Assessing the Impacts of Magnetogram Projection Effects on Solar Flare Forecasting and Extending the SWAN-SF Dataset

Griffin T. Goodwin[1], Viacheslav M. Sadykov[1], Petrus C. Martens[1], Dustin Kempton[2], Reet Gupta[2]
[1]Physics & Astronomy, Georgia State University, Atlanta, GA, USA
[2]Computer Science, Georgia State University, Atlanta, GA, USA

I. What Are Projection Effects?

Near the solar limbs, active region (AR) vector magnetograms and their properties suffer from systematic observational trends unrelated to the evolution of the AR itself (see Figure 1). This complicates flare forecasting and may even introduce unnecessary false positive and negative predictions.

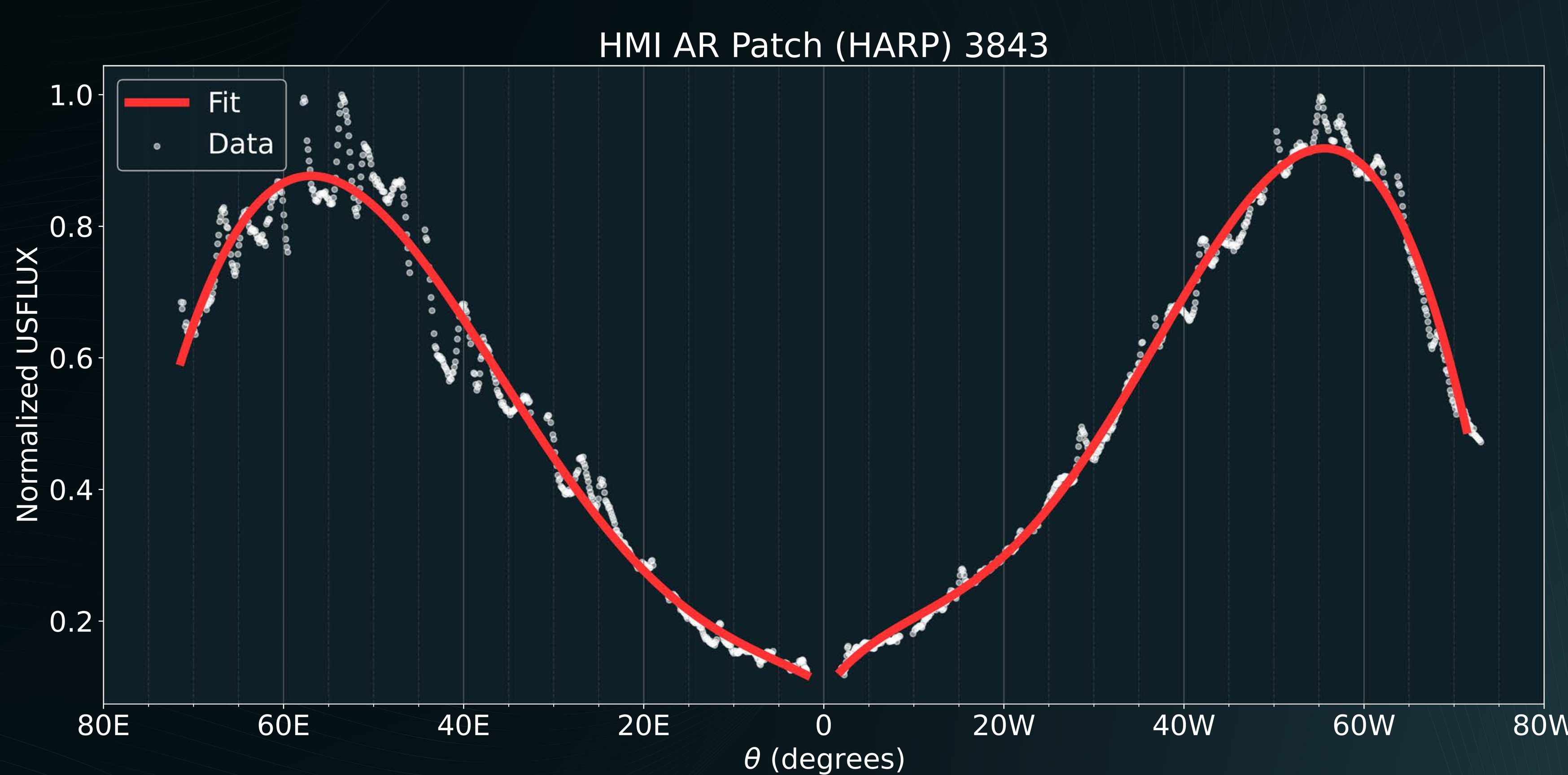

Figure 1: The unsigned magnetic flux versus the angle between the line-of-sight and the surface normal (heliocentric angle, θ) for HARP 3843. Note the synthetic double-peaked trend in the data.

Goal: Mitigate the projection effects present in each of the 24 physics-based magnetogram parameters in Georgia State University's Space Weather Analytics for Solar Flares (SWAN-SF) benchmark dataset. Then, compare performance metrics when forecasting using corrected and uncorrected data.

II. How Do We Correct the Data?

To correct for projection effects, we follow a similar methodology to Falconer et al. (2016). By normalizing the magnetic field data for a given AR to its value near the central meridian and combining the results for over 260 ARs, we can produce a polynomial fit that describes the severity of the projection effects as a function of *θ*. New ARs can then be corrected by rescaling the data back to unity based on the value of the fit at the appropriate angle.

