Evaluating plasma sheet properties with in-situ observations and machine learning – Recent advancements and limitations

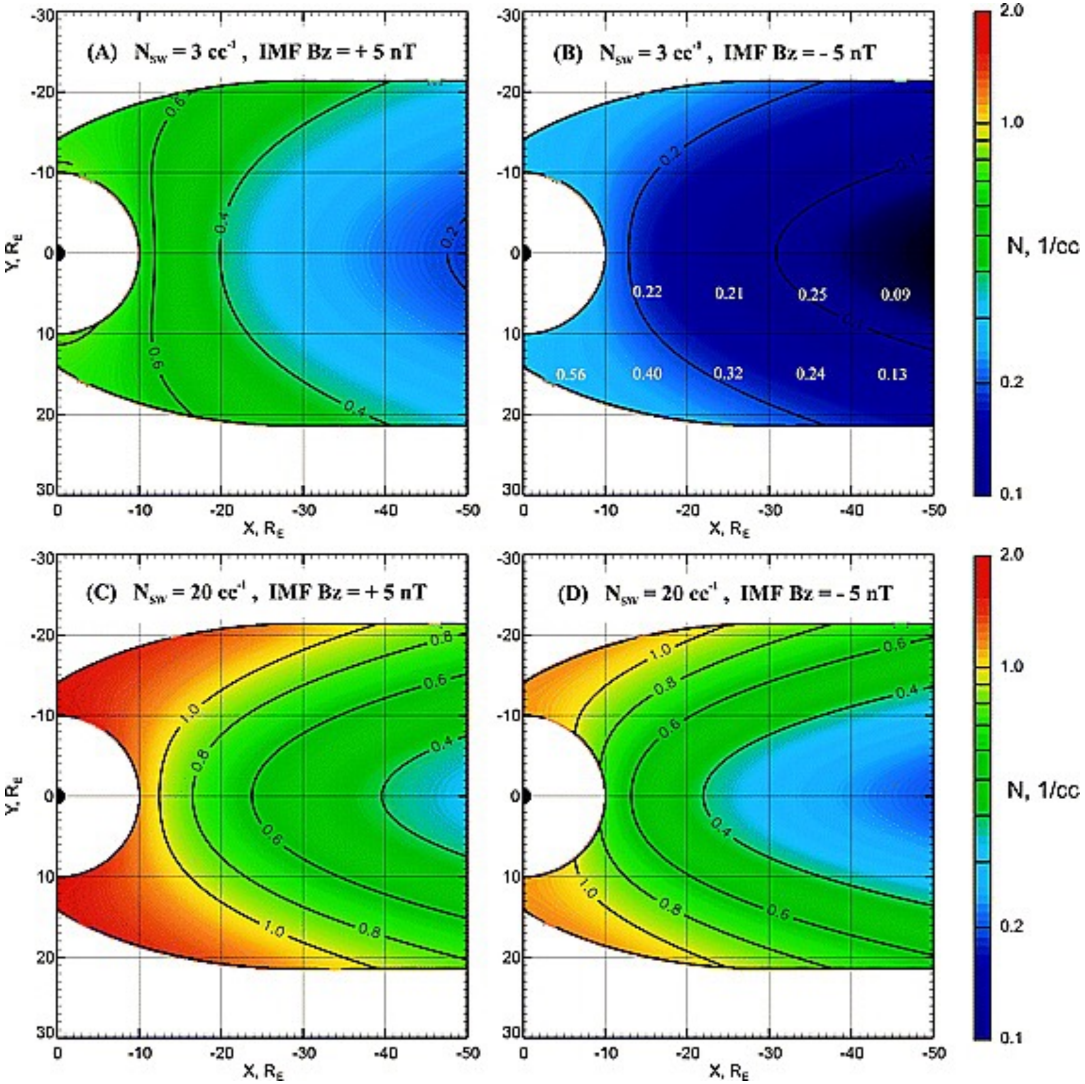
Savvas Raptis

APL/JHU, Laurel, MD, USA

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

