Check for updates

JGR | Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2026JH001589

Machine Learning Modeling of Earth's Plasma Sheet Using Multi-Spacecraft Observations

Savvas Raptis[1], Connor O'Brien[1], Kareem Sorathia[1], Viacheslav Merkin[1], Shinichi Ohtani[1], Louis Richard[2,3], Anusree P. Devanandan[4], and Simon Wing[1]

[1]Johns Hopkins University Applied Physics Laboratory, Laurel, MD, USA, [2]Swedish Institute of Space Physics, Uppsala, Sweden, [3]Rudolf Peierls Centre for Theoretical Physics, Clarendon Laboratory, University of Oxford, Oxford, UK, [4]Department of Physics and Astronomy, University of New Hampshire, Durham, NH, USA

Abstract Accurate modeling of Earth's plasma sheet is essential for global simulations and understanding magnetospheric dynamics, yet empirical models often fail to capture its complex spatial structure. Leveraging decades of in situ observations from Geotail and the Magnetospheric Multiscale mission, we employ gradient boosting and neural networks, benchmarked against linear regression and empirical baselines, to model plasma sheet ion density, temperature, and ion-to-electron temperature ratios. Our findings reveal that predictive accuracy is highly sensitive to data set curation, particularly the classification of the plasma sheet region and the accuracy of the plasma moments used for training. The presented data-driven models successfully reproduce global asymmetries that empirical baselines miss, including dawnside density enhancements, duskward ion temperature elevations, and T_i/T_e variations driven by charge-dependent drifts. To test generalization, we evaluate the models against Time History of Events and Macroscale Interactions During Substorms observations during a severe geomagnetic storm and demonstrate that incorporating high-density intervals, often removed as outliers, enables the model to capture storm-time plasma sheet density enhancements, reducing the mean absolute error by ∼30% relative to the empirical baseline for this event. The remaining discrepancies arise primarily from the sparsity of extreme events in historical records, suggesting that hybrid physics-based and data-driven modeling is a promising path toward improved storm-time prediction.

Plain Language Summary The plasma sheet is a region of Earth's nightside magnetic environment filled with hot, charged particles. Changes in this region drive space weather events that can affect satellites and power grids, yet predicting its behavior remains a challenge. We used machine learning to analyze decades of satellite measurements and build a new predictive model for the density and temperature of these particles. A major success of our approach is its ability to reproduce previously observed physical asymmetries that older models miss: the plasma sheet is denser on the dawn side, and on the dusk side, ions are hotter relative to electrons than on the dawn side. We also tested the model against a severe geomagnetic storm and found that including rare, high-intensity data points in the training set reduced the density error by roughly 30% relative to an established empirical model and allowed the model to track storm-time density enhancements that simpler models cannot reproduce. The remaining gaps in accuracy point not to algorithmic limitations but to the scarcity of historical data recorded during these rare, extreme events, a challenge that future multi-mission data sets and simulation-observation fusion can address.

Key Points:

- Data-driven models reproduce dawn-dusk density, temperature, and T_i/T_e asymmetries that previous empirical models miss
- Data set curation, namely plasma sheet classification and moment contamination correction, critically controls model performance
- Including high-density intervals in training improves prediction during a severe storm, pointing to data sparsity as a key bottleneck

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

S. Raptis,
savvas.raptis@jhuapl.edu;
savvasraptis@pm.me

Citation:

Raptis, S., O'Brien, C., Sorathia, K., Merkin, V., Ohtani, S., Richard, L., et al. (2026). Machine learning modeling of Earth's plasma sheet using multi-spacecraft observations. *Journal of Geophysical Research: Machine Learning and Computation*, *3*, e2026JH001589. https://doi.org/10.1029/2026JH001589

Received 13 JUL 2026
Accepted 9 NOV 2026

© 2026 The Author(s). *Journal of Geophysical Research: Machine Learning and Computation* published by Wiley Periodicals LLC on behalf of American Geophysical Union.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Earth's plasma sheet is a highly dynamic plasma environment where cross-scale phenomena, fast flows, plasma instabilities, and magnetotail reconnection create a complex and difficult-to-characterize system (Sitnov et al., 2019). Even without transient processes, the global configuration of the plasma sheet can drastically change depending on geomagnetic conditions, with storms exhibiting different plasma properties and associated fields compared to quiet times (Devanandan et al., 2026; Ohtani & Mukai, 2008; Raptis et al., 2024). To characterize this region globally, one must invoke either physics-based simulations or data-driven solutions based on the statistical picture from multiple decades of observations.

