Journal of the Korean Astronomical Society
50: 21 ∼ 27, 2017 April

https://doi.org/10.5303/JKAS.2017.50.2.21
pISSN: 1225-4614 · eISSN: 2288-890X
http://jkas.kas.org

Performance of the Autoregressive Method in Long-Term Prediction of Sunspot Number

Jongchul Chae[1] and Yeon Han Kim[2]

[1]Department of Physics and Astronomy, Seoul National University, Gwanak-gu, Seoul 08826, Korea; jcchae@snu.ac.kr
[2]Solar and Space Weather Group, Korea Astronomy and Space Science Institute, Daejeon 34055, Korea

Received November 25, 2016 ; accepted January 6, 2017

Abstract: The autoregressive method provides a univariate procedure to predict the future sunspot number (SSN) based on past record. The strength of this method lies in the possibility that from past data it yields the SSN in the future as a function of time. On the other hand, its major limitation comes from the intrinsic complexity of solar magnetic activity that may deviate from the linear stationary process assumption that is the basis of the autoregressive model. By analyzing the residual errors produced by the method, we have obtained the following conclusions: (1) the optimal duration of the past time for the forecast is found to be 8.5 years; (2) the standard error increases with prediction horizon and the errors are mostly systematic ones resulting from the incompleteness of the autoregressive model; (3) there is a tendency that the predicted value is underestimated in the activity rising phase, while it is overestimated in the declining phase; (5) the model prediction of a new Solar Cycle is fairly good when it is similar to the previous one, but is bad when the new cycle is much different from the previous one; (6) a reasonably good prediction of a new cycle can be made using the AR model 1.5 years after the start of the cycle. In addition, we predict the next cycle (Solar Cycle 25) will reach the peak in 2024 at the activity level similar to the current cycle.

Key words: solar activity; sunspot number

1. Introduction

One can better prepare for the future if one knows what will happen in the future. Since the magnetic activity of the Sun dominates the space environment, planning space missions requires predictions of future solar magnetic activity. For instance, the plan of a low-altitude satellite mission needs anticipating the orbital decay by the atmospheric drag and the quality level of the payloads for safe operation is determined by the knowledge of the space radiation hazard. Both the atmospheric drag and the radiation hazard strongly depend on the solar magnetic activity.

Predicting the future is not an easy task. For a successful prediction, one should have belief that there exist a pattern of the nature describing not only the past, but also the future, and that one can find the pattern. One also has to have data from the past. Using the past data and the pattern, one can make a prediction of the future. But what if there is no pattern or if the pattern is too complex to be discovered or if the past data are incomplete?

Solar magnetic activity is usually predicted using the sunspot number (SSN) data accumulated since 1749. There is an implicit belief among scientists that the SSN follows a pattern. The pattern, however, is not simply described. A number of scientific methods have been proposed for the long-term prediction of the SSN, and there exist large differences in the SSN prediction among the different methods (Pesnell 2008, 2012).

Corresponding author: J. Chae

The autoregressive(AR) method is one of the methods often used for the prediction of the SSN. Yule (1927) supposed that the SSN at a year t is given by the linear combination of the SSNs at two previous years and determined the linear coefficients from the SSN data. This approach was generalized by Walker (1931) as the AR model. The approach used by Yule (1927) corresponds to the AR model of order 2. Moran (1954) argued that a higher order is needed for the AR method. Recently Werner (2012) proposed that the AR model of order 9 is the most suited for predicting SSN data.

The objective of the present work is to examine the performance of the AR method more in detail. A similar study was done by Werner (2012) using the V1.0 data. Our study is based on V2.0 data published in 2015 and emphasizes the error analysis and revealing the origin of systematic errors. We expect our study to contribute to a better understanding of both the AR method itself and the physical nature of the solar magnetic activity.

The applicability of the AR method to the SSN data could shed light on the physics of the solar dynamo. If the AR method turns out to be successful in modeling the SSN data, one could conclude that the solar dynamo behaves like a harmonic oscillator where the future values linearly depend on the past values. If the AR method fails totally, the solar dynamo should be understood as a highly non-linear system where the future values are not a simple linear combination of the past values. In this case, either the nonlinear dependence on the past values or the dependence on other physical parameters of the Sun might be important.