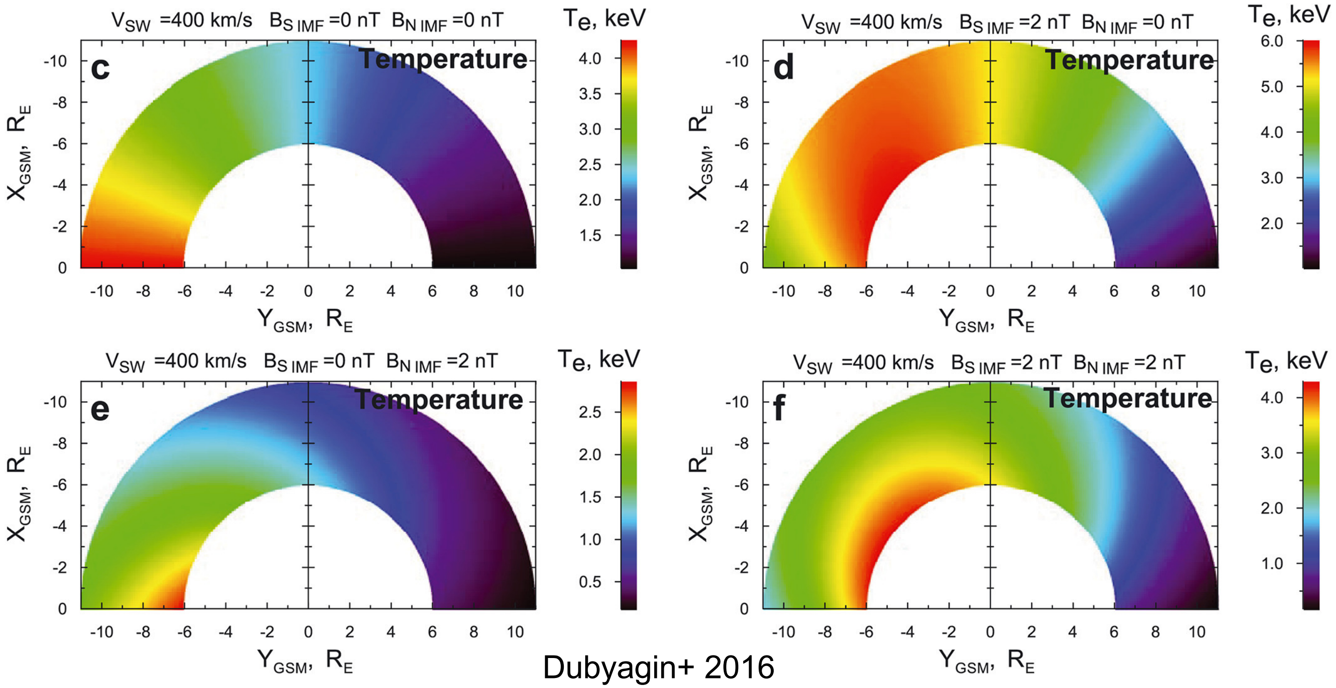
Baseline empirical models for Ti and Te

Modelled with Geotail



Tsyganenko & Mukai 2003

Modelled with THEMIS



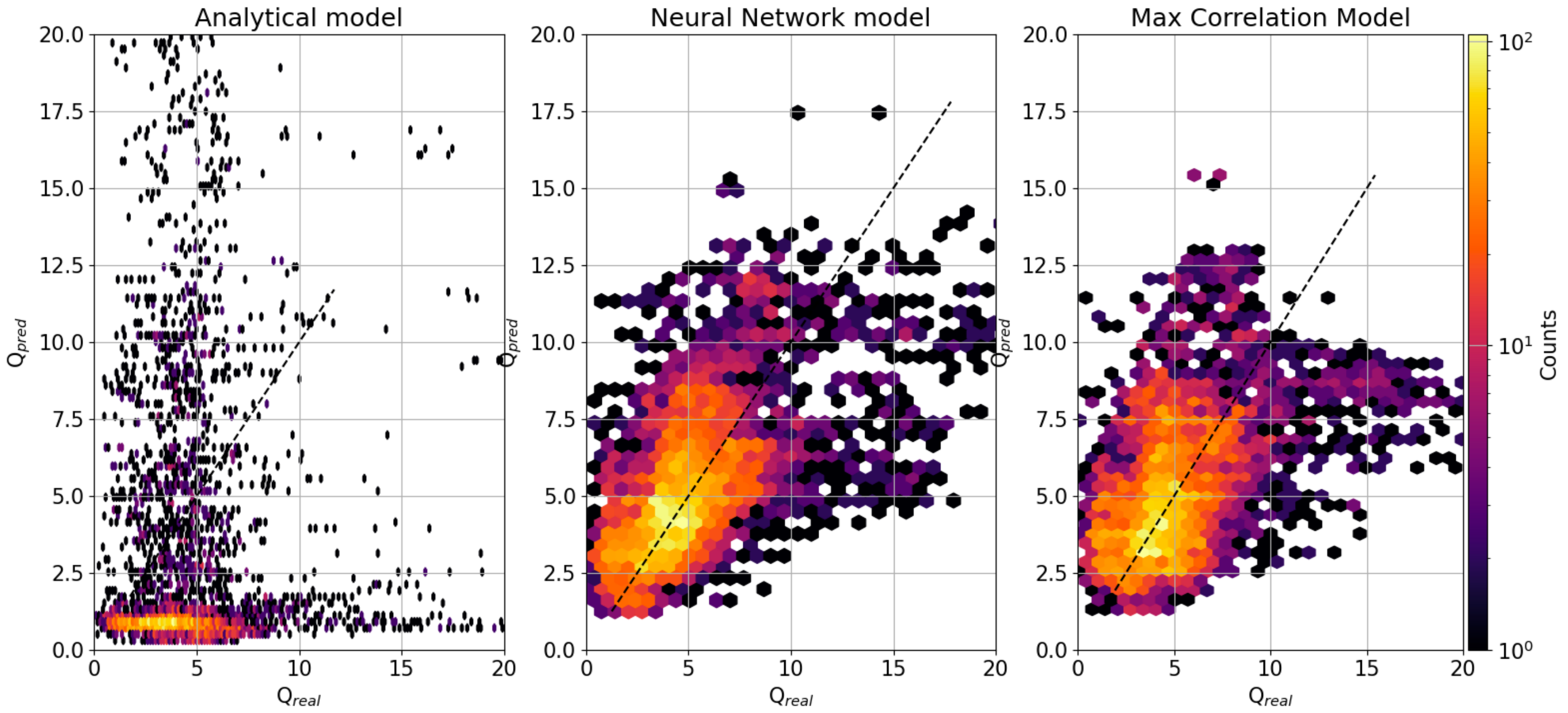
Dubyagin+ 2016

Why work on this?

