Characterizing the Time Evolution of Free-Energy Proxies to Forecast West Limb Flares, CMEs, and SEPs

Ayla Weitz[1], David Falconer[2, 3]

[1]University of California, Berkeley; [2]NASA Marshall Space Flight Center; [3]University of Alabama in Huntsville

THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

CENTER FOR SPACE PLASMA & AERONOMIC RESEARCH

Introduction

Predicting an active region's (AR) tendency to produce major flares, coronal mass ejections (CMEs), and solar energetic particle events (SEPs) is essential for ensuring astronaut safety. MagPy predictions are derived from free-energy proxies from HMI vector magnetograms. Due to projection effects, we show magnetic measures in JSOC deprojected cylindrical magnetograms are over-estimated as a function of distance from disk center. To better forecast eruptions from regions far from disk center, most notably the west limb, we want to accurately predict free-energy proxies several days into the future. **Figure 1** illustrates how measuring far from disk center gives rise to measurement challenges. We determine that a correction must be applied to more accurately predict AR free-energy proxies west of central meridian. With the corrected data, we investigate multiple methods of using observations of eastern longitudes to predict the west for a variety of lead times. Because of the similarity we found between the free-energy proxies we investigated, the results shown are for the gradient weighted neutral line measure defined as $WL_{SG1} = \int |\nabla B_z|^1 \, dl$.

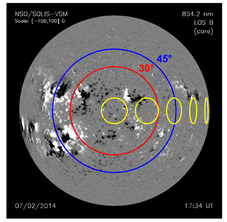

Figure 1. Full disk magnetogram (nso.edu/data/nisp-data/magnetograms/). Yellow circles are overlaid to depict how the area of an AR becomes increasingly distorted the farther it is from disk center.

Methods

Code was developed in Python and we used a hour cadence MagPy output spanning from May 2010 to June 2020 (sharp_cea_720s).

We worked with only large AR which we defined as AR with total magnetic flux greater than 10^{22} Mx at closest approach to central meridian. This gave us a population of 606 AR. The following procedure was performed on various free-energy proxies:

- Normalized data by dividing all of an AR's measures by the AR's measure value when closest to central meridian.
- Plotted normalized measures as a function of degree distance from disk center. We notice a pronounced dependence on radial distance. Because distance from disk center is simply a product of Earth's perspective of the Sun, we deduce this dependence must be an artifact.
- To correct for this projection error, we fit the data with Chebyshev fits[4]. We are interested in the measure's time evolution from east to west, so we used only even degree Chebyshev fits to preserve east-west symmetry.

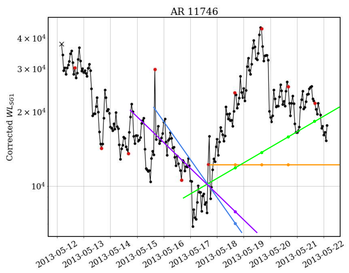

Figure 2. WL_{SG1} corrected by 6th degree Chebyshev fit (see Figure 3) as a function of time for AR 11746. Graphical representations of linear fit predictions for a variety of prediction times are overlaid: a persistence model—only using central meridian as predictor (orange), using central meridian and up to one day before (green), using central meridian and up to two days before (blue), using central meridian and up to 3 days before (purple). These lines are overlaid for visualization and are not true depictions of the actual models used.

- After applying the corrections, we investigate the free-energy proxies' time evolutions for multiple time steps (24, 48, 72, 96 hours) using a variety of methods. We varied the following parameters:
  - Time before central meridian used for prediction (24, 48, 72, and 96 hours)
  - Prediction function (persistence, linear, quadratic, and cubic)

  For a persistence model, the measure value at the time of closest to central meridian approach is used as a prediction for all future times. For linear, quadratic, and cubic models, the function is fit to the central meridian measure value and all values before that are within the time used for prediction. All prediction functions are fit in linear space. A graphical depiction of linear prediction functions using different prediction times is shown in **Figure 2**.
- To determine how well a method predicts, we construct histograms of the ratios of observed values to values the model gives. We fit third order Gaussians to the histograms and conclude that the most accurate method should minimize the second order term (standard deviation).

