JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Modeling Earth's Plasma Sheet using Machine Learning

Savvas Raptis[1], Connor O' Brien[3], Louis Richard[2], Slava Merkin[1], Kareem Sorathia[1], Simon Wing[1]

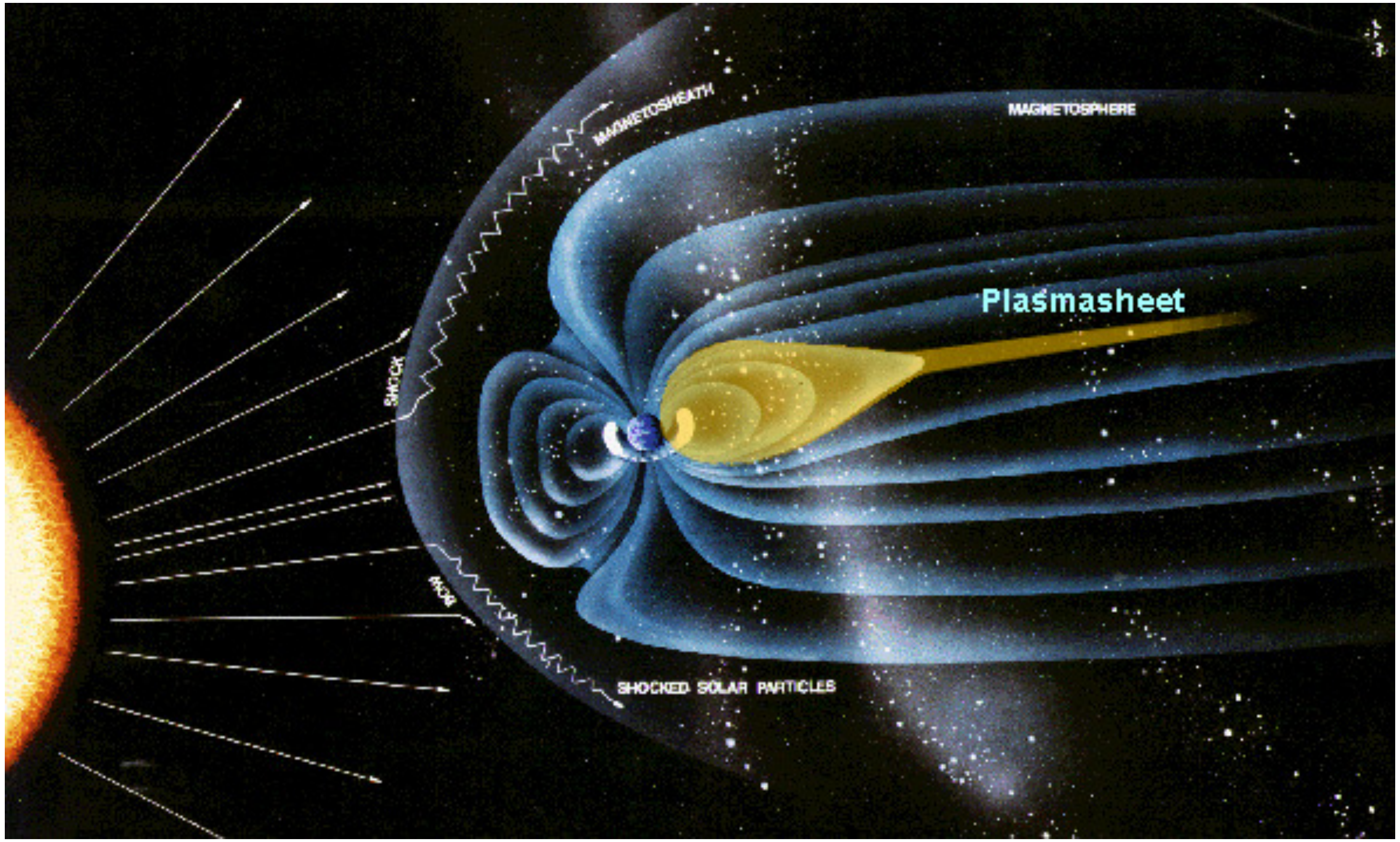
[1] APL/JHU, Laurel, MD, US
[2] Swedish Institute of Space Physics, Uppsala, Sweden
[3] Center for Space Physics, Boston University, Boston, MA, USA

Supported by John Hopkins University Applied Physics Laboratory independent R&D fund

savvas.raptis@jhuapl.edu
https://savvasraptis.github.io

CENTER FOR GEOSPACE STORMS

Earth’s plasma sheet



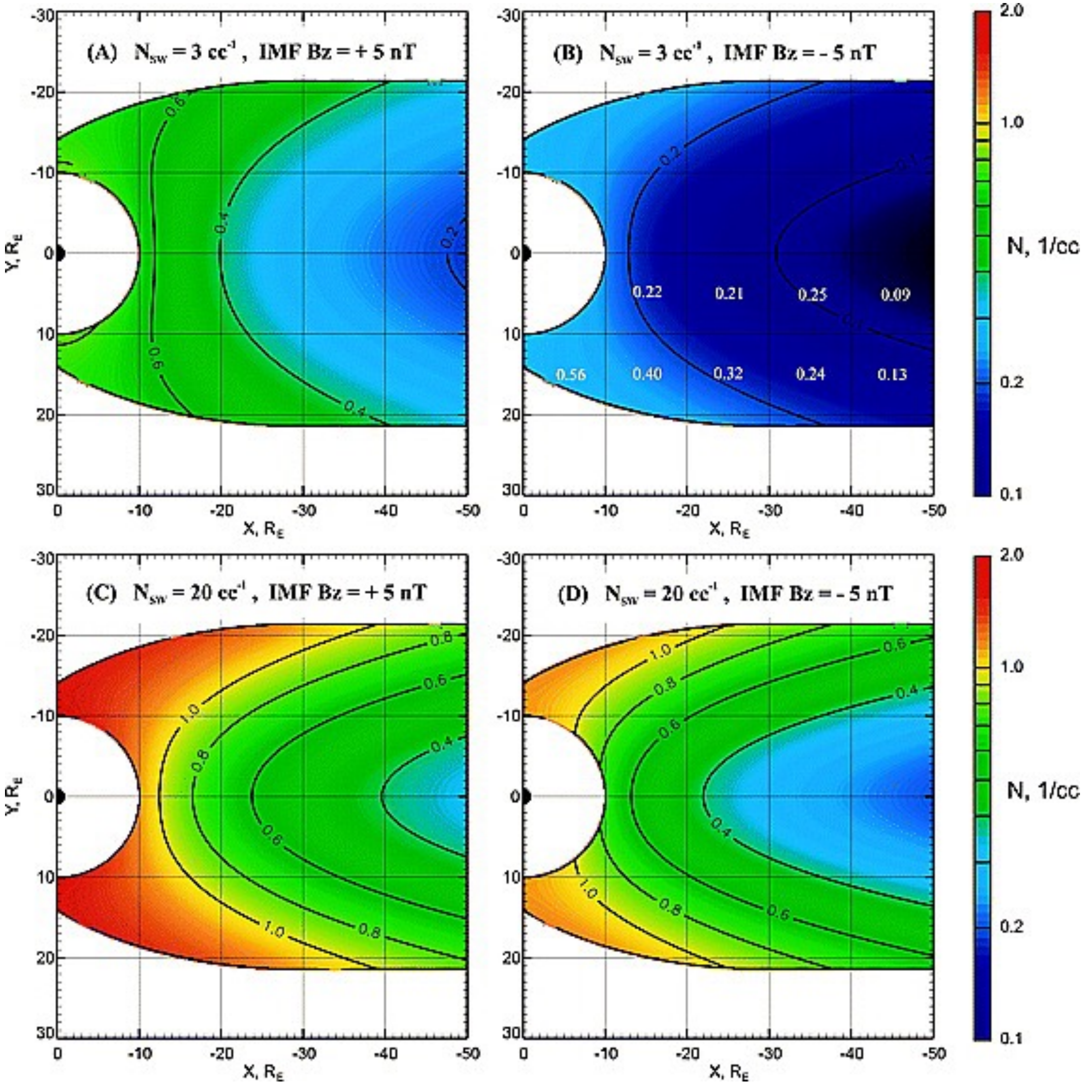
- Magnetotail reconnection
- Bursty Bulk Flows (BBFs)
- Global Convection Patterns
- Ring current

Modeling PS is useful for:

(a) Understanding storm/substorm dynamics
(b) Explain ring current configuration
(c) Facilitate space weather modeling
(d) Understand inner magnetosphere
(e) Source for radiation belts