1. Understand plasma sheet properties and how different populations get heated
2. Important quantity for global modeling to initialize distribution of particles in the inner magnetosphere

Results shown 2 days ago in ML session

Trying to model Ti/Te in the plasmasheet (input Solar wind, Output MMS)



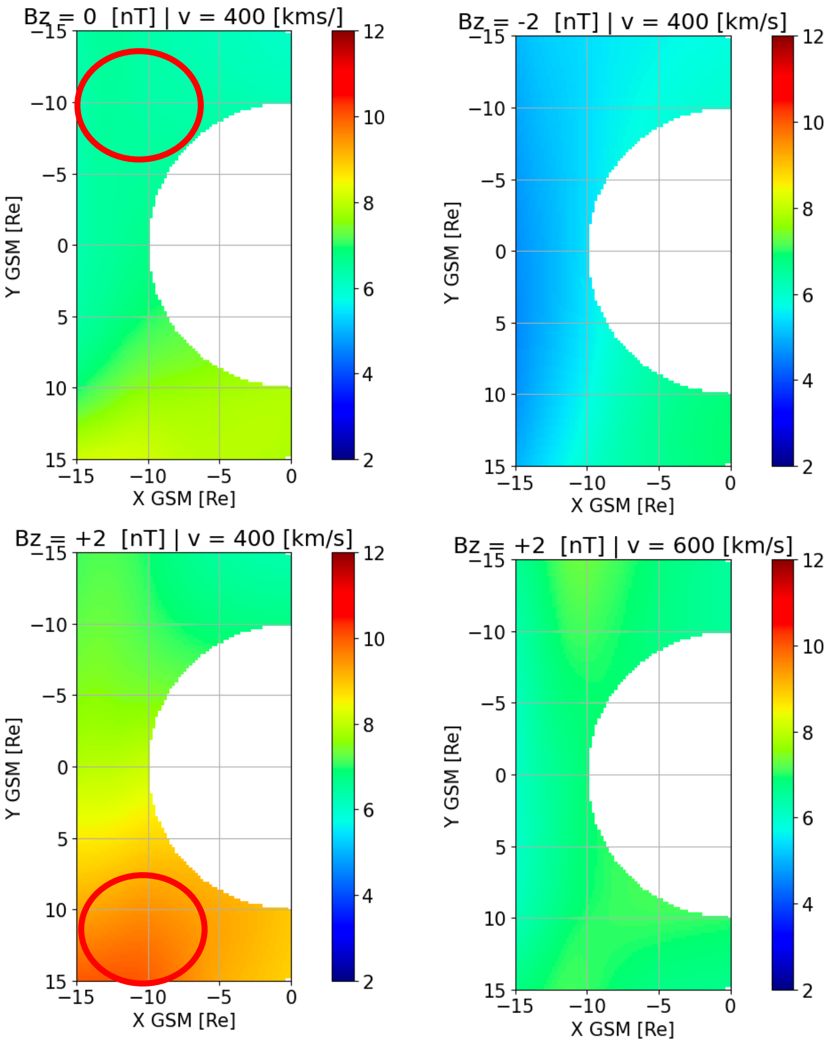
	NN	Linear
R2	**0.32**	0.12
MAE	**1.8**	2.1
RMSE	**2.5**	2.9
r (cor)	**0.6**	0.45

TLDR: Analytical model bad, neural network good? Yes kinda.

Note: To be fair to the analytical models, they are not trained on similar radial distances or datasets

