COMPARING SOLAR FLARE IRRADIANCE IN GOES X-RAY AND SDO/AIA EUV DATA VIA MACHINE LEARNING REGRESSION

Kiera van der Sande[1], Natasha Flyer[2], Varad Deshmukh[3], Thomas Berger[4]

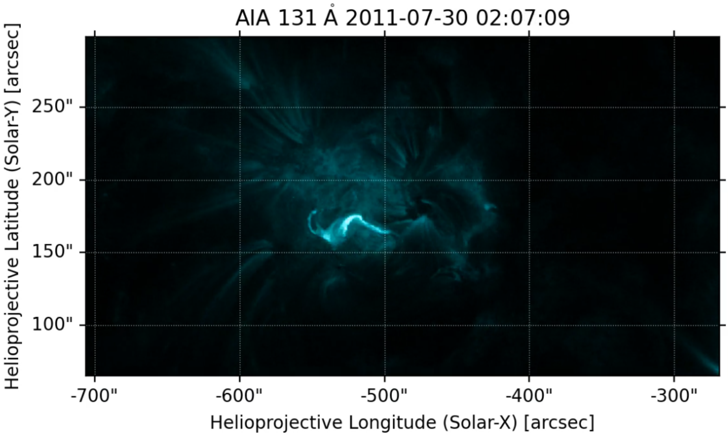
March 21, 2022

[1]Department of Applied Mathematics, University of Colorado Boulder
[2]Flyer Research LLC, Boulder CO
[3]Department of Computer Science, University of Colorado Boulder
[4]Space Weather Technology, Research, and Education Center, University of Colorado Boulder

Funded by NASA Space Weather OR Grant Number 80NSSC20K1404

Current ML flare forecasting models primarily use photospheric magnetic field data, i.e., SDO/HMI data.

Solar flare labels are derived from the NOAA GOES instrument with manual identification of location using other data sources.



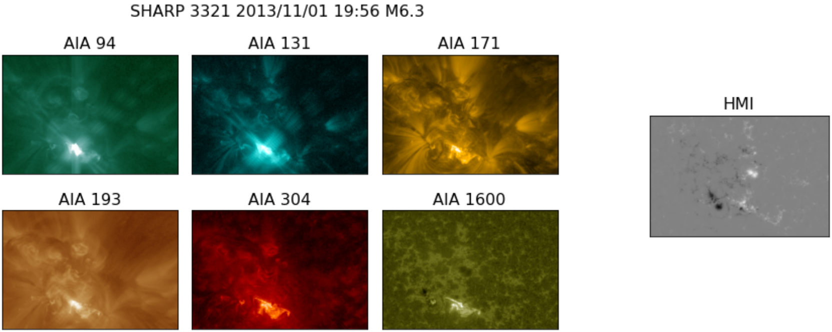
[1]

[1]From the Solar Dynamics Observatory (SDO) Atmospheric Imaging Assembly (AIA)

Use cutouts of SDO/AIA images to match SHARPs (Space weather HMI active region patches).

AIA subsampled at 60s cadence, while HMI is only available at 12 minute cadence.

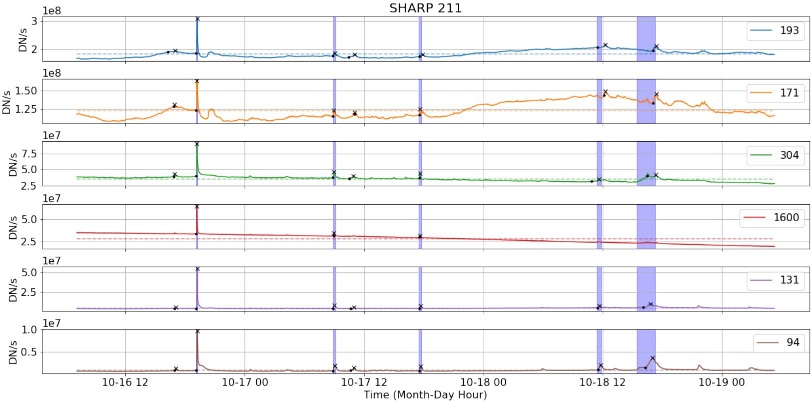
Preprocess to normalize by exposure time and correct for instrument degradation.



