequentially generated observations over time, with data inherently organized chronologically. Successive observations within the time series are typically interdependent, and this non-

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

independence is harnessed to formulate reliable predictions. The goals of time series analysis extend beyond mere observation; they involve crafting an accurate depiction of the distinctive characteristics of the underlying process. This analysis further entails constructing a model to interpret and elucidate the chain's behaviour in relation to other variables, establishing connections between observation values and the governing rules and dynamics of the chain.

In this paper, we applied ARIMA (autoregressive integrated moving average) Models on the previous sunspots data observed by SILSO World Data Center during the period 1749 – Aug. 2024 to predict sunspot number for the current solar cycle 25. Our results of sunspot number prediction were compared with published prediction of NOAA.

The data used in this study is introduced in section data used. Next sections present the deduced model and discuss the results. The concluding remarks are drowning in section conclusion.

DATA USED

We utilized Sunspot data from the World Data Center SILSO, Royal Observatory of Belgium, Brussels (https://www.sidc.be/SILSO/datafiles). The data is collected during the period from 1749 to Aug. 2024. A sample of this data is presented in Table 1.

Table (1) sample of observed sun spot number data

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2000	133.1	165.7	217.7	191.5	165.9	188	244.3	180.5	156	141.6	158.1	143.3
2001	142.6	121.5	165.8	161.7	142.1	202.9	123	161.5	238.2	194.1	176.6	213.4
2002	184.6	170.2	147.1	186.9	187.5	128.8	161	175.6	187.9	151.2	147.2	135.3
2003	133.5	75.7	100.7	97.9	86.8	118.7	128.3	115.4	78.5	97.8	82.9	72.2
2004	60.6	74.6	74.8	59.2	72.8	66.5	83.8	69.7	48.8	74.2	70.1	28.9
2005	48.1	43.5	39.6	38.7	61.9	56.8	62.4	60.5	37.2	13.2	27.5	59.3
2006	20.9	5.7	17.3	50.3	37.2	24.5	22.2	20.8	23.7	14.9	35.7	22.3
2007	29.3	18.4	7.2	5.4	19.5	21.3	15.1	9.8	4	1.5	2.8	17.3
2008	4.1	2.9	15.5	3.6	4.6	5.2	0.6	0.3	1.2	4.2	6.6	1
2009	1.3	1.2	0.6	1.2	2.9	6.3	5.5	0	7.1	7.7	6.9	16.3
2010	19.5	28.5	24	10.4	13.9	18.8	25.2	29.6	36.4	33.6	34.4	24.5

- **Statistical description of the sun spots sample**

Table 2 presents the count of sunspots from January 1749 to August 2024 about 3308 months (approximately 276 years). Notably, the data reveals that the highest number of sunspots occurred in May 1778, reaching 398.2 sunspots, while the lowest count was recorded in 67 months, with zero sunspots. On average of sample, there were approximately 82 sunspots, with a standard deviation of 68, offering insights into the variability and distribution of sunspot numbers during this period.

Table (2): the descriptive statistics of observed sunspots number series

Variable	Sample size	Mean	STD	Min.	Max.
Sun spots no.	**3308**	**81.95**	**67.7**	**0**	**398.2**

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Tables (3, 4) have the zero sunspots number (67 months), very weak intervals, and the largest sunspots number equal or over 200 (226 months), very active time of cycles, with years and corresponding months.

Table (3): 67 months for zero sunspots months and years

Year	Month	Sn	Year	Month	Sn	Year	Month	Sn	Year	Month	Sn
1754	1	0	1810	2	0	1811	8	0	1824	6	0
1755	5	0	1810	3	0	1812	4	0	1824	7	0
1755	6	0	1810	4	0	1813	1	0	1824	11	0
1775	2	0	1810	5	0	1822	1	0	1855	9	0
1798	5	0	1810	6	0	1822	9	0	1856	5	0
1798	6	0	1810	7	0	1822	11	0	1867	1	0
1798	7	0	1810	8	0	1822	12	0	1878	8	0
1799	8	0	1810	9	0	1823	1	0	1879	3	0
1799	9	0	1810	10	0	1823	2	0	1901	4	0
1800	4	0	1810	11	0	1823	4	0	1901	12	0
1807	12	0	1810	12	0	1823	5	0	1902	2	0
1808	1	0	1811	1	0	1823	6	0	1902	4	0
1808	3	0	1811	2	0	1823	8	0	1912	2	0
1809	10	0	1811	3	0	1823	9	0	1913	5	0
1809	11	0	1811	4	0	1823	10	0	1913	6	0
1809	12	0	1811	5	0	1823	11	0	2009	8	0
1810	1	0	1811	6	0	1824	3	0			