Machine learning (ML) models, particularly supervised regression techniques, have been applied to a wide range of tasks in heliophysics: modeling radiation belt fluxes, phase space density, and diffusion coefficients (Aminalragia-Giamini et al., 2023; Chu et al., 2021; Ma et al., 2022; Smirnov et al., 2020; Wing et al., 2022), predicting geomagnetic conditions (Hu et al., 2022; Xu et al., 2020), modeling the inner magnetosphere (Hua et al., 2026),

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

characterizing solar wind under various magnetospheric processes (Cao et al., 2023; Li et al., 2023) and obtaining near-Earth dayside conditions in the solar wind and magnetosheath (O'Brien et al., 2023, 2024). However, despite decades of research on plasma sheet properties, ML applications in this region remain relatively limited. Earlier data-driven efforts have produced models for bulk plasma sheet properties based on solar wind conditions and nightside location (Dubyagin et al., 2016; Tsyganenko & Mukai, 2003), while other work has used data mining to reconstruct full 3D magnetic field configurations that describe field lines in the plasma sheet region (Sitnov et al., 2021). Nevertheless, modeling plasma properties remains challenging due to the complexity of solar wind entry into the plasma sheet and the multiple temporal and spatial scales involved (Wing et al., 2006, 2014). Capturing the variability of plasma sheet properties, especially at short timescales and under different upstream conditions, has proven extremely difficult using functional models.

In this work, we revisit the problem of formulating a data-driven plasma sheet model using modern ML tools and multi-mission observations from the ISAS/NASA Geotail mission (Nishida, 1994) and the NASA Magnetospheric Multiscale (MMS) mission (Burch et al., 2016). The Geotail data set spans a broader range of plasma sheet conditions than the data sets used in prior models, while MMS' state-of-the-art instrumentation remains largely unexplored for this purpose. The main goal of this paper is to illustrate current data-driven statistical modeling (i.e., ML) capabilities for plasma sheet modeling by investigating different methodologies, target variables, and input–output configurations. We also investigate the influence of time history, which most previous plasma sheet models either omit or represent only through lagged averages of the solar wind (e.g., Dubyagin et al., 2016), despite evidence that past solar wind driving shapes the magnetospheric system. Finally, beyond producing ML models, we evaluate the relative importance of each input parameter under different driving regimes, including extreme events such as geomagnetic storms, and propose strategies to address the challenges they present in future efforts.

2. Data

For the analysis, we use Geocentric Solar Magnetospheric (GSM) coordinates aberrated by 4° (AGSM). The target output of our models consists of observations from the Geotail mission (1994–2022) and MMS mission (2015–2024). MMS observations comprise ion and electron plasma moments in fast mode from the Fast Plasma Investigation (FPI) (Pollock et al., 2016). Specifically, we use partial moments as generated by the FPI team, which integrate the distribution over a restricted energy range that excludes potentially contaminated channels, rather than over all measured energies, making them more accurate when the spacecraft resides in the plasma sheet region (Gershman et al., 2019; Graham et al., 2023; Liu et al., 2025). To generate temperature moments for ions and electrons, we integrate the particle distribution starting from the 19th (≈1.5 keV) and 6th (≈65 eV) energy bins, respectively. This simple approach effectively removes penetrating radiation and focuses on the statistically expected plasma sheet bulk population. It should be noted that, while this is a necessary post-processing step, it significantly changes the target output and alters its statistical properties relative to the level 2 (L2) product, a point we discuss further at the end of the manuscript. Data for both electrons and ions are taken every 4.5 s from the L2 survey/fast product. From Geotail, the ion plasma moments are taken from the Low-Energy Particle experiment (LEP) (Mukai et al., 1994), corresponding to 12-s resolution measurements; in its magnetosphere (EA) mode, the sensor covers energies from 32 eV/Q to 39 keV/Q.

