May 2019

Understanding Solar Activity During the Last 400 Years

Sam Q. Hollenbach
Macalester College, shollenb@macalester.edu

Abstract

The solar cycle has a profound effect on both terrestrial and extra-terrestrial operations. Understanding the history of the solar cycle is necessary for studying long term trends, however older data is difficult to calibrate due to the sparsity of observations. In this paper we propose a new method for calibrating sunspot number data that does not rely on observational overlap. Initial testing shows promise in this method's success, though more work must be done to ensure calibration consistency across all observers.

Follow this and additional works at: https://digitalcommons.macalester.edu/mjpa

Part of the Physics Commons, and the The Sun and the Solar System Commons

Recommended Citation

Hollenbach, Sam Q. (2019) "Understanding Solar Activity During the Last 400 Years," *Macalester Journal of Physics and Astronomy*: Vol. 7 : Iss. 1 , Article 4.
Available at: https://digitalcommons.macalester.edu/mjpa/vol7/iss1/4

Footer Logo

Understanding Solar Activity During the Last 400 Years

A Preprint

Sam Q. Hollenbach
Macalester College
Oakland, CA 94610
samhollenbach@gmail.com

Andés Muños-Jaramilla
Southwest Research Institute
Boulder, CO, 80301
amunozj@boulder.swri.edu

April 17, 2019

1 Introduction

Due to its proximity and importance to human existence, the sun is one of the most studied and well understood objects in astronomy. Since telescopes have been around, diligent records have been kept of solar activity. As early as 800 B.C., astronomers began to notice small areas on the sun that appeared darker than their surroundings, and called them "sunspots".[1] It was later discovered that these regions are indicators of solar regions with concentrated magnetic fields, and often show up in pairs due to magnetic polarity.[2] A pair or closely located group of sunspots is now often called as an "active region". If we plot the number of active regions over time we see a very clear trend form. The number of active regions on the sun varies quite regularly on an 11 year cycle, where at the peak (solar maximum) there will be numerous active regions, and at the trough (solar minimum) there will be close to none.[3] This cycle has persisted since at least the 1750's, when active region count (sometimes known as Wolf number) records started to be kept. The solar cycle has a significant impact on space weather, and must be taken into account when considering many physical scenarios in space, the upper atmosphere, and even the surface of the earth.

Examining older active region data also shows some peculiar trends. One famous feature of the solar cycle is known as the Maunder Minimum, and is a period of roughly 70 years (over 6 solar cycles) starting at 1645 where the reported sunspot numbers were extremely low.[4] This feature has been studied extensively, though there is disagreement as to *how low* the sunspot numbers actually were, since we have not seen such an event in the time since. Another trend that some solar cycle reconstructions illuminate is a general increase in the total number of active regions at solar maximum. The current alleged higher than normal sunspot count has been labeled the "Modern Maximum".[5] Some anthropogenic climate change deniers believe that this increase in solar activity could actually be the true cause of climate change.[6] Many attribute to this trend to historical observers having less powerful tools to measure sunspots, effectively lowering the counts by being unable to see some smaller active regions.

Although the historical records all show that the same 11-year cycle, the actual number of active regions, especially during solar maximum, is hard to verify. Many groups have made attempts to calibrate the older active region number data, but there is serious disagreement among methods. Much of the disagreement comes from the fact that most observations from observers before 200 years ago have very little to no overlap with other observers, as seen in Figure 1. Our goal is to calibrate historical active region observer records as accurately as possible, by eliminating the need for observational overlap.