Table (4): 226 months for sunspots number over or equal 200

Year	Month	Sn	Year	Month	Sn	Year	Month	Sn	Year	Month	Sn
1778	5	398.20	1871	4	270.90	1990	8	252.10	1788	3	238.80
1957	10	359.40	1958	3	270.00	1778	11	250.50	1871	3	238.70
1836	12	343.80	1948	4	268.50	1849	2	250.30	1779	7	238.30
1957	12	339.00	1847	9	268.20	1979	12	249.60	1989	11	238.20
1957	9	334.00	1947	8	267.40	1958	5	248.20	2001	9	238.20
1837	1	313.40	1870	4	266.90	1769	9	248.00	1836	4	238.00
1959	1	307.70	1979	9	266.90	1957	4	248.00	1770	2	237.50
1847	10	300.60	1848	12	265.80	1980	12	246.90	1948	6	237.50
1957	11	298.60	1958	12	265.70	1769	11	246.80	1788	10	236.70
1849	1	298.30	1848	1	265.20	1948	5	246.40	1979	1	235.90

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

1778	1	295.50	1957	7	265.10	1870	11	246.00	1787	11	235.80
1870	5	293.60	1749	11	264.30	1956	9	245.40	1788	7	235.80
1837	2	292.60	1769	10	263.70	1981	9	245.30	1788	9	235.00
1787	12	290.00	1979	10	263.60	1870	10	244.30	1838	3	234.50
1958	1	286.70	1837	6	263.40	2000	7	244.30	1847	8	234.30
1778	9	286.20	1959	3	263.00	1959	5	243.60	1957	1	233.70
1778	6	286.00	1870	3	262.70	1777	11	243.30	1958	2	233.60
1958	9	285.10	1777	12	262.20	1958	6	242.90	1778	8	233.30
1947	5	285.00	1787	9	262.20	1871	5	242.60	1779	5	233.30
1956	11	285.00	1787	10	261.70	1981	8	242.50	1980	10	233.30
1989	6	284.50	1778	10	260.50	1937	7	241.80	1957	5	233.00
1957	6	284.30	1979	11	259.50	1778	4	241.70	1980	4	232.20
1958	8	283.50	1949	2	258.00	1779	4	241.70	1947	6	232.10
1959	8	282.60	1917	8	257.70	1838	1	241.50	1848	7	232.00
1958	4	277.60	1788	6	257.00	1991	8	240.80	1989	8	232.00
1779	2	276.20	1958	10	256.90	1991	2	240.30	1947	10	231.70
1938	7	275.60	1870	8	256.50	1956	8	240.20	1847	11	231.50
1956	12	272.00	1778	7	255.00	1991	7	240.20	1959	4	231.30
1837	7	271.20	1980	5	254.70	1947	9	239.90	1988	12	231.20
1958	7	271.00	1771	5	254.50	1959	6	238.90	1982	2	230.90

Table (4): continued

Year	Month	Sn	Year	Month	Sn	Year	Month	Sn	Year	Month	Sn
1992	2	230.70	1839	9	221.10	1938	5	212.30	1860	5	203.50
1837	4	230.30	1982	3	221.10	1989	10	212.20	1949	11	203.20
1788	1	230.00	1848	8	220.90	1947	4	212.10	1859	8	203.00
1937	8	229.50	1870	7	220.90	1787	4	212.00	1946	12	202.90
1838	5	229.40	1937	1	220.90	1959	7	211.90	1948	9	202.90
1836	10	228.90	1848	10	220.60	1979	6	211.70	2001	6	202.90
1787	8	228.70	1770	11	220.30	1989	12	211.40	1959	2	202.60
1990	1	227.40	1956	10	219.90	1838	4	211.10	1836	11	201.50
1870	9	226.80	1980	9	219.60	1770	8	210.50	1948	7	201.40
1788	8	226.70	1980	2	219.40	1989	1	210.10	1979	8	201.40
1870	6	226.30	1839	8	218.80	1980	11	209.50	1859	9	200.90
1980	1	226.10	1859	10	217.80	1871	2	209.00	1769	8	200.50
1789	12	225.80	2000	3	217.70	1789	2	208.80	1872	2	200.30
1979	7	225.70	1937	6	217.10	1989	2	208.70	1851	2	200.10
1981	4	225.30	1870	12	216.70	1937	10	208.20	1789	3	200.00
1989	9	225.10	1981	10	216.20	1949	4	208.10	1789	6	200.00

