Analysis of Solar Cycle and predicting next Solar Maximum

Pramod Kadagattor

Dept. of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, Maryland
cq71296@umbc.edu

Aarsh Oza

Dept. of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, Maryland
aoza1@umbc.edu

Abstract— **Pockmarks on the sun's surface is always shifting. These phenomena is evidently visible in roughly 11 years. The 11 year cycle of solar activity is characterized by the rise and fall in the numbers of sunspots. A number of other solar activities also vary during this cycle. The major activity is reversal of the sun's magnetic field which causes the poles to interchanging their places and variations in solar geomagnetic activity. Other major variations occurring are 10.7 cm radio flux, the total solar irradiance. Huge quantities of plasma known as coronal mass ejections fly into space which can disrupt satellites and other electronic signals if they reach earth. More solar activity during the cycle also amplifies auroras and warms the earth's temperature slightly. A solar cycle is characterized by their maxima and minima, cycle period and amplitudes. The sun can also be erratic, making it tricky for physicists to predict future sunspots as there is always an element of randomness. Solar activity prediction methods have been wide-ranging, mostly numerical, and essentially curve fitting. We conclude with analyzing all 24 solar cycle and predicting the number of sunspots in 25th cycle along with the year of occurence. We have used regression to fit 2 curves and we conclude with comparing different orders of regression.**

Keywords— Sunspots, Solar cycle 24, Prediction, Regression, Solar maximum and minimum

I. Introduction

The Sun's surface, the photosphere radiates more actively when there are more sunspots. Satellite monitoring of solar luminosity (radiant flux of sun) revealed a direct relationship between the Schwabe cycle and luminosity. As each cycle begins, sunspots appear at mid-latitudes, and then move closer and closer to the equator until solar minimum is reached.

Various methods have been used in the past to predict the amplitude (solar maxima) and time of maximum of a sunspot cycle. Some of them are based on sound physical principles. There were many predictions for solar cycle 24, but very few of them used *solar forerunners* to help in predictions. The importance of using solar forerunners is that they have existed in the past and their physical explanations could have implications for dynamo models (models used for prediction). One of the useful prediction methods, using solar forerunners, is the dependence of solar maximum on the preceding solar minimum. Another approach used is the relation between rise time (time period between occurrence of solar maximum and solar minimum) and current solar maximum is used to predict the next solar maximum.

Use of solar maxima versus solar minima, among other forerunners methods for prediction purposes was not in use for some time because of high correlation values of solar maximum with other foreruners such as geomagnetic index for single variate and combination of minimum amplitude and solar minimum for bivariate models. However, in the recent solar cycle these models could not perform well with their predictions and therefore it was very important to look for a forerunner that has real physical significance for the amplitude of the next cycle like solar maxima and solar minima.

II. Background Work

A. Solar Cycle

The rotation of the sun on its own axis generates a powerful magnetic field. The magnetic field goes through a cycle. Approximately every eleven years the magnetic field of the sun completely flips, the north and south pole flip. The solar cycle causes activity on the sun such as occurrences of sunspot, solar flares, coronal mass ejections. The solar cycle and the activities on the sun affects the earth in several ways. Various kinds of radiations are produced during the solar cycle which affect the electronic devices, satellites. The occurrences of auroras are also an effect of solar activity. Scientists try to predict these occurrences to safeguard devices from damage.

B. Sunspots

Sunspots are temporary occurrences on the Sun's photosphere that appear as spots darker than the surrounding areas. They are areas of reduced surface temperature caused by concentrations of magnetic field flux that reduce convection. Sunspots usually appear in pairs of opposite magnetic polarity. Sunspots may last for a few days to a few months. All Sunspots do eventually decay and disappear.

Sunspots expand and contract as they move across the surface of the Sun, with diameters ranging from 16 km to 160,000 km

C. *Variations of Sunspots due to Solar Cycle*

The Sunspots vary in number and size during the solar cycle. The beginning of the cycle is called the solar minima where the number of sunspots are low. The middle of the cycle is called the solar maxima where the number of sunspots are very high. The number of sunspots die down from the solar maxima and reach the solar minima, completing one solar cycle. The time taken for the sunspots to raise to the maxima from minima is less than the time taken to drop from maxima to minima[1]

D. Analysis and prediction

Several scientists have analyzed the variations of sunspots versus the solar cycle, a clear way to understand the variations is by using the Fast Fourier Transform tool on the sunspot data over a large period of time, a frequency peak can be observed around the eleven year mark indicating maximum number of sunspots can be seen approximately in eleven years[2]. The number of sunspots in the solar maxima can be predicted based on the number of solar minima and the time taken in the previous cycles to reach the solar maxima from minima.