2. DATA AND TRANSFORMATION

Our work is based on V2.0 sunspot data from the World Data Center SILSO (http://www.sidc.be/SILSO/), Royal Observatory of Belgium, Brussels. The main difference between V1.0 data and V2.0 data is that the V2.0 data does not multiply the conventional 0.6 Zürich scale factor, to be consistent with the level of modern sunspot counts made since 19th century. In fact, each observing station used its own conversion factor that was around 0.6 and depended on the telescope. For instance, Sim et al. (2001) estimated the conversion factor at 0.72 for one telescope, and at 0.56 for another telescope. The new conversion factor of each observing station to match with the V2.0 SILSO data is now expected to be around 1.0.

Specifically we use the time series of monthly averaged SSN (v2.0) sampled every month from the year 1749. By convolving this data with the 13-element kernel $K = [0.5, 1, 1, \cdots, 1, 0.5]/12$ and re-sampling, we obtain the time series of annually averaged SSN X_j sampled at every $\delta t = 0.5$ yr.

It is convenient to apply the AR model to

$$x_j \equiv \sqrt{X_j} \tag{1}$$

rather than to X_j directly at time t_j for $j = 0, 1, \cdots, N-1$. This transformation obviously guarantees the positivity of X_j. It also make it easy to estimate the standard error. Since X_j is a positive-definite integer number, it may have a Poisson noise where the standard error is proportional to $\sqrt{X_j}$ or $\sigma(X) = a\sqrt{X}$ with a constant a. If error is sufficiently small, the analysis of error propagation using Equation (1) then leads to the relation $\sigma(x) = a/2$, which indicates that the error $\sigma(x)$ is independent of the signal x while $\sigma(X)$ is not, and hence can be more easily estimated from the data than $\sigma(X)$. Once $\sigma(x)$ is known from the data anyway, the standard lower and upper bounds of $X = x^2$ are known to be $(x \pm \sigma(x))^2$, respectively, that are much asymmetric about X if $\sigma(x)$ is not much smaller than x.

Figure 1 presents the plots of X_j and x_j. We find that the asymmetry between the maxima and the minima clearly seen in the plot of X_j is much reduced in the plot of x_j, so the plot of x_j is more like a sinusoidal function than that of X_j.

3. MODEL

We choose t_i for some $i \leq N-1$ as the "supposed" present time. The past data x_j for $j \leq i$ are known from the measurements, and their mean value is denoted as $\bar{x}$. The model value of the next step in the future y_{j+1} can be determined from the past data using the autoregressive (AR) model of order P

$$y_{i+1} - \bar{x} = \sum_{k=0}^{P-1} \alpha_k (x_{i-k} - \bar{x}) \tag{2}$$

which states that the next step value is given by a linear combination of P values in the past. Once P is specified, the coefficients α_k can be determined from the past data by minimizing the functional

$$\chi^2 = \sum_{j=i-M}^{i-1} (x_{j+1} - y_{j+1})^2 \tag{3}$$

where M is the number of data points used for the determination of the coefficients. Equations (2) and (3) indicate that the prediction y_{i+1} directly depends on the near past of short duration $P\delta t$, and indirectly on the extended past of longer duration $M\delta t$. Once y_{i+1} is determined using Equation (2), it can be considered as the addition of data point, and hence it is possible to proceed another step for y_{i+2}. Thus for $l \geq 2$, we can write the forecast AR model as

$$y_{i+l} - \bar{x} = \sum_{k=0}^{l-2} \alpha_k (y_{i+l-1-k} - \bar{x}) + \sum_{k=l-1}^{P-1} \alpha_k (x_{i-k} - \bar{x}) \,. \tag{4}$$

For $i+l \leq N-1$, the real data in the supposed future at times t_{i+l} are available and can be compared with the models. The error of prediction in the supposed future is then known as

$$\epsilon_{i,l} = y_{i+l} - x_{i+l} \tag{5}$$

which is given as a function of the "present" t_i for $i = j_1, .., j_2$ and the length of extension into the "future" $l\delta t$. Here j_1 and j_2 are the indices of the first and last data points used for the construction of $\epsilon_{i,l}$ depending on l. The mean standard error of prediction is given by

$$\sigma_l = \sqrt{\frac{1}{j_2 - j_1 + 1} \sum_{i=j_1}^{i=j_2} \epsilon_{i,l}^2} \,. \tag{6}$$

