DOI: 10.31401/WS.2024.proc.01

Prediction of Solar Cycle 25 Using Simplex Projection Method: A Long-Term Analysis Based on F10.7

Gerceker K.[1], Kilcik A.[1]

[1]Akdeniz University, Institute of Science, Department of Space Sciences and Technologies
07070, Antalya, Turkiye, e-mail: kemalhangerceker@gmail.com

Abstract

This study aims to predict the amplitude and time of solar activity for the solar cycle 25 by applying the Simplex Projection method to the 27-day averaged 10.7 cm Solar Radio Flux (F10.7) data. For this purpose, F10.7 data from 1963 to 2019 (up to the end of Solar Cycle 24) were used, and the current values of the Solar Cycle 25 were excluded from the dataset. Consequently, the dataset consists of 746 points (approximately 700 months), with the first 460 points selected as the library set and the remaining 286 points as the prediction set. Simplex Projection is a method that analyzes the complexity and chaotic properties of dynamic systems and can generate predictions. Subsequently, it compares the data points in the prediction set with the library set for each embedding dimension, identifying similar points and producing forward predictions (i.e., similarities) according to the desired time to prediction (Tp). To obtain better predictions and partially eliminate unsuccessful ones, the Mean Absolute Error (MAE) was calculated between the observed first 71 data points of the current cycle (Solar Cycle 25). Predictions with an error of less than 15% were considered successful predictions. When evaluating successful predictions, a double-peaked pattern is expected for Solar Cycle 25. Accordingly, the first peak of solar activity for Solar Cycle 25 was observed around mid-2023, with the second peak, stronger than the first, anticipated to occur in early 2025 (January 2025), and the minimum of the next cycle expected in the middle or later stages of the year 2030 (July 2030). Finally, the effect of the initial conditions of the library set and the prediction set on the prediction performance is revealed.

Keywords: Solar activity; Solar cycle prediction; Space weather.

Introduction

The Sun can be defined as an active star considering its proximity to the Earth and this activity exhibits quasi-periodic behaviors. Various solar activity parameters/indicators (Sunspot Numbers (SSN), Sunspot Areas, Solar Flare (SF) Numbers, 10.7 cm Solar Radio Flux (F10.7), Coronal Mass Ejection (CME) Numbers and Speeds, Maximum Coronal Mass Ejection Speed Index (MCMESI), etc.) are used for tracking, temporal behavior. These structures/events are highly dynamic and affect life on Earth by changing the space around the Earth and space weather conditions. As of 2024, given the technological sophistication of civilization, the impact of space weather conditions on the Earth has become more important than in the past. Satellite systems, internet networks, electrical systems, and all technological and electronic systems used in every moment of life are extremely vulnerable to being affected by space weather. Therefore, it is critical to understand the Sun-Earth-Space Weather triangle and its interactions.

Understanding solar activity and making predictions for the future are directly related to the fate of our planet and the development of our civilization. One of the urgent problems in the research of Sun-Earth relations today is to analyze and predict the space weather caused by solar activity (Tlatov et al., 2017). Analyzing the nonlinearity/chaoticity of this relationship and related parameters is instructive for future predictions. From this point of view, a good prediction and understanding of the solar activity indices of the considered dynamical structure allows a more accurate presentation of the impact of space weather on the Earth and living life.

Solar radio flux at 10.7 cm (2800 MHz) is a very good indicator of solar activity (Tapping, 2013). This index, often abbreviated as F10.7, is one of the important representative records of solar activity. In addition, the F10.7 index is known to be very valuable in determining and predicting space weather. The measurement advantage of F10.7 is remarkable because it is not affected by atmospheric distortions. For example, observing sunspots can be a problem when the weather is cloudy and atmospheric conditions are challenging.

