

## Summary

- SEP event classification based on the GOES >10 MeV channel :-
  - Large:** peak proton fluxes > 10 pfu
  - Weak:** peak proton fluxes > 0.1 and < 10 pfu.
- We focused on understanding the statistical relationships of parent solar eruptions, namely, coronal mass ejections (CMEs) and solar flares with SEP events.
- We analyze and discuss spatio-temporal properties such as flare magnitudes, locations, rise times, and speed and width of CMEs.
- Feature importance study shows both the CME speed and flare intensity are top contributors to the corresponding SEP event peak fluxes.
- We implemented supervised machine learning (ML) models to classify SEP events as large and weak.

## The GSEP Dataset

- From 1986 to 2017, there are 433 SEP events identified in the GSEP list. In that, 244 events are **large**.
- The headers in the list describe physical descriptors and carry relevant indicators for each event.
- We have put effort into identifying for each SEP event in our data set with at least one eruption, either a flare or CME.
- We also include a "Flag" column to indicate "1" or "0" representing "large" or "weak" events, respectively.



## Statistical Properties

- 383 (220-large and 163-weak) events have GOES class information available in our dataset. Of them, 121 (32%) are associated with X-class flares, while 182 (48%) and 80 (20%) are associated with M and C- classes, respectively.
- Eastern hemisphere events typically have gradually rising proton profiles, whereas western events reach peak fluxes within a few hours of the parent flare eruption.

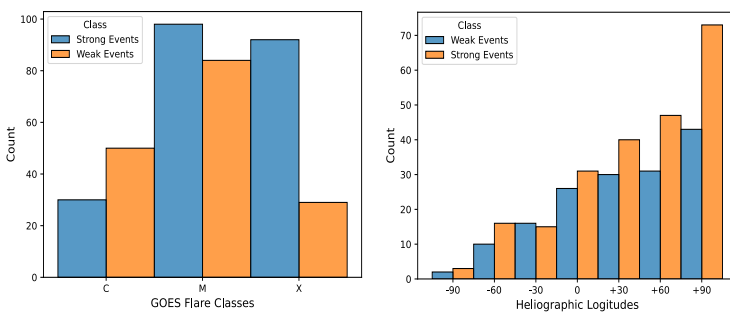


Figure 1: Distribution of (left) solar flare classes based on NOAA scale, and (right) longitudinal variation of solar flare locations, that are associated with SEP events.

- We find most (66%) of the CMEs associated with SEP events are halo in angular extents and fast with a median speed of 1200 km.s<sup>-1</sup>

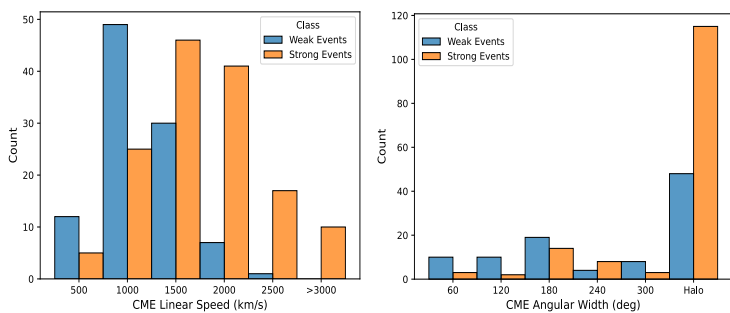


Figure 2: Distribution of (left) linear speed and (right) angular width of CMEs that are associated with SEP events.

## Feature importance

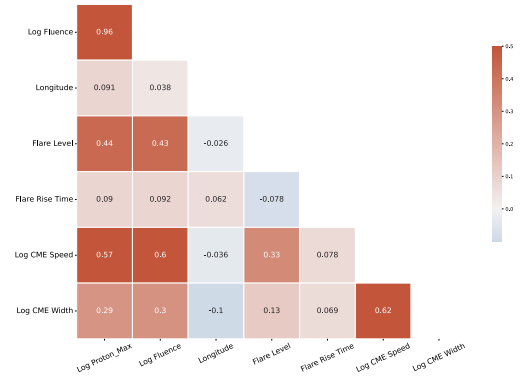


Figure 3: Correlations coefficients of source solar eruptions for peak proton fluxes and fluences of SEP events. CME Speed shows a good correlation with SEP events, while flare intensities are moderately correlated.

## ML Binary Classification

Table 1. Performance of supervised classifiers on the data set.

Model	F1_score	TSS	MCC	HSS
XGBoost	0.85	0.41	0.43	0.43
RF	0.80	0.40	0.38	0.38
LR	0.77	0.58	0.52	0.46
SVM	0.71	0.58	0.52	0.46

**Model names:** Random forest (RF), Extreme gradient boosting (XGBoost), Logistic regression (LR) and Support vector machines (SVM)

**Metrics:** Matthews correlation coefficient (MCC), True skill statistics (TSS), Heidke skill score (HSS)

## Acknowledgements

We acknowledge the use of data from: (i) NOAA-GOES missions, and (ii) NASA-CDAW Halo CMEs list. This work was supported by NASAs FINESST Grant 80NSSC21K1388 and SWR202R Grant 80NSSC22K0272.