For model input, we use solar wind information from two sources. The first approach uses the OMNI database (King & Papitashvili, 2005), which originates mainly from the Advanced Composition Explorer (Stone et al., 1998) and Wind spacecraft (Wilson et al., 2021) residing at the Sun-Earth L1 point, with data automatically time-shifted to Earth's bow shock nose. The time-shifted data we use are at 5-min resolution and account for bow shock location and shape. The second approach uses Wind data directly, allowing the neural network (NN) to evaluate its propagation from L1 to Earth's magnetospheric environment. This approach can be valuable, as the variability of local conditions with the OMNI data set can be significant (B. Walsh et al., 2019). We additionally supplement the input space with geomagnetic conditions via the SYM-H/SMR, AE/SME and AL indices. To associate the target outputs (MMS and Geotail plasma sheet data) with the inputs (OMNI and Wind observations), we resample via averaging the target output data to match the available Wind resolution of 100 s, resulting in a uniform data set between Wind, Geotail, and MMS at 100-s cadence across our input and output data set. For PRIME-PS (A probabilistic recurrent NN; see details below), 200 observations prior to each target time are used as input (333 min; see Section 3.5.2 for the physical motivation), whereas the other models use the prior twelve 5-min OMNI observations (1 hr).

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

We selected this 1-hr window after testing longer horizons (up to 4–5 hr). For the gradient boosting (GB) and shallow NN models, no statistically significant improvement was observed: the best-performing configuration fell within the same range as the shorter windows. Longer windows also reduce the usable sample size, because OMNI data gaps eliminate entire input sequences, although the relative contribution of this effect is minor. The long lags reported by Dubyagin et al. (2016) were derived at 6–11 R_E, where the plasma sheet reflects hours of earthward transport; most of our Geotail and MMS samples lie farther tailward, and solar wind parameters remain correlated over several hours, so a 1-hr history already carries much of the relevant information. For PRIME-PS, performance similarly saturated beyond approximately 1 hour of history. We nevertheless retain the physically motivated 5.5-hr window, because the GRU architecture can learn lag-dependent structure that tabular models cannot, which should matter most closer to Earth.

Finally, we make an additional comparison of our Geotail- and MMS-based model outputs against plasma sheet observations from the Time History of Events and Macroscale Interactions During Substorms (THEMIS) mission (Angelopoulos, 2008b) during a geomagnetic storm. From THEMIS, we specifically use the THEMIS-D probe, accessing L2 particle moments for ions and electrons mainly from the Electrostatic Analyzer (ESA) instrument (McFadden et al., 2008) to obtain precomputed ion and electron temperature moments at a cadence of ∼3.3 s. Specifically, the density and temperature observations from the THEMIS spacecraft were obtained using calculated combined ESA and Solid State Telescopes moments (gmoms) that are anticipated to be more accurate due to the inclusion of the high energy instrument. Additional post-processing of the raw observations is performed to make the storm-time data-modeling comparison meaningful. Data intervals exhibiting rapid, non-physical fluctuations associated with the spacecraft's transition from the plasma sheet to the plasma sheet boundary layer (PSBL) were flagged and excluded. This process was assisted by using a lower-bound threshold to exclude density measurements below 1.0 cm^{-3}. The gaps resulting from these exclusions were reconstructed for visual purposes using linear interpolation, while the reconstructed time series was smoothed using a moving mean window filter of 60 s.