Although the indices of solar activity seem to follow a certain order/linearity, there is a disorder/chaos within the existing order and an order within this disorder. This situation is explained by "deterministic chaos". Therefore, it seems logical to try to predict the parameters of solar activity, which are dynamic and chaotic, based on linearity. The nonlinear approach has proven its capabilities in applications to chaotic (nonlinear) systems (Sugihara 1994; Maye et al., 2007; Tsonis et al., 2015) and a new method called Empirical Dynamic Modeling (EDM) and the Simplex Projection method based on EDM have made significant progress in such research (Sugihara et al., 2012; Deyle et al., 2013; Ye et al., 2015). EDM is an equation-free approach for modeling the nonlinear interaction of different variables of the same complex system. Therefore, it seems appropriate to use this approach to predict the solar activity parameters.

In this study, the chaotic nature and behavior of the 27-day average F10.7 data were investigated. Accordingly, the Simplex Projection method was used to predict the amplitude and time of Solar Cycle 25. The results and the performance of the prediction method as well as the importance of initial conditions/points are presented.

Data

In this study, 27-day average 10.7 cm Solar Radio Flux (F10.7) data were used. The F10.7 data were obtained from NASA-OMNIWeb (https://omniweb.gsfc.nasa.gov). The data from 1963 to the end of 2018 (the end of the solar cycle 24) were considered and the current values of the solar cycle 25 were excluded from the prediction dataset.

The data set consists of 746 points (approximately 700 months) and the first 460 points are selected as the library set and the remaining 286 points as the prediction set.

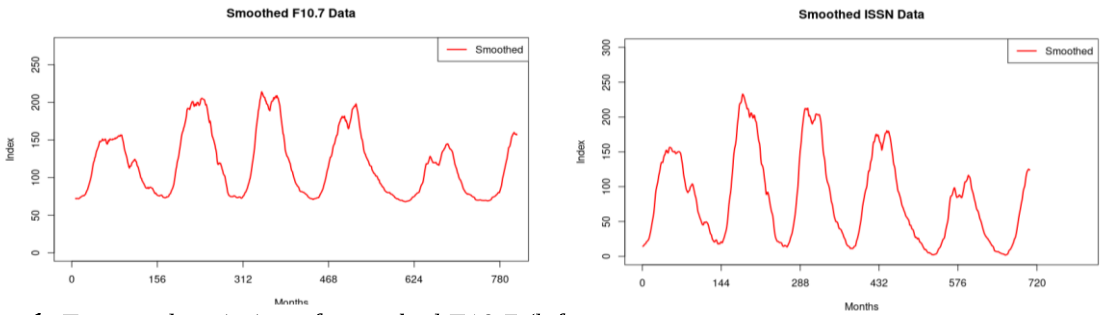

Figure 1. *Temporal variation of smoothed F10.7 (left panel) and smoothed SSN (right panel) for the last five cycles.*

As shown in Figure 1, both indicators show similar temporal variations.

SSN is a very advantageous data. Because, compared to other indicators, it has the oldest records (since the 1700s). F10.7, on the other hand, can be considered a short data set compared to SSN and others since its records started in 1963. Therefore, the fact that the data to be used for predicting is so limited and short-term should make us think about how the prediction performance will be for this data. If the prediction performance is confirmed in short-term data, it is thought that this will be an important finding for the used prediction method.

Methods

3.1. Simplex Projection

Simplex Projection is part of Empirical Dynamic Modelling (EDM) for analyzing nonlinear and chaotic systems developed by Sugihara. The method is used specifically to analyze time series data and predict future trends. It does this by using mathematical and geometric techniques to predict the future state from the past state of data.

The aim of EDM, and by extension Simplex Projection, is to reconstruct system dynamics from time series data. Time series are sequential observations of system behavior and therefore information about the rules governing system behavior (i.e. system dynamics) is hidden/embedded in the data. Takens Theorem (1981) expresses a way to recover this information using only a single time series.

$$X(t) = \{x(t), x(t-\tau), x(t-2\tau), \ldots, x(t-(E-1)\tau)\} \quad (1)$$

Where X is a vector of the original state space, x is a sample in the time series t, τ (tau) is a time delay, and E represents the embedding dimension used to reconstruct the phase space.