2 Approach

The lack of overlap in observing records makes traditional calibration techniques extremely difficult. Our approach differs from others in that it allows for independent observer calibration, comparing only to a single **reference observer** with high accuracy. For clarity, in this section we refer to the observer we want to calibrate as the **calibrating observer**.
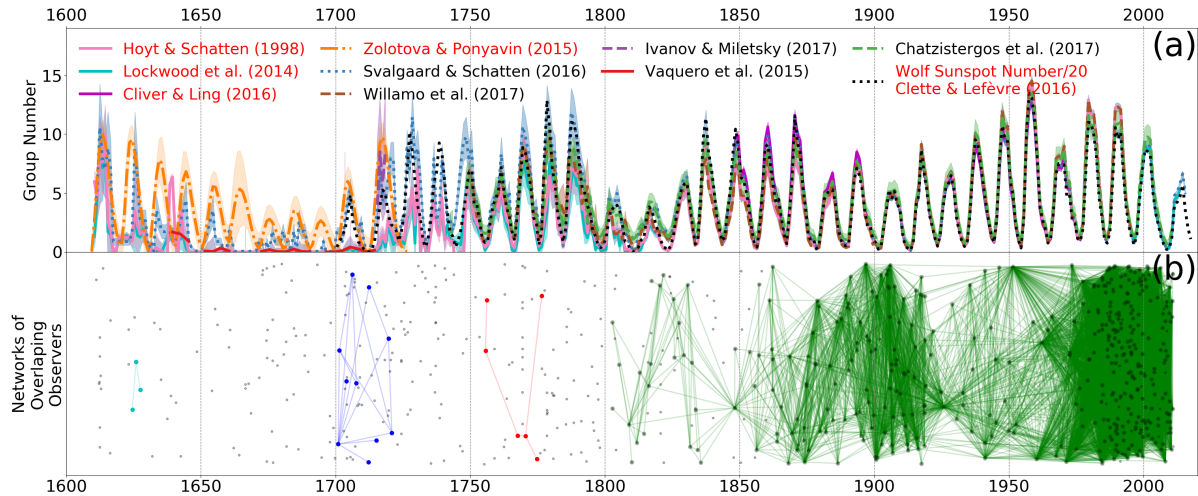
Figure 1: (a): Solar cycle reconstruction attempts. Notice the disagreement among methods increases before 1800. (b): Each node represents a single observer for which we have data. An edge connects two observers if they have more than 30 days of observing in common. Notice the lack of connection between observers before 1800, which is the main cause of the disagreement in the top panel.

2.1 Observer Thresholds

In order to calibrate observers without any observational overlap, we make an assumption that the amount of sunspots recorded over time by an observer is directly related to the quality of their observing tools. For example, an observer from the 1700s with a rudimentary telescope may not have been able to see smaller active regions that modern observatories can now record. We do not have a record of every observers tools, so to solve this problem, we introduce the idea of an observational **threshold**. The threshold is defined as the size of the smallest active region that this observer could have seen, in MSH (millionth of solar hemisphere). Our task is to calculate the appropriate threshold for each observer, which lets us calibrate the observer's data properly against a **reference observer**. Since most historical data does not have sizes accompanying the active region count data, we calculate the threshold by applying it to the reference observer, and compare the amount of **active** and **quiet days**.

2.2 The Reference Observer

Instead of only comparing observations from similar time frames to each other, we instead compare all observations from the observer we want to calibrate to a **reference observer**. The reference observer can be any modern observatory that has a significant span of complete active region data, including both counts and sizes of active regions per day. We choose to calibrate against this data because we know the active region count calibration is self-consistent throughout the data set. Currently we are using the Royal Greenwich Observatory (RGO) as our reference observer. Figure 2 shows a visualization of a few reference observer data sets.