4. ANALYSIS

4.1. Parameters

After a number of experiments, we find the choice of $M\delta t = 120$ yr leads to the smallest error prediction error. As $M\delta t$ deviates from this value, the error increases slowly. This value is larger than 92 yr reported by Kim (1991) as the period of the long-term modulation of SSN. More important in the model is $P\delta t$, the duration of the near past that directly affects the forecast. We have chosen $P\delta t = 8.5$ yr based on the investigation of the prediction errors. Figure 2 shows that the prediction errors of long-term prediction ($l\delta t$ =3, 10 yr) are more sensitive to P than those of short-term prediction ($l\delta t$ =0.5, 1 yr). The long-term prediction errors become minima at $P\delta t = 8.5$ yr.

Figure 3 illustrates the plot of α_k determined at two specific times t_i. Its sign alternates between even and odd values of k, and its absolute value exponentially decreases with k. The coefficients slowly vary with time t_i.

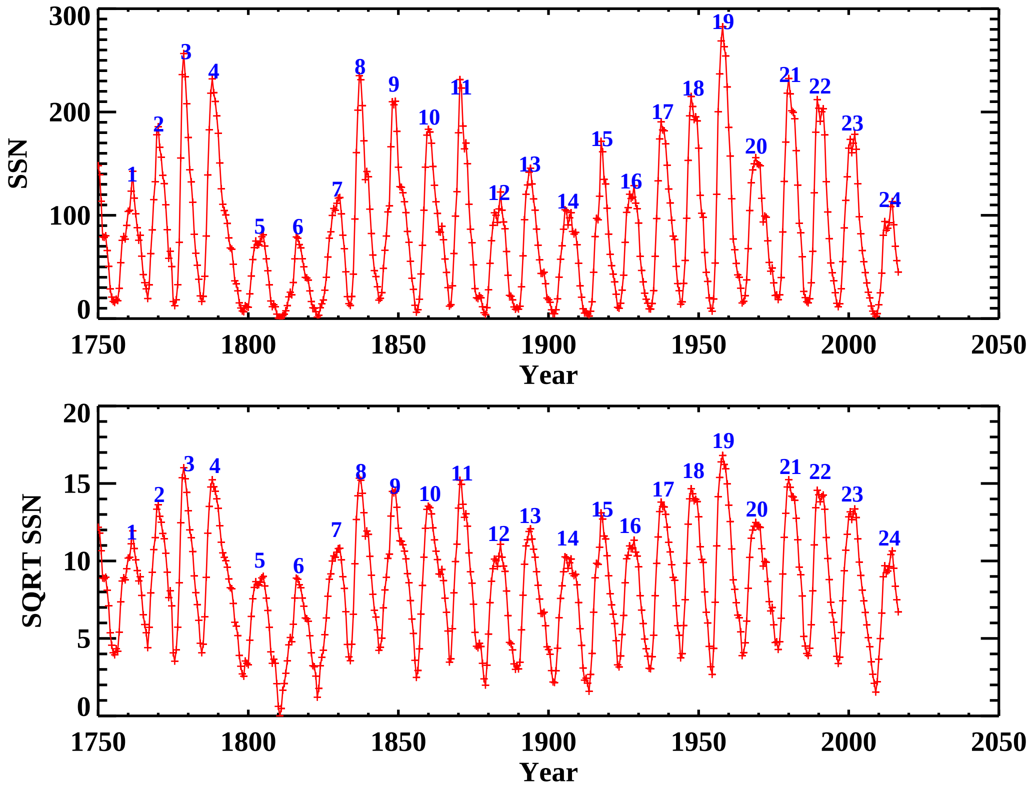

Figure 1. Plots of SSN data X_j and their square root values x_j used for the present study.