The method proceeds as follows:

- First, the time series, i.e. the data, is placed/embedded in a phase space. This is made using the Takens Embedding Theorem. The state space is constructed accordingly. Takens embedding theorem is the reconstruction and representation of the state space of a dynamical system using time series data.
- Here, the method involves creating vectors that represent the state of the system for a given time after the phase space has been constructed. These vectors are generated using a given time delay (tau) and embedding dimension (E). This is called the embedding space.
- Next, in the embedding space, the nearest neighbors of the current state of the time series are identified. These neighbors are past states that were similar to the current state. This similarity is found by calculating the distances between the current vector of the state and all other vectors in the embedding space.
- In the last step, a projection (simplex) is created using similar neighbors. The weighted average of the distances of these similar neighbors is used to generate future values, i.e. predictions.

In simpler terms, the model created for F10.7 data looks at the past dynamics/states of the data, captures similarities, and produces predictions to predict future behavior.

3.2. Attractor Visualisation

The Simplex Projection method is known to be used for chaotic and non-linear time series and the solar activity indicators can be argued to meet this criterion. However, how does one know that F10.7 or any other data is non-linear and chaotic, or not too chaotic to make prediction possible? The attractor visualization helps with this and provides an understanding of the structure and nature of the time series.

This method is used to observe and analyze the trajectory of a dynamic system in state space. This visualization involves the projection of data from a time series onto a high-dimensional state space.

Attractor visualization is a technique used to understand the dynamic behavior of a time series. This technique aims to visually assess whether the time series has a chaotic or regular (patterned) structure.

Here, a two-dimensional (2D) attractor was used to verify this dynamic structure and to understand the chaotic nature of the system (see Figure 2).

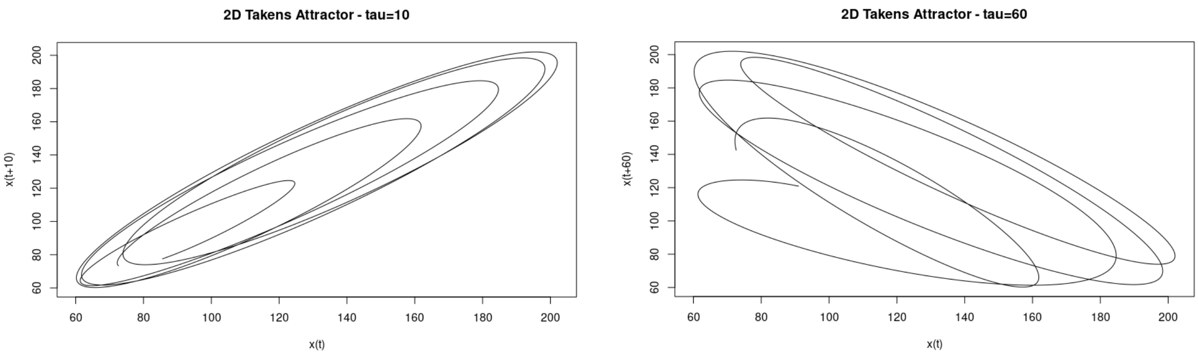

Figure 2. *Attractor Visualizations for F10.7 at 10 (left panel) and 60 (right panel) time delays.*

Although F10.7 was used for this analysis, unlike the prediction data, we used F10.7 data with low-pass filtering based on Fourier analysis (this also means moving the system from state space to frequency space). The reason for this is that the visualization is desired to be obtained with a clearer and more interpretable shape/structure.

Attractor results reveal the chaotic structure of F10.7. The circular features visible especially for 10 reveal the chaotic nature of the system. For 60, there are parallel stripes and this appearance reveals that the chaoticity is high. In the analysis, the aim is only to understand the nature and behavior of the system/data. Therefore, perhaps tau 60 and around may be an appropriate value for the original data and the prediction model.

Results

4.1. Prediction Results

The Simplex Projection method was used to generate predictions for each embedding dimension (E, ranging from 2 to 8) at various time delays. There were multiple successful predictions at varying time delays for various embedding dimensions.

Mean Absolute Error (MAE) was used as a success criterion and the first 71 known points of Solar Cycle 25 were compared with the first 71 points of the predictions. Predictions were considered successful if the MAE was below 15%.