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

1787	5	224.70	1837	12	216.10	1836	6	207.90	
1991	6	224.70	1788	12	215.80	1960	1	207.20	
1787	1	224.50	1917	9	215.60	1999	6	207.20	
1837	3	224.30	1958	11	215.60	1779	10	206.70	
1948	8	223.70	1917	12	215.50	1860	6	206.30	
1957	8	223.70	2024	8	215.50	1946	11	206.30	
1837	8	223.50	1893	8	215.40	1837	10	206.20	
1947	7	223.50	1788	2	215.30	1789	5	205.80	
1778	3	223.30	1848	6	214.90	1949	9	205.80	
1789	11	223.30	1937	2	214.10	1936	12	205.60	
1949	3	223.00	2001	12	213.40	1959	9	205.60	
1957	3	222.80	1787	7	213.30	1789	4	205.50	
1980	6	222.70	1991	12	212.60	1981	7	205.30	
1860	7	221.90	1790	2	212.50	1938	11	203.60	

- **The statistical Method**

We employed the Box-Jenkins method (Box, et al., 2016) for time series analysis, specifically using the same variable. This approach, known as autoregressive integrated moving average models (ARIMA), focuses on extracting expected changes in observed data. ARIMA provides a systematic way to identify the suitable model for analyzing time series phenomena. The estimated parameters play a crucial role in assessing the model's accuracy and subsequently predicting future values for our studied phenomena.

The Autoregressive model (AR) establishes a relationship between our current observation and a set number of lagged observations. To make the time series stationary, these observations are integrated, involving the difference between consecutive raw observations. On the other hand, the Moving Average model (MA) relies on the connection between an observation and the residual error derived from a moving average model applied to lag observations.

In the context of ARIMA, a standard notation (p, d, q) is employed, with the parameters represented by integer values to indicate the specific ARIMA model in use. The lag order (p), referring to the number of lag observations included in the model, is determined. The degree of differencing (d), or the number of times raw observations are differenced, plays a role in making the time series stationary. Additionally, the order of the moving average (q) determines the size of the moving average window (Box, et al., 2016). We can explain the categories of Box-Jenkins time series models as follows:

1- Autoregressive Models

AR(1) of the 1st degree is of the following form:

$$y_t = C + \varphi y_{t-1} + e_t \quad (1)$$

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

$$y_t = C + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + e_t \quad (2)$$

2- Moving Average Models

MA(1) of the 1st order model is of the following form:

$$y_t = C + e_t + \theta e_{t-1} \quad (3)$$

In General, the formula of (q) order MA(q) is:

$$y_t = C + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q} \quad (4)$$

3- Mixed model (ARMA)

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

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

$$y_t = C + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q} \quad (5)$$

Model (5) is generalized when we take the differences between the values of the series until stationary, if it is non-stationary. The mixed integrated model is called the Autoregressive Integrated Moving Average ARIMA (p,d,q) model, where the letter (d) refers to the number of times the differences of the series until stationary, i.e. d=1, 2, 3, ...

Where,

y_t: refers to the time series or the values of the phenomenon in the different time periods t= 0, 1, 2, ..., m.

φ_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 different time periods t = 0, 1, 2, ..., m.

C: refers to constant.

- **Model identification**

In our pursuit of identifying the most suitable model for the time series data, we plotted the observed sunspot number (SSN) data. The time series graph illustrates its regular nature, representing the original monthly mean sunspot number (SSN) data observed by SILSO, World Data Center, from 1759 to August 2024, as shown in Figure 1.

Fig. 1: The original time series of mean monthly sunspot number (SSN) during 1749–August 2024.

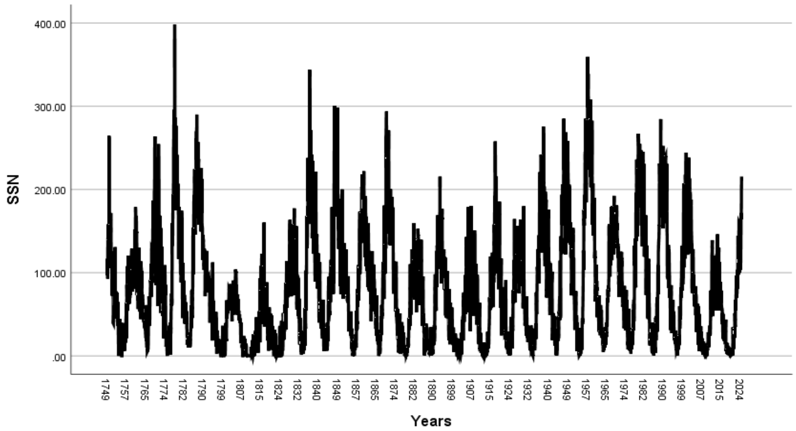


Fig. (1): Shows that the original time series of mean monthly sunspot number (SSN) data observed by SILSO, World Data Center, Royal Observatory of Belgium, Brussels during 1749 until August 2024.