Can look at summed-pixel intensity for each AIA channel as a timeseries [2,3]

Detect peaks in AIA channels [4], corroborate between wavelengths to verify.

Correlate with the GOES flare catalog, GOES flares below shown as purple bands.



---

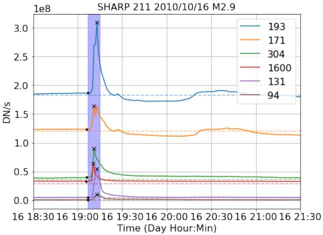
[2] Aschwanden, M. J., Zhang, J., & Liu, K. 2013. The Astrophysical Journal, 775, 23.

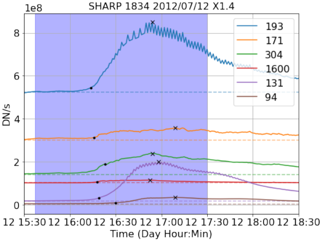
[3] Mahajan, K. K., Lodhi, N. K., Upadhayaya, A. K. 2010. Journal of Geophysical Research: Space Physics, 115, A12.

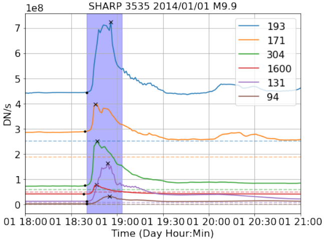
[4] Kraaikamp, E., & Verbeeck, C. 2015, Journal of Space Weather and Space Climate, 5, A18.

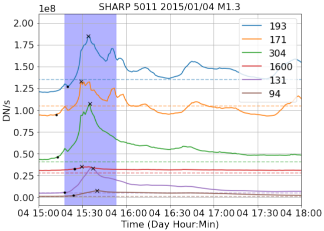
GOES flare duration shown here as purple bands.

Dots represent start times defined by derivative conditions and x's represent peaks.



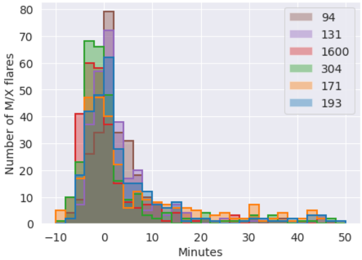




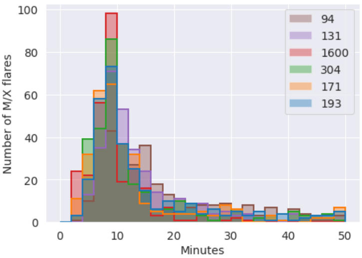


Some previous work has looked at timing of flares in EUV/UV [5], [6].

Here we compare between wavelengths as well as GOES for 325 M/X flares.



(a) Time from GOES start to AIA start

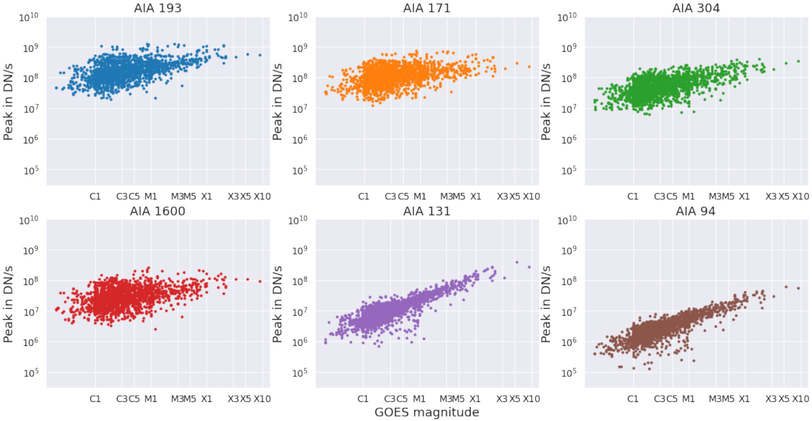


(b) Time from AIA start to AIA peak

[5] Verbeeck, C., Kraaikamp, E., Ryan, D. F., & Podladchikova, O. 2019. The Astrophysical Journal, 884, 50.
[6] Mahajan, K. K., Lodhi, N. K., & Upadhayaya, A. K. 2010. Journal of Geophysical Research: Space Physics, 115, A12.