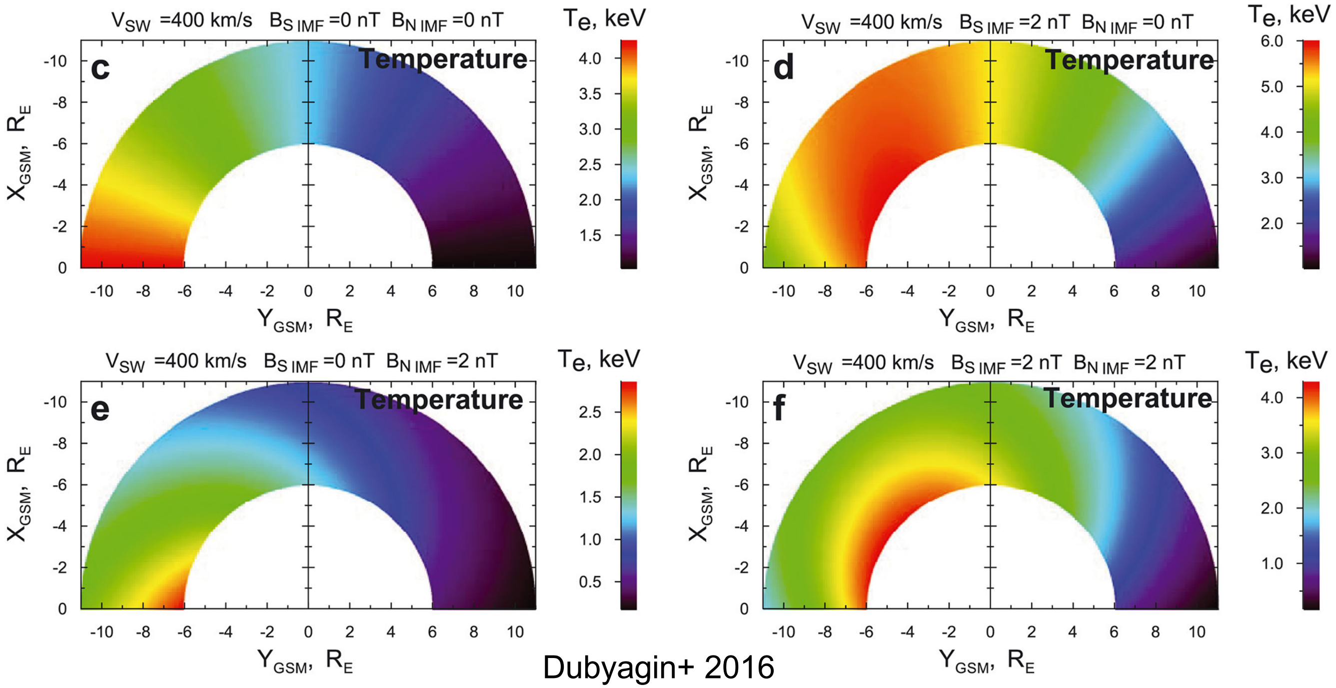
Baseline empirical models

Modelled with Geotail



Tsyganenko & Mukai 2003

Modelled with THEMIS

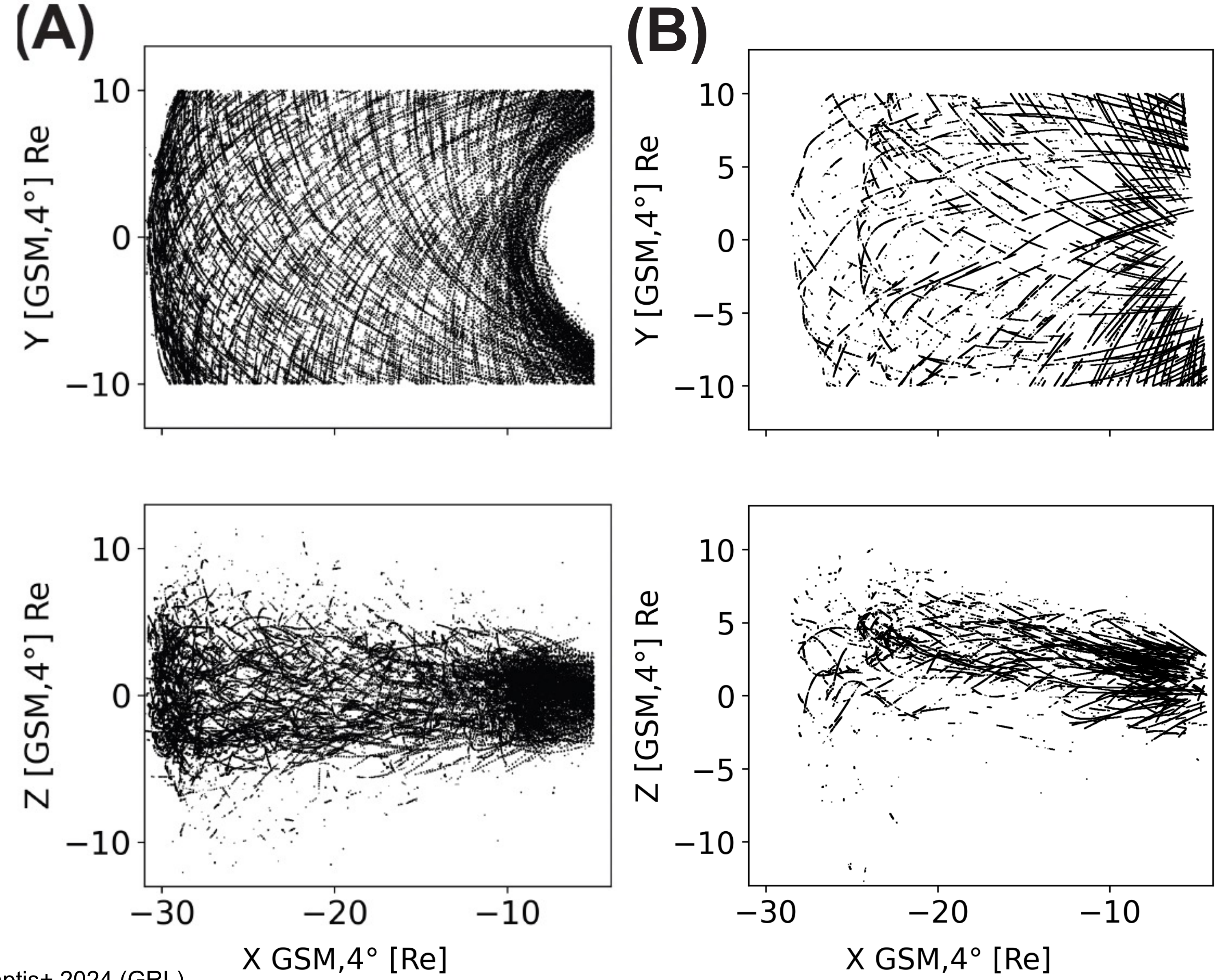


Dubyagin+ 2016

<u>Why then work on this?</u>

1. More data under different conditions
2. MMS was never used with its state of the art instrumentation
3. These models don't include time history
4. ML methods can reveal non-linear relationships easily

The dataset (output – Central Plasma Sheet)



Raptis+ 2024 (GRL)

(A) Geotail (1994 - 2022)
>1 million points (~12s res)

(B) MMS (2015 – 2024)
~ 250k points (~12s res)

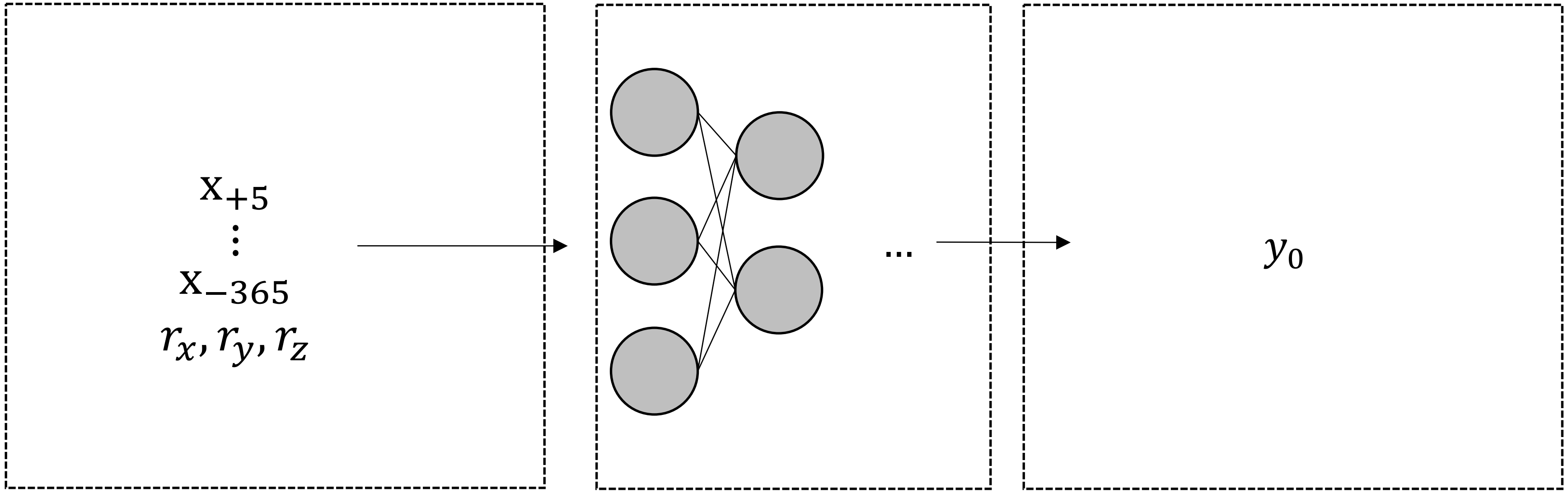
<u>Output</u>:
Anything locally measured
(In this example plasma moments)

Data Scientist POV
(i.e., Input, output & regression)

Input
SW properties +
Geo conditions

Dense Neural Network
Regression

Output
Plasma Sheet
Quantity



Input:
x: Different solar wind features (e.g., n, B, etc.) + geomagnetic indices including time history up to 6h
r: Location of SC measuring output

Output:
y: Different quantities at plasma sheet (e.g., n, B, T etc.)

SW Input tried: Wind (1min res) – OMNIweb (1/5 min res)

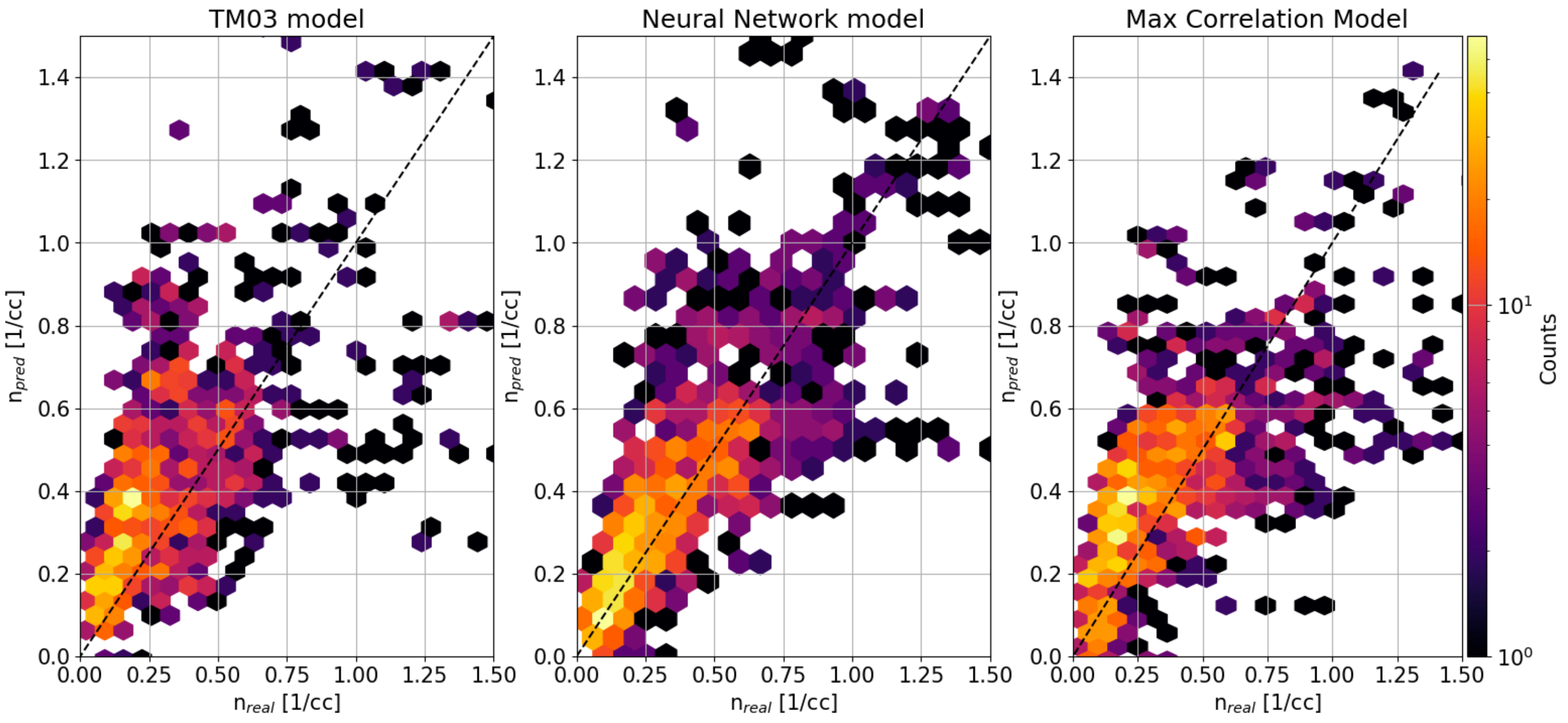
Results here shown for 1min averaged quantities for output