We study both the autocorrelation function (ACF), and the partial autocorrelation function (PACF). The ACF defines how data points in a time series are correlated, on average, to the preceding data points (Box, G.E.P. et al 2016). On the other hand, the PACF reveals the partial correlation of a stationary time series with its own lagged values, offering insights into the specific contributions of each lag to the current observation. Autocorrelation and partial autocorrelation coefficients are crucial indicators for identifying suitable models, and measuring the degree of correlation between data points. Figs. 2 and 3, ACF takes the exponential form while PACF finished after the first order. So, the primary model is the autoregressive from first order, AR(1).

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Fig. (2): autocorrelation function of main series

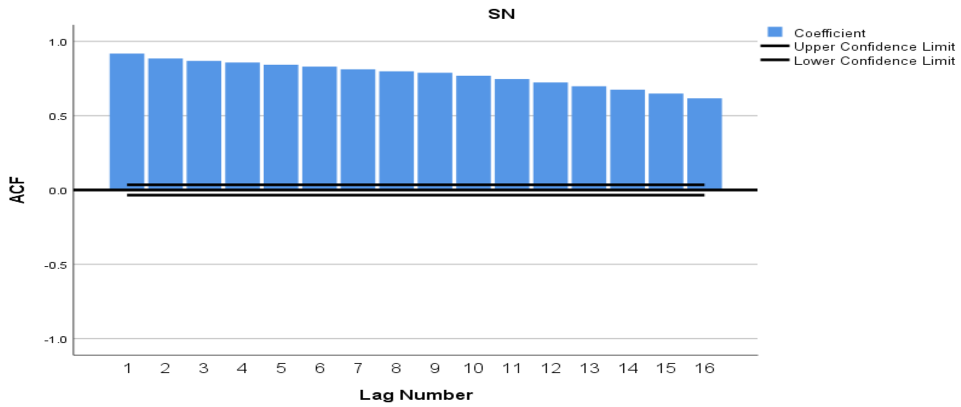
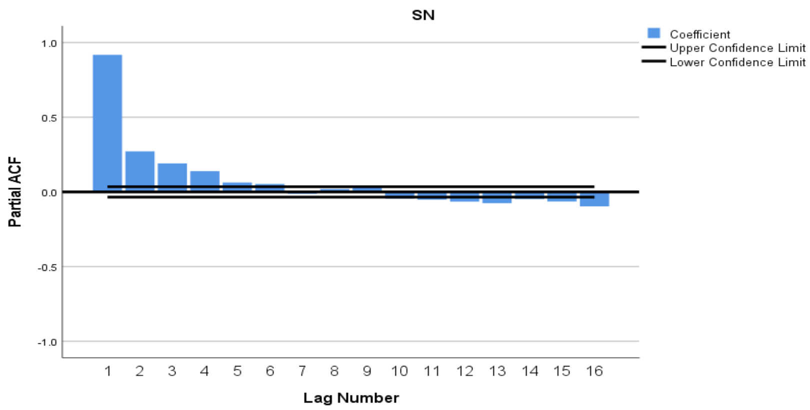
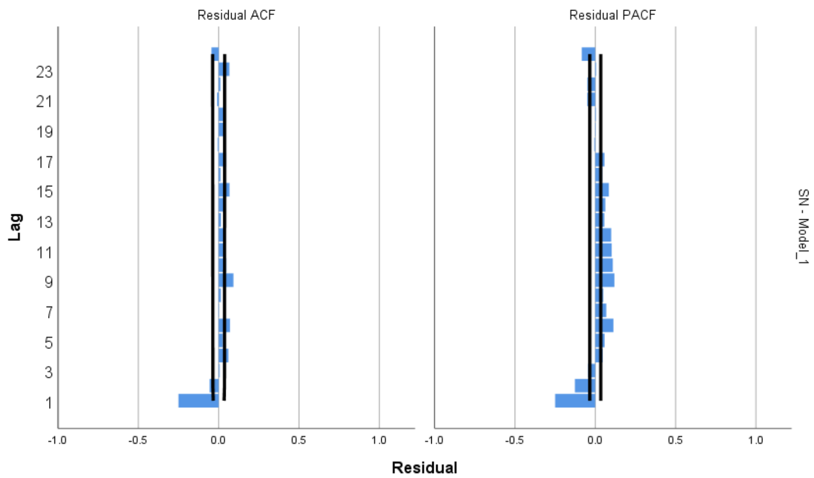


Fig. (3): partial autocorrelation function of main series



To test the proposed model, the residuals of the ACF and PACF functions must be examined for any defect in the conditions required for the validity of this model, and these conditions are that the ACF and PACF functions of the residuals are not auto correlated and do not have a specific pattern. By examining Fig. 4 for the residuals, we find that some of them are correlated and have a specific pattern in several lags in the ACF and PACF, so the proposed model is acceptable but not in the desired form for prediction. Therefore, other models must be tested and compared according to other statistical measures.