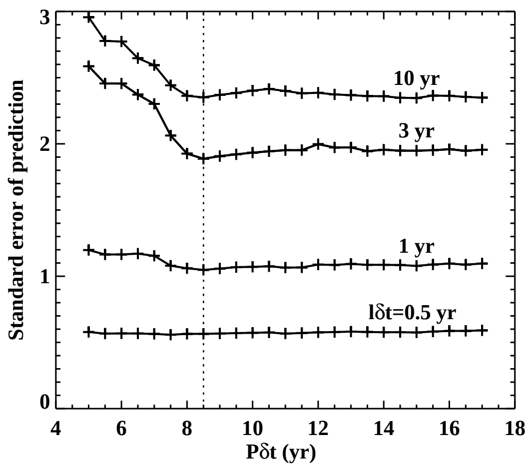

Figure 2. Dependence of prediction error σ_l on P for the different case of $l\delta t$.

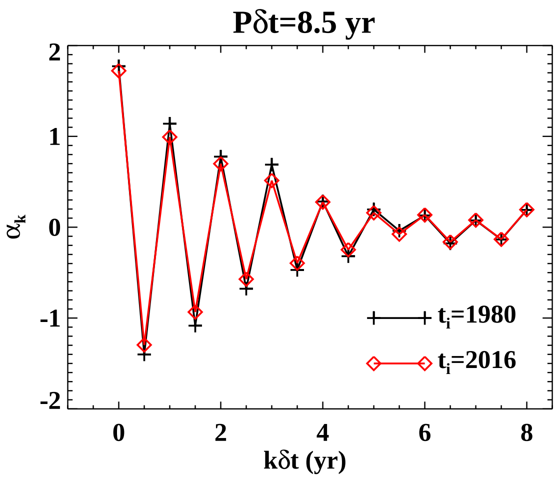

Figure 3. Plots of the coefficient α_k as a function of $k\delta t$ with $P\delta t = 8.5$ yr at $t_i = 1980$ and 2016, respectively.

4.2. Dependence of Standard Error on Prediction Horizon

Figure 4 presents the error of prediction $\epsilon_{i,l}$ as a function of time for the three different values of prediction horizon $l\delta t$. It is obvious from the figure that the mean magnitude of error increases with $l\delta t$. A careful comparison of the error ϵ and the signal x reveals that the error is anti-correlated with the signal, with the absolute value of the correlation coefficient increasing with $l\delta t$. In case of $l\delta t \geq 3$ yr, the correlation coefficient is found to be around -0.6. The strong anti-correlation indicates that most of the error may be a systematic one arising from the limitation of the prediction model. The model tends to underestimate the SSN during the maxima, and overestimate it during the minima.

Figure 5 clearly indicates that the standard error increases with prediction horizon. It increase rapidly with $l\delta t$ for about two years. Then, the standard error either slowly increases or remains constant as $l\delta t$

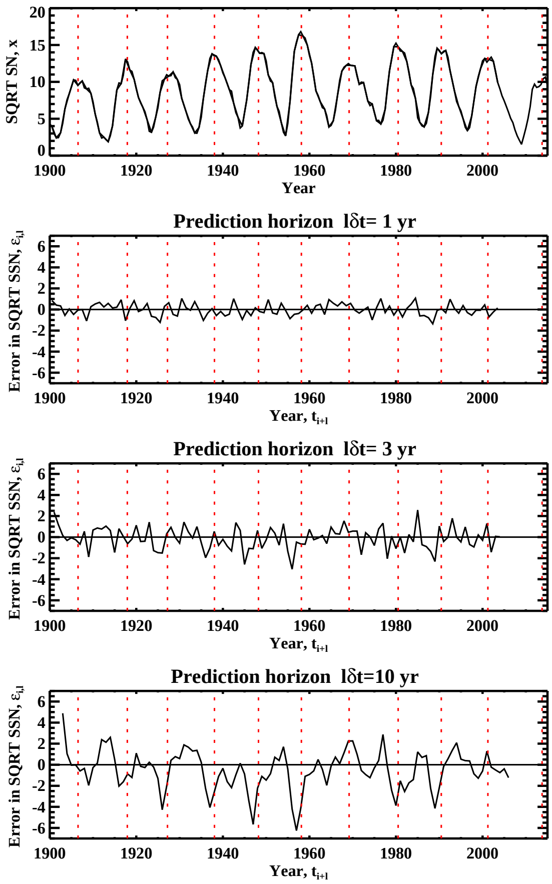

Figure 4. Plots of error $\epsilon_{i,l}$.