Statistical Results

Modeling Density | Predictions vs Observations

Results from Last GEM
Key Message: NN > Baseline> TM03

Model maximizing correlation for input and output (replace for linear regression)



	TM03	NN	Base
R2	0.17	**0.68**	0.32
MAE	0.19	**0.11**	0.18
RMSE	0.27	**0.2**	0.27
r (cor)	0.58	**0.83**	0.57

Geotail data

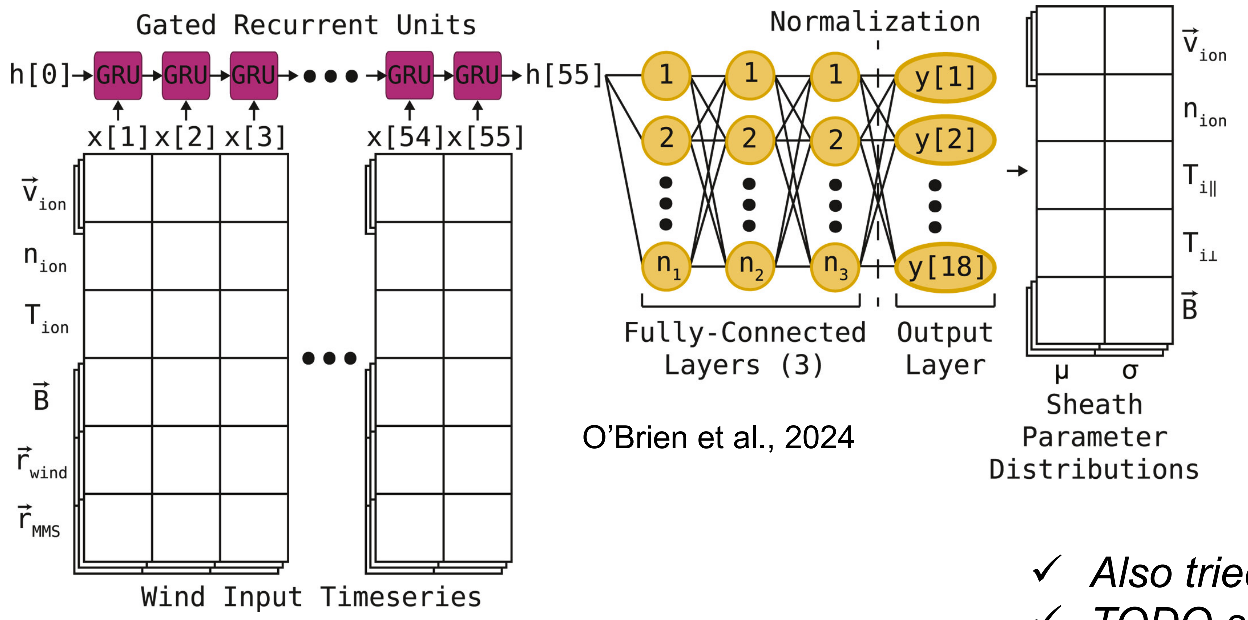
- Presented Testing of NN → Prone to data leakage
- Harder test set (i.e., 5 years of out of sample test data) gives R2 ~0.3-0.4

Raptis, O'Brien et al., 2024 (under prep.)

Disclaimer: not exactly same input of original empirical models

More methodologies & input space

- **PRIME**: GRU architecture, non-propagated Wind values tried up to several hours of history time



O'Brien et al., 2024

Time History	Type of Input	Architectures
1-10h	Wind (L1)	Linear Reg
	OMNIweb	Gradient Boosting
		Neural Network
		RNN/LSTM/GRU (PRIME-PS)

✓ *Also tried different error functions, optimizers, hyperparameters etc.*
✓ *TODO some different imbalanced techniques*

Key Takeaway:
To quantify our method's impact, we tested diverse variations of the problem.

Raptis, O'Brien et al., 2024 (under prep.)

Updated results (Test set, last 20% of data)

Method	Strict CPS MAE	R^2	r	CRPS
LightGBM	0.145	0.242	0.631	–
Neural Net	0.152	0.325	0.603	–
Linear Reg	0.173	0.265	0.620	–
PRIME-PS	**0.113**	**0.453**	**0.707**	0.083
TM03	0.163	0.208	0.570	–

Key Results:

- PRIME-PS demonstrates a performance edge (~30% MAE from TM03 and ~15% from other ML).
- This advantage can get quite low (from cross-validation | not shown).
- Different input, method, time-history, and hyperparameter tuning etc. had overall a statistically marginal effect.
- Why is this the case?

Raptis, O'Brien et al., 2024 (under prep.)

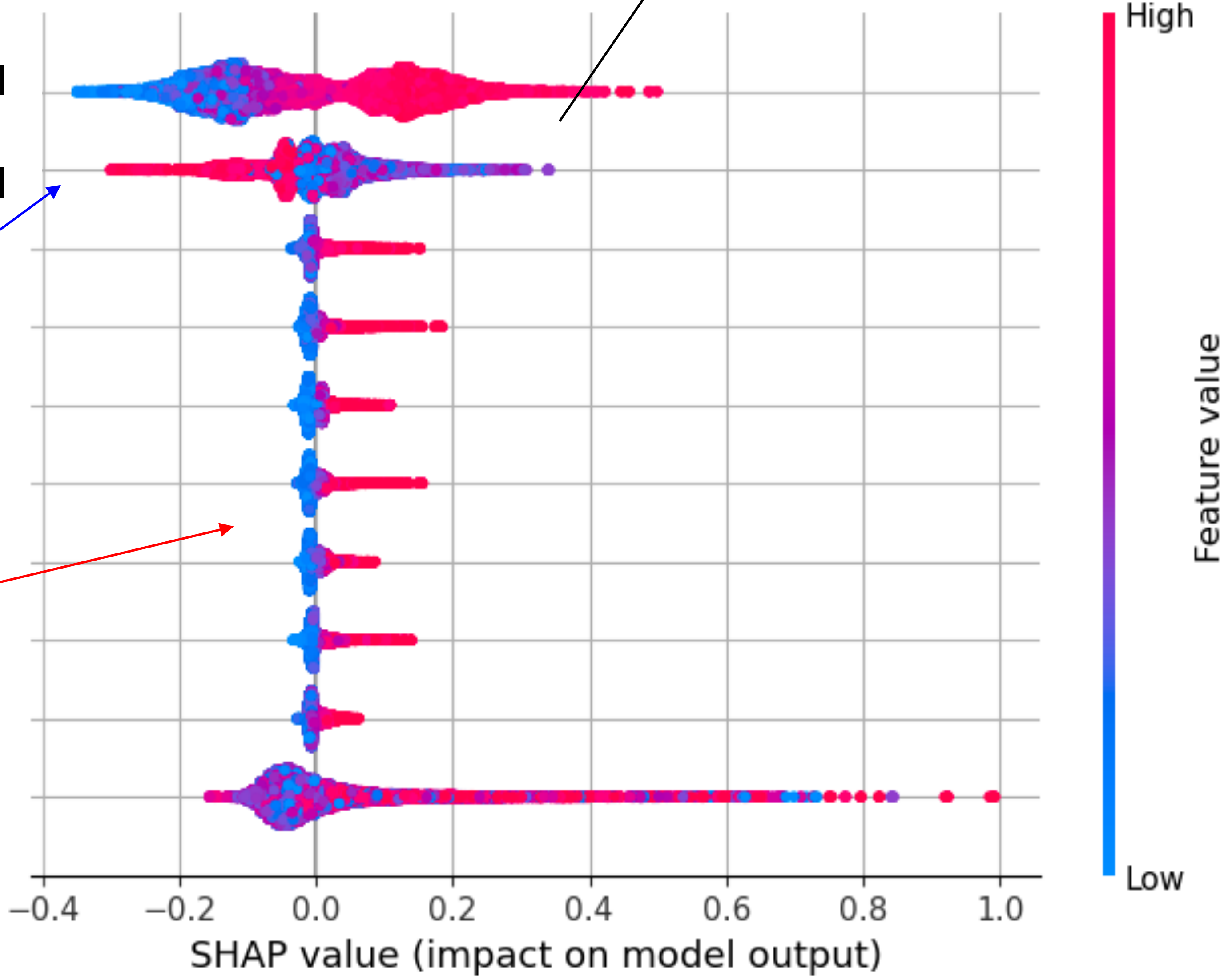
Preliminary feature Importance Analysis

More density close to earth and at dawn

Answer: In most cases (statistically):