2.3 Active and Quiet Days

Without observational overlap, we must disregard the reported sunspot numbers, and instead mark each observed day as either an **active day** or a **quiet day**. In this formulation, an **active day** is defined as a day with 1 or more active regions recorded, while a **quiet day** is a day with 0 recorded active regions. For the observer we are attempting to calibrate there is only one way to calculate this without knowing the sizes of each active region. However, for our reference observer, setting different thresholds will produce different results. For example, if a reported day has a single active region of size 25 MSH, originally this would be counted as an **active day**. Though if we set a threshold of 30 MSH, since the only active region is now below the threshold, this day would be reported as a **quiet day**. What is interesting to us is not any individual day, but the proportion of active and quiet days over a time span, or as we call it the **Active Day Fraction**.
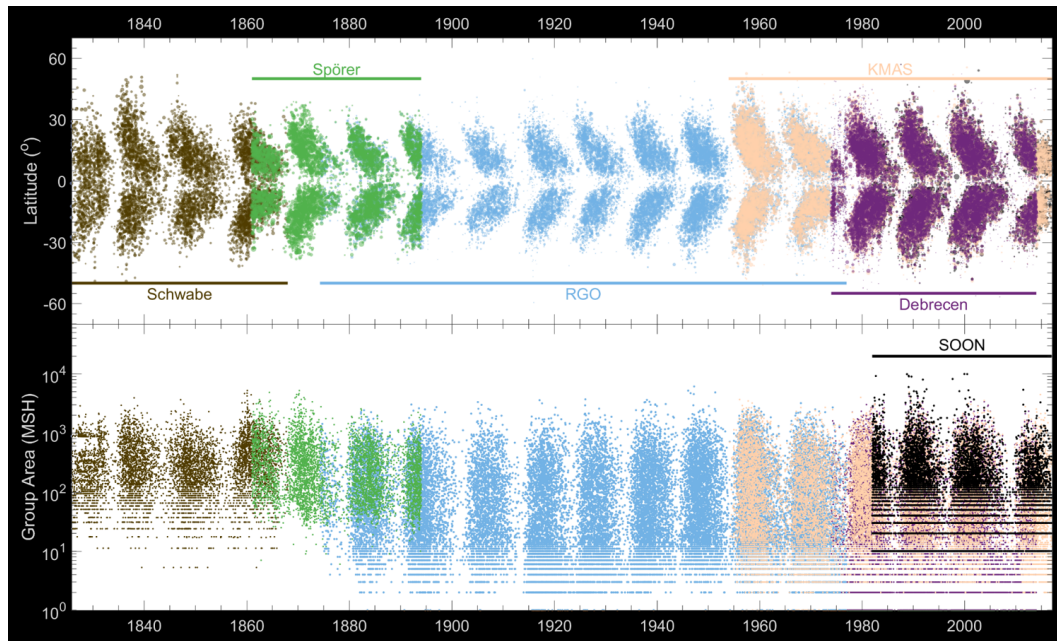
2

Figure 2: Position and size of sunspots within our reference observer data sets. We are using the RGO (blue) as our main reference observer. The top panel shows the latitudinal position of of the recorded sunspots over time, often referred to as a butterfly diagram. The bottom panel shows the size of each recorded sunspots. One can imagine an observational threshold as eliminating any sunspots below a certain horizontal line on the bottom panel.

2.3.1   Active Day Fraction

The idea behind the **Active Day Fraction**, or **ADF**, is to examine the fraction of active days or quiet days over a discrete time span (on the order of a month). During solar minimum, the amount of quiet days should be higher, while during solar maximum the amount of active days should be higher. When looking at Figure 3, which shows a section of the reference observer data converted to active and quiet days, the solar cycle is still very apparent. In this figure each column is 30 days of observations from our reference observer, and the different colors represent the amount of active and quiet days recorded within those 30 days. When the observational threshold is increased (moving up panels) the amount of quiet days increases, and changes the ADF for each month. We can apply the same ADF formulation to the observer that we want to calibrate and make a similar plot, though again we cannot apply a threshold because we do not have size data. However, we can look at the active day fractions for our calibrating observer and find a portion of time where our reference observer ADFs looks similar. The easiest way to compare the two ADF portions is by looking at the **ADF distribution**.

