Validation of Machine Learning Models for Classification of Solar Wind

Lois J. Landwer[1,2], Hazel Bain[1,2], Mark Miesch[1,2], George Millward[1,2], Enrico Camporeale[1,2], Eric Adamson[2]

[1]Cooperative Institute for Research in Environmental Sciences (CIRES)

[2]NOAA Space Weather Prediction Center

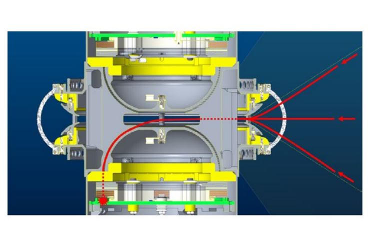
Abstract

Space Weather Prediction Center(SWPC) in NOAA plans to operationalize machine learning(ML) model for use with the Space Weather Follow On Lagrange 1 (SWFO-L1) data to enhance forecaster situational awareness. SWFO-L1 mission is a deep-space mission operating in a Lissajous orbit at the Sun-Earth Lagrange 1 (L1) point, enabling upstream measurements of solar wind disturbances before they reach Earth. It will provide continuous measurements of the sun's corona and of the solar wind at the L1 point and transmit continuous real-time data to Earth. It is scheduled for launch in 2025. We validated a Gaussian process ML model developed by Camporeale et al. (2017) using DSCOVR and ACE data. SWFO-L1 will replace ACE's and DSCOVR's monitoring of solar wind, energetic particles, and the interplanetary magnetic field. This ML model is a four-category classification algorithm for the solar wind, previously adopted in Xu and Borovsky (2015): ejecta, coronal hole origin plasma, streamer belt origin plasma, and sector reversal origin plasma. The algorithm is trained and tested on a labeled portion of the OMNI data set identifying the wind regime based on several parameters, including in-situ observations of wind speed, proton temperature standard deviation, temperature ratio, proton specific entropy, and Alfven speed, as well as non-in-situ data on sunspot number and solar radio flux.

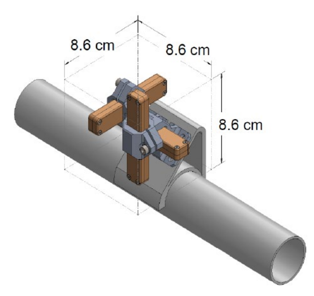
SWFO-L1 (Space Weather Follow On Lagrange 1)

- The importance of understanding the near-Earth space environment cannot be overstated, as it impacts a wide range of users. The absence of accurate space weather predictions and forecasts results in reduced efficiency, increased costs, and heightened risk for many industries that depend on satellite and related services.
- Since several monitoring satellites, such as SOHO, ACE, and DSCOVR, have a limited remaining lifespan, it is important for NOAA to plan for follow-on missions for solar, heliospheric, and other observations to continue providing space weather information to its users. The SWFO-L1 observatory is scheduled to launch in 2025.

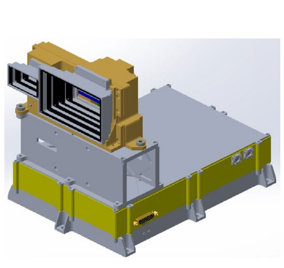
SWFO-L1 : Heliophysics Instruments



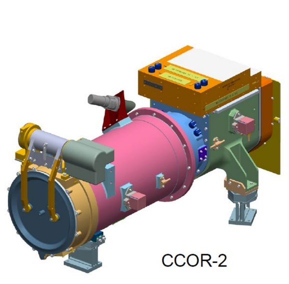
Solar Wind Plasma Sensor (SWiPS): Built by Southwest Research Institute (SwRI), it will measure properties of the solar wind flowing past SWFO-L1, such as density, velocity, and temperature.



Solar Wind Plasma Sensor (SWiPS): Built by Southwest Research Institute (SwRI), it will measure properties of the solar wind flowing past SWFO-L1, such as density, velocity, and temperature.