Model is predominantly driven by spacecraft location

Solar wind input only marginally affects performance

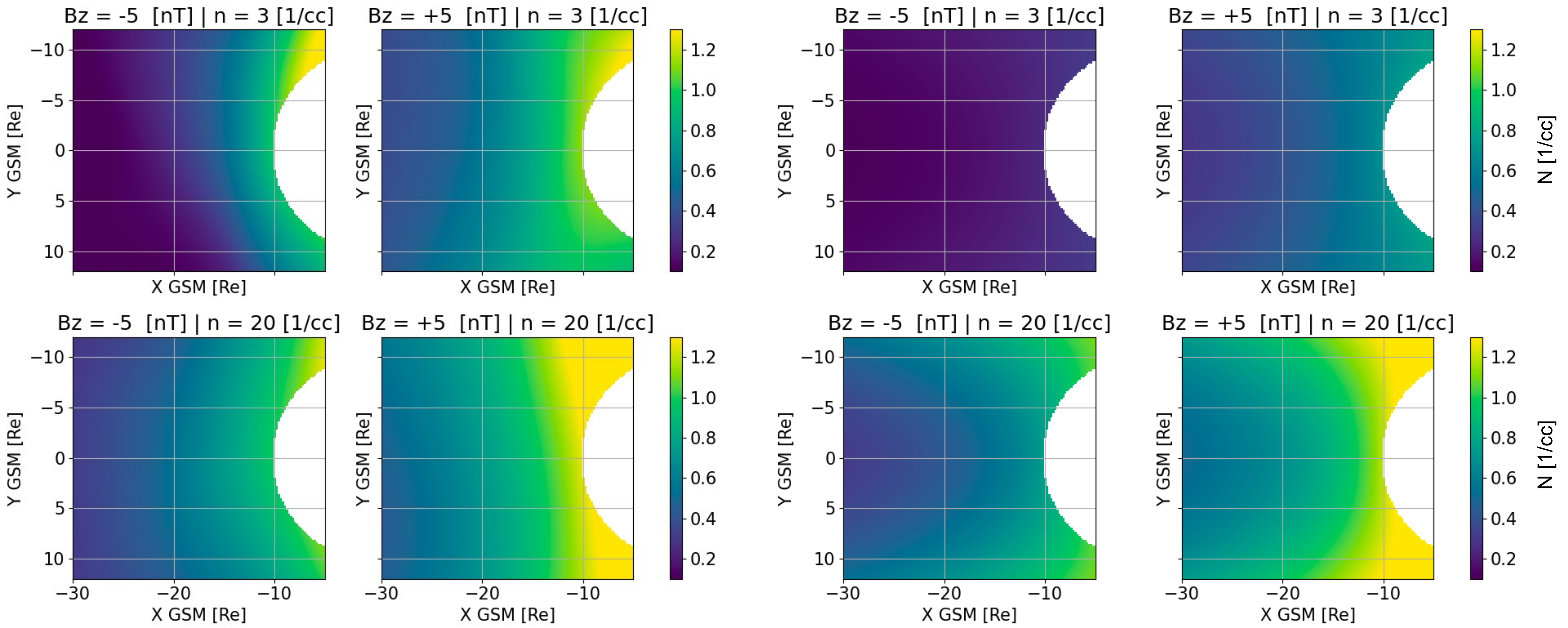


SHAP Values explain why a model made a specific prediction, by showing each feature's impact.

Modeling Efforts

Modeling Density | 2D Maps

Asymmetries introduced



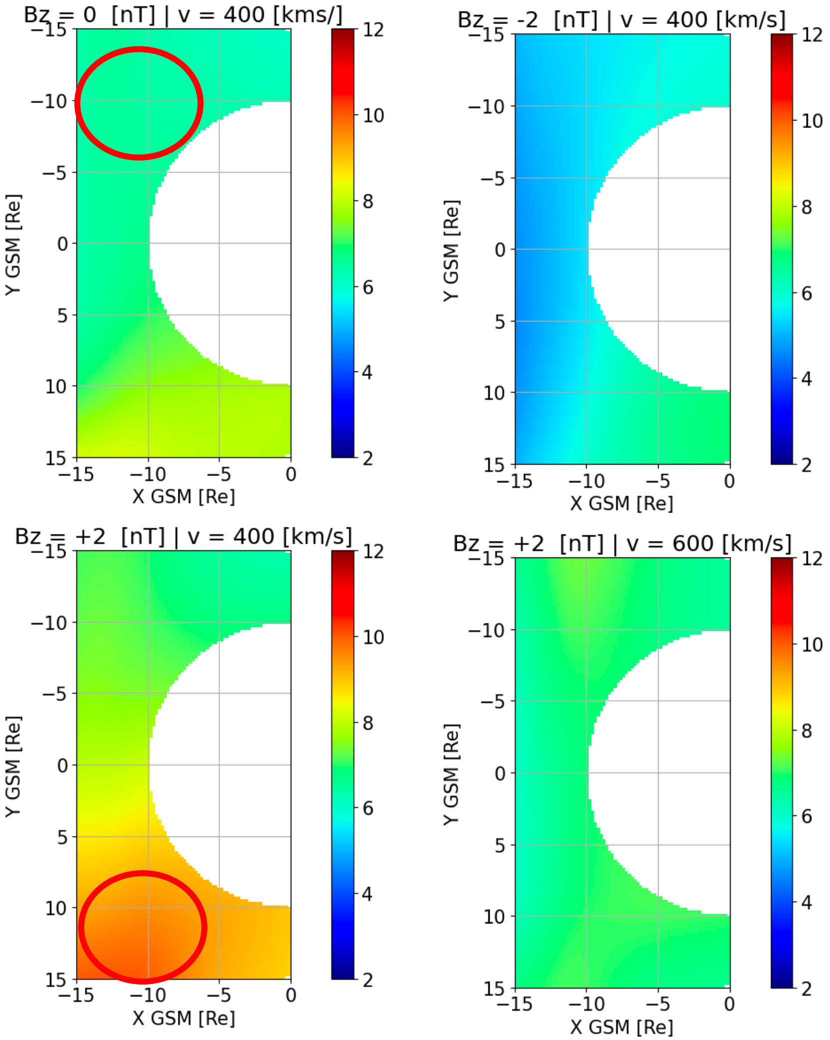
Neural Networks modeling

Empirical modeling (TM03)

Raptis, O'Brien et al., 2024 (under prep.)

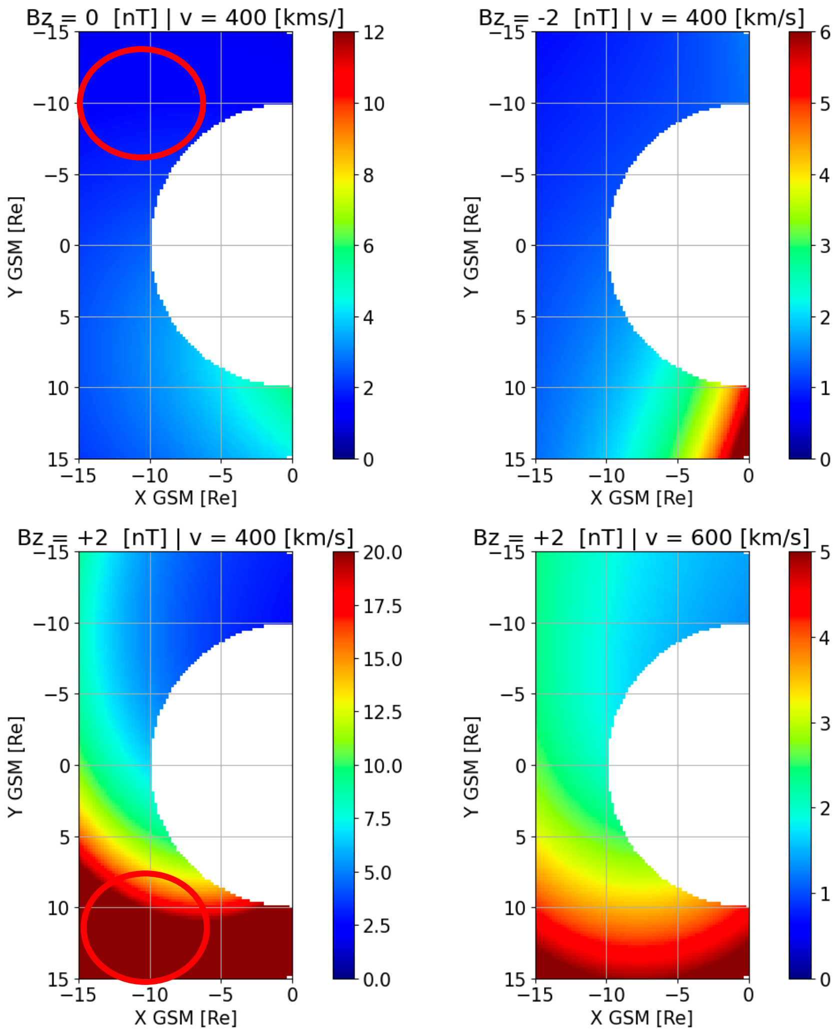
Modeling Temperature Ratios with MMS | 2D Maps

Reproducing: Wang et al., 2009 with dusk Ti/Te much higher than dawn



Neural Networks modeling

+No extreme values
+Asymmetries are shown
+ Coherent physical picture



Empirical modeling (TM03/DSGR16)

Criterion	Strict CPS	Flexible CPS	High density
$\beta > 1$	yes	—	—
$\beta > 0.5$	—	yes	—
$\sqrt{B_x^2 + B_y^2} < 2\lvert B_z \rvert$	yes	—	—
$N < 6$	yes	—	—
$N > 6$	—	—	yes
EA1SW0 = EA	yes	yes	yes
$-31 < R_x < -5$	yes	yes	yes
$\lvert R_y \rvert < 15$	yes	yes	yes
$\lvert R_z \rvert < 10$	yes	yes	—
$V_x > -20$	—	—	yes

Table 1. Plasma sheet classification thresholds for the strict CPS, flexible CPS, and high-density subsets. *beta* is the ion plasma beta parameter, density (N) is in 1/cc units, V_x is in km/s, and all the locations ($R_{x,y,z}$) are in Earth radius. The coordinate system for all vectors is the aberrated Geocentric Solar Magnetospheric (GSM) coordinates