III. Our Approach

Many scientists and analysts uses R and MATLAB to perform their analysis. Apart from these, python is also the tool which is increasingly being used due to its specialized modules, like Numpy, Scipy, Matplotlib, Pandas and so on. So we decided to go ahead with python V3 for our analysis and prediction.

We are performing two predictions in our approach. The first prediction is the next solar maxima of the 25th solar cycle. Then we also predict the next solar cycle maxima's occurrence time.

In numerical analysis, there are two methods for predicting next data point based on all the previous data points:

1. *Extrapolation*
    - This method tries to fit the curve perfectly on all the data points available.
    - The problem of variance or noise does not happen in extrapolation.

2. *Regression*
    - This method tries to fit the curve approximately on all the data points available to us.
    - Variance or noise is the main concern in regression.
    - The main objective of regression is to reduce the variance.

As observed by solar physicist there is always an element of randomness in solar cycle and thus it cumbersome and inaccurate to predict solar maxima using linear or polynomial extrapolation curve . Thus, we chose to go ahead with regression to predict the next solar maxima and its occurence approximately.

As there are two predictions that we are making we have done two regression two predict those two values. The first regression is between solar maxima and solar minima to predict the solar maxima of next cycle. The second regression is between solar maxima and time duration to predict the approximate time at which the next solar maxima will occur.

To predict the future solar maximas, we use previously collected sunspot data and perform a statistical regression analysis on the data. The data was obtained from sidc.oma.be/silso/INFO/snmtotcsv.php. The NGDC(National Geophysical Data Center) maintains the database.

The data is spread over a period of 270 years, from 1749 to 2019. This data contains the date and the number of sunspots observed on that date [3]. This data is used to observe the sunspot variations over the period. The data contains sunspots variations over 24 solar cycles.

A. *Data Cleaning and Smoothing*

To abate the observational errors and manual errors occurred during the collection of data, a running average is computed on the data. First a 3 month running average is computed on the data to reduce the *monthly errors* observed, then a 13 month running average is computed over the data to reduce the *yearly errors* observed. While other methods have been used to smooth the data[4] we used this method to find the number of sunspots during a maxima(Smax) and minima(Smin).

The Smoothed data is then used to perform analysis and predict the 25th solar maxima.

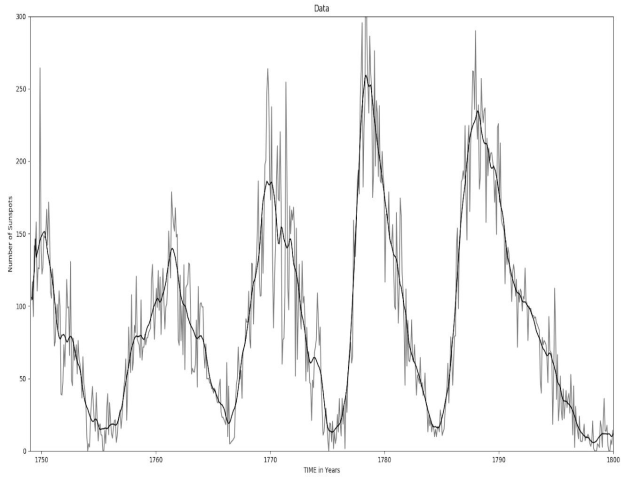

Fig 1. The gray lines indicate the original data plotted time vs number of sunspots. The black line represents the smoothed data. A 13 month running average applied on a 3 month averaged data.

B. *Predict the number of sunspots in the next solar maxima*

As per the data, it is seen that a solar minimum was observed in the year 2018. To predict the next solar maxima we perform a statistical regression analysis between the number of sunspots observed in a minima vs the number of sunspots observed in a maxima. The continuous line in Figure 2 depicts the linear fit of the form Smax = A + B*Smin (A: intercept on the ordinate and B: slope of the fitted line). The regression coefficient B turns out to be 4.20386. Using this regression line we can predict the Smax of the upcoming solar maxima. The Smin observed in 2018 is 4.7974 sunspots. using this in the equation A+B*4.7974 we predict the number of sunspots in the solar maxima of the 25th cycle, that is 155.469 sunspots.