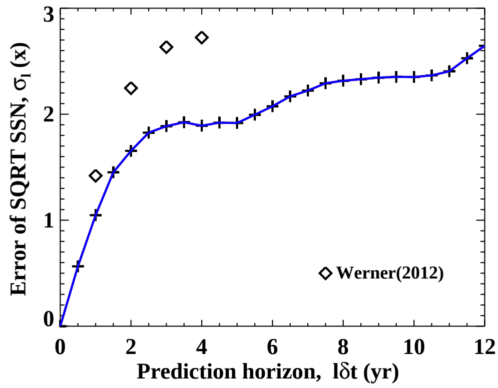

Figure 5. Plot of root-mean-square error σ_l. as a function of prediction horizon. The errors of Werner (2012) obtained from the V1.0 data have been multiplied by the factor of $\sqrt{1/0.6}$ to be compared with our results from the V2.0 data.

increases. This pattern is qualitatively compatible with the result obtained by Werner (2012). The difference is that the errors in our implementation of the AR model are smaller than those of Werner (2012). This may be partly because we have determined the model coefficients separately for each time of prediction while Werner (2012) used the same set of coefficients for all the prediction. The finer sampling of our data (0.5 yr) may also have contributed to the better performance of our implementation.

We find that the standard error $\sigma(x)$ has a value between 2 and 2.5 in the long-term prediction with prediction horizon $l\delta t = 3$ to 10 yr. Taking $\sigma(x) = 2.2$, we can estimate the upper and lower bound of SSN X when x is known. Let's consider, for instance, the case $X = 83$, which is the mean of the SSN data, or of $x = 9.1$. The upper and lower bounds are given by $X_u = (9.1 + 2.2)^2 = 128$, and $X_l = (9.1 - 2.2)^2 = 48$, respectively. The difference between the upper bound and the prediction value is $X_u - X = 45$, while that between the prediction value and the lower bound is $X - X_l = 35$. These values are as big as 54% and 42% of the prediction value.

For shorter-term prediction with $l\delta t = 1$ yr,the error is much smaller. With $\sigma(x) = 1.0$ in this case, the upper bound and the lower bound are calculated to be $X_u = 102$ and $X_l = 66$, respectively. The differences are as small as 23% and 20% of the prediction value $X = 83$, respectively.

4.3. Error in Peak SSN and Time

The maximum SSN and the time of maximum occurrence are two important parameters characterizing a Solar Cycle. We have examined how well the AR model can predict these parameters.

Figure 6 indicates that the goodness of the model for the prediction of the maximum SSN and peak time depends on the time for which the prediction is sought. If the prediction is made before the start of the cycle, which is identified by the time of the minimum SSN, the standard error of the maximum SSN ranges between 60 and 70. The error is mostly systematic, as can be seen from the fact that the mean error ranges between -40 and -30, being much below zero. As for the peak time prediction, the standard error is around 1 yr and the mean error is about 0.7 yr. The AR model thus tends to underestimate predictions for the maximum SSN by an amount of about 35 at a delayed peak time by the amount of 0.7 yr.

The figure also indicates that the AR model prediction is quite good when the prediction is made more than 1.5 year after the start of the cycle. The standard errors of the maximum SSN and peak time are reduced to below 35 and 0.7 yr, respectively. Moreover, their mean errors become very close to zero, which indicating the suppression of systematic errors. There seem to be two reasons why the AR model well operates in this case. First, most of the information about a new cycle seems to be accumulated 1.5 year after the beginning of the new cycle. Second, the prediction horizon for the maximum (time difference between the peak time and the time of prediction) is small.