Fig. 4 The autocorrelation and partial autocorrelation functions for the residuals of AR(1).



Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Our model selection process extends beyond relying solely on autocorrelation functions. We incorporate other statistical measurements to identify the optimal model for predicting the sunspot numbers of the 25th cycle. Several ARMA models were tested and the model selection criteria. The best model is verified using the Bayesian Information Criterion (BIC) and Root Mean Square Error (RMSE) method, where we choose the lowest in these two measures as well as the highest in R-squared (R^2). After applying these measures, we reached the results below in Table (5).

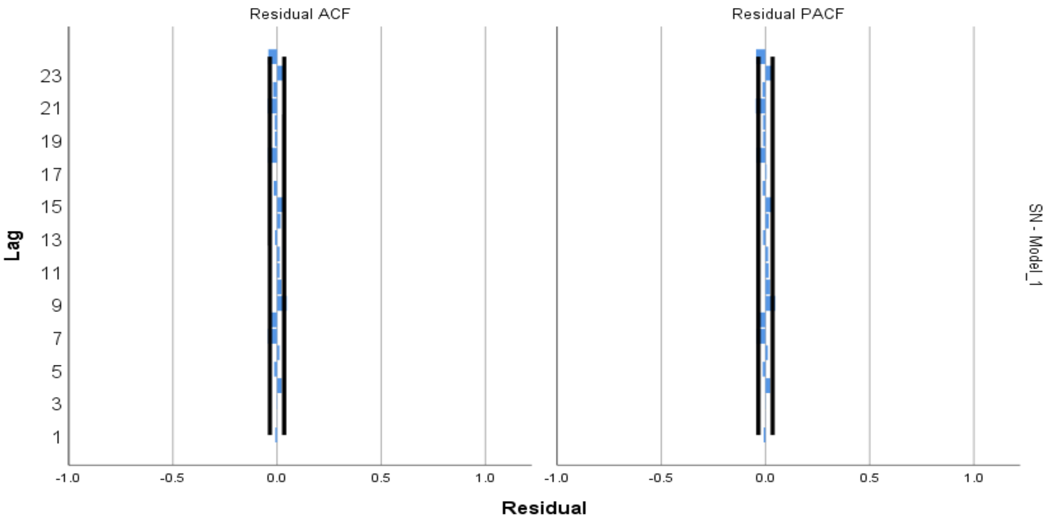
Table (5) shows the models and the values of their different scales:

Model	R^2	BIC	RMSE
ARMA(1,0)	0.843	6.586	26.829
ARMA(2,0)	0.693	7.260	37.524
ARMA(2,1)	0.863	6.455	25.068
ARMA(1,1)	0.861	6.465	25.222
ARMA(1,2)	0.863	6.455	25.066
ARMA(2,2)	0.862	6.470	25.215
ARMA(3,2)	**0.870**	**6.411**	**24.453**
ARMA(3,0)	0.860	6.476	25.331
ARMA(3,3)	0.863	6.461	25.050
ARMA(0,1)	0.529	7.686	46.502
ARMA(0,2)	0.679	7.303	38.346

The proposed model stands out with the lowest values among all tested models and a high determination coefficient (R^2). We ascertain that the most appropriate model for predicting sunspot numbers is ARMA(3,2).

To assess the suitability of our obtained model ARMA(3,2), we examine the residuals, as shown in Fig. (5). We observe that the residuals are independent of each other. Consequently, we conclude that this model is well-suited for estimating the predicted sunspot numbers.

Fig. 5 The autocorrelation and partial autocorrelation functions for the residuals of ARMA(3,2).



Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- **Parameters Estimation**

The general form equation of the primary ARMA (3,2) model is expressed as follows:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varphi_3 y_{t-3} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} \quad (6)$$

Where,

$(\varphi_1, \varphi_2, \varphi_3)$ and (θ_1, θ_2) are the parameters of the model while e_{t-1}, e_{t-2} are the lags for the residuals of the estimation of variable y_t, $e_t \sim N(0, \sigma_t^2)$.

The parameters in Eq. (6) were estimated using nonlinear least squares, resulting in the following equation:

$$y_t = -1.56 + 2.341\, y_{t-1} - 1.699\, y_{t-2} + 0.357 y_{t-3} + e_t + 1.803\, e_{t-1} + -0.82\, e_{t-2} \quad (7)$$

RESULTS AND DISCUSSIONS

Using Equation (7), we can obtain the prediction of the number of sunspots for 76 months from September 2024 to December 2030, the end of the cycle, and the forecast data of the National Oceanic and Atmospheric Administration (NOAA) as shown in Table 7.