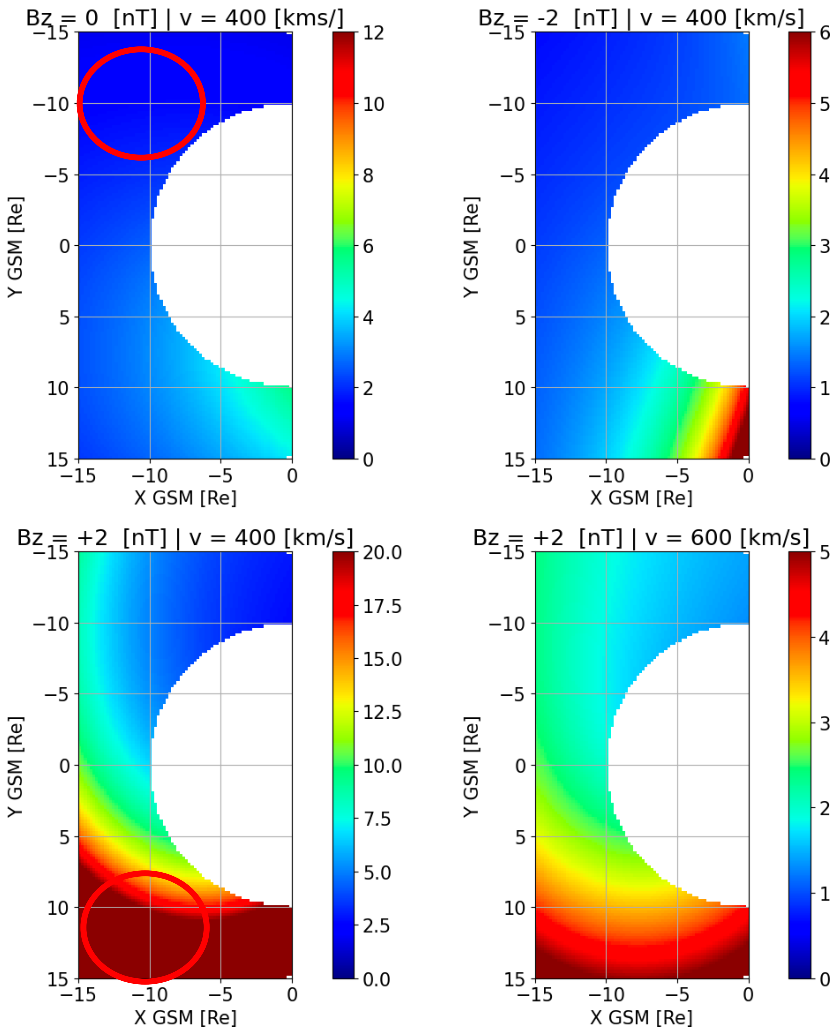
Modeling Temperature Ratios | 2D Maps

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)

But….there are many issues ☺

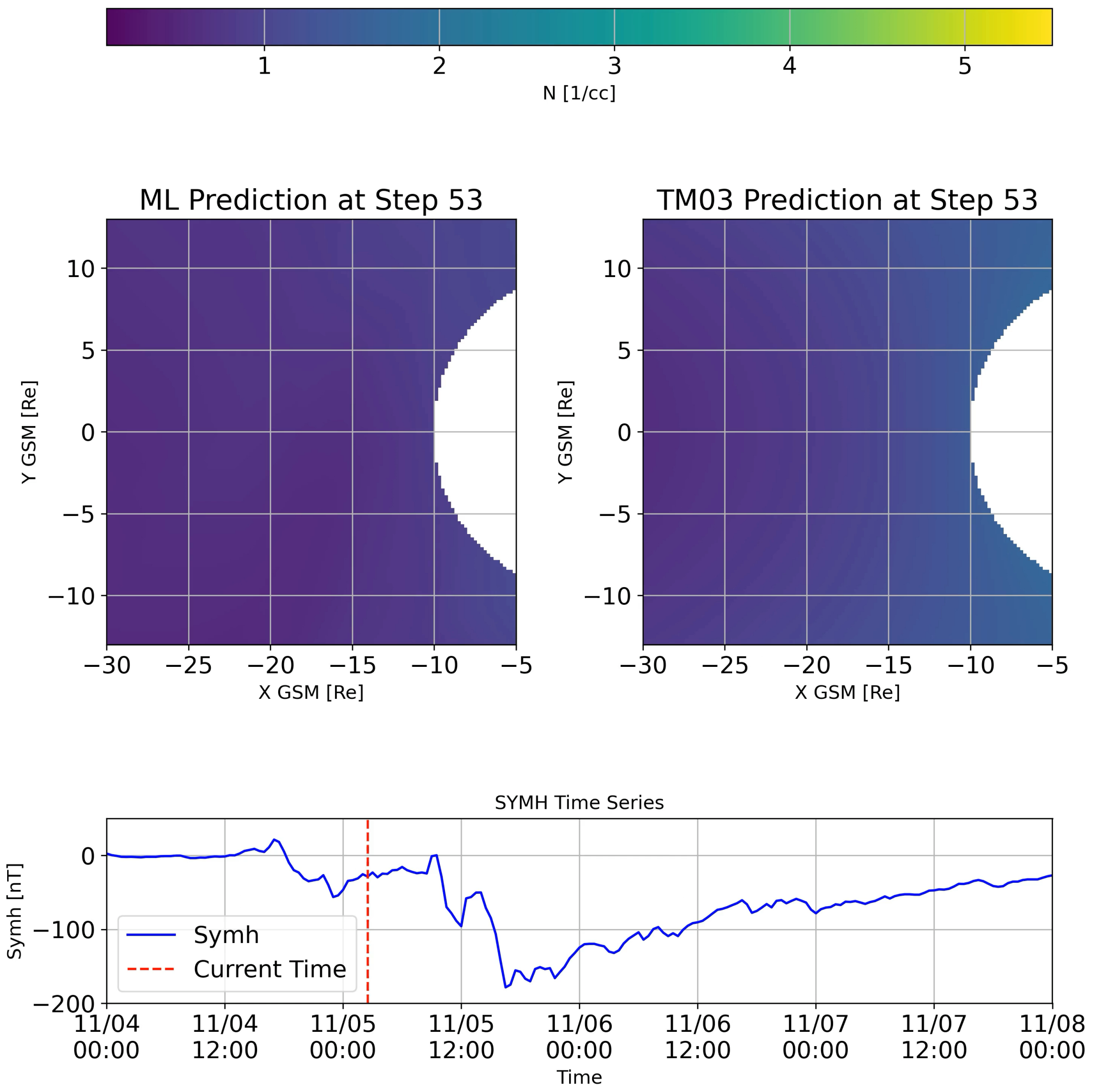
- **Data Preprocessing**
- **Data Sparsity & Extreme Events**
- **Statistical Metrics & Pitfalls**
- **Modeling Challenges for Storms**