5. Examples of Long-term Prediction

Figure 7 presents long-term predictions of SSN made in the different cycles. The time of each prediction was

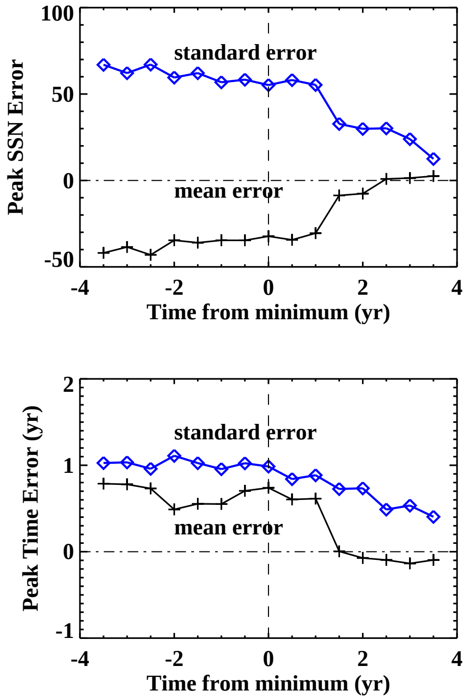

Figure 6. Dependence of prediction errors of peak SSN (top) and peak time (bottom) on the time of prediction.

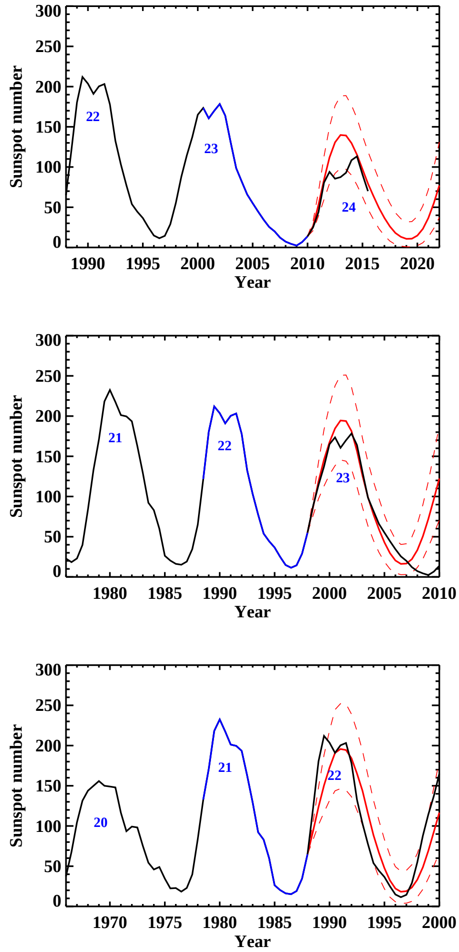

Figure 7. Test of long-term prediction of SSN made for Solar Cycles 24, 23, and 22, respectively. The black curve refers to the real data or true values. The blue colored part is the near-past data directly used for the prediction. The solid red curve is the prediction profile, and the two dashed curves represent the upper bound and lower bound for the prediction.

taken to be 1.5 year after the start of each cycle. We first find that in all the cases the model prediction is fairly consistent with the actual data within the range defined by the upper and lower bounds of prediction. The model particularly well describes the increasing phase and declining phase in Solar Cycles 24 and 23. The prediction of Solar Cycle 22 is a little worse, but still satisfactory.

Figure 7 clearly indicates that the AR method is not able to account for a solar maximum with two peaks. In all the cases, the model predicts a single regular peak of SSN, while the data show a complex variation with two peaks. This failure to predict the complex variation of the SSN is partly responsible for the errors in the maximum SSN and the peak time shown in Figure 6.

Figure 8 presents the test predictions for Solar Cycles 21, 20 and 18. These predictions are worse than those for Solar Cycles 24, 23 and 22. The prediction for Solar Cycle 20 was the worst. It seems that the prediction error becomes large when the amplitude of the new cycle is very different from the previous one because the AR method tends to predict that the new cycle is similar to the previous one. The successful prediction for Solar Cycle 23 shown in Figure 7 may be attributed to the fact that Solar Cycle 23 was similar to Solar Cycle 22. The prediction for Solar Cycle 20 was not good since this cycle was not similar to Solar Cycle 21.