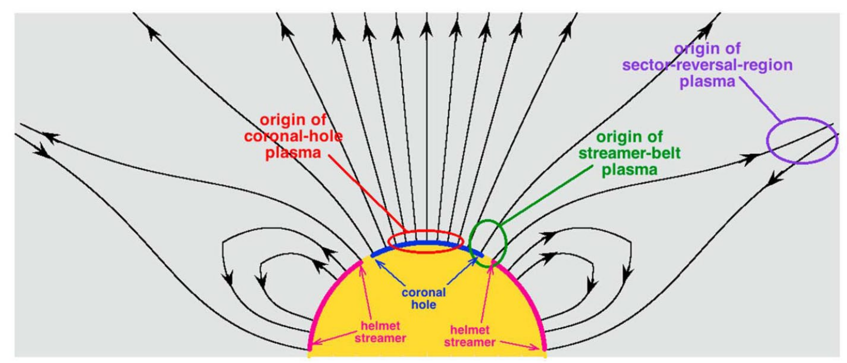


Solar Wind Plasma Sensor (SWiPS): Built by Southwest Research Institute (SwRI), it will measure properties of the solar wind flowing past SWFO-L1, such as density, velocity, and temperature.



Compact Coronagraphs (CCORs): Developed by the Naval Research Lab (NRL), the telescope will be used to observe the solar corona and detect coronal mass ejections (CMEs),CIRs and other structures. CCOR-1 will fly on the GOES-U satellite (2024) and a nearly identical CCOR-2 on SWFO-L1 (2025).

Classification of Solar Wind with Machine Learning



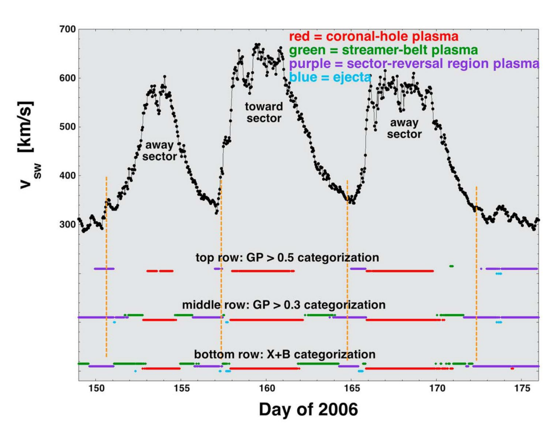
Xu & Borovsky (2015) described a categorization scheme for that divided the solar wind into four components

- ❑ Ejecta
- ❑ Streamer-belt-origin plasma
- ❑ Sector-reversal-region plasma
- ❑ Coronal-hole-origin plasma

List of Attributes

Attribute	Symbol
Solar wind speed	V_{sw}
Proton temperature standard deviation	σ_T
Sunspot number	R
Solar radio flux (10.7 cm)	$F_{10.7}$
Alfven speed	v_A
Proton specific entropy	S_p
Temperature ratio	T_{exp}/T_p

- In 2017, Camporeale et al. developed a machine learning model that utilized Gaussian processes to identify wind regimes based on multiple parameters.
- This ML model is a four-category classification algorithm for the solar wind, previously adopted in Xu and Borovsky (2015).



Median Percentage of Number of Events Correctly Categorized Over 100 Tests, for Different Ratios of Training to Test Data Size

	Ratio of training to test data			
Category	10%	15%	20%	25%
Ejecta	90.8	93.1	94.9	96.1
Coronal hole origin	99.2	99.5	99.7	99.7
Sector reversal origin	92.2	95.1	96.9	97.5
Streamer belt origin	93.2	96.8	97.8	98.7

Training and Testing Data Using OMNI Data from 1995 to 2017

Confusion Matrix for the Case of 20% Training Set, When Only Probabilities Larger Than 50% Are Considered