Storm Time Behavior and Importance of Outliers

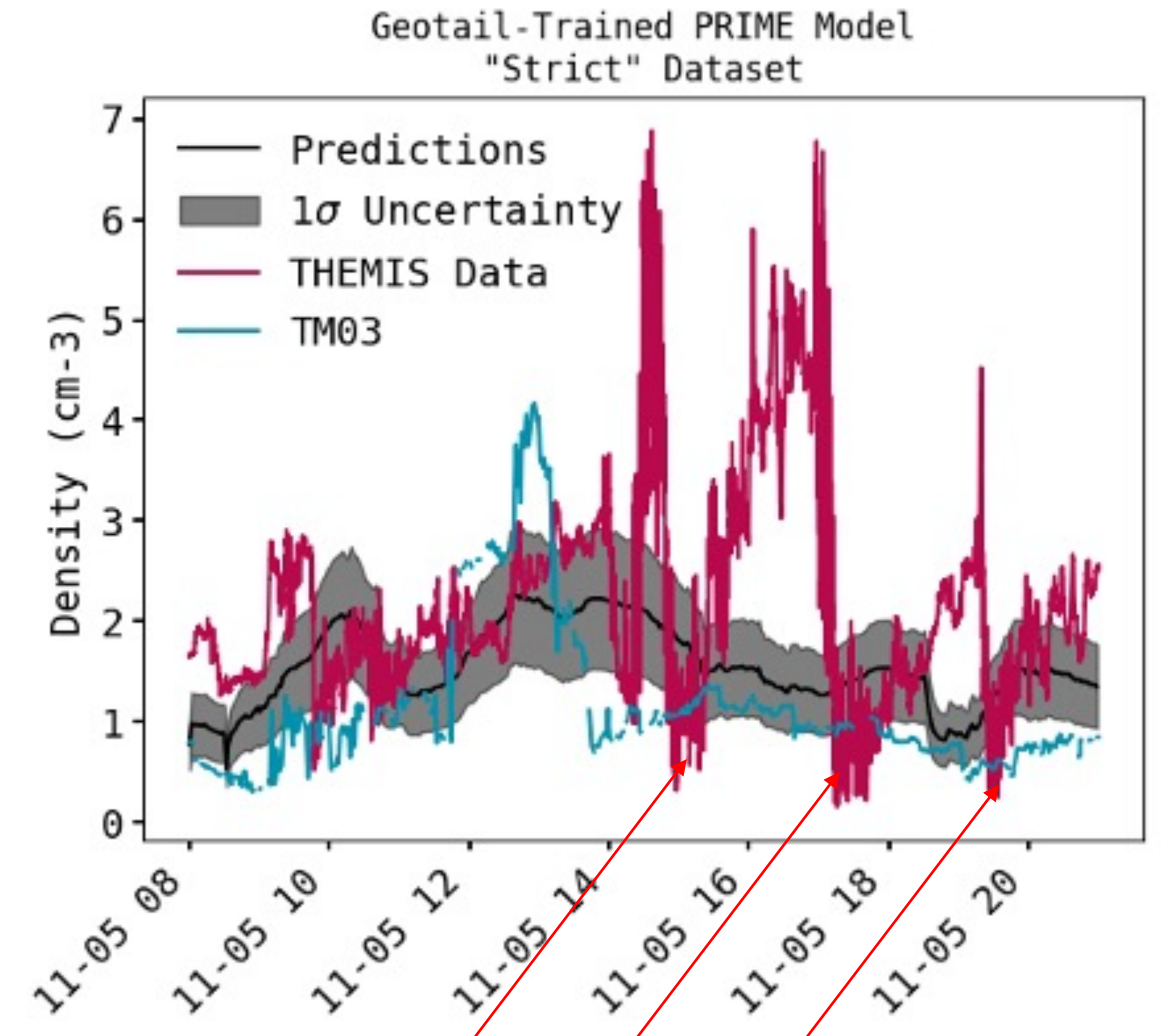
The Problem: We use static thresholds for dynamic environents.

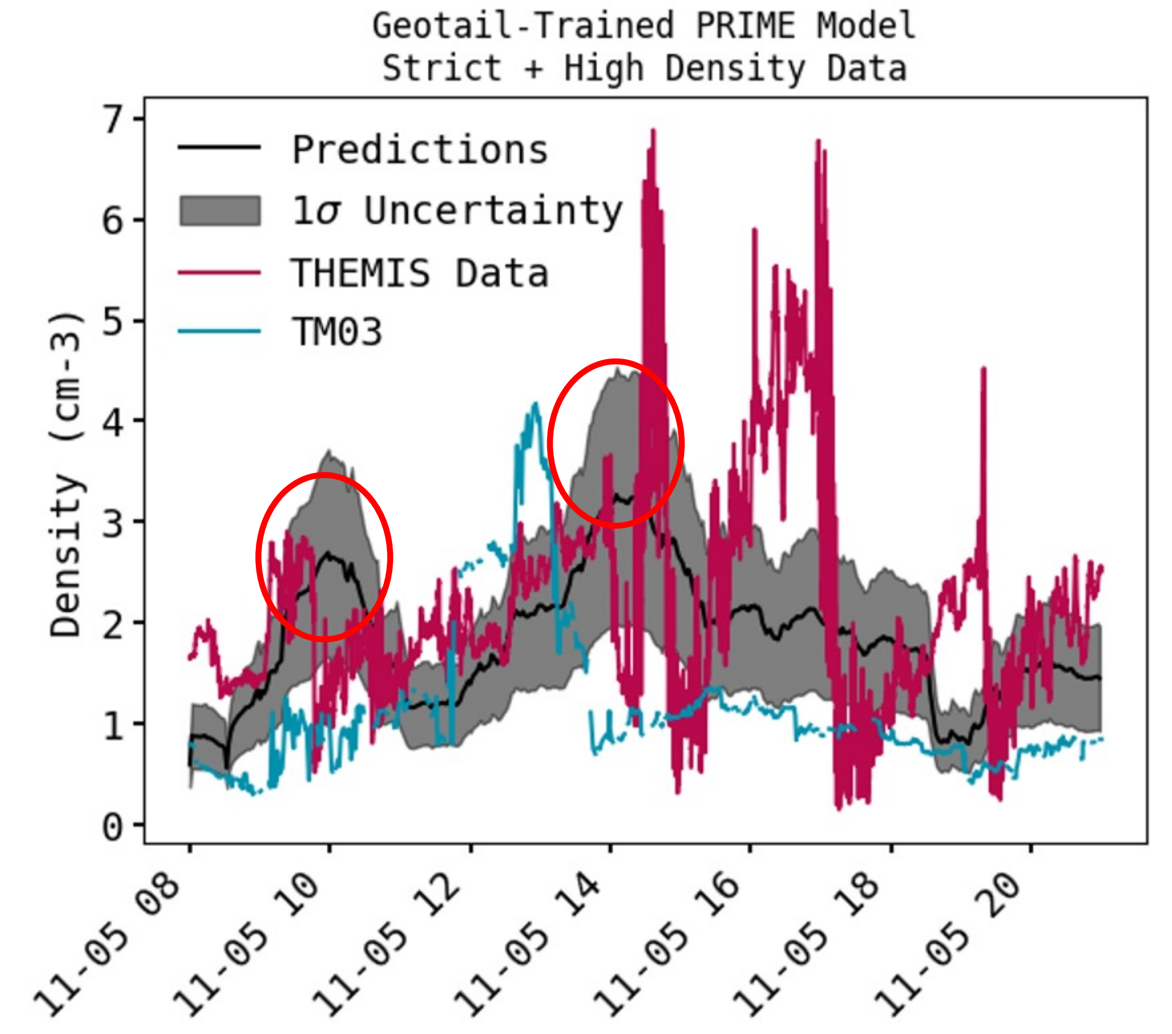
The Risk: Therefore we can mistakenly remove the crucial "stormtime plasmasheet."

The "Solution": Manually find the missing data and add it to the dataset.

Strict CPS (e.g., Ohtani et al., 2008 Raptis et al., 2024) & Flexible CPS (e.g., Richard et al., 2022)

Test case of a storm (05 Nov 2023)





<u>MAE (>40% improvement)</u>

ML model: 0.7 [1/cc]

TM03: 1.22 [1/cc]

Note: values <1 cm^{-3}, are transitions to the lobe/BL (will filter them out).

Summary & Discussion

Results

✅ **Marginal Gains:** ML models overall outperform analytical methods and show hidden asymmetries.

❌ **Mediocre Predictability:** We only capture "boring" conditions, not the critical rare events.

🧠 **Core Problem:** Our training data is biased. Extreme events, which are not captured by simple thresholds, must be included, but even then, they are very rare...

Future Work

Understand the output: How can we use these output to understand more about the physical processes?

Simulations to the Rescue (?): Try use simulations to generate extreme event data that *in-situ* observations struggle to provide.

Advertisement: LMAG25 (13 – 17 OCT 2025 JHU/APL)

When: 13–17 October 2025

Where: JHU/APL, Laurel, MD (primarily in-person)

Remote access: Zoom participation available

Format: ~20 minute talks plus short Q&A. Emphasis on interaction and collaborative problem-solving

Topics: See the LMAG2025 site for science themes; topic suggestions and ideas welcome

Audience: Heliophysics and geospace researchers, data scientists and computer scientists experts

No registration fee

RSVP today!