We compared the observed data from 1749 to August 2024 with the results from ARMA(3,2) for the same time period and Calculating the correlation coefficient between our predictions and the actual SSN observations (over 3308 months), we found a strong correlation of **93.3%**, as shown in Table (6). This indicates that the model is suitable for predicting the 25th solar cycle up to its end in December 2030. Figure (6) illustrates the main observation series, the fitted model, and the forecasted values for this cycle and we see identical in the curve between them.

Fig. (6): The observed series, fitted model and forecasting values (after the reference line) for 25th cycle

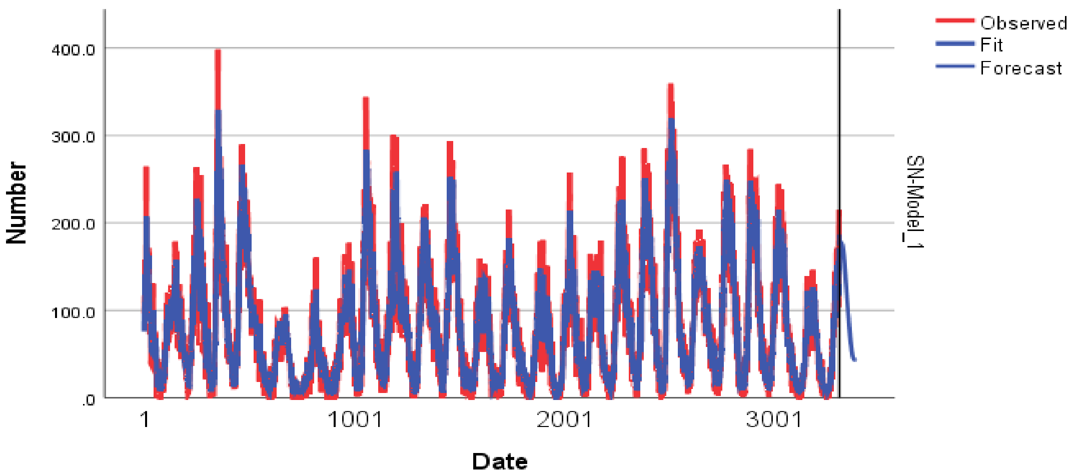


Table (6) the correlation coefficient of our predicted data and observed data from 1749 to Aug. 2024

		Observed SN	Predicted value from Model ARMA(3,2)
SSN	Pearson Correlation	1	0.933**
	Sig. (2-tailed)		0.000
	N		3308

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

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table (7) our predicted data and NOAA predicted data of 25th solar cycle.

Year	month	Our prediction	NOAA prediction	Year	Month	Our prediction	NOAA prediction	Year	month	Our prediction	NOAA prediction
2024	Sep	186.8	108.4	2027	Jun	125.5	89.2	2030	Mar	45.4	29.2
2024	Oct	177.5	109.6	2027	Jul	122.2	87.3	2030	Apr	44.8	27.9
2024	Nov	175.0	110.7	2027	Aug	118.8	85.4	2030	May	44.3	26.6
2024	Dec	175.0	111.6	2027	Sep	115.4	83.4	2030	Jun	44.0	25.3
2025	Jan	175.6	112.4	2027	Oct	112.0	81.5	2030	Jul	43.7	24.1
2025	Feb	176.3	113.1	2027	Nov	108.7	79.5	2030	Aug	43.6	23.0
2025	Mar	176.9	113.6	2027	Dec	105.3	77.5	2030	Sep	43.6	21.8
2025	Apr	177.4	114.0	2028	Jan	102.1	75.5	2030	Oct	43.7	20.8
2025	May	177.5	114.3	2028	Feb	98.8	73.5	2030	Nov	43.9	19.7
2025	Jun	177.5	114.5	2028	Mar	95.6	71.5	2030	Dec	44.3	18.7
2025	Jul	177.2	114.6	2028	Apr	92.4	69.4				
2025	Aug	176.7	114.5	2028	May	89.2	67.4				
2025	Sep	176.0	114.3	2028	Jun	86.2	65.4				
2025	Oct	175.1	114.1	2028	Jul	83.2	63.4				
2025	Nov	174.0	113.6	2028	Aug	80.3	61.5				
2025	Dec	172.7	113.1	2028	Sep	77.5	59.5				
2026	Jan	171.2	112.5	2028	Oct	74.7	57.6				
2026	Feb	169.6	111.8	2028	Nov	72.1	55.6				
2026	Mar	167.7	110.9	2028	Dec	69.5	53.7				
2026	Apr	165.7	110.0	2029	Jan	67.1	51.9				
2026	May	163.5	109.0	2029	Feb	64.8	50.0				
2026	Jun	161.2	107.9	2029	Mar	62.6	48.2				
2026	Jul	158.8	106.7	2029	Apr	60.5	46.4				
2026	Aug	156.2	105.4	2029	May	58.5	44.7				
2026	Sep	153.5	104.1	2029	Jun	56.6	43.0				
2026	Oct	150.7	102.7	2029	Jul	54.9	41.3				
2026	Nov	147.8	101.2	2029	Aug	53.2	39.6				
2026	Dec	144.8	99.6	2029	Sep	51.7	38.0				
2027	Jan	141.8	98.0	2029	Oct	50.4	36.4				
2027	Feb	138.6	96.3	2029	Nov	49.1	34.9				
2027	Mar	135.4	94.6	2029	Dec	48.0	33.4				
2027	Apr	132.1	92.8	2030	Jan	47.0	32.0				
2027	May	128.8	91.0	2030	Feb	46.2	30.6				

