Draft version December 15, 2023
Typeset using LaTeX default style in AASTeX631

Term Paper for ASTR 513

Malhar Dave

(Dated: December 14, 2023)

1. INTRODUCTION

Every 11 years or so , the sun's magnetic field flips so north becomes south and south becomes north. Changes in the sun's magnetic field affect the amount of activity on the solar surface, and the time it takes for the whole process to occur is termed as a solar cycle.

One of these activities are Sunspot numbers, a feature long observed on the photosphere of the Sun, These are temporary dark spots that arise on the Sun's surface. Sunspot numbers have been recorded since 1755, providing a historical record of solar cycles.

Thus, a sunspot cycle is classified as the periodic variations in the number of these sunspots over a 11-year solar cycle. We are currently in the 25th measured sunspot cycle.

Since the continuously recorded sunspot cycle data is so old, there is inconsistency in the recorded data set, and there have been attempts to rectify errors from older datasets for sunspot cycles, often grouped under the topic of sunspot number calibration. A global viewpoint of the current consensus and progress of the community is summarized in Pesnell (2016)

However, such discrepancies are not taken into account by Yu et al. (2012) , and from the multiple available avenues of recorded sunspot cycle datasets available, they used the monthly smoothed sunspot number from the Solar Influences Data Analysis Center in Belgium (http://sidc.oma.be).

The research conducted by Yu et al. (2012) delves into the correlations between these sunspot cycles and establishes predictive models for future cycles, determining key parameters such as amplitude, duration, and rise time for forthcoming cycles.

Better prediction of the solar cycle has multiple practical benefits, from having an idea on the occurrence of various types of space weather storms, to better management of low earth orbit satellites, and future exploration.

Building upon the success of Yu et al. (2012) predictive model for the 24th solar cycle, efforts were made to explore the characteristics of the 25th solar cycle, one of which was carried by Yan et al. (2021), which uses the exact same methodology as the former one. There is a large range of possible predictive models which can be used, a comprehensive list of them for Solar Cycle 24 can be found in Pesnell (2016).

2. METHODOLOGY

The solar cycle depends on a lot of variables. It would be challenging to fit for all of them individually, and thus, an easier model is utilized to approximate the sunspot numbers, labelled as U_t. This is is further split up into the rising and declining phase to model the data before and after the maximum sunspot number count.

In our code, we took the time of maximum sunspot number count at the halfway point of the solar cycle 24.

The amplitude for a given solar cycle is usually around a 100, usually between 90 and 120. Although there have been outliers to the dataset, and the curve isn't always a smooth increasing and then decreasing function as shown in the model. This can be seen in **??** for the 24th solar cycle, which has a maximum sunspot number near 120, on the higher side of the range. It also shows a dip in the first half of the cycle, which our parameterized function fails to account for. However, such dips aren't extremely common amongst all the solar cycles, and thus, it is not taken into account.

The methods used by Yu et al. (2012) can briefly be summarized into the following steps:

2.1. *Cycle modelling*

Three things are measured here, with one result:

1. amplitude

- for the rising phase $t < t_{\max}^{(i)}$

$$U_t = c_i \left(1 - \left(\frac{t_{\max}^{(i)} - t}{t_{\max}^{(i)} - t_0^{(i)}}\right)^{\alpha_1}\right) \quad \text{and}$$

- for the declining phase $t > t_{\max}^{(i)}$

$$U_t = c_i \left(1 - \left(\frac{t - t_{\max}^{(i)}}{t_1^{(i)} - t_{\max}^{(i)}}\right)^{\alpha_2}\right),$$

Figure 1. i here is the cycle, t(i) max is the time where cycle maximum occurs, t (i)1 is the end time, c is the amplitude, and Ut is a parameter that captures the "average solar activity level" at time tSSN (for cycle i) is a strictly increasing function for the rising phase and a strictly decreasing function for the falling phase . Alpha values are the shape parameters, assumed to be constants. NOTE: The given data is then used to approximate the solar cycle Yu et al. (2012)

2. rise time
3. fall time

Result: Cross correlation function fit

The data set also utilizes Spearman's rank coefficient , which is always between -1 and 1.

After analysing the modeled data, a combination of parameters is used to define the statistical model for the next cycle's metrics. The correlation measured with the rank coefficient between and within cycles in the data-set are observed for the amplitude , rise time and fall time.

Notably, within a cycle, there exists a robust anti-correlation between amplitude and rising time, as well as between rising time and fall time, suggesting that more such cycles ascend rapidly and descend more gradually.

?? and **??** , showcasing the relationship of a cycle with the previous and next one, show the continuity of amplitude correlation across cycles, and a weak inter cycle relationship. The dataset also functions as a proof for the the Waldmeier effect, which correlates a cycle's rising time and amplitude in an inverse relationship.

Accounting for this, the rise time, fall time and amplitude is corrected on at the end of this step, and taken into account along with the parameterized function to create a predictive model.