The **Active Day Fraction distribution** is simply the distribution of the ADFs for each month we are examining. Once we create the ADF distribution for the data we are calibrating, we want to find a period of time within the reference observer in which the ADF distribution closely matches. This is based on the assumption that the shape of the solar cycle is predictable, so while the overall level of solar activity may change or trend over time, the motions of the rise and fall of solar activity over each cycle are similar. This assumption allows us to compare historical data to our reference by finding a period of time where the shape of the rise and fall periods of the two cycles are similar. To find the appropriate time period we scan along the reference observer data in chunks (lowering chunk length will increase accuracy but also increase computation speed). To avoid artificially increasing accuracy when running validation on this method (see section 3), we ignore the time section where the calibrating data and reference data overlap. For each reference section we apply a number of thresholds ranging from 0 to 100 MSH, and create a ADF distribution for each. We compare each distribution with our calibrating distribution, and compute an earth mover's distance (EMD), a simple measure of distribution similarity.[1] Minimizing the EMD will give us the most similar reference distribution to the distribution we are calibrating (See Figure 4). This distribution has the threshold that we will assign to our calibrating observer. Finding the proper threshold for this calibrating observer will finally allow us to properly calibrate their data. This is done by determining the fraction of active regions that were eliminated from our reference data by applying the threshold, and scaling the calibrating observer active region count data up by this fraction.

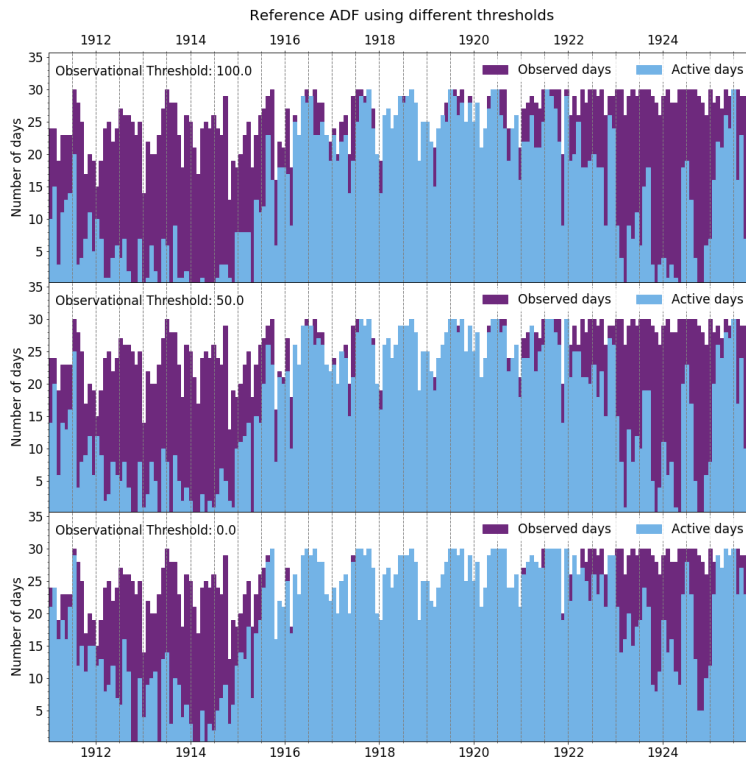[1]See https://github.com/wmayner/pyemd for more explanation of the EMD

Figure 3: Active day (blue) and quiet day (purple) visualization for our reference observer, with 0.0 (bottom), 50.0 (middle) and 100.0 (top) MSH thresholds. Each column represents a month of observations. Notice the solar cycle is visible, as solar minimum shows many more quiet days, while solar maximum has nearly all active. Furthermore, the amount of quiet days increases as we increase the observational threshold.
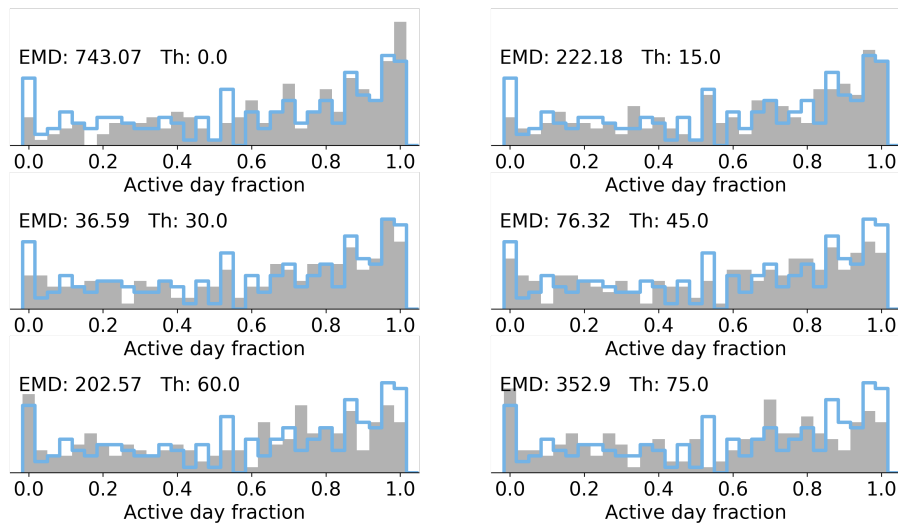