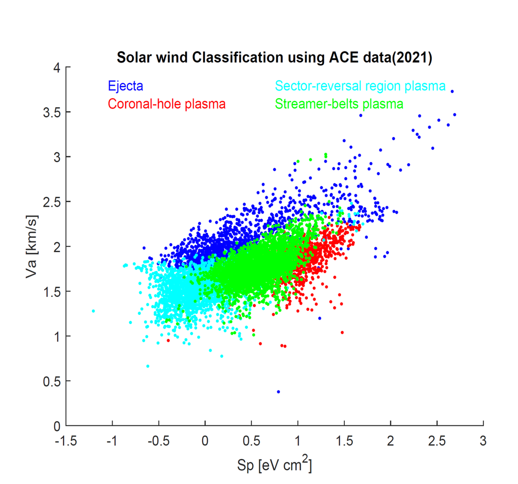
	Observed category			
Prediction	Ejecta	Coronal hole	Sector reversal	Streamer belt
Ejecta	97.9	0	0.8	0.6
Coronal hole origin	0.2	100	0.1	0.1
Sector reversal origin	1.0	0.0	98.5	0.3
Streamer belt origin	1.0	0.0	0.5	99.0

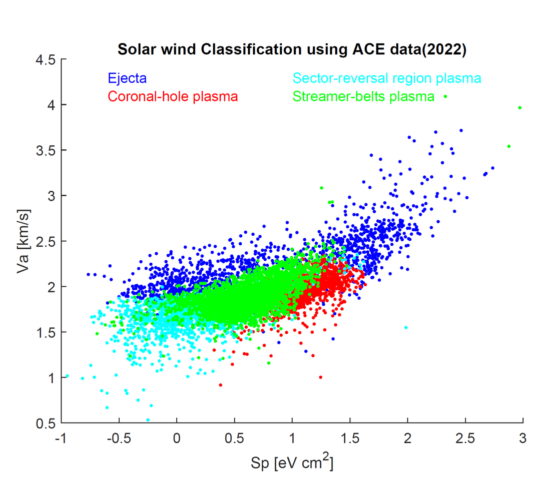
Note. Probabilities are conditioned on the observed category.

	GP>0.3	GP>0.5
Percentage of the classification s is the same	90.77%	100%
Mean absolute value difference s	0.0486	0.0527
Number of data in 2017	OMNI = 5743 , DSCOVR=5104 OMNI=DSCOVR=5038	OMNI=1762, DSCOVR=1415 OMNI=DSCOVR=1226

Comparison between OMNI and DSCOVR classification for 2017

Solar wind classification using ACE data

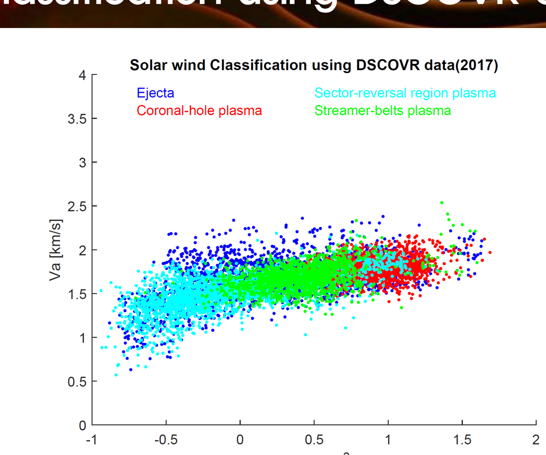




$$\text{Alfven speed } (V_A) = \frac{21.8\times\beta}{\sqrt{n_p}}$$ β → Magnetic field strength; n_p → Proton density