Results

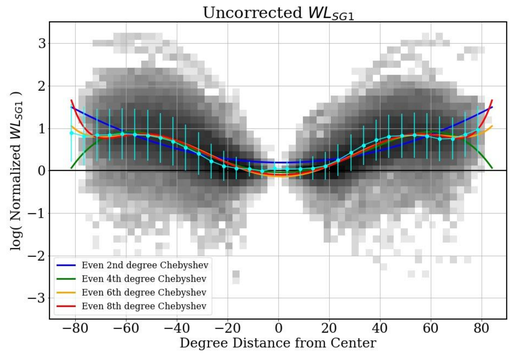

Figure 3. Normalized WL_{SG1} as a function of degree distance from disk center. Means of each 5 degree bins are shown (cyan) along with the bin's standard deviation (vertical cyan lines). Even 2nd, 4th, 6th, and 8th degree Chebyshev fits are applied to characterize radial dependence.

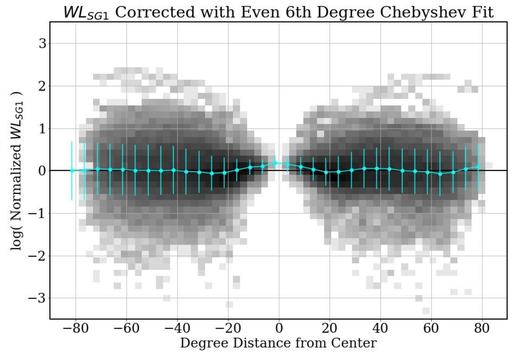

Figure 4. Normalized WL_{SG1} detrended with the even 6th degree Chebyshev fit (orange line in Figure 3). Means and standard deviation of the corrected data are shown in cyan.

- **Figure 3** shows that for WL_{SG1} there is a strong dependence on degree distance from disk center. To characterize this behavior, we fit the data with even degree Chebyshev fits. Using even degree polynomials gave similar fits.
- After correcting the data with an even 6th degree Chebyshev function, this effect was effectively removed as seen in **Figure 4**. Because of the lack of data near disk center, we overcorrect in that region. Investigating other weighted neutral line measures show a similar dependence on radial distance. In the following figures and analysis, corrected data is in reference to this even 6th degree Chebyshift applied.
- **Figures 5** and **6** show ratio histograms of WL_{SG1} for a persistence prediction model for two lead times: 24 and 72 hours. We expect the ratio distributions to be roughly centered about zero as AR should, on average, be growing and and decaying by similar amounts.
- These histograms were made for all lead times of interest (24, 48, 72, and 96 hours) and for other free-energy proxies. Other lead times show a similar trend in the uncorrected data as the two shown—the longer the lead time, the farther the fit peak moves from zero.

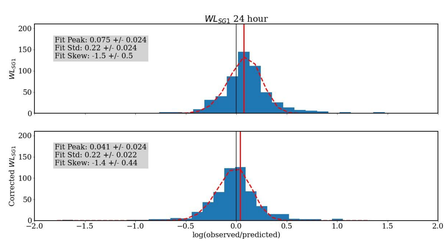

Figure 5. Ratio histograms of WL_{SG1} for a 24 hour lead time for the uncorrected data (top) and the corrected data (bottom). The 3rd order Gaussian fit (dashed red) and peak location (vertical red) are shown. For this lead time, both the uncorrected and corrected data show a very similar distribution.

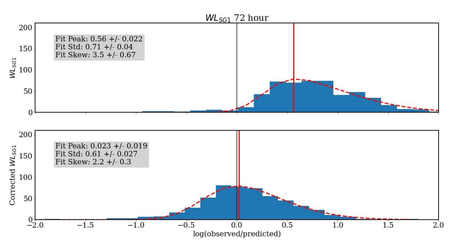

Figure 6. Ratio histograms of WL_{SG1} for a 72 hour lead time for the uncorrected data (top) and the corrected data (bottom). The 3rd order Gaussian fit (dashed red) and peak location (vertical red) are shown. The uncorrected data shows that, on average, AR grow by roughly a factor of 3 in 72 hours. The corrected data shows a distribution peaking close to zero.