CORRELATION OF AIA PEAKS WITH GOES FLARE INTENSITY

AIA peak flux is weakly correlated with GOES flare magnitude, verifying previous results.

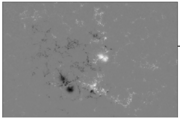


We use an extremely randomized trees (ERT) model for regression with mean absolute error (MAE) as the splitting criterion.

26 AIA SHARPs-based Features
λ = 94, 131, 171, 304, 193, 1600

λ_magnitude	Peak intensity of AIA channel λ	DN/s
λ_rel_magnitude	Difference between peak and start intensity of AIA channel λ	DN/s
λ_est_size	Maximum number of saturated pixels for AIA channel λ	pixels
λ_duration	Time from start to end of flare for AIA channel λ	seconds
Nx	Width of the AIA SHARP cutout	pixels
Ny	Height of the AIA SHARP cutout	pixels

Closest SHARP HMI image



20 HMI SHARPs-based Features

ERT Model

Flare intensity

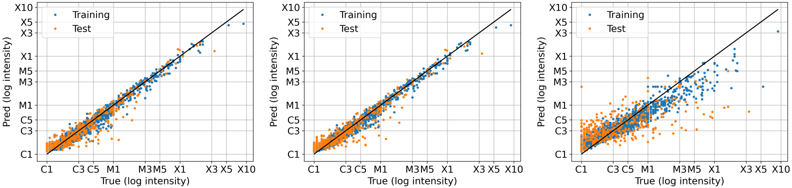
Trained the ERT on 80% of the data, tested on 20% for 10 different random splits.

5-fold cross validation on each random split for hyperparameter tuning:

- Number of trees: 100
- Minimum impurity decrease: 5e-5

Ran tests of different input feature sets:

- AIA SHARPs
- HMI SHARPs
- AIA and HMI SHARPs



(a) AIA SHARPs only (b) AIA and HMI SHARPs (c) HMI SHARPs only

Metric	AIA SHARPS	AIA and HMI SHARPS	HMI SHARPS
RMSE	0.135 ± 0.007	0.133 ± 0.007	0.329 ± 0.011
MAE	0.094 ± 0.003	0.092 ± 0.003	0.230 ± 0.007
R2	0.911 ± 0.007	0.913 ± 0.007	0.472 ± 0.036
% L2 Error (M/X only)	3.504 ± 0.620	3.531 ± 0.605	13.478 ± 0.694

Can estimate GOES flare magnitudes with high accuracy from AIA alone.

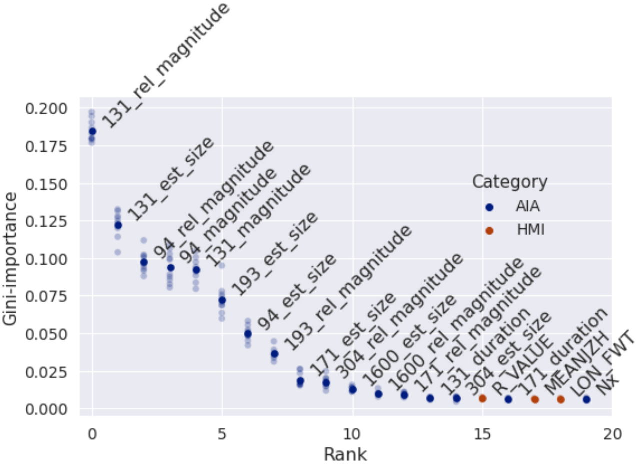
Little difference between input features of AIA SHARPs alone and AIA and HMI SHARPs.

FEATURE RANKING USING GINI IMPORTANCE

Gini importance measures how much each feature reduces the criterion for measuring quality of the decision tree splits (MAE in our case).

AIA 94 and 131 are the most important wavelengths.

HMI SHARPs features are of minimal importance.