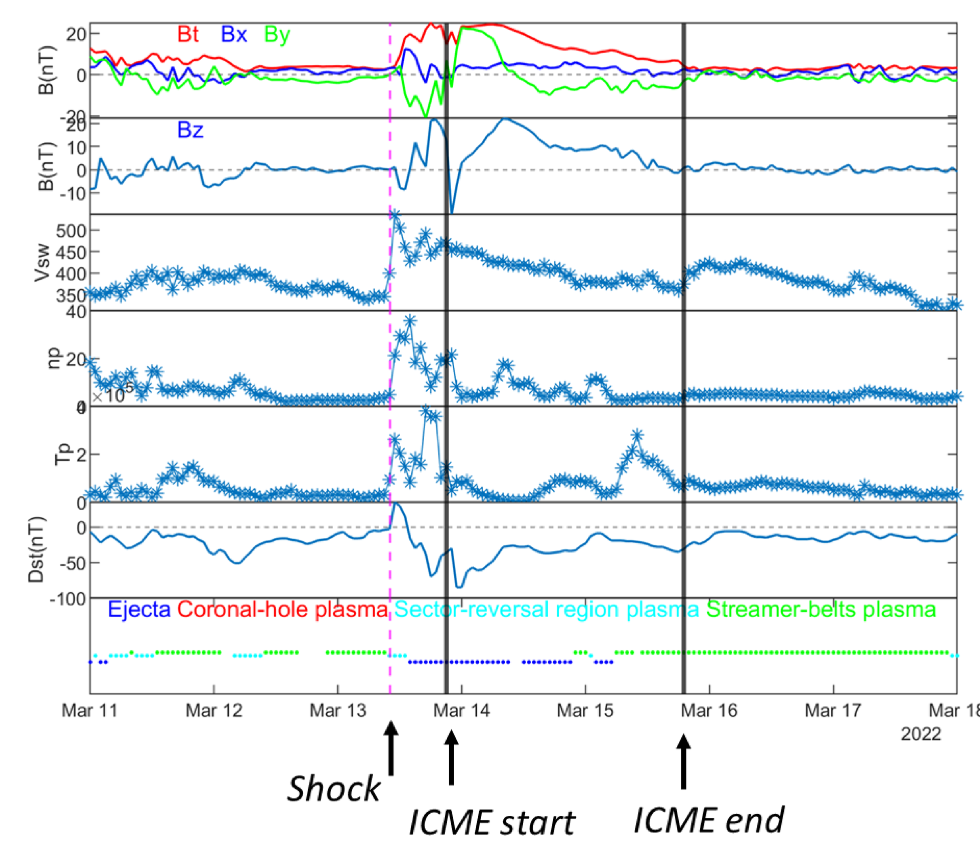
$$\text{Proton specific entropy } (S_p) = \frac{T_p}{n_p^{2/3}}$$ T_p → Proton temperature

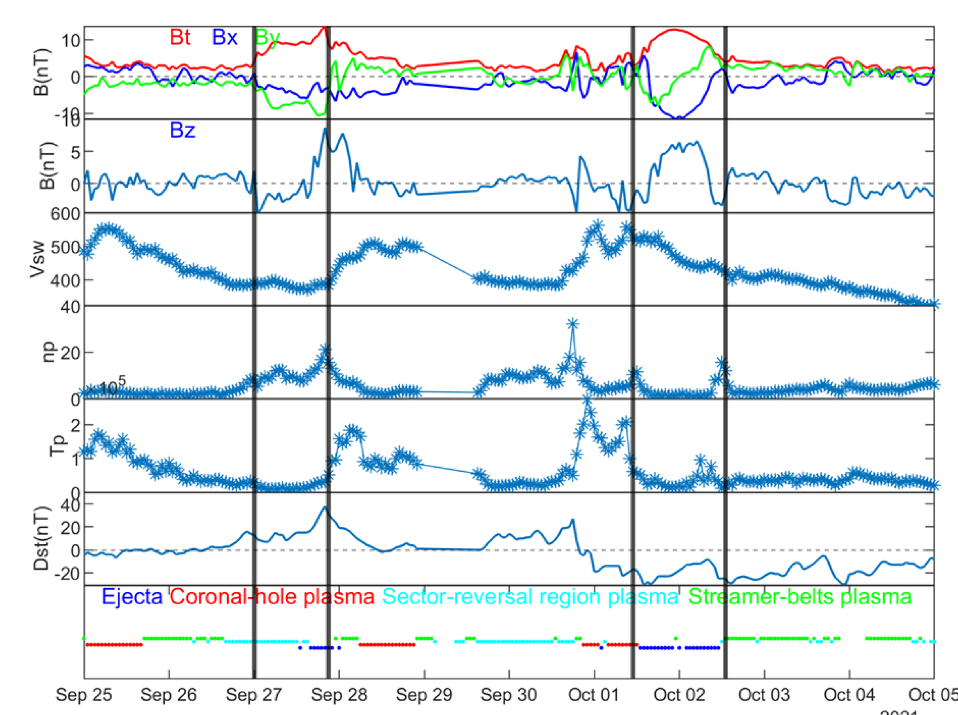
Solar wind classification using DSCOVR data