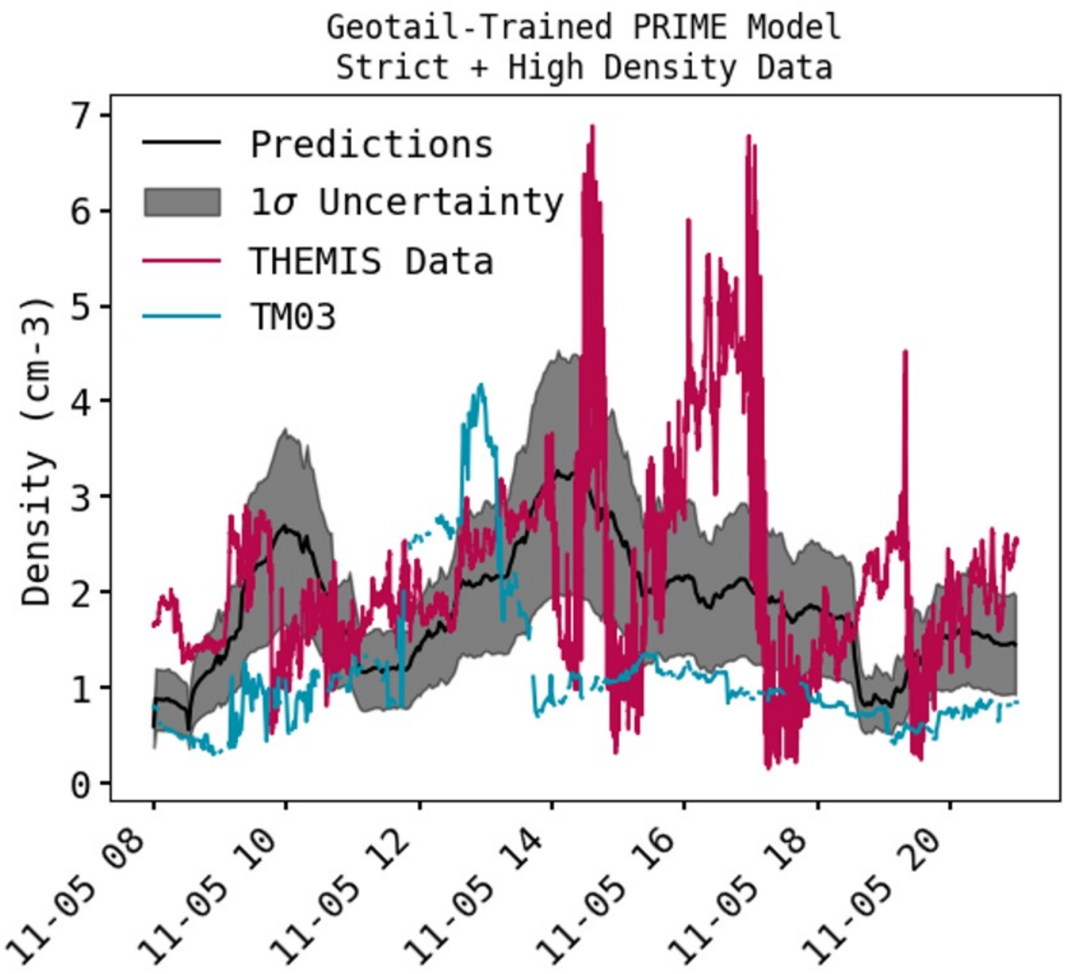
My initial plan was to talk about all of them…

- **Data Preprocessing (Ti/Te)**
  - ➢ Better moments can yield ~30% difference (Q~4 rather than 5.5)
- **Data Sparsity & Extreme Events (General)**
  - ➢ Number of unique events and distribution of driving conditions is more insightful than statistical metrics.
- **Statistical Metrics & Pitfalls (Ti/Te)**
  - ➢ Constant-value models (baseline) can outperform complex models.
  - ➢ Use (adjusted) R^2 instead of correlation for model evaluation.
  - ➢ Binning data can artificially increasing metrics.
  - ➢ Always compare to baseline models (*linear* / *persistence* model)
- **Modeling Challenges for Storms (Ti/Te)**
  - ➢ Storms are difficult to model due to sparse observations and limited information

- **Data Preprocessing (Ti/Te)**
  - Better moments can yield ~30% difference (Q~4 rather than 5.5)
- **Data Sparsity & Extreme Events (General)**
  - Number of unique events and distribution of driving conditions is more insightful than statistical metrics.
- **Statistical Metrics & Pitfalls (Ti/Te)**
  - Constant-value models (baseline) can outperform complex models.
  - Use (adjusted) R^2 instead of correlation for model evaluation.
  - Binning data can artificially increasing metrics.
  - Always compare to baseline models (*linear* / *persistence* model)
- **Modeling Challenges for Storms (Ti/Te)**
  - Storms are difficult to model due to sparse observations and limited information