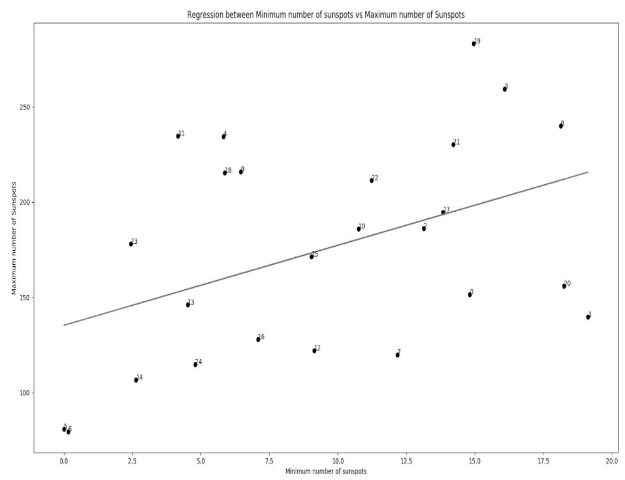

Fig 2: This displays a scatter plot of Smin vs Smax and the line is the regression line between Smin and Smax.

C. *Trends of Smax vs Smin in solar cycles*

On observing the trends in figure 2, it was seen that the number of sunspots in maxima (Smax) is higher when the sunspots in minima (Smin) is high. On observing the details of the plot in Fig 2, an anomaly was noted in the 19th cycle. The number of sunspots in minima(Smin) was low but the sunspot in maxima (Smax) was abnormally high.

D. *Predict the next Solar Maxima*

The time of the next solar maxima is predicted by fitting a regression line between the time taken for the sunspots to rise from minima to maxima vs the Smax value. Predicting the time of maximum of the upcoming solar cycle is an issue having been discussed for several decades. It was first Waldmeier [5] who formulated the inverse correlation between rise time and the cycle amplitude. Cameron and Schüssler[6] have pointed out that the high correlation between the rate of rise of the cycle and Smax can serve as a better tool to predict the time of maximum.

We use the Smax value obtained from the regression fit between Smin vs Smax. Three different orders of polynomials were fit for the data (time rise vs Smax)

(i) ***Quadratic fit***

The data time rise vs Smax was fit in a quadratic curve (A*x^2 + B*x +C) where the values of A B and C were 6.13e-05, -3.711e-02 and 8.879 respectively. From the plot we find the time when the next solar maxima can be observed, that is September of 2023.

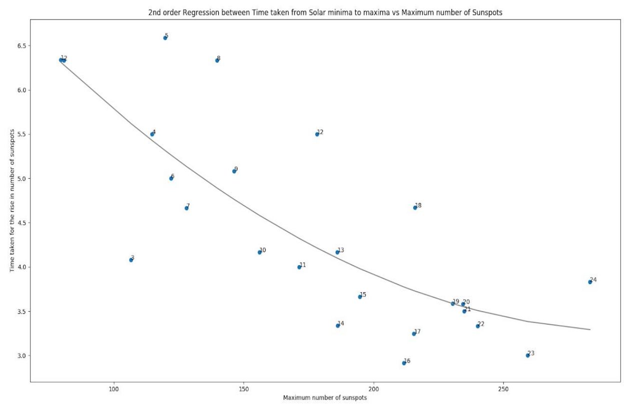

Fig 3: A scatter plot between the Smax vs Time rise is seen along with the quadratic regression fit between the same.

(ii) ***Cubic fit***

The data time rise vs Smax was fit in a quadratic curve (A*x^3 + B*x^2 +C*x + D) where the values of A B C and D were 6.837e-07, -3.054e-04, 2.42e02 and 5.7433 respectively. From the plot we find the time when the next solar maxima can be observed, that is October of 2023.

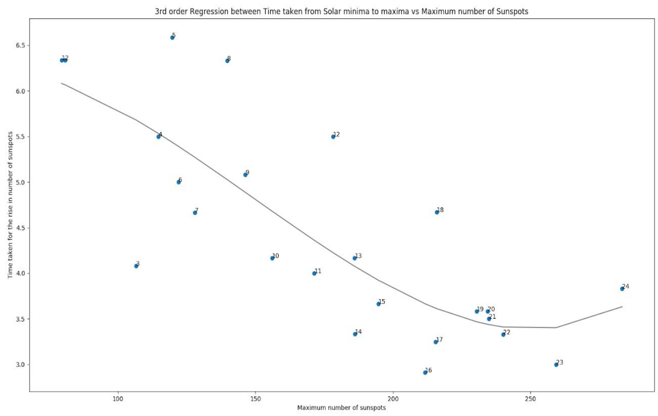

Fig 4: A scatter plot between the Smax vs Time rise is seen along with the cubic regression fit between the same.