Figure 4: ADF distribution for calibrating observer (blue) and reference observer with numerous thresholds (grey). Notice the EMD is minimized at 30.0 MSH threshold, showing that this is the proper threshold to assign to our calibrating observer.
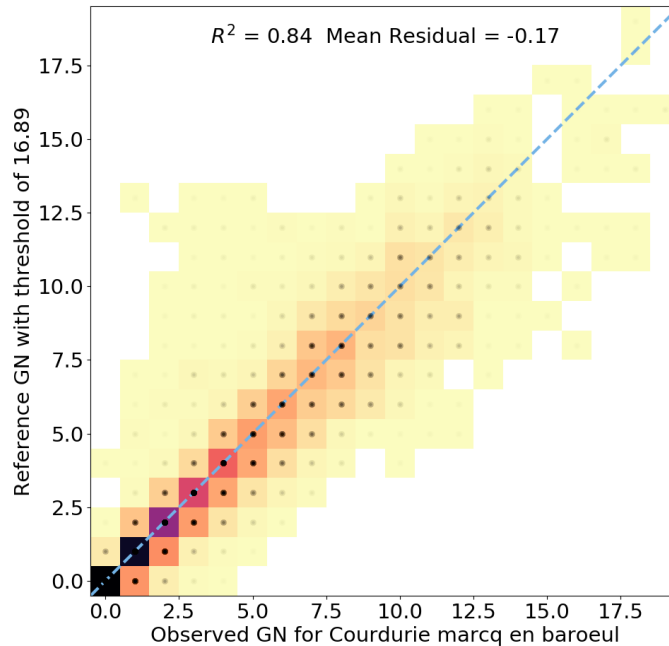
4

Figure 5: Checking accuracy of computed threshold for one observer. The horizontal axis shows the reported group numbers (active region count) from the calibrating observer, while the vertical axis shows the newly calculated group numbers from the reference observer with the computed threshold. The density of data points is shown by color. If a data point is marked on the identity line it means that the calibrating group number and threshold-applied reference group number of a single day of observer is equal. However, if the group number calculations for a single day disagree, the point will show up outside of the identity line. With perfect method accuracy, all points would appear on the identity line, though most examples show some disagreement.

3 Method Accuracy

Once we find the appropriate threshold for our calibrating observer, we must make sure that using this method is worthwhile. We have a number of observers that overlap with our reference observer, which we can use to test the method accuracy. Applying to ADF distribution method detailed throughout section 2, we can compare calibrated results with our reference observer (which we assume is completely accurate) to see how well our method performs. Figure 5 shows an example of how we measure performance and obtain an R^2 value for each calibration. One key feature of our method that can have a massive effect on performance is how we calculate the Active Day Fraction. In section 2.3 we discussed the basic concept of the Active Day Fraction, however in order to properly diagnose performance we need to examine how we calculate the ADF more closely.

3.1 ADF Calculation Types

The ADF is a simple fraction, where we can choose to alter the numerator and the denominator. The basic format of the ADF is:

$$ADF = \frac{\text{Active Day Calculation}}{\text{Month Length Calculation}}$$

We initially specified two calculations for our numerator and two calculation for our denominator, totally four unique ways to compute the ADF. The formulations are as follows, with labels for each that we will refer to throughout this section.