2.2. *Prediction model*

Yu et al. (2012) uses baysian inference via PyStan and MCMC methods to sample from the posterior distribution of the model parameters to estimate latest cycle

This involves implementing a two-step hierarchical model to estimate correlations and using Markov Chain Monte Carlo (MCMC) for sampling the posterior distribution, aiming to create a predictive model based on the current dataset.

Building upon the prior distribution, they introduce hyperparameters each with its own hyperprior resulting in a multilevel model structure, to account for the cross correlation functions, which ultimately depends the original parameterized function.

From this, MCMC is used to draw out the posterior distribution. Since the dataset deals with a lot of dimensions, it combines Gibbs sampling and M-H strategy, updating each parameter using Gibbs sampling, and uses the M-H strategy for the process of proposing and accepting/rejecting new values for each parameter at each iteration within the Gibbs cycle, taking advantage of its conditional draw strategy (Yu et al. (2012))

3. RESULTS AND FUTURE STEPS

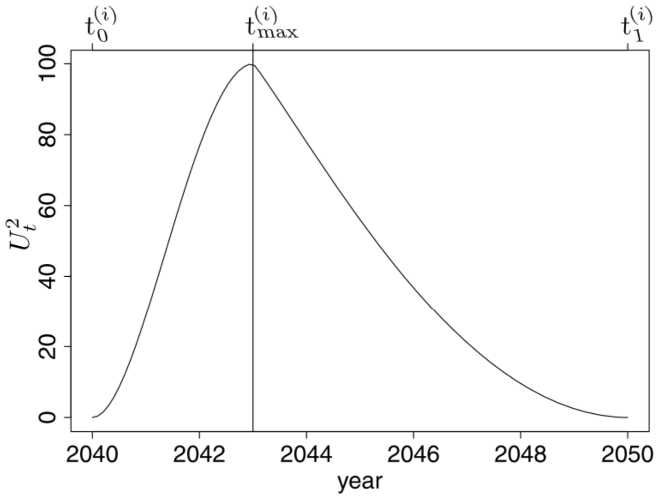

Figure 2. This is the parameterized form of the solar cycle from Yu et al. (2012). Using the equations and values highlighted in 1. The given dataset has U_t squared as the initial value is analogous to the sqyare root of Sunspot numbers (SSN)

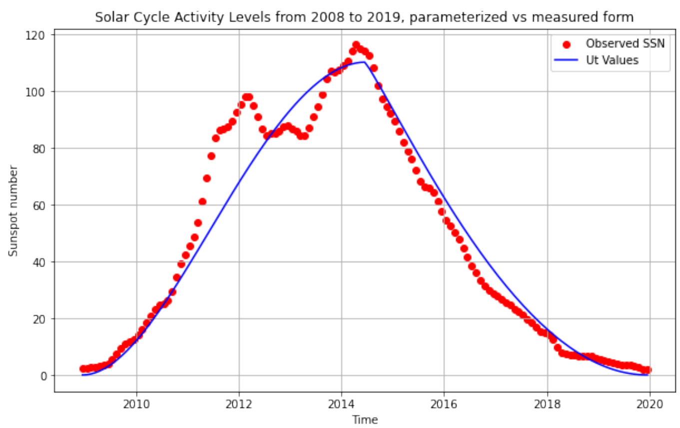

Figure 3. This is the parameterized form of the solar cycle in comparison to the data from the actual 24th solar cycle as shown.

Ultimately, it was observed that more data-points on current cycle correlate to more accurate graphs, and their predictions for the 24th Solar Cycle are given in 6, which measure favorably when compared to actual observation records Yan et al. (2021).

The two step baysian heirarchial modelling could be replicated in order to produce a better predictive model to see how it stacks up compared to the model created by Yu et al. (2012). Since more data points correlate to a better prediction for a given solar cycle, it will be interesting to see if our data-set still aligns with this expectation. It would also be useful to compare our results with the follow up study done on solar cycle 25 by Yan et al. (2021), as we are further along solar cycle 25 as of this point.

Basic regression analysis taking the parameterized equation was done, however, it yielded insufficient results as shown in 7 , highlighting the importance of created better models for the given data set and using more complex analysis.

On this note, it could also be interesting to see if there is a function which could work well with even basic fitting techniques, and if accounting for the gaps between cycles is imperative, as while Yu et al. (2012) take it into account, many other predictive models do not.