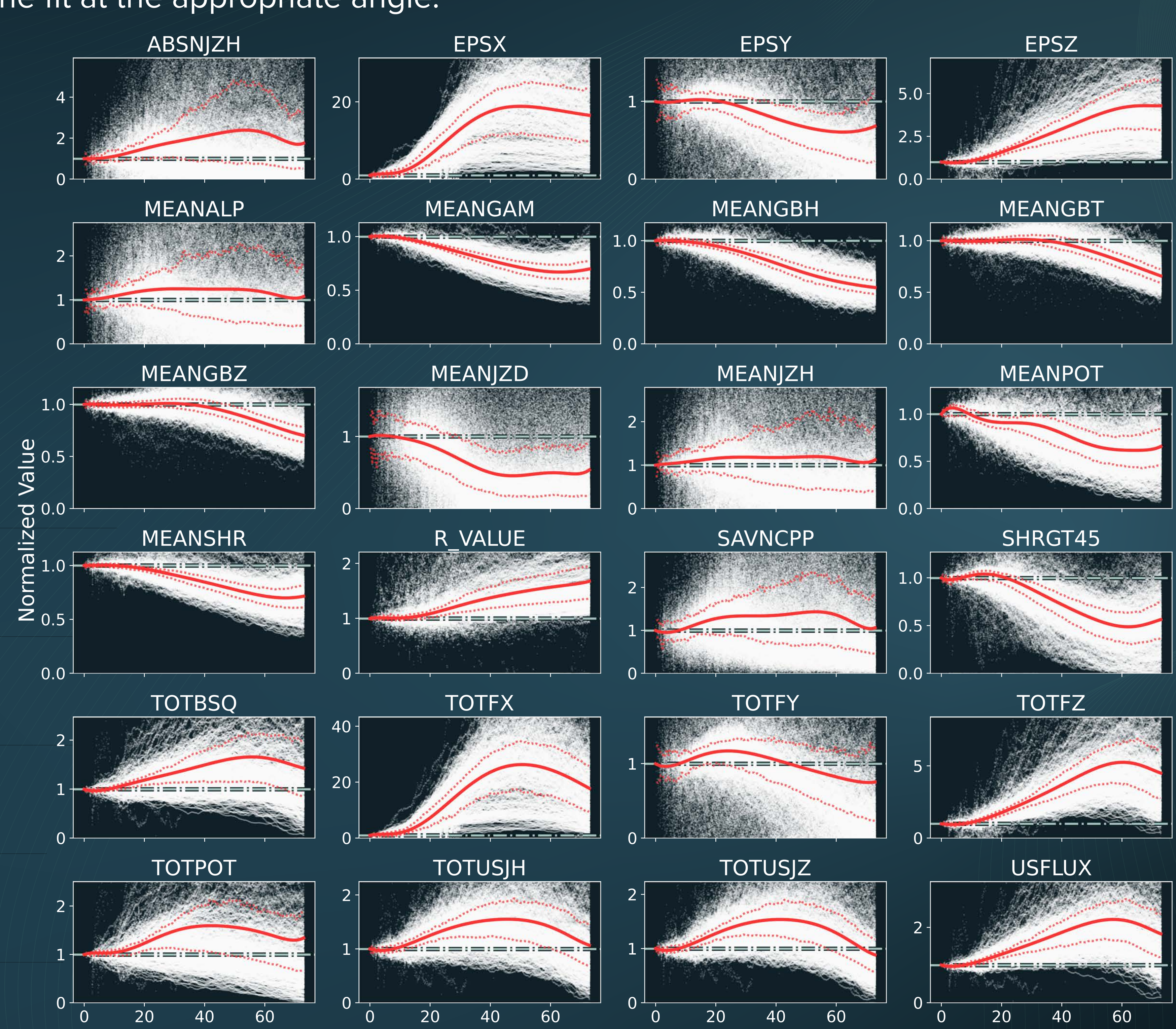

Figure 2: The correction polynomials for the 24 magnetic field parameters in SWAN-SF. The solid red line is the fit to the data in white. The dashed red lines highlight the median absolute deviation of the data above and below the fit. The blue dashed-dotted line emphasizes the point where the correction factor is 1.