In Table (7), we present our predicted results alongside NOAA's predictions for the 25th solar cycle from Sept. 2024 to December 2030. The correlation coefficient between our predictions and NOAA's is 98.7%, indicating very strong correlation and good agreement, as summarized in Table (8).

Table (8) correlation coefficient between our prediction and NOAA predicted data of 25th solar cycle from Sept. 2024 to Dec. 2030

SSN	SN_Prediction		SN_NOAA
	Pearson Correlation	1	0.987**
	Sig. (2-tailed)		0.000
	N		76

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

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

Fig. 7 shows the comparison between our predictions of sunspot numbers from September 2024 to December 2030 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.

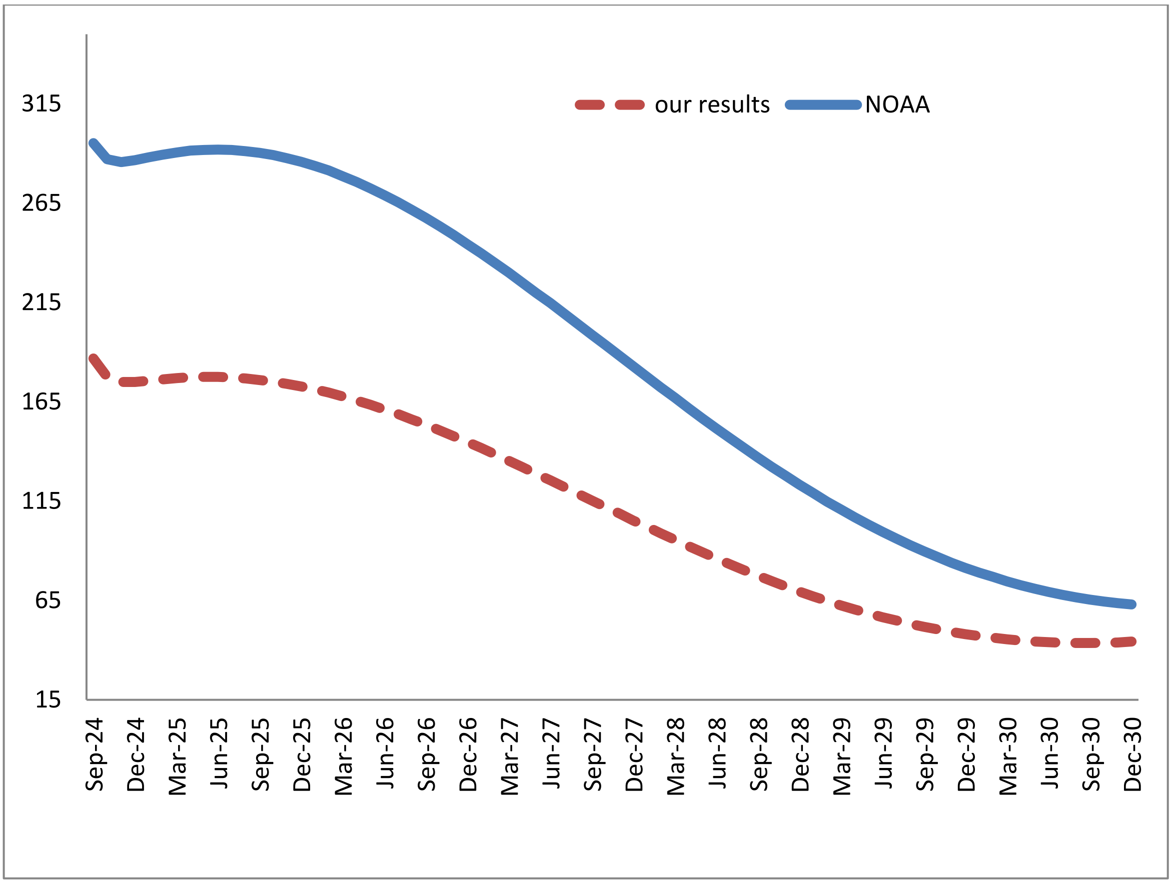


Fig. 7 Comparison between our predicted values of monthly mean sunspot numbers with that predicted by NOAA from Sept. 2024 to Dec. 2030.