REFERENCES

Clette, F., Svalgaard, L., Vaquero, J. M., & Cliver, E. W. 2014, SSRv, 186, 35, doi: 10.1007/s11214-014-0074-2

Pesnell, W. D. 2016, Space Weather, 14, 10, doi: https://doi.org/10.1002/2015SW001304

		Amplitude+	Rise time+	Fall time+	Gap+
Amplitude	ρ	0.46 ± 0.02	-0.21 ± 0.06	0.39 ± 0.05	-0.06 ± 0.07
	p	0.03	0.34	0.07	0.80
Rise time	ρ		0.05 ± 0.10	-0.27 ± 0.07	-0.06 ± 0.13
	p		0.83	0.21	0.37
Fall time	ρ			0.20 ± 0.07	0.19 ± 0.07
	p			0.36	0.39
Gap	ρ				0.25 ± 0.10
	p				0.25

Figure 4. Correlation of one cycle with the next

		Amplitude	Rise time	Fall time	Gap
Amplitude−	ρ	0.46 ± 0.02			
	p	0.03			
Rise time−	ρ	-0.17 ± 0.08	0.05 ± 0.10		
	p	0.44	0.83		
Fall time−	ρ	0.07 ± 0.05	-0.08 ± 0.06	0.20 ± 0.07	
	p	0.76	0.72	0.36	
Gap−	ρ	-0.37 ± 0.07	0.29 ± 0.07	-0.25 ± 0.09	0.25 ± 0.10
	p	0.08	0.18	0.25	0.25

Figure 5. Correlation of one cycle with the previous cycle

Yan, Y., Stenning, D. C., Kashyap, V. L., & Yu, Y. 2021, Research Notes of the American Astronomical Society, 5, 192, doi: 10.3847/2515-5172/ac1ea0

Yu, Y., van Dyk, D. A., Kashyap, V. L., & Young, C. A. 2012, SoPh, 281, 847, doi: 10.1007/s11207-012-0090-x

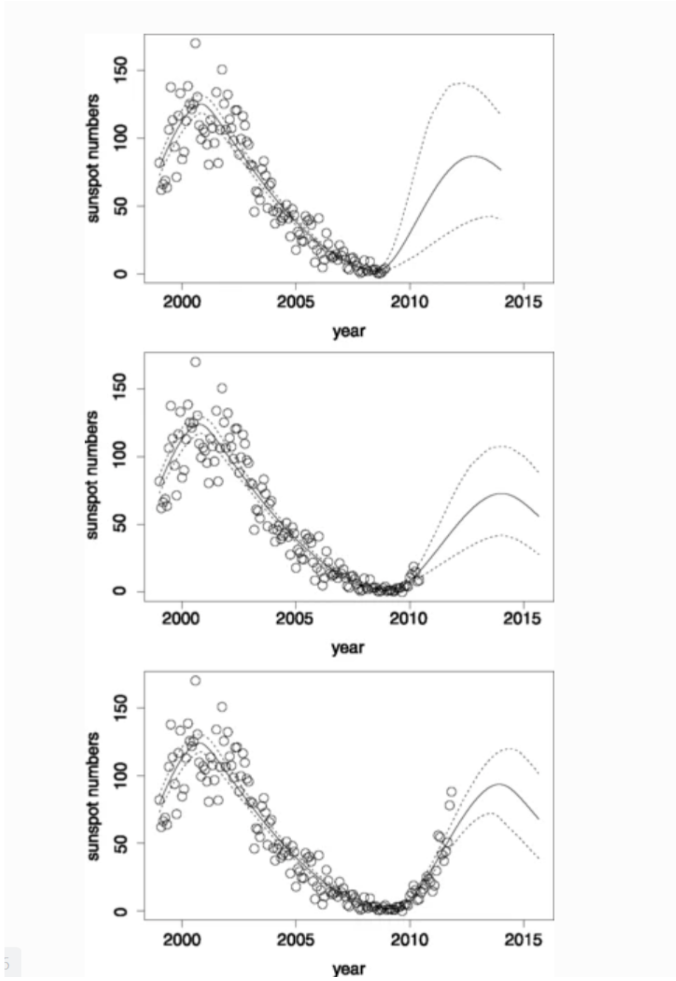

Figure 6. Predictions for Solar Cycle 24, based on data up to November 2008, May 2010, and October 2011. There is decreasing uncertainty as more current cycle data is included , with the top figure representing a new cycle.

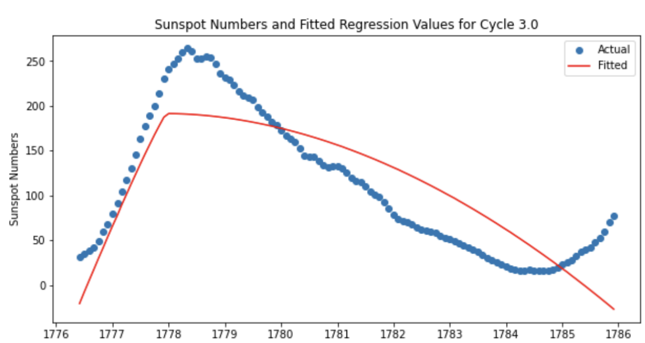

Figure 7. Given is a basic regression model for the parameterized function to the actual dataset using the values of amplitude and time given in Yu et al. (2012). The model breaks down for subsequent cycles, highlighting the need for a model robust model.