Numerator: **Active Day Calculation**

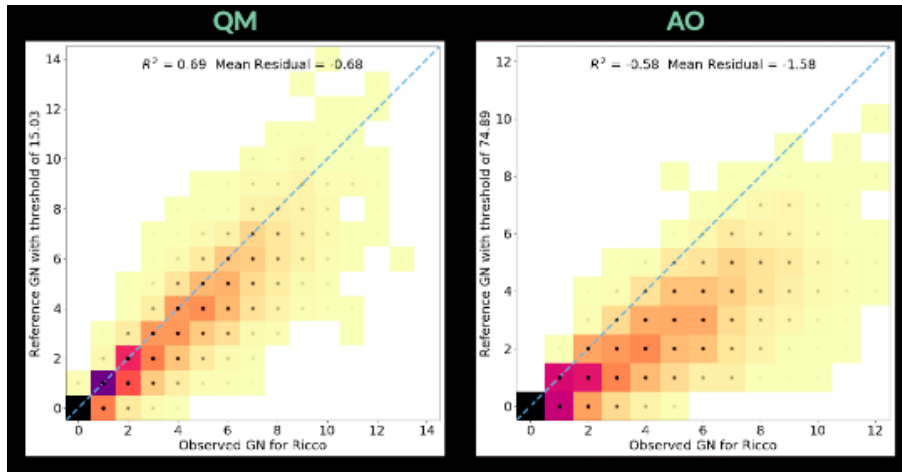| Label	Meaning
A	Total Active Days
Q	Inverse of Total Quiet Days

5

Figure 6: Measuring the performance of the same observer with different ADF calculation methods can reveal extremely different results. This example shows the performance plots (See Figure 5) with the same observer run with ADF methods QM and AO respectively. The observer performs very well with the QM method, but very poorly with the AO method. Not all observers show this much disparity, but it is important to keep track of which methods perform best for certain observers.

Denominator: **Month Length Calculation**

Label	Meaning
M	30 Days
O	Days with Observations

Combining these pieces into our four Active Day Fractions we achieve the following calculations with corresponding effects.

ADF Calculation Methods and Effects

Label	Definition	Effect
A/M	Total Active Days / 30 Days	Assumes missing days are Quiet days
A/O	Total Active Days / Observed Days	Assumes missing days follow similar ratio as observed days
Q/M	1 - Quiet Days / 30 Days	Assumes missing days are Active days
Q/O	1 - Quiet Days / Observed Days	Assumes missing days follow similar ratio as observed days

When applying all four ADF methods across our test observers (observers we can compare to reference), we noticed that some observers performed much better with certain ADF methods than others (See Figure 6). In order to determine which method an observer should use, we compared observers that perform significantly better on each specific method with each other, in an attempt to uncover patterns that could point us to which method to use for a specific observer. In the process we discovered that the **A/O** and **Q/O** methods were performing very similarly, and chose to replace the **Q/O** method completely with **A/O**. This reduced the amount of choices from four to three, which helped our fitting attempts. Figure 7 shows that each of these three methods now does perform significantly better than others for some observers.