3. Methods

3.1. Plasma Sheet Classification

To derive a plasma sheet model, we must first ensure that spacecraft observations from the broader magnetotail region correspond specifically to the central plasma sheet (CPS). Classifying magnetotail regions is not trivial and typically involves establishing a series of thresholds on plasma quantities to exclude magnetosheath, PSBL, and lobe intervals. While several classification efforts exist (Ohtani & Mukai, 2008; Raptis et al., 2024; Vo et al., 2023), it is very difficult to define a general set of criteria that simultaneously retains a sufficient number of data points (necessary for ML applications) while maintaining a negligible number of false positives.

To address this sensitivity, we derive two sets of criteria for the Geotail data set and one for MMS, listed in Table 1. The flexible set (e.g., Richard et al., 2022) uses relatively permissive criteria that may admit data outside the plasma sheet, while the strict set (e.g., Ohtani & Mukai, 2008) may remove many points that are valid plasma sheet observations. We complement the Geotail data set with an additional set of approximately 10,000 data points (at 12-s resolution) corresponding to “high-density” intervals. These intervals, while representing atypical values, exhibit elevated densities ($n_i > 6$ [1/cc]) during geomagnetic storms driven by upstream pulses. Figure 1 shows a scatter plot of the training and test set spatial distribution for Geotail throughout the plasma sheet region. The MMS distribution is similar, with the primary difference being the relatively fewer measurements (see Figure 1 of Raptis et al., 2024).

3.2. Linear Regression

Linear regression (LR) serves as our baseline supervised learning model to establish the performance of a purely linear relationship between input and output variables. Coefficients were determined using ordinary least squares, providing a benchmark against which more complex models can be compared.

3.3. Empirical Models

For comparison with previous empirical models, we adopt the Tsyganenko and Mukai (2003) plasma sheet model (hereafter TM03) to provide prescriptions of CPS ion density and temperature. TM03 is based on in situ particle

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Plasma Sheet Classification Thresholds for the Strict (e.g., Ohtani & Mukai, 2008), Flexible PS (e.g., Richard et al., 2022), and High-Density Subsets

Criterion	Strict	Flexible PS	High density
$\beta > 1$	Yes	–	–
$\beta > 0.5$	–	Yes	–
$\sqrt{B_x^2 + B_y^2} < 2\lvert B_z\rvert$	Yes	–	–
$n_i < 6$	Yes	–	–
$n_i > 6$	–	–	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

Note. β is the ion plasma beta parameter, density (n_i) 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 (AGSM) coordinates. EA1SW0 = EA is a Geotail specific criterion describing the exclusive use of EA (Energy Analyzer) in magnetosphere mode for the LEP instrument which provides the optimal data set for magnetospheric studies.

data collected by the Geotail spacecraft at geocentric radial distances of 10–50 R_E and parameterizes the two-dimensional (along- and cross-tail) distributions of ion temperature, density, and pressure as functions of upstream solar wind and interplanetary magnetic field (IMF) conditions. The model employs the Shue et al. (1998) magnetopause model to delimit the valid region of the CPS, thus offering self-consistent estimates of plasma parameters across varying geomagnetic conditions.

For electron temperature, we employ the empirical nightside electron plasma sheet model of Dubyagin et al. (2016) (hereafter DGSR16), which provides electron density and perpendicular temperature (T_e) distributions at radial distances of 6–11 R_E based on ∼400 hr of THEMIS spacecraft measurements during geomagnetic storms. DGSR16 expresses T_e as a sum of second-order polynomials in solar wind proton density, velocity, and IMF B_z components, each averaged over optimally lagged intervals determined by correlation analysis. By pairing these T_e estimates with TM03's ion temperature (T_i) profiles derived from Geotail particle data, we can compute the T_i/T_e ratio in the near-Earth CPS.