We are able to define flare timing and magnitude using multi-wavelength AIA data

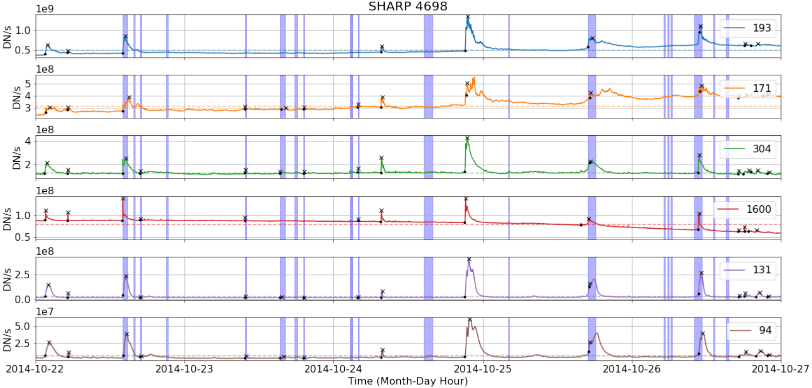
- Define start, peak and end times for each AIA channel
- Estimate flare magnitudes using an ERT regression model on AIA features alone
- Achieve 3.5% error on regression to GOES magnitude for large M/X flares

Moving on to flare prediction

- Using a new AIA-based flare catalog for labeling
- Using AIA and HMI data in ML flare prediction models
- Focus on short term flare predictions: 3-12 hours

QUESTIONS

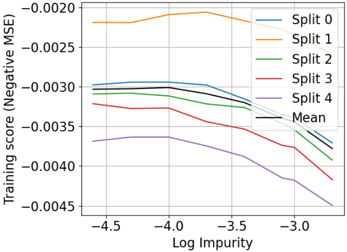
AIA SHARPS TIMESERIES WITH UNLABELLED FLARES



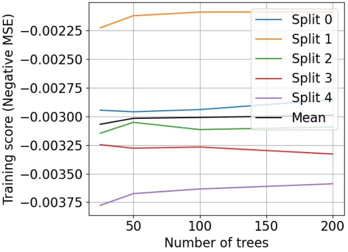
Acronym	Description	Units
λ_magnitude (x6)	Peak intensity of AIA channel $\lambda = 94, 131, 171, 304, 193, 1600$	DN/s
λ_rel_magnitude	Difference between peak and start intensity of AIA channel $\lambda = 94, 131, 171, 304, 193, 1600$	DN/s
λ_est_size	Maximum number of saturated pixels for AIA channel $\lambda = 94, 131, 171, 304, 193, 1600$	pixels
λ_duration	Time from AIA channel $\lambda = 94, 131, 171, 304, 193, 1600$ start to end	seconds
Nx	width of the AIA SHARP cutout	pixels
Ny	height of the AIA SHARP cutout	pixels
LAT_FWT	Latitude of the flux-weighted center of active pixels	degrees
LON_FWT	Longitude of the flux-weighted center of active pixels	degrees
AREA_ACR	Line-of-sight field active pixel area	micro-hemispheres
USFLUX	Total unsigned flux	Mx
MEANGAM	Mean inclination angle, gamma	degrees
MEANGBT	Mean value of the total field gradient	G/Mm
MEANGBZ	Mean value of the vertical field gradient	G/Mm
MEANGBH	Mean value of the horizontal field gradient	G/Mm
MEANJZD	Mean vertical current density	mA/m^2
TOTUSJZ	Total unsigned vertical current	A
MEANALP	Total twist parameter, alpha	1/Mm
MEANJZH	Mean current helicity	G^2/m
TOTUSJH	Total unsigned current helicity	G^2/m
ABSNJZH	Absolute value of the net current helicity	G^2/m
SAVNCPP	Sum of the absolute value of the net currents per polarity	A
MEANPOT	Mean photospheric excess magnetic energy density	$ergs/cm^3$
TOTPOT	Total photospheric magnetic energy density	$ergs/cm^3$
MEANSHR	Mean shear angle (measured using B_{total})	degrees
SHRGT45	Percentage of pixels with a mean shear angle greater than 45 degrees	percent
R_VALUE	Sum of flux near polarity inversion line	G
NACR	The number of strong LOS magnetic field pixels in the patch	N/A
SIZE_ACR	Projected area of active pixels on image	micro-hemispheres
SIZE	Projected area of patch on image	micro-hemispheres

[7]

[7] HMI SHARPS parameters: see Bobra, M. G., Sun, X., Hoeksema, J. T., et al. 2014, Sol Phys, 289, 3549.



(a) Error vs. minimum impurity decrease
(number of trees set to 100)



(b) Error vs. number of trees
(minimum impurity decrease set to 1e — 4)