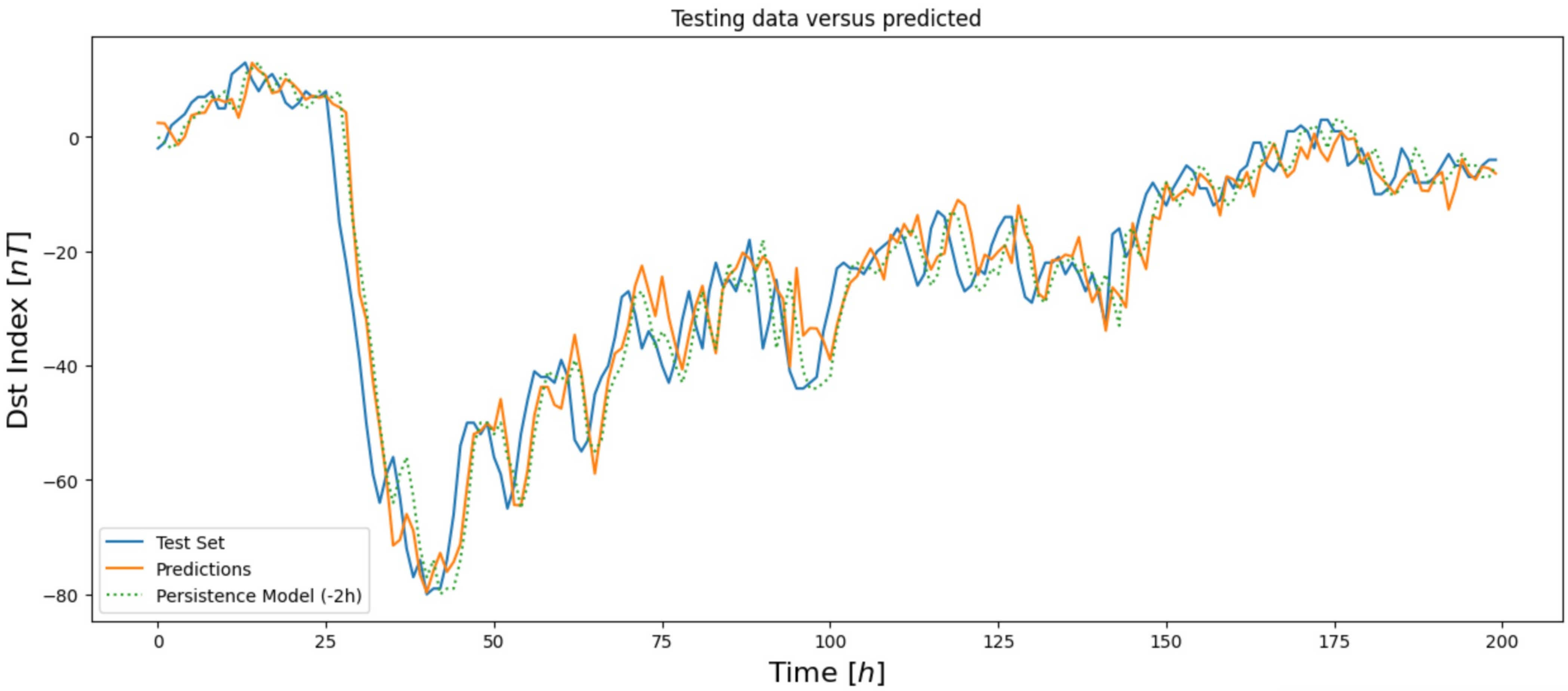
Extras

Forecasting DST index 3h in advance

explained_variance: 0.849
median absolute error: 3.758
r2: 0.848
MAE: 5.183
RMSE: 7.472

https://mybinder.org/v2/gh/SavvasRaptis/machine-learning-examples/HEAD

Forecasting DST index (3h) vs baseline model (2h)

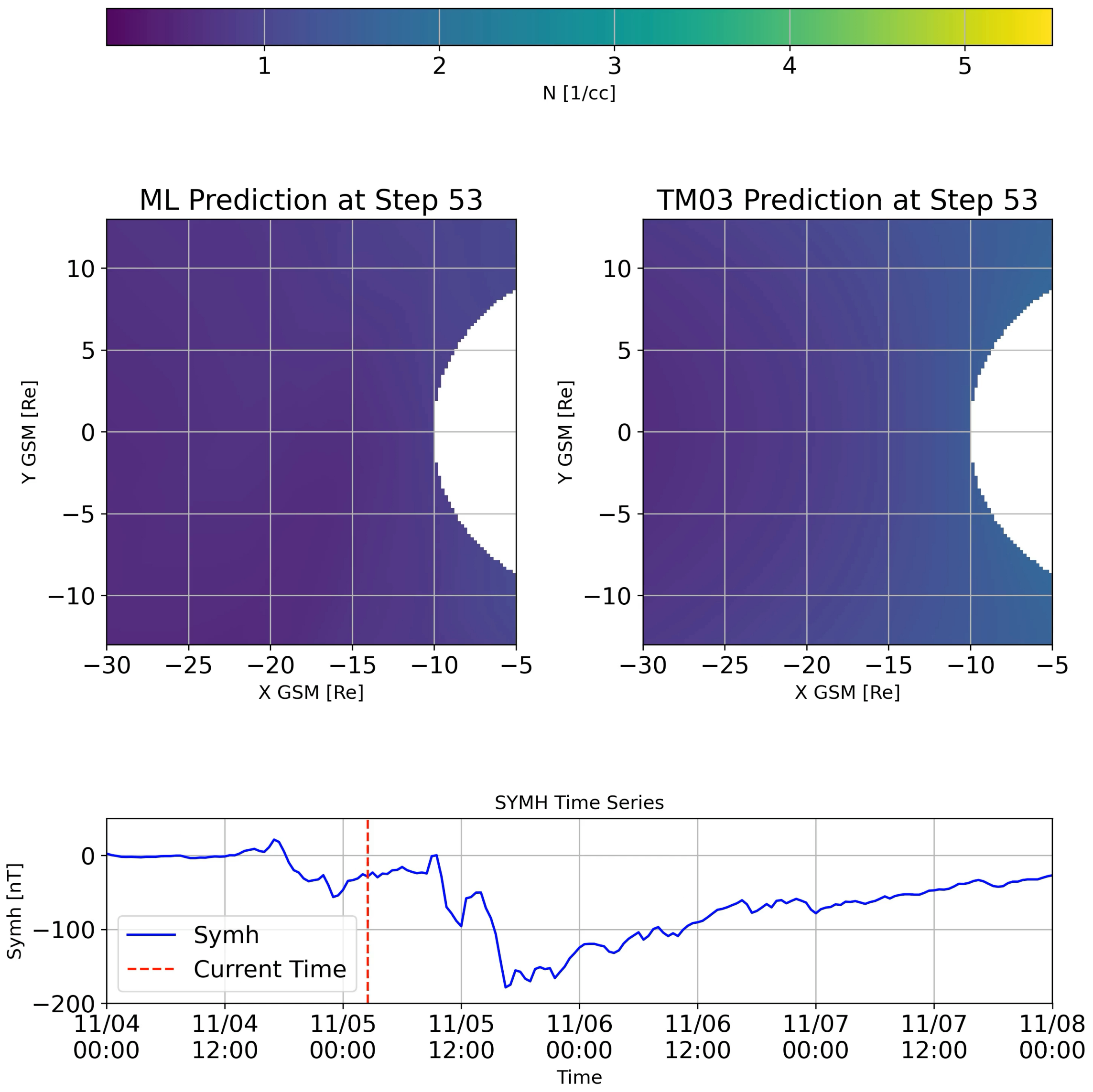


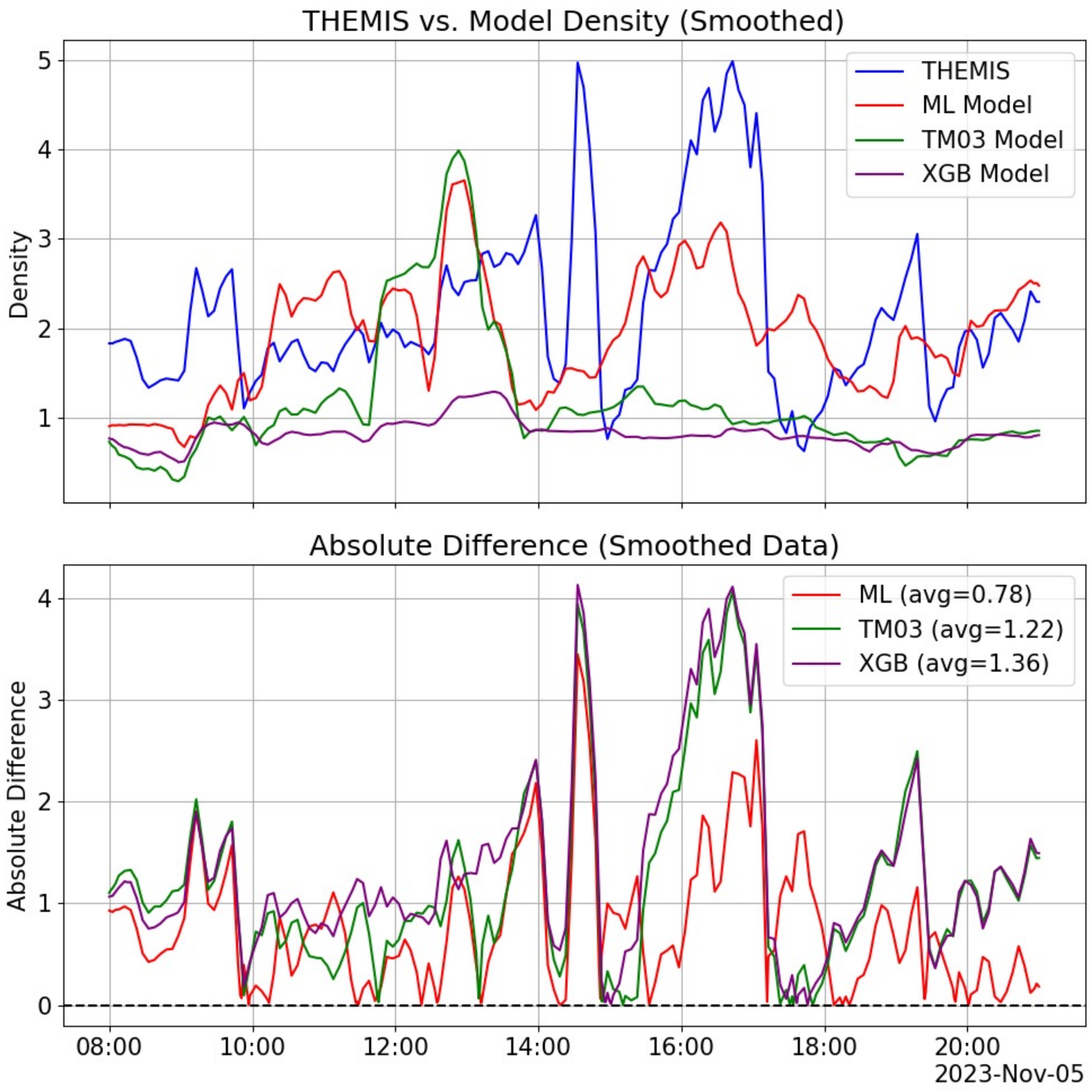
Predictions (3h)
explained_variance: 0.849
median absolute error: 3.758
r2: 0.848
MAE: 5.183
RMSE: 7.472