Figure 9 further demonstrates that the success of the AR prediction of a new cycle depends on the similarity between the new cycle and the old one. The 1994 prediction of Solar Cycle 23 is very successful even though the prediction was made a few years before the start of Solar Cycle 23. This success should be attributed to the similarity of Solar Cycle 23 to Solar Cycle 22. The 1974 prediction of Solar Cycle 21 is far from satisfactory. It seriously underestimates the SSN at the solar maximum, and overestimates it during the minimum. This unsatisfactory behavior results because Solar Cycle 21 is much different from Solar Cycle 20. It appears that Solar Cycle 21 did not strongly depend on the previous cycle, making the prediction very difficult. In a sense, the development of Cycle 21 was like an impulsive event which occurs independent of the past. The proper prediction of this kind of Solar Cycle can be made only when enough information of this cycle itself has been accumulated. The 1978 prediction made about 1.5 yr after the beginning (shown in Figure 8) was fairly good, because the SSN data acquired until

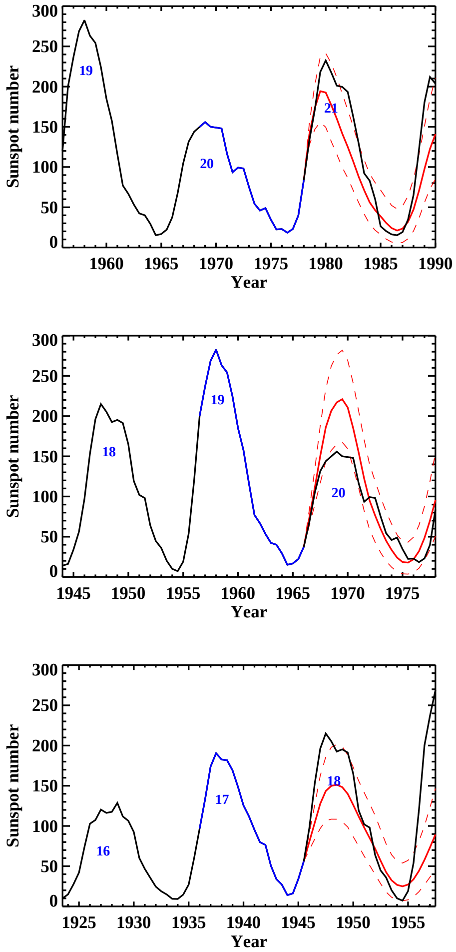

Figure 8. Test of long-term prediction of SSN made for Solar Cycles 21, 20, and 18 (from top to bottom), respectively.

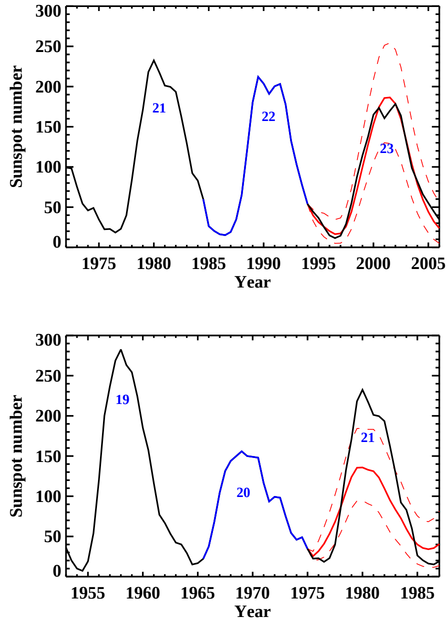

Figure 9. Test of long-term prediction of SSN made much before the start of Solar Cycles 23 and 21, respectively.

1978 contained much information on Solar Cycle 21.

Figure 10 presents the long-term prediction of SSN for Solar Cycle 25 made at the year 2016.5. According to this prediction, Solar Cycle 25 will begin around at 2019.5, in 3 years, and will reach the peak at 2024.5, in 8 years. Solar Cycle 25 looks similar to Solar Cycle 24, and will have the peak sunspot number of about 118. If we correct the predictions of the peak SSN, and the occurrence time for the mean systematic errors of 35 and 0.7 yr, we reach the conclusion that Solar Cycle 25 will have the peak SSN of 153 in the year 2024.0, which is still within the upper and the lower bounds of the AR prediction.

6. SUMMARY AND DISCUSSION

Our results are summarized as follows.