Can we really predict extreme events with data? Density



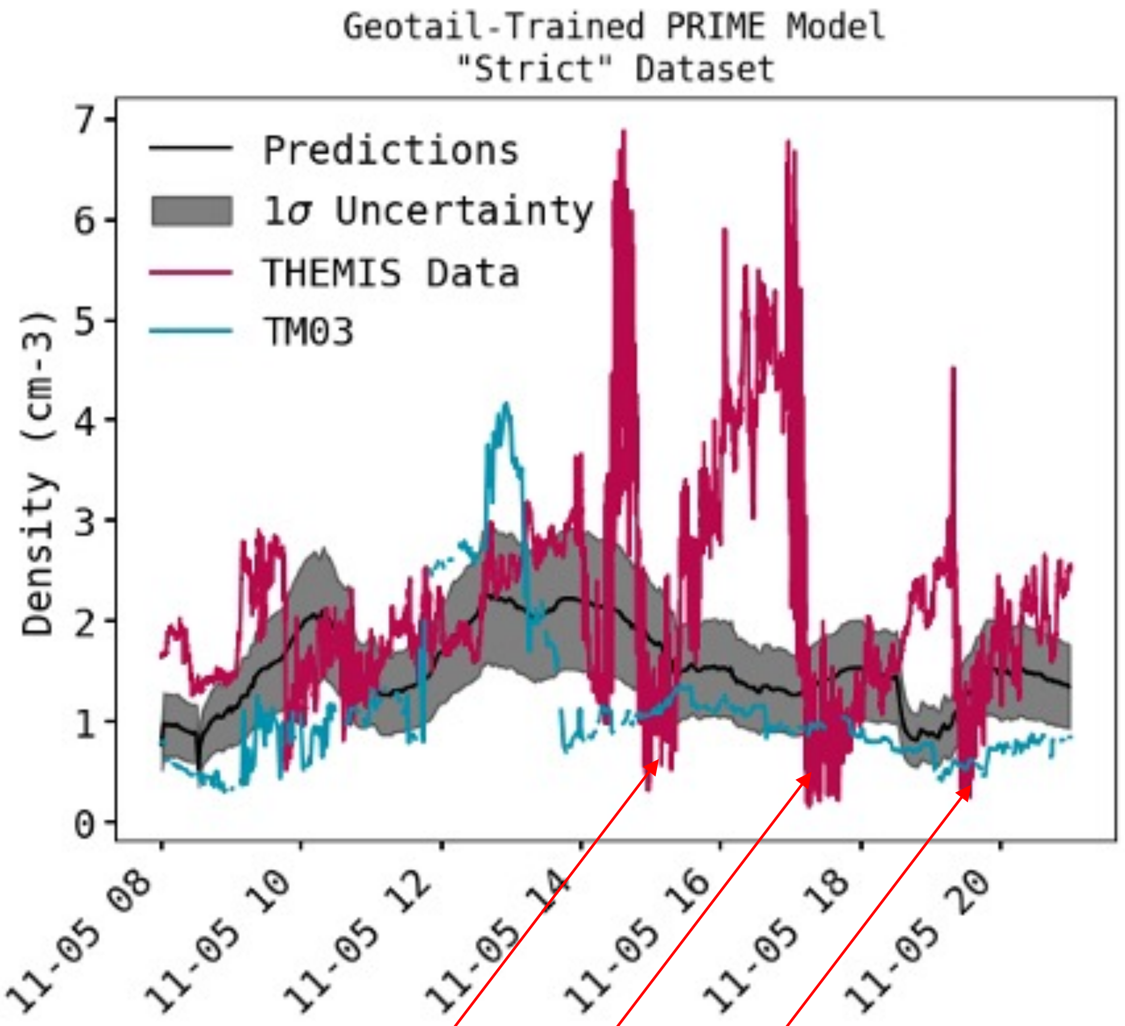


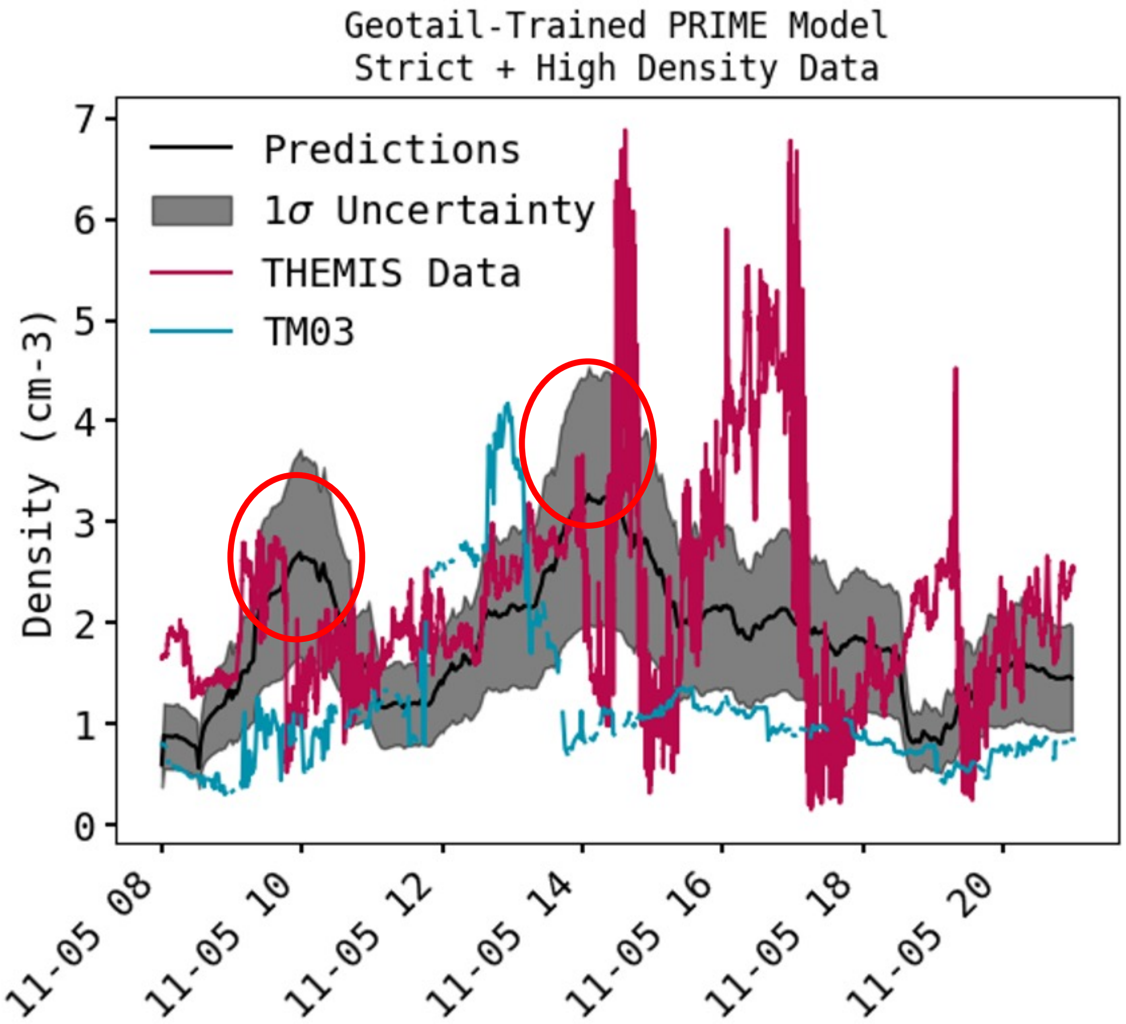
<u>MAE (>40% improvement)</u>

PRIME-PS: 0.7 [1/cc]

TM03: 1.22 [1/cc]

The importance of "Rare Events"