Persistance model (2h)
explained_variance: 0.864
median absolute error: 3.0
r2: 0.864
MAE: 4.71
RMSE: 7.076

https://mybinder.org/v2/gh/SavvasRaptis/machine-learning-examples/HEAD

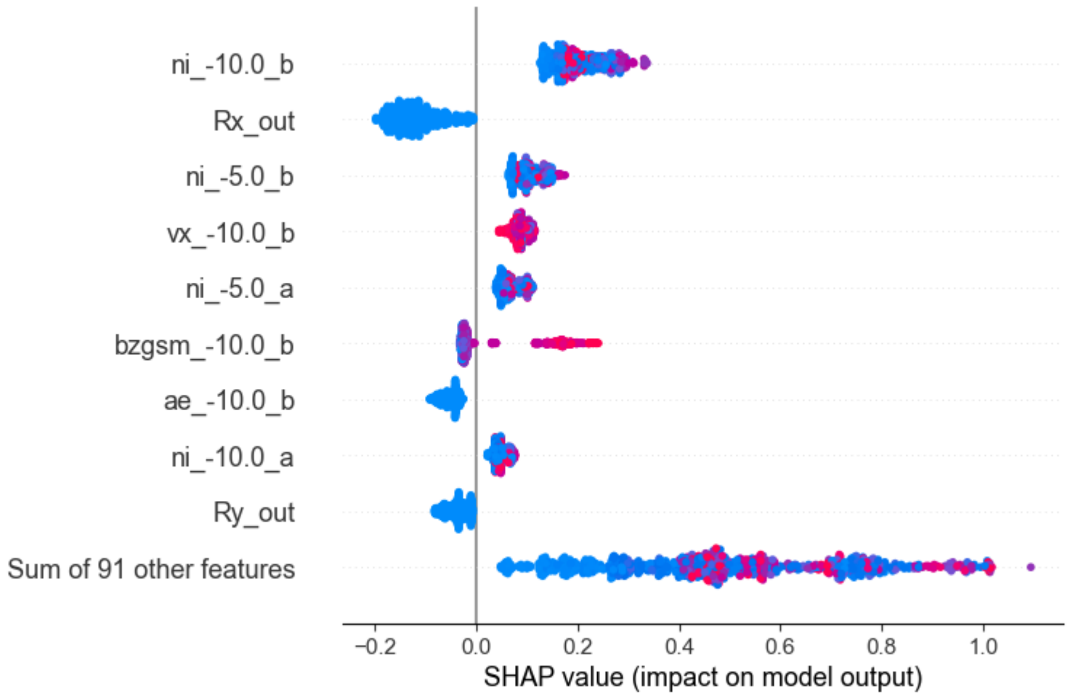
ML storm time density modeling



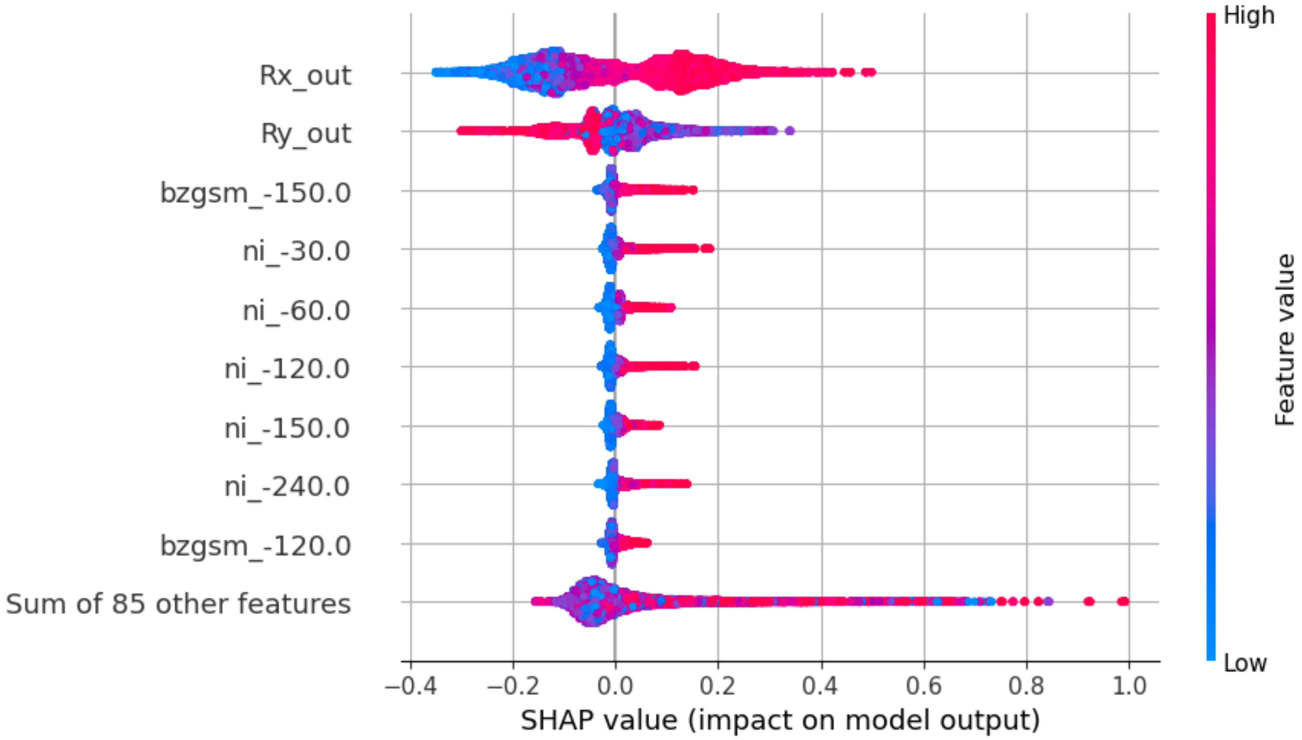


Model Feature importance storm vs quiet

In other words, the increased upstream density had a greater impact during the storm than the SC location.



X_test_storm



X_test_total

Updated results

Two different tests: (20% last data) – fixed dates 2005-2010 (after TM03 was published)

Table 2. Performance metrics for each method when evaluating plasma density (n_i) in units of [1/cc]

Testing: Last 20% of data								
	Strict CPS				Flexible			
Method	MAE	R^2	r	CRPS	MAE	R^2	r	CRPS
LightGBM	0.145	0.242	0.631	–	0.330	0.002	0.072	–
Neural Net	0.152	0.325	0.603	–	0.353	0.000	0.055	–
Linear Reg	0.173	0.265	0.620	–	0.367	0.001	0.066	–
PRIME-PS	**0.113**	**0.453**	**0.707**	0.083	**0.307**	**0.02**	**0.086**	0.278
TM03	0.163	0.208	0.570	–	0.339	0.000	0.055	–

Testing: Fixed dates 2005–2010								
	Strict CPS				Flexible			
Method	MAE	R^2	r	CRPS	MAE	R^2	r	CRPS
LightGBM	0.158	0.201	0.587	–	0.185	0.032	0.176	–
Neural Net	0.169	0.175	0.518	–	0.175	0.115	0.374	–
Linear Reg	0.197	0.162	0.543	–	0.256	−0.051	0.318	–
PRIME-PS	**0.146**	**0.286**	**0.658**	0.106	**0.142**	0.117	**0.454**	0.106
TM03	0.160	0.234	0.523	–	0.167	**0.159**	0.425	–

Table 3. Performance metrics for each method when evaluating ion temperature (T_i) in units of [KeV]

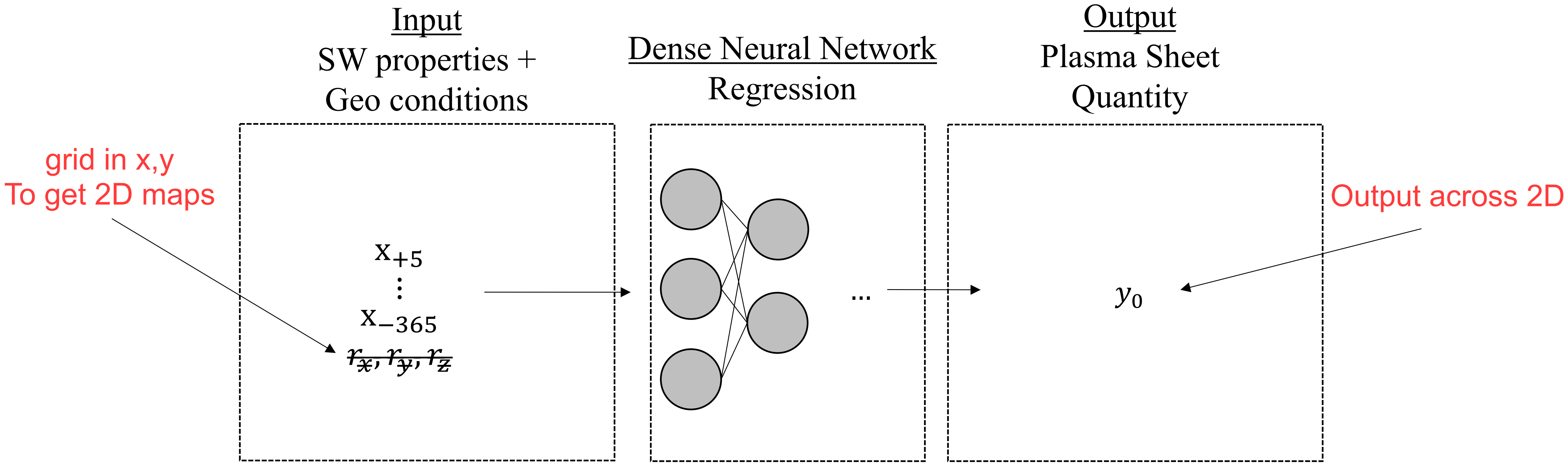
Last 20% of data								
	Strict CPS				Flexible			
Method	MAE	R^2	r	CRPS	MAE	R^2	r	CRPS
LightGBM	1.49	0.113	0.667	–	1.51	0.166	0.670	–
Neural Net	XXX	XXX	XXX	–	XXX	XXX	XXX	–
Linear Reg	1.26	0.344	0.609	–	1.29	0.377	0.635	–
PRIME-PS	**1.223**	**0.383**	**0.699**	0.859	**1.098**	**0.519**	**0.746**	0.779
TM03	4.08	-3.906	0.545	–	3.55	-2.766	0.587	–

Fixed dates 2005–2010								
	Strict CPS				Flexible			
Method	MAE	R^2	r	CRPS	MAE	R^2	r	CRPS
LightGBM	1.65	0.054	0.692	–	1.65	0.111	0.685	–
Neural Net	XXX	XXX	XXX	–	XXX	XXX	XXX	–
Linear Reg	1.36	0.380	0.619	–	1.35	0.425	0.655	–
PRIME-PS	**1.258**	**0.442**	**0.682**	0.9054	**1.246**	**0.486**	**0.709**	0.893
TM03	4.27	-3.409	0.555	–	3.78	-2.535	0.601	–

<u>Key Results:</u>

- PRIME-PS demonstrates a performance edge (~30% MAE from TM03 and ~10% from NN).
- This advantage is relatively low (from cross-validation).
- Different input, method, history, etc. had a marginal effect.
- Why is this the case?