Solar wind classification using ACE data

TP (True Positive)

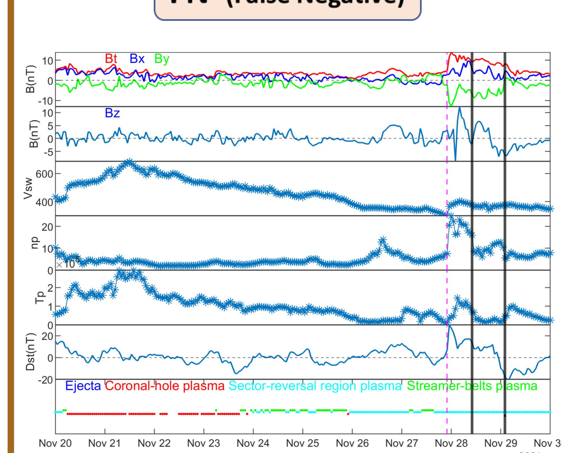




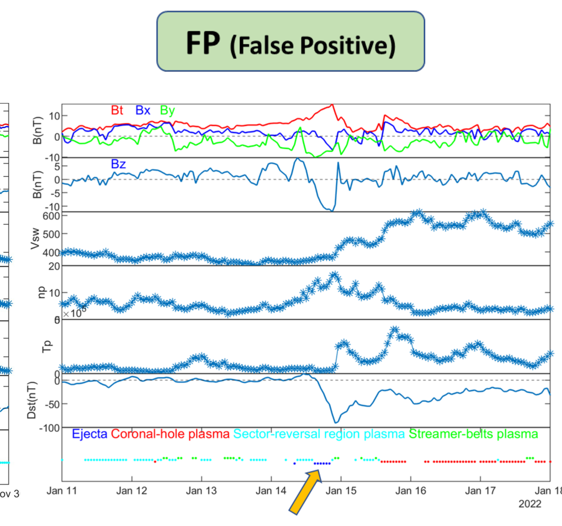
ICME=Richardson and Cane

Shock= CfA Interplanetary Shock Database - Wind

FN (False Negative)



FP (False Positive)



Ejecta Classification Compared to Richardson and Cane ICME List

Ejecta Classification (ACE data 2021~2022) vs Richardson and Cane (26 ICMEs 2021~2022)

		Predicted condition	
	Total population = P + N	Positive	Negative
Actual condition	Positive (P)	18 True positive (TP)	8 False negative (FN)
	Negative (N)	103 False positive (FP)	- True negative (TN)

Refercences

E. Camporeale, A. Care & J. E. Borovsky 2017, "Classification of Solar Wind with Machine Learning", J. Geophys. Res. Space Physics, 122, 10910-10920

F. Xu & J. E. Borovsky 2015, "A New Four-Plasma Categorization Scheme for the Solar Wind", J. Geophys. Res. Space Physics, 120, 70-100