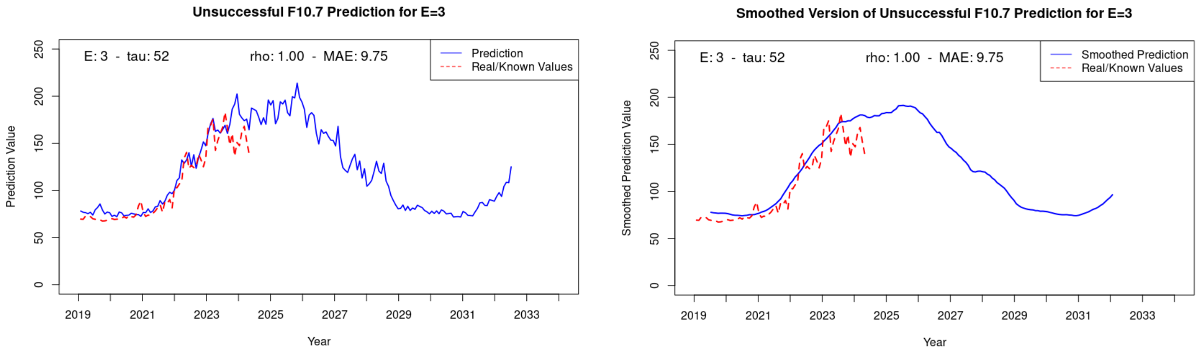

Figure 3. *Original prediction (left panel) and smoothed version (right panel) of F10.7 with optimal tau (52) at embedding dimension E=3.*

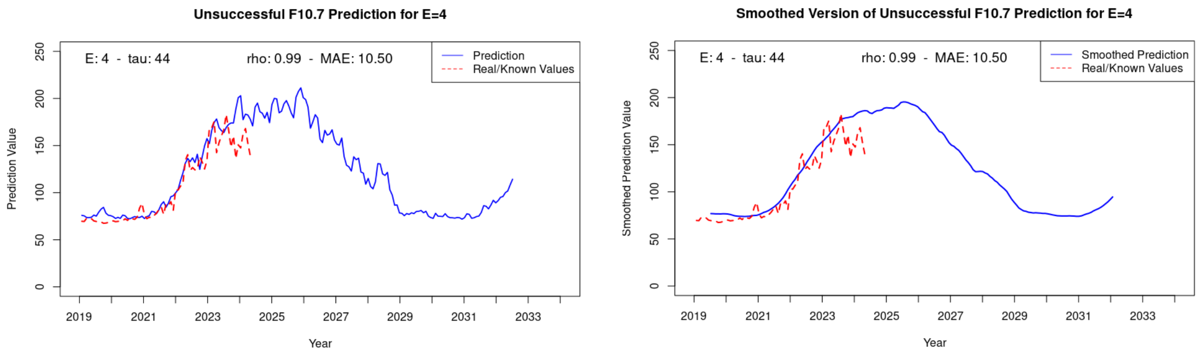

Figure 4. *Original prediction (left panel) and smoothed version (right panel) of F10.7 with optimal tau (44) at embedding dimension E=4.*

Figures 3 and 4 show the most successful prediction results for different embedding dimensions. As can be seen, these results are similar to each other. However, it should be noted that when evaluating the results, the focus is mostly on E=3 and E=4. This is because higher embedding dimensions (such as E=6-7-8-9) are impractical for short data sets. With limited data, as the embedding dimension increases, the system dynamics are over-analyzed and the results become smoother.

For the predicted amplitude and times of maximum for Solar Cycle 25:

1. Solar Cycle 25 is expected to be stronger than Solar Cycle 24.
2. Predictions indicate that a double peak structure is more likely for Solar Cycle 25.
3. The first peak of solar activity for Solar Cycle 25 was observed around mid-2023. The second peak, which is expected to be stronger than the first one, is predicted in the early months of 2025 (January 2025). The end of the cycle/starting of the next cycle is predicted in mid-2030 (July 2030).