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

ML model: 0.7 [1/cc]

TM03: 1.22 [1/cc]

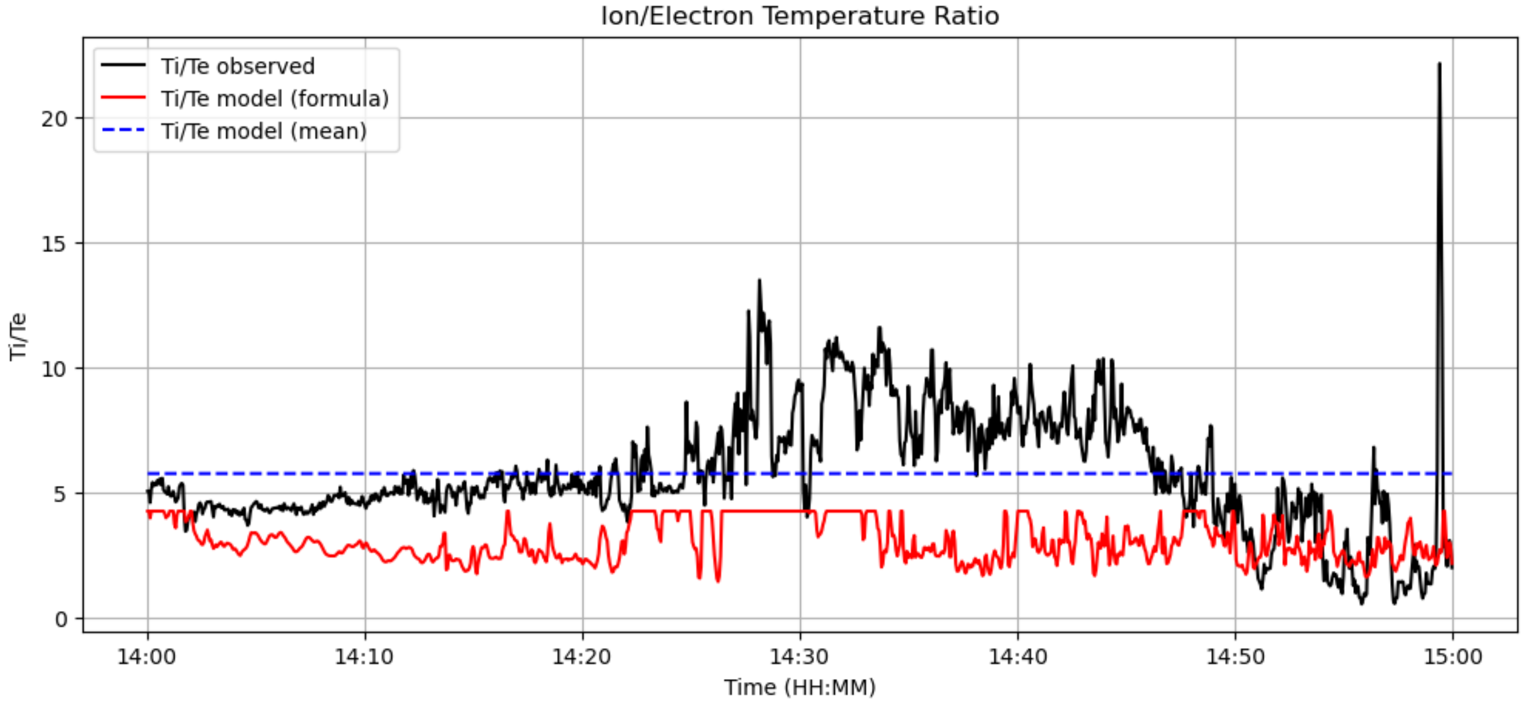
Note: values <1 cm^{-3}, are transitions to the lobe/BL (will filter them out).