III. Can We Improve Forecasting?

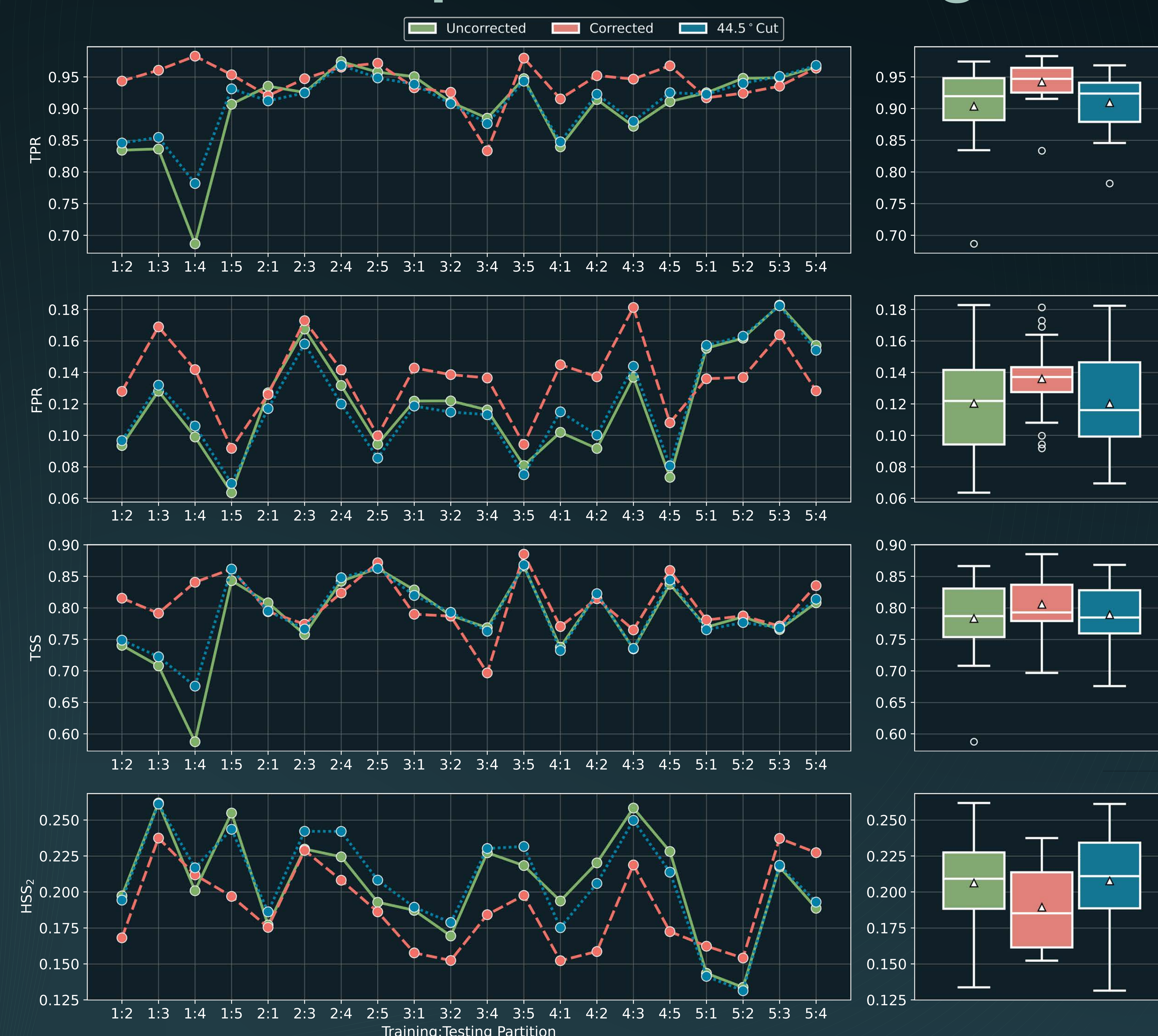

Figure 3: The true positive rate (TPR), false positive rate (FPR), true skill statistic (TSS), and Heidke skill score (HSS_2) for every possible training/testing partition combination in SWAN-SF. Results are color-coded to indicate whether the support vector machine classifier was trained/tested with uncorrected data (green), corrected (coral), or a combination of the two (blue - "44.5° Cut"). The "44.5° Cut" results use the uncorrected forecasts below θs of 44.5° and the corrected forecasts above this angle.

Key Takeaway: Correcting for projection effects has *minimal* impacts on forecasting performance.

IV. Extending the SWAN-SF Dataset

We have identified two main limitations of the SWAN-SF dataset:

1. It only contains a single data product (derived magnetogram features)
2. It only covers a portion of Solar Cycle 24 (May 2010 - August 2018)

Goal: Integrate Solar Dynamics Observatory Atmospheric Imaging Assembly (SDO/AIA) data into SWAN-SF and expand the dataset with the most recent observations. Ultimately, we would like to develop a pipeline for daily dataset updates.