4.2. Adherence of model sets to initial conditions/points and unsuccessful predictions generated to demonstrate this

Chaos theory studies dynamical systems that are extremely sensitive to initial conditions. In chaotic systems, small changes in initial conditions can lead to large changes in the future behavior of the system.

In chaotic systems and models, the importance of the starting point of the prediction (or the end point of the data set) is known. While generating the predictions in this study, it was observed that the prediction performance does not only depend on the starting point of the prediction. The library set and the points where the prediction set starts and therefore ends are also important for the prediction performance.

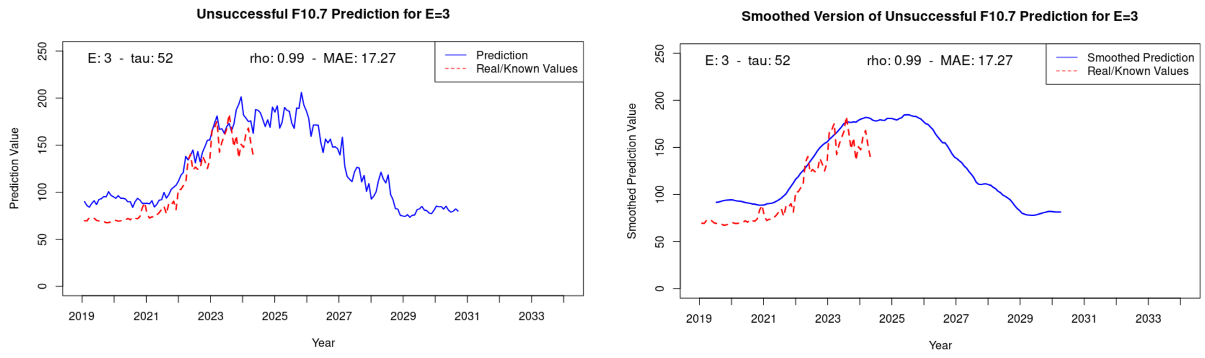

Figure 5. *Original (left panel) and smoothed version (right panel) of unsuccessful F10.7 prediction with non-optimal initial/starting point of sets with embedding dimension E=3.*

As can be seen in Figure 5, an unsuccessful prediction was produced. It should be noted that the same parameters were chosen for this prediction as for the successful prediction in Figure 3 (tau=52 for E=3). The reason for the unsuccessful prediction is that only one condition was changed: The point where the library set ends and the prediction set starts.

In Section 2, it was stated that the data set consisted of 746 points, the first 460 points were the library set and the remaining 286 points were the prediction set. For the unsuccessful prediction, this condition was changed and the sets were shifted. Accordingly, the optimal parameters from the successful prediction were taken, but the library set was set to 400 and the prediction set to 346. In other words, the last 60 points of the library set were shifted and added to the prediction set.

This case for the effect of the initial conditions of the sets on the prediction performance reveals an important fact. This can be seen in Figure 5, especially at the beginning (ascending phase) and at the end (descending phase) of the cycle prediction. There is an upward shift in amplitude in these regions. Also, on closer inspection, there is a downward shift in the center of the cycle (at maximum). The increase in the MAE (almost 10%) demonstrates this theoretically. As a result, the importance of the initial and end conditions of the data sets is thus revealed and specified.

Conclusion and Discussion

In this study, the Simplex Projection method is used to predict the amplitude and time of maximum for Solar Cycle 25. In addition, the initial conditions for the sets are analyzed and presented.

The findings of the study are as follows:

- Our predictions indicate that Solar Cycle 25 will be stronger than Solar Cycle 24.
- The predictions suggest that a double-peak maximum is more likely for Solar Cycle 25. The second peak of Solar Cycle 25, which is expected to be stronger

than the first, is predicted to occur in the early months of 2025 (January 2025), while the end of the Solar Cycle 25 is predicted to occur in the mid-2030s (July 2030).

- The success of the prediction directly depends on the initial condition/point of the sets. Changes in the initial and end points of the sets can significantly change the prediction. Therefore, it is very important to choose the optimal initial point.