- Because we want to find a prediction method that best fits the majority of AR, the best method should minimize the second order term of the Gaussian fit. A summary of this fit term for the various method we used is shown in **Table 1**.
- The persistence model dominates for 24 and 48 hour lead times. For lead times longer than 48 hours, we see linear and quadratic models giving smaller second order terms than the persistence model. Yet these come with great uncertainty, and thus we cannot confidently say these prediction functions give the best estimates for those lead times.

Second Order Terms for Various Prediction Methods

Prediction Fit Function	Time used for prediction	24 Hour Lead Time	48 Hour Lead Time	72 Hour Lead Time	96 hour Lead Time
Persistent	-	0.22 ± 0.02	0.44 ± 0.04	0.61 ± 0.03	0.61 ± 0.03
Linear ($y = a + bx$)	24 hour	0.43 ± 0.05	0.64 ± 0.06	0.82 ± 0.04	0.85 ± 0.10
	48 hour	0.42 ± 0.05	0.60 ± 0.07	0.49 ± 0.20	0.55 ± 0.10
	72 hour	0.38 ± 0.04	0.60 ± 0.06	0.54 ± 0.10	0.79 ± 0.07
	96 hour	0.35 ± 0.02	0.58 ± 0.05	0.68 ± 0.07	0.77 ± 0.07
Quadratic ($y = a + bx + cx^2$)	24 hour	-	-	-	-
	48 hour	0.60 ± 0.04	0.49 ± 0.06	0.53 ± 0.10	0.50 ± 0.20
	72 hour	0.56 ± 0.04	0.69 ± 0.05	0.68 ± 0.05	0.60 ± 0.09
	96 hour	0.46 ± 0.04	0.46 ± 20	0.67 ± 0.30	0.70 ± 0.20
Cubic ($y = a + bx + cx^2 + dx^3$)	24 hour	-	-	-	-
	48 hour	0.64 ± 0.05	0.55 ± 0.05	0.62 ± 0.06	0.64 ± 0.06
	72 hour	0.50 ± 0.07	0.48 ± 0.10	0.63 ± 0.03	0.54 ± 0.05
	96 hour	0.59 ± 0.05	0.70 ± 0.05	0.58 ± 0.04	0.64 ± 0.07

Table 1. Table of the second order fit terms to ratio histogram for the various methods we investigated. The prediction fit functions were applied in linear space. The minimum value for a lead time is highlighted in yellow. Rows are empty if there were too few data points for the prediction function to be fit.

Conclusions and Future Work

There are significant projection errors in JSOC deprojected cylindrical magnetograms for several free-energy proxies that need to be corrected. The even 6th degree Chebyshev fit effectively removes the artificial radial distance dependence. Whether there are other dependencies that need to be corrected for should be investigated.

Of the methods we analyzed, persistence seems to overall give us the best predictions for short lead times (24 and 48 hours). Other techniques gave slightly better results for lead times longer than 48 hours, but with large uncertainties, we cannot conclude that these models are better predictors. Testing more advanced function fitting techniques possibly with machine learning could prove to be fruitful.

Similar analysis should be performed on a MagPy output with thresholds optimized for flare forecasting. Determining these thresholds is currently ongoing. With optimized corrected data, further analysis into determining which free-energy proxy is the best predictor of flares, CMEs, and SEPs should be done.

References

[4]Falconer, David & Tiwari, Sanjiv & Moore, Ronald & Khazanov, Igor. (2016). A new method to quantify and reduce projection error in whole-solar-active-region parameters measured from vector magnetograms. The Astrophysical Journal. 833. 10.3847/2041-8213/833/2/L31.

Acknowledgements

This work was supported by funding from NSF grant AGS-1950831 for the UAH CSPAR/NASA MSFC Heliophysics REU program and NASA 18-2HSWO218_2-0003.

I would like to thank David Falconer for all his patience and guidance throughout the summer. Also I am grateful for the 2021 UAH REU team and all the other interns for all their wonderful support.