CONCLUSION

In this study, we developed a new statistical model for predicting the sunspot numbers of the 25th solar cycle, utilizing observed sunspot data spanning 276 years from 1749 to August 2024. The original time series of monthly mean sunspot number (SSN) data, observed by SILSO, World Data Center.

After examining the autocorrelation (ACF) and partial autocorrelation functions (PACF) of the observed series data, we found that the data goes to the first-order autoregressive, AR(1), but after examining ACF and PACF for the residuals, we found that there is a defect in the conditions, so we found that this model will have a large error in its predictions.

Since the autocorrelation and partial autocorrelation functions were not the only ones in determining the optimal model for prediction, we tried several models shown in Table (5) and examined them according to accurate statistical measures such as (BIC, RMSE and R-squared) to help us determine the correct model.

Based on these measures, the optimal model was chosen and we found it to be the ARMA(3,2) model. After examining the (ACF and PACF) functions, we found that there is no autocorrelation in the residuals and they do not have a specific pattern, which indicates that the model is optimal in prediction.

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The parameters of our deduced model were estimated using nonlinear least squares, resulting in the following equation:

$$y_t = -1.56 + 2.341\, y_{t-1} - 1.699\, y_{t-2} + 0.357 y_{t-3} + e_t + 1.803\, e_{t-1} + -0.82\, e_{t-2}$$

We validated our model by comparing the estimated sunspot numbers with observed SSN data from 1749 to August 2024 (3308 months), yielding a robust correlation coefficient of 93.3%. Furthermore, we assessed the correlation between our predicted data and NOAA's published predictions, revealing an 98.7% correlation indicating strong agreement.

Finally, our predicted model ARMA(3,2) demonstrates its efficacy in estimating sunspot numbers, showcasing very strong correlations with observational data and strong agreement with NOAA predictions.

Data availability

The data that support the findings of this study are openly available (http://www.sidc.be/silso/) as well as the comparison of sunspot number prediction from NOAA (https://www.noaa.gov/).

REFERENCES

[1] Abdel Rahman, H. I. & Marzouk, B. A., 2018, "Statistical method to predict the sunspots number". NRIAG Journal of Astronomy and Geophysics, Volume 7, 2018 - Issue 2.
[2] Baranovski, A.I.; Clette, F.; and Nollau, V.; 2008, Nonlinear solar cycle forecasting theory and perspectives. Ann.Gephys. J.26,231-234.
[3] Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994), Time Series Analysis; Forecasting and Control. 3rd Edition, Prentice Hall, Englewood Cliff, New Jersey.
[4] Hathaway, D. H., 2009, Solar cycle forecasting, Space Sci. Rev., 144 (2009), pp. 401-412.
[5] Kane, R. P., 2008, Prediction of solar cycle maximum using solar cycle lengths, Sol. Phys., 248 (1), pp. 203-209.
[6] Podladchikova, T. and Van der Linden, R., 2011, J. Space Weather Space Climate, 1, p. A01.
[7] Sabry M. A., Issa, I. A., Abdel Rahman, H. I. and A. A. Shaker, 2012, The tree clustering technique and the physical reality of galaxy groups, NRIAG Journal of Astronomy and Geophysics 1, 81-86.
[8] Sarp, V; Kilcik, A; Yurchyshyn, V; Rozelot, J P; Ozguc, A (December 2018). "Prediction of solar cycle 25: a non-linear approach". *Monthly Notices of the Royal Astronomical Society*. **481** (3): 2981–2985.
[9] SILSO, World Data Center, Sunspot Number and Long-term Solar Observations, Royal Observatory of Belgium, *Monthly Report on the International Sunspot Number,* on-line Sunspot Number catalogue: http://www.sidc.be/SILSO/, '1749 to August 2024'.
[10] Upton, Lisa A.; Hathaway, David H., (28 August 2018), An updated solar cycle 25 prediction with AFT: The modern minimum. *Geophysical Research Letters*. **45** (16): 8091- 8095. arXiv:1808.04868. Bibcode:2018GeoRL..45.8091U. doi:10.1029/2018GL078387.S2CID 118913702
[11] https://www.sidc.be/SILSO/datafiles
[12] https://www.noaa.gov/