Prediction of maximum amplitude of solar cycle 25 using machine learning

T Dani¹ and S Sulistiani¹

¹Space Science Center, LAPAN, Indonesia

Email: tiar.dani@lapan.go.id

Abstract. Prediction of maximum amplitude of Solar Cycle 25 is obtained by using four different machine learning regression methods, i.e. Linear Regression (LR), Random Forest (RF), Radial Basis Function (RBF) and Support Vector Machine (SVM). Monthly mean sunspot number data during the 1856-June 2018 (solar cycles 10-24) from the World Data Center SILSO, Royal Observatory of Belgium, Brussels are used as machine learning inputs. According to LR, RF, RBF and SVM, the maximum of Solar Cycle 25 is predicted to occur in September 2023 (sunspot number of 159.4 ± 22.3), in December 2024 (sunspot number of 110.2 ± 12.8), in December 2024 (sunspot number of 95.5 ± 21.9) and in July 2024 (sunspot number of 93.7 ± 23.2), respectively. The prediction using LR method suggested that Solar Cycle 25 maximum will be slightly higher than the current cycle, while RBF and SVM suggested much lower cycles. RF prediction suggested a lower maximum with well-constructed double-peak. It was also found that the Solar Cycle 25 is predicted to begin in the late 2019 or early 2020 according to all four methods.

Keywords: solar cycle and machine learning

1. Introduction
The prediction of solar activity in term of sunspot number is very important in the mitigation of space weather effects on the Earth’s technology systems. The prediction of peak time and value of the solar cycle can give an overview of the solar activity in the next solar cycle compared to the current cycle. The solar activity in term of sunspot number during Solar Cycle 24 was considerably low. NOAA suggested that Solar Cycle 24 has reached its maximum in April 2014 with a mean sunspot number of 82 and is currently in the declining phase. This value is lower than Solar Cycle 23 maximum with sunspot number around 120.

Sunspot number is one of the parameters of solar activity describing the measure of the number of sunspots and groups of sunspots on the visible surface of the Sun. The larger the number, the bigger the probability of flare and CME occurrences on the Sun. These events can cause the geomagnetic and ionospheric storms on Earth and eventually can disrupt the ground-based and space-based technology systems. Previous studies have predicted the peak time and value of Solar Cycle 25 by means of various methods, e.g. [1-8]. This study is predicting the peak time and value of Solar Cycle 25 by using four different machine learning methods, i.e. Linear Regression (LR), Random Forest (RF),...
Radial Basis Function (RBF) and Support Vector Machine (SVM). The comparison of this study and previously published works were also performed.

2. Methods
Our study is using the SILSO (Sunspot Index and Long-term Solar Observations) [9] data of monthly mean total sunspot number from 1855 (Solar Cycle 10) through May 2018 prepared by the Royal Observatory of Belgium, Brussels as inputs to machine learning program for prediction. We also made use of the Waikato Environment for Knowledge Analysis (WEKA) software to predict the peak time and value of the sunspot number. WEKA is developed by the University of Waikato, New Zealand, equipped by various machine learning algorithm written in Java programming language [10][11]. By using Time-Series Forecasting package available in WEKA, we utilized four different techniques, i.e. LR, RF, RBF, and SVM, to predict the monthly sunspot number for Solar Cycle 25.

3. Result and Discussion
WEKA provides time-series forecasting packages for regression schemes which include lagged variables and date-derived periodic variables. It is also able to do closed-loop forecasting. In this paper, we used four base learner machine learning algorithms for the monthly sunspot number forecast, i.e. Linear Regression, Radial Basis Function for Regression, Support Vector Machine for Regression and Random Forest. We used sunspot number historical data from 1855 until May 2018 as input.

3.1. Linear Regression (LR)
Figure 1 shows the prediction of the monthly sunspot number for Solar Cycle 25 by using the linear regression technique. In WEKA, this technique is using the Akaike criterion for model selection. The root means squared error for this prediction is 22.3.

![Figure 1. Monthly sunspot number during Solar Cycle 24 (red) and Solar Cycle 25 (black). The latter values are the prediction based on linear regression technique.](image_url)
Based on this technique, Solar Cycle 25 will start in early 2020 and reach the maximum sunspot number of 159.4 ± 22.3 in September 2023. The value is higher than the previous cycle and this prediction was unable to reconstruct the common double-peak feature.

3.2. Random Forest (RF)
In WEKA, this technique constructed a tree with random attributes K for each node which perform no pruning and backfitting option to allow estimation for class probabilities. Figure 2 shows the prediction of the monthly sunspot number for Solar Cycle 25 by random forest technique. The root mean squared error for this prediction is 12.8.

![Figure 2](image)

**Figure 2.** Monthly sunspot number during Solar Cycle 24 (red) and Solar Cycle 25 (black). The latter values are the prediction based on random forest technique.

Based on this technique, Solar Cycle 25 will start in early 2020 and reach the maximum value of 110.2 ± 12.8 in December 2024. This value is lower than that of Solar Cycle 24. The prediction of Solar Cycle 25 using random forest technique shows the double-peak feature commonly found in solar cycles. The first peak is predicted to reach a value of 103 ± 12.8 in April 2023.

3.3. Radial Basis Function (RBF)
RBF in WEKA uses the k-means clustering algorithm to provides a logistic regression or linear regression learning model. Figure 3 shows the prediction of the monthly sunspot number for Solar Cycle 25 by radial basis function technique. The root mean squared error for this prediction is 21.9.
Based on this technique, Solar Cycle 25 will start in early 2020 and reach the maximum value of 95.5 ± 21.9 in December 2024. This value is lower than that of Solar Cycle 24 and even lower than that predicted by random forest technique. The prediction of Solar Cycle 25 by radial basis function technique unable to construct double-peak feature commonly found in solar cycles.