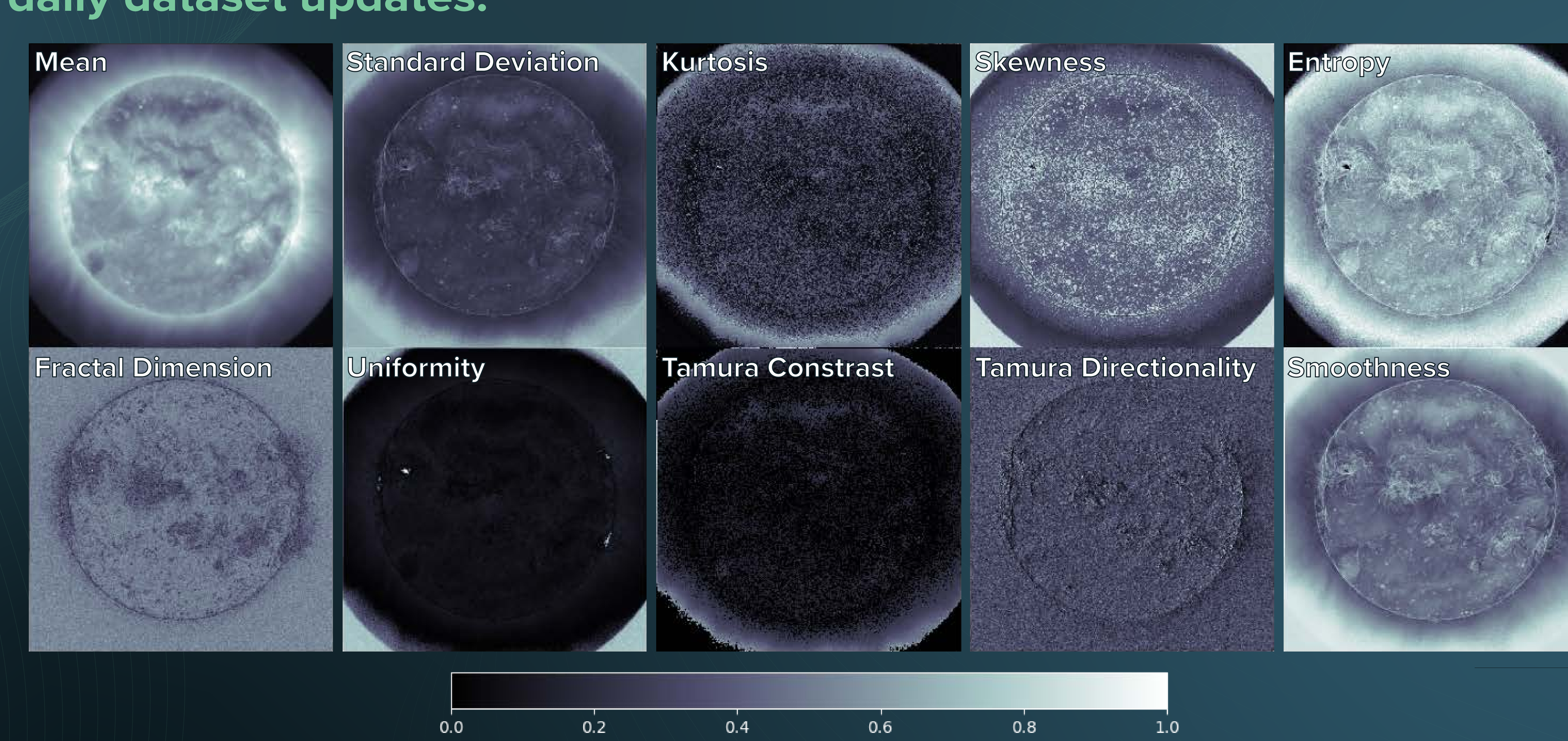

Figure 4: The data products we plan to include are derived texture parameters of AR cutouts taken at the nine available SDO/AIA wavelengths. This figure presents heatmaps of the calculated texture parameters for a full-disk solar image at 193Å, with a 16 x 16 pixel binning applied. In the final data product, we will apply the calculation across the full AR cutout, yielding a single value per parameter for the entire cutout.

The projection effects work was supported by NASA FINESST (80NSSC23K1639), NASA LWS (80NSSC22K0272/80NSSC25K7235), NSF FDSS (1936361), and NASA COFFIES DSC (NNH18ZDA001N-DRIVE). The SWAN-SF extension is funded by NASA's Multidomain Reusable AI Tools program (80NSSC23K1026) and Heliophysics AI/ML-ready Data program (80NSSC24K0238).

ApJ Paper SWAN-SF LinkedIn