In chaotic systems, the importance of the point where the prediction starts (or the point where the data set ends) is known when creating a model and obtaining a prediction. The importance of the initial/starting point of the prediction for solar activity data was first shown by Sarp, V. et al. (2018) using Simplex Projection. In our study, in addition to this finding for initial conditions, we revealed the sensitivity of the start and end points of the sets to the model and prediction. In addition, in the mentioned study in 2018, it was aimed to find the time and amplitude of Solar Cycle 25 using SSN and in parallel with our findings, it was predicted that Solar Cycle 25 would be stronger than Solar Cycle 24 with its double peak structure.

Simplex Projection (referred to as the Sugihara-May algorithm) was first applied to solar activity data by Kilcik et al. (2009). They predicted the amplitude and time of Solar Cycle 24 as of December 2012 with 87 sunspot units and first presented a successful example of the method.

The success of the Simplex Projection approach in nonlinear chaotic systems is demonstrated by its successful prediction performance for the relatively limited F10.7 data (approximately 700 months of data). Remarkably, this approach can obtain accurate predictions without the need for large amounts of data (e.g. more than 10 cycles).

Studies in the literature and various methods have been used for many years and can be considered successful. In this paper, we argue that the Simplex Projection approach is an alternative, for good prediction of solar cycle. Besides the advantages of the F10.7 data and its good representation of solar activity, we chose this data because it is limited data. The performance and success of the method are demonstrated by making a one-cycle prediction from only five cycles of information. In the future, we plan to continue prediction studies by improving both the method and our perspective.

Acknowledgements

This work was supported by Project 124N011 awarded by the Scientific and Technological Research Council of Turkey. All data used in the study were taken from the NASA-OMNIWeb (https://omniweb.gsfc.nasa.gov) database, and we are gratefully acknowledged. We would also like to thank the Bulgarian Academy of Sciences for their partial support.

References

Deyle, E. R., Fogarty, M., Hsieh, C. H., Kaufman, L., MacCall, A. D., Munch, S. B., ... & Sugihara, G. (2013). Predicting climate effects on Pacific sardine. Proceedings of the National Academy of Sciences, 110(16), 6430-6435.

Kilcik, A., Anderson, C. N. K., Rozelot, J. P., Ye, H., Sugihara, G., & Ozguc, A. (2009). Nonlinear prediction of solar cycle 24. The Astrophysical Journal, 693(2), 1173.

Maye, A., Hsieh, C. H., Sugihara, G., & Brembs, B. (2007). Order in spontaneous behavior. PloS one, 2(5), e443.

Sarp, V., Kilcik, A., Yurchyshyn, V., Rozelot, J. P., & Ozguc, A. (2018). Prediction of solar cycle 25: a non-linear approach. Monthly Notices of the Royal Astronomical Society, 481(3), 2981-2985.

Sugihara, G. (1994). Nonlinear forecasting for the classification of natural time series. Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences, 348(1688), 477-495.

Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. science, 338(6106), 496-500.

Takens, F. (2006, October). Detecting strange attractors in turbulence. In Dynamical Systems and Turbulence, Warwick 1980: proceedings of a symposium held at the University of Warwick 1979/80 (pp. 366-381). Berlin, Heidelberg: Springer Berlin Heidelberg.

Tapping, K. F. (2013). The 10.7 cm solar radio flux (F10. 7). Space weather, 11(7), 394-406.

Tlatov, A. G., Shramko, A. D., Chernov, Y. O., Strelkov, M. A., & Naga Varun, E. (2017). Space weather parameters: Modeling and prediction from the data of groundbased observations of solar activity. Geomagnetism and Aeronomy, 57, 854-858.

Tsonis, A. A., Deyle, E. R., May, R. M., Sugihara, G., Swanson, K., Verbeten, J. D., & Wang, G. (2015). Dynamical evidence for causality between galactic cosmic rays and interannual variation in global temperature. Proceedings of the National Academy of Sciences, 112(11), 3253-3256.

Ye, H., Beamish, R. J., Glaser, S. M., Grant, S. C., Hsieh, C. H., Richards, L. J., ... & Sugihara, G. (2015). Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. Proceedings of the National Academy of Sciences, 112(13), E1569-E1576.