3.4. Gradient Boosting

Gradient boosting algorithms are efficient, lightweight, and typically outperform other ML approaches for tabular data (McElfresh et al., 2023). However, as tree-based techniques, they can be limited in their ability to generalize to out-of-sample input spaces, essentially finding the nearest known point. This is largely solvable by implementing an additional LR task as a secondary technique, but it overall limits their applicability for operational use in this application since data set sparsity prevents comprehensive coverage of the input space along with well-sampled spacecraft locations. Nevertheless, GB techniques are particularly well-suited for interpretation, as feature importance methods such as SHAP (SHapley Additive exPlanations) values (Delgado-Panadero et al., 2022; Lundberg & Lee, 2017) are seamlessly implemented. In our work, we use the LightGBM model (Ke et al., 2017) with OMNI data as input to establish a baseline result and supplement our discussion. A sensitivity analysis over 31 configurations, varying tree depth (4–10), leaf count (32–256), learning rate (0.005–0.1), number of estimators (200–1,000), and feature fraction (0.7–1.0), showed that the best configuration improves the Mean Absolute Error (MAE) by less than 1% relative to the adopted setting, and that the spread across all configurations is about 10% of the MAE. This sweep was evaluated on the 2005–2010 test set (Section 3.7); because all reported results use the adopted setting rather than the best configuration, it does not bias them. Performance is therefore insensitive to hyperparameter choice.

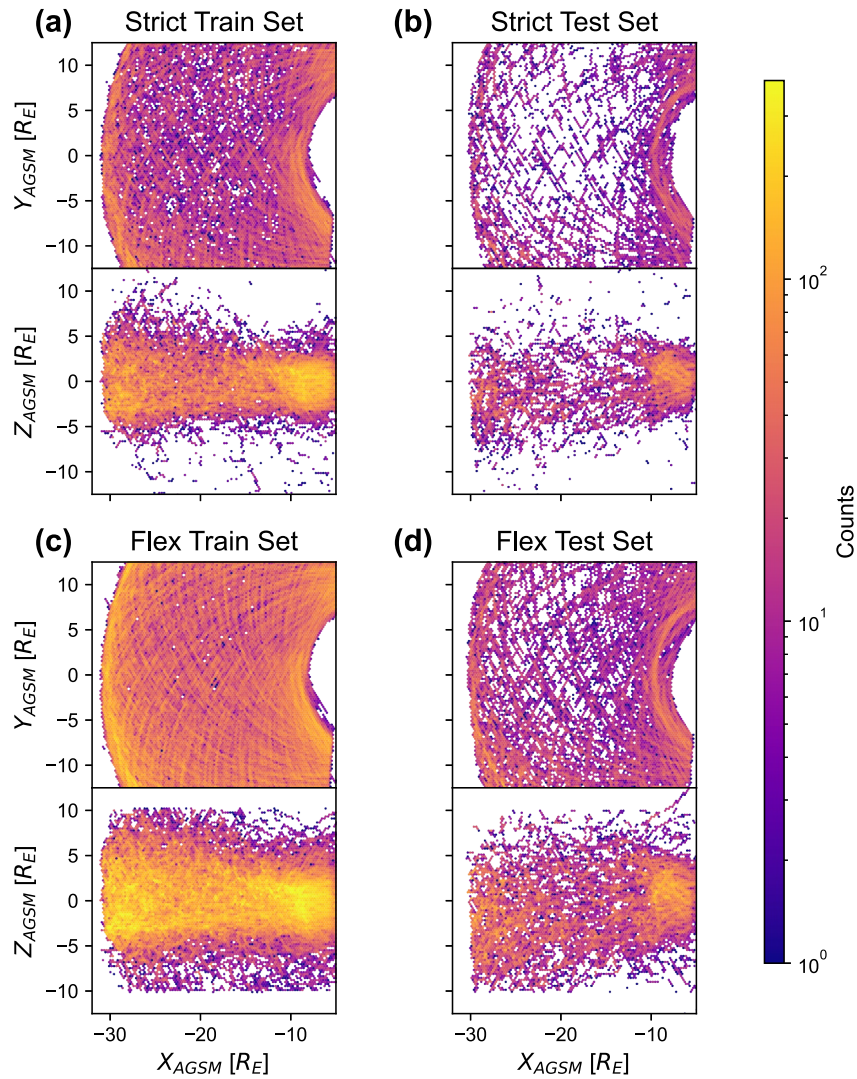


Figure 1. Spatial distribution of Geotail observations projected onto the aberrated Geocentric Solar Magnetospheric X–Y (top panels) and X–Z (bottom panels) planes. The data are split chronologically into training (80%) and testing (20%) subsets. Panels display the density of (a) the strict training set, (b) the strict testing set, (c) the flexible training set, and (d) the flexible testing set. The color scale indicates the logarithmic count of data points per bin.