1. We have applied the AR model to the square root of the annually-averaged SSN data sampled every $\delta t = 0.5$ yr.
2. We found that the AR model of order $P = 17$ works the best. The corresponding value of $P\delta t = 8.5$ yr is quite compatible with $P\delta t = 9$ yr of Werner (2012) who chose $P = 9$ and $\delta t = 1$ yr. The 17 coefficients can be determined from the fitting of the SSN data obtained for 120 years before the time of prediction.
3. We have estimated the standard error in the predicted square root of SSN as a function of prediction horizon. The errors are mostly systematic ones resulting from the incompleteness of the model.
4. There is a tendency that the predicted value is underestimated in the activity rising phase, while it is overestimated in the declining phase.
5. For a reliable prediction of a new Solar Cycle, the prediction should be made at least 1.5 years after the start of the new cycle. In this case, the prediction profile is fairly compatible with the real SSN data within the upper and lower bounds of the prediction.
6. The performance of the AR prediction is good when the new cycle is similar to the previous one, and is bad when the new cycle is much different from the previous one.
7. The AR model predicts that Solar Cycle 25 will reach the maximum SSN of 118 during the middle of the year 2024. If the mean systematic errors inferred from the past data are corrected for, Solar Cycle 25 will reach the maximum SSN of 153 at the beginning of the year 2024.

It is known that the AR model is applicable to the time series of stationary data where the mean and standard deviation of the data do not change with time. This explains why the model is successful in making a long-term prediction of SSN when the new cycle is

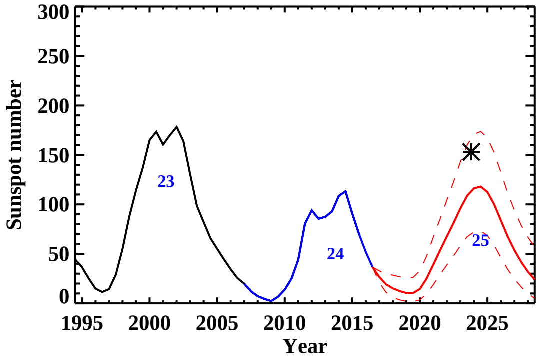

Figure 10. The AR model prediction for Solar Cycle 25. The symbol mark the time of peak occurrence and the peak SSN obtained by subtracting the mean systematic errors of the AR method.

similar to the previous one, and is not successful in predicting the new cycle that is much different from the previous one.

It is very difficult to determine in advance whether the new cycle will be similar to the previous one or not. In many cases, it seems that a cycle developed independently of the previous cycle, behaving like an stochastic event. It is doubtful whether there is any pattern of SSN in the previous cycles that can be used to foresee the coming cycle. It is only much after the new cycle starts that one can determine whether the new cycle is similar to the previous one or not. For this reason, the AR model prediction made 1.5 years after the start of the new cycle is reasonably good, being within the upper and lower bounds defined by the standard errors.

Based on our results that the AR method is partially successful in predicting SSN, we suggest that the solar dynamo may be regarded as a quasi-linear system or a weakly non-linear system where both the linear dependence of future values on the past values and the stochastic modulation operate.

Acknowledgments

We appreciate the referee's constructive comments. A part of J. Chae's work was carried out during his visits to KASI. J. Chae thanks the KASI staff for the hospitality provided to him. Y.-H. Kim is supported by the development of models for analyzing solar images and for predicting long-term solar activities, a project of Korean Space Weather Center of Radio Research Agency (RRA), and the KASI basic research fund.

References

Kim, K.-T. 1991, Revisit to the Sunspot Cycle, JKAS, 24, 117

Moran, P. A. P. 1954, Some Experiments on the Prediction of Sunspot Numbers, J. Royal Statist. Soc. Ser. B16, 112

Pesnell, W. D. 2008, Predcitions of Solar Cycle 24, Solar Phys., 252, 209

Pesnell, W. D. 2012, Solar Cycle Predictions (Invited Review), Solar Phys., 281, 507

Sim, K. J., Moon, Y.-J., Lee, C.-W., Chang, B. H., & Woo, H. S. 2001, The Relative Sunspot Numbers in 2000, JKAS, 34, 119

Walker, G. 1931, On Periodicity in Series of Related Terms, Proceedings of the Royal Society of London, Ser. A, 131, 518

Werner, R. 2012, Sunspot Number Prediction by an Autoregressive Model, Sun and Geosphere, 7, 75

Yule, G. U. 1927, On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers, Phil. Trans. Royal Soc. Ser. A, 226, 267