3.2 Determining Best ADF Method

Once we reduced the number of ADF calculation methods from four to three, we made attempts to fit the each observer with the ADF method that optimized accuracy, using a number of machine learning techniques. We needed to develop a system to assign ADF methods based only on the initial data, since we cannot test individual ADF method performance for observers with data that does not overlap with our reference observer. We hypothesized that machine learning could be a good way to tackle this problem since it is a three group classification problem. In addition to approaching this as a classification problem, we furthermore tried to decide what method an observer should use by predicting how well a certain method will perform (predicting R^2). We used Support Vector Machines (SVM) [SciKit Learn], and Neural Networks (NN) [TensorFlow + Keras], two of the most common machine learning classification methods. However neither of these methods achieved above 0.50 classification accuracy. There is ongoing work to better these results through improving feature selection and hyper-parameter tuning.
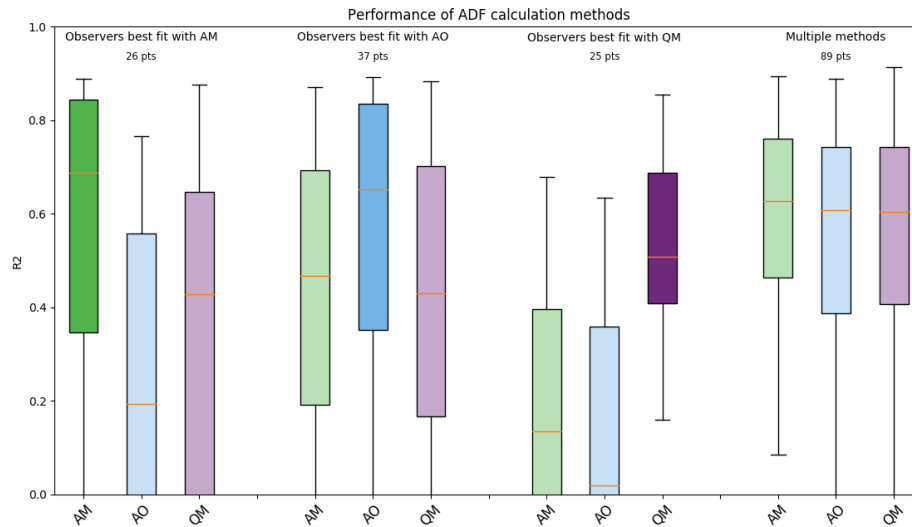
6

Figure 7: Each section of 3 boxplots shows a different ADF method run on the same group of observers. The vertical axis shows performance. These groups were split up by determining which ADF method performed the best for each (the best performing method's boxplot is darkened). The purpose of this visualization was to determine if it was worthwhile to continue using multiple different ADF methods. We found that when a method performs well for an observer, it generally performs significantly better than the other methods. This prompted us to keep using these three methods for our ADF calculations.

4 Next Steps

The current focus of this project is diagnosing performance, and finding ways to improve method accuracy. We are not yet satisfied with the level of accuracy we are achieving across all test observers, however some observers do show high levels of accuracy, which gives us confidence in this model. Observers that exhibit very low accuracy are disregard under the assumption that their data is inaccurate and not self-consistent in reported activity level. Significant improvements to accuracy have been made since the beginning of this project since applying different ADF. As we uncover patterns within observers we expect this trend to continue.

The last step in this project is to apply our calibration method to the full set of active region number data. This will produce a full solar cycle history plot, akin to those seen in 1. We believe once we achieve high accuracy with our test observers that our solar cycle reconstruction will be extremely trustworthy for calibrating historical observers, since the method does not rely on observational overlap. We hope to eventually use our reconstruction to evaluate interesting solar cycle features like the Maunder Minimum, and resolve disagreement on long term trends.

References

[1] High Altitude Observatory. "History of Solar Physics: A Time Line of Great Moments: 1223 BC–250 BC". University Corporation for Atmospheric Research.

[2] NOAA. "Sunspots/Solar Cycle." https://www.swpc.noaa.gov/phenomena/sunspotssolar-cycle

[3] Tribble, A. "The Space Environment, Implications for Spacecraft Design." Princeton University Press (2003)

[4] Usoskin; et al. "The Maunder minimum (1645–1715) was indeed a grand minimum: A reassessment of multiple datasets". Astron. Astrophys. 581: A95. (2015).

[5] Solanki, Sami K.; Usoskin, Ilya G.; Kromer, Bernd; Schüssler, Manfred; Beer, Jürg "Unusual activity of the Sun during recent decades compared to the previous 11,000 years" Nature. 431 (2004)

[6] Swiss National Science Foundation "Sun's impact on climate change quantified for first time" https://phys.org/news/2017-03-sun-impact-climate-quantified.html, March 27, 2017

7