Next step: 2D modeling

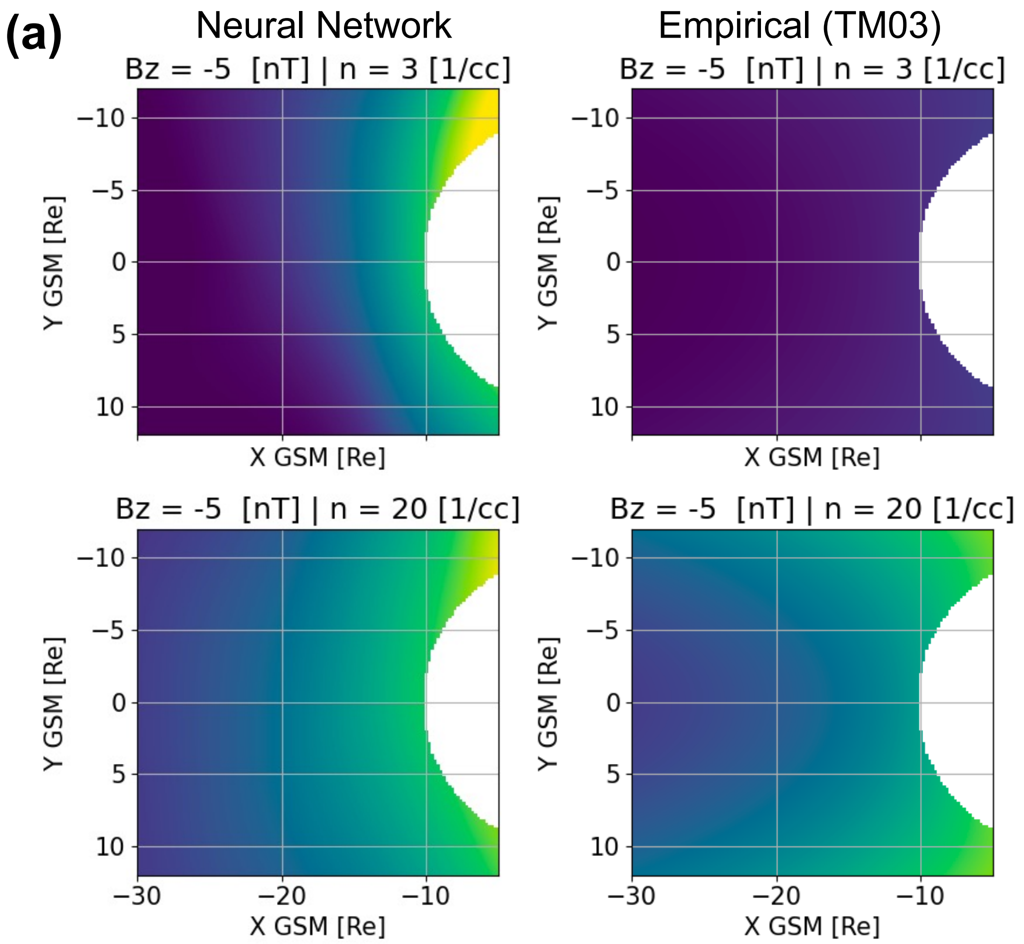


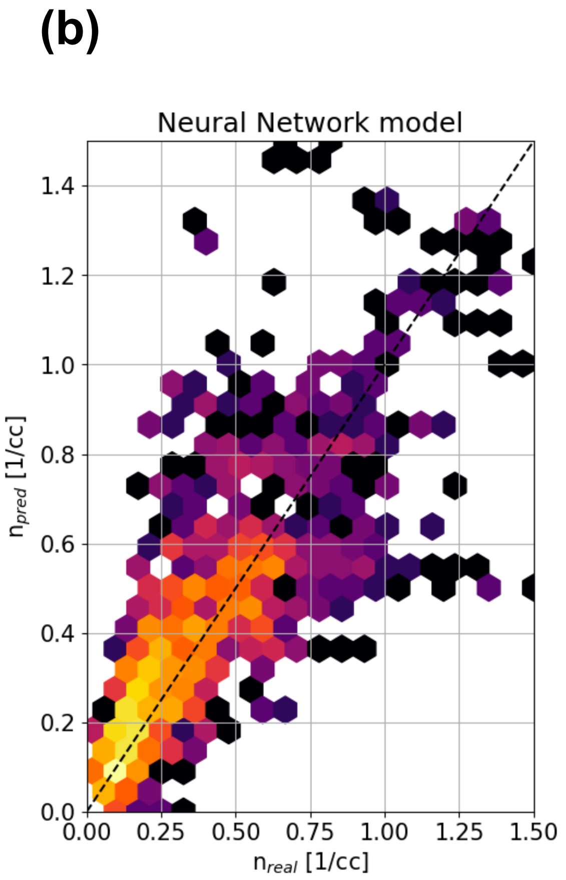
Input:
x: Different solar wind features (e.g., n, B, etc.) + geomagnetic indices including time history up to 6h
r: Location of SC measuring output

Output:
y: Different quantities at plasma sheet (e.g., n, B, T etc.)

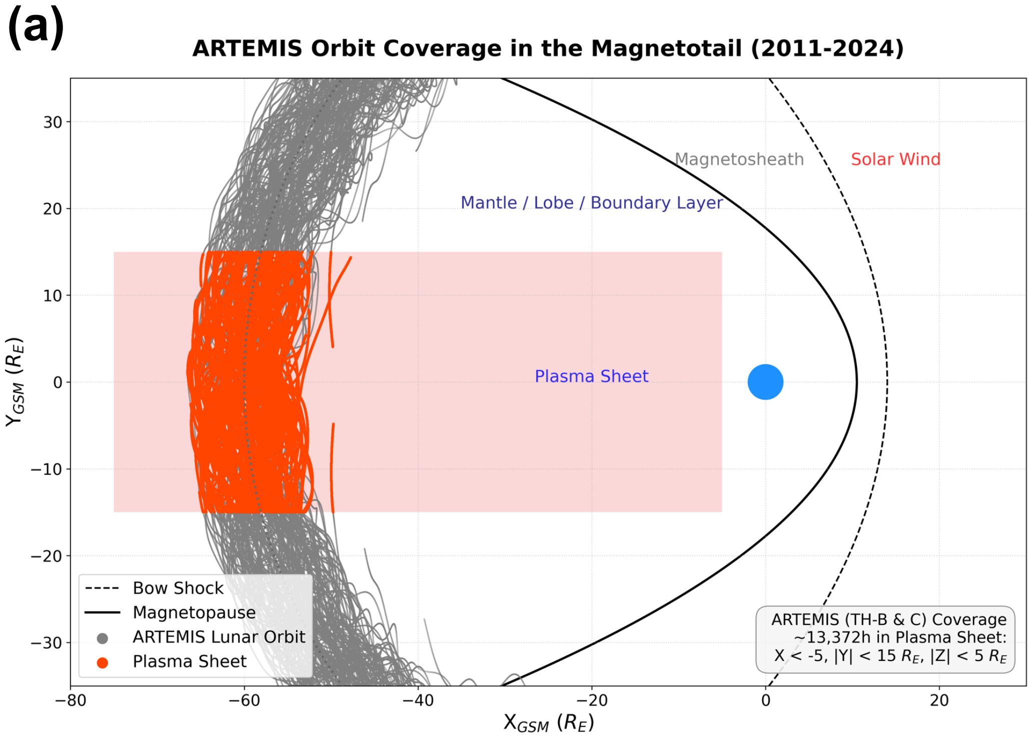
SW Input tried: Wind (1min res) – OMNIweb (1/5 min res)

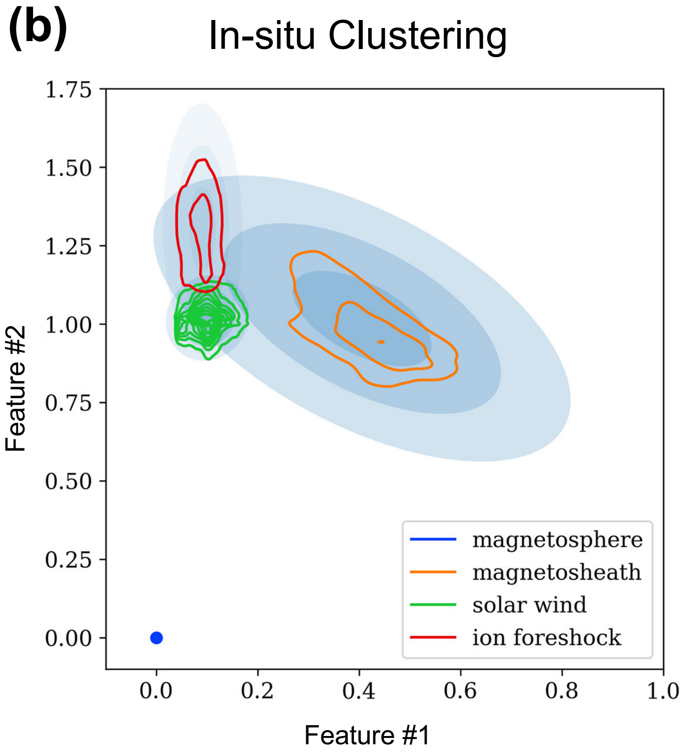
Results here shown for 1min averaged quantities for output





	TM03	NN
R2	0.17	**0.68**
MAE	0.19	**0.11**
RMSE	0.27	**0.2**
r (cor)	0.58	**0.83**


(a)
ARTEMIS Orbit Coverage in the Magnetotail (2011-2024)
Magnetosheath
Solar Wind
Mantle / Lobe / Boundary Layer
Plasma Sheet
Bow Shock
Magnetopause
ARTEMIS Lunar Orbit
Plasma Sheet
ARTEMIS (TH-B & C) Coverage
~13,372h in Plasma Sheet:
X < -5, |Y| < 15 R_E, |Z| < 5 R_E
Y_{GSM} (R_E)
X_{GSM} (R_E)


(b)
In-situ Clustering
Feature #2
Feature #1
magnetosphere
magnetosheath
solar wind
ion foreshock