How is Ti/Te doing there?

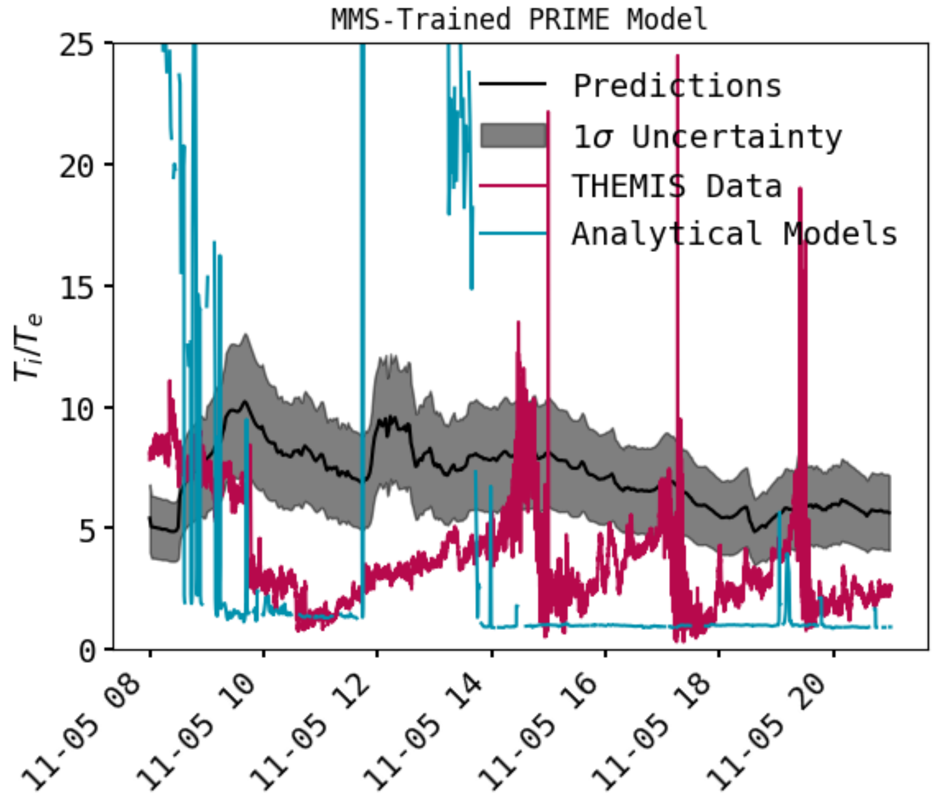
Mean Model: MAE = 1.67

Formula Model (a polynomial fit from another period): MAE = 3.21

Formula vs Mean: MAE = -92.64%



Old model shown, new is actually better, there is still hope



Key Message: Modeling (in-situ or from SW) struggle during extreme events like storms

Ending on a relatively positive note

What We Achieved: Better Performance ("shrug")

By incorporating time history and more diverse input, our ML models improved plasma sheet density predictions by up to +40% for quiet times and for a case study storm.

What We Learned: The Limits of Data-Driven models

- 💡 Data-driven **ML models** are powerful but **can be matched by simpler baseline models** if the input data is not sufficient to describe the system
- 💡 The "Rare Event" Problem: **Adding more "rare" events can help, but isn't a silver bullet**. We need a better data strategy.

Path Forward

- 🔭 Adopt **rigorous validation and transparent assessment** as a core practice. **e.g., a model with R^2 of 0.054 had a r of 0.7.**
- 🔭 Build Better methodologies: **Develop hybrid simulation-observation methods** to create **representative datasets that include more extreme events**.

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!