(iii) ***Fourth order fit***

The data time rise vs Smax was fit in a quadratic curve (A*x^4 + B*x^3 +C*x^2 + D*x + E) where the values of A B C D and E were 9.944e-09, -6.457e-06, 1.517e-03 and -1.699e-01 respectively. From the plot we find the time when the next solar maxima can be observed, that is November of 2023.

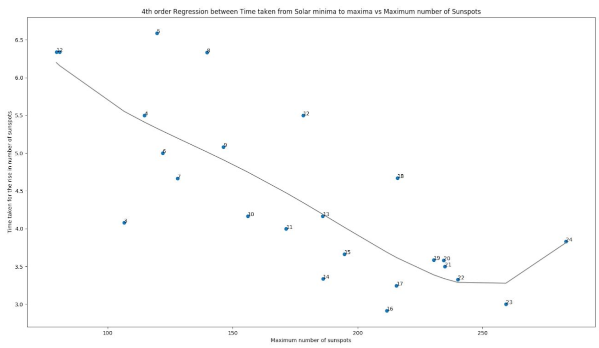

Fig 5: A scatter plot between the Smax vs Time rise is seen along with the fourth order regression fit between the same.

E. *Trends of Time rise vs Smax in solar cycles*

On observing the trends in figure 3,4 and 5, it was seen that the time rise is higher when the sunspots in maxima (Smin) is low. On observing the details of the plot in the three figures, the anomaly noted in the 19th cycle can clearly be observed. The time rise is high though the sunspot in maxima (Smax) was abnormally high.

F. *Errors*

The three polynomial fits were compared using the sum of least squares. On observing, quadratic fit resulted in a least square sum (time rise) of 11.402, Cubic fit resulted in a sum of 10.984 and the fourth order fit resulted in a sum of 10.747. On comparing the three values it can be concluded that the fourth order fit is more accurate than the other two.

The method sum of least squares is a standard approach in statistical regression analysis which is used to approximate the solution of overdetermined systems by minimizing the sum of the squares of the residuals made in the results of every single equation.The fit with least of this measure is considered to be the best fit for the data.

IV. Conclusion

After performing both regression we got satisfactory results in predicting both solar maxima for current solar cycle and the time at which it will occur.

According to our result of the first regression, this solar cycle will have 155.46 sunspots as maximum number. For the second regression we have used 3 different orders for polynomial fit:

(i) Quadratic fit: This fit gives the prediction that solar maxima will occur in 2023.71 which is approximately September of 2023.

(ii) Cubic fit: This fit gives the prediction that solar maxima will occur in 2023.815 which is approximately October of 2023.

(iii) Fourth order fit: This fit gives the prediction that solar maxima will occur in 2023.88 which is approximately November of 2023.

As discussed earlier Least Square measure of error, implies that a fit with least measure is the best fit. The least square measure for Quadratic fit is 11.402, for cubic it is 10.984, while for 4th order it is 10.747. As you can see the error is least for 4th order fit. It is obvious to extrapolate that 4th order fit is the best fit.

REFERENCES

[1] Du, Z., Wang, H., Zhang, L.: 2009a, Correlation function analysis between sunspot cycle amplitudes and rise times. Solar Phys. 255, 179 – 185.

[2] Hathaway, D.H., Wilson, R.M., Reichmann, E.J.: 1994, The shape of the sunspot cycle. Solar Phys. 151, 177 – 190.

[3] sidc.oma.be/silso/INFO/snmtotcsv.php.

[4] Hathaway, D.H.: 2010, The solar cycle. Living Rev. Solar Phys. 7, 1 – 65.

[5] Waldmeier, M.: 1935, Neue Eigenschaften der Sonnenfleckenkurve. Astron. Mitt. 14, 105 – 130.

[6] Cameron, R., Schüssler, M.: 2008, A robust correlation between growth rate and amplitude of solar cycles: consequences for prediction methods. Astrophys. J. 685, 1291 – 1296.