3.4. Support Vector Machine (SVM)

SVM in WEKA using SMOreg function which its parameters can be set using a various algorithm. Figure 4 shows the prediction of the monthly sunspot number for Solar Cycle 25 by support vector machine. The root means squared error for this prediction is 23.2. As can be seen from Figure 4, based on this technique Solar Cycle 25 is predicted to start in the early 2020 and to reach the maximum value of 93.7 ± 23.2 in July 2024. This value is the lowest among all machine learning technique used in this study. Similar to linear regression and radial basis function techniques, this one also unable to construct the double-peak feature on the prediction of Solar Cycle 25.
3.5. **Comparison with other studies**

In this study, we tried to compare our results with the prediction of Solar Cycle 25 obtained from various methods by previous works, as shown in Table 1. Most of the studies suggested that Solar Cycle 25 will be similar to or lower than the current cycle, while our result using linear regression technique suggested otherwise. Furthermore, all studies considered here predicted the peak time of Solar Cycle 25 to range between the year of 2023-2024.

![Figure 4. Monthly sunspot number during Solar Cycle 24 (red) and Solar Cycle 25 (black). The latter values are the prediction based on support vector machine technique.](image-url)
Table 1. Comparison of Solar Cycle 25 peak predictions from several works using different methods.

| References                      | MAX       | Methods                                              |
|---------------------------------|-----------|-----------|
|                                 | Month     | Year      | SSN       |                                                      |
| This study                      | Dec       | 2024      | 110.2 ± 12.8 | Random forest                                       |
| This study                      | Sep       | 2023      | 159.4 ± 22.3 | Linear Regression                                    |
| This study                      | Dec       | 2024      | 95.5 ± 21.9 | Radial Basis Function                               |
| This study                      | Jul       | 2024      | 93.7 ± 23.2 | Support Vector Machine                              |
| [1]                             | Apr - Jun | 2023      | 112.3 ± 33.4 | Correlation Between Cycle Parameters                |
| [2]                             | Apr       | 2023      | 132.1      | Extrapolation of sunspot number spectral components |
| [3]                             | 2022-2023 | 118.2     |            | Spotless                                             |
| [4]                             | 2022-2023 | 42 ± 13   |            | Sunspot Group                                        |
| [5]                             | Oct       | 2023      | 109.1      | Classical statistical relations among feature       |
|                                 |           |           |            | parameters of solar cycle profiles                  |
|                                 |           |           |            | Cycle 25 will be similar in size to Cycle 24        |
|                                 |           |           |            | 116 (southern hemisphere - smoothed)                |
|                                 |           |           |            | Solar polar microwave brightness temperature        |
|                                 |           |           |            | Ludendorff coronal flattening index of the          |
|                                 |           |           |            | Total Solar Eclipse (TSE) on March 9, 2016          |

4. Conclusion

Predictions of peak time and value of Solar Cycle 25 is done by using four machine learning technique, i.e. LR, RF, RBF and SVM. According to LR, RF, RBF, and SVM, the maximum of Solar Cycle 25 is predicted to occur in September 2023 (sunspot number of 159.4 ± 22.3), in December 2024 (sunspot number of 110.2 ± 12.8), in December 2024 (sunspot number of 95.5 ± 21.9) and in July 2024 (sunspot number of 93.7 ± 23.2), respectively. RF prediction is able to predict the double-peak feature of solar cycle. Comparison between this study and from other studies with different methods was also done. It was found that generally all predictions suggested that Solar Cycle 25 will be similar to or lower than the current cycle. Moreover, all works predicted Solar Cycle 25 will peak around the year of 2023-2024.

5. Acknowledgments

The sunspot records are courtesy of WDC-SILSO, Royal Observatory of Belgium, Brussels. We would like to thank the Indonesian Minister of Research, Technology and Higher Education (Kemenristekdikti RI) for funding this work through Program Insentif Riset Sistem Inovasi Nasional (InSINas) 2018.

References

[1] Pishkalo M I 2008 Kinematics and Physics of Celestial Bodies 24 242–247
[2] Rigozo N R, Echer M P S, Evangelista H, Nordemann D J R and Echer E 2011 Journal of Atmospheric and Solar-Terrestrial Physics 73 1294-99
[3] Helal J H R, and Galal A A 2013 Journal of Advanced Research 4 275–278
[4] Javaraiah J 2015 New Astronomy 34 54-64
[5] Li K J, Feng W and Li F Y 2015 Journal of Atmospheric and Solar-Terrestrial Physics 135 72–76
[6] Hathaway D H and Upton L A 2016 Journal of Geophysical Research:Space Physics 121 11
[7] Gopalswamy N, Mäkelä P, Yashiro S and Akiyama S 2018 Journal of Atmospheric and Solar-Terrestrial Physics
[8] Dani T, Priyatikanto R and Rachman A 2016 J. Phys.: Conf. Ser. 771 012005
[9] SILSO, World Data Center 1855-2018 International Sunspot Number Monthly Bulletin and online catalogue http://www.sidc.be/silso/
[10] Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P and Witten I H 2009 SIGKDD Explorations 11 1
[11] Frank E, Hall M A and Witten I H 2016 The WEKA Workbench (Morgan Kauffman) 4th Edition