3.5. Neural Networks

Our NN approach consists of two different models representing two distinct input data sets and complexity levels. First, we use a simple shallow NN with OMNI data as input, then we use PRIME-PS with Wind data input along with time history, while the output remains the same.

3.5.1. Shallow Neural Network

For the shallow NN with OMNI input, we build on the approach of Raptis et al. (2020) and Aminalragia-Giamini et al. (2021). Since the original model was designed for classification, we made several modifications. First, the final layer activation function was changed from sigmoid to linear to accommodate the regression task. Second, the loss function for backpropagation was changed from binary cross-entropy to MAE and Huber loss (Huber, 1973). These choices yielded no statistically significant difference in model performance, so the final model uses Huber loss, which is well-suited for robust regression owing to its reduced sensitivity to outliers. The final architecture consists of three hidden layers (128, 64, and 32 units with ReLU activation functions), each followed by batch normalization, and a linear output layer. A simpler two-layer variant (40 and 30 units, with

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ReLU and batch normalization) and the original architecture of Raptis et al. (2020) both produced statistically indistinguishable results, indicating that the problem does not require deep architectures. The implementation uses the Keras library (Chollet et al., 2015) with PyTorch (Paszke et al., 2019) as the backend, and the Adam optimizer (Kingma & Ba, 2017) with default parameters. Both training and test set performance saturate after approximately 25 epochs, as confirmed by early stopping experiments that monitored test loss with a patience of 20 epochs.

3.5.2. Probabilistic Recurrent Neural Network (PRIME-PS)

For the highest-cost model we adopt a modification of the Probabilistic Regressor for Input to the Magnetosphere Estimation (PRIME) recurrent NN architecture dubbed PRIME-PS (PS denoting plasma sheet). The network is broadly similar to prior implementations (O'Brien et al., 2023, 2024). The PRIME-PS architecture has two advantages. First, a recurrent encoder feeds its output into a fully connected network that performs the propagation, which lets the model use the time history of the solar wind at L1. Second, a specialized output layer produces the mean and standard deviation of a Gaussian distribution for each parameter, rather than a single value. This last advantage is also enabled by using the Continuous Ranked Probability Score (CRPS) as the model's loss metric, a generalization of the MAE to probabilistic predictions.

We have made some modifications to the PRIME architecture for PRIME-PS. Notably, we encode the target spacecraft position separately from the solar wind timeseries using Fourier features (Rahimi & Recht, 2007). This technique maps each coordinate of the target spacecraft position to a higher-dimensional spectral space using the equation

$$\gamma(x) = \left[\sin\left(2^0\pi x\right), \cos\left(2^0\pi x\right), \ldots, \sin\left(2^{L-1}\pi x\right), \cos\left(2^{L-1}\pi x\right)\right] \quad (1)$$

where x is a generic coordinate and L is the number of frequency bands (here we use four). This encoding captures low and high frequency variations in spatial coordinates, traditionally a very difficult task for neural networks. This technique has been shown to improve the accuracy and fidelity of models trained on in situ solar wind data (Martin et al., 2025). The position encoding is concatenated with the history vector from the recurrent section of the NN before being passed to the fully connected network. The final input and output space for all combinations of the methodology can be visualized in Figure 2, while more information regarding their architecture can be found in Supporting Information S1.

We use the Adam algorithm to optimize PRIME-PS's weights (Kingma & Ba, 2017) with a learning rate of 10^{-5}, as our testing confirmed that a constant learning rate resulted in the most stable and fastest convergence. A more complex training regimen is likely unnecessary given PRIME-PS's relatively small size (311,000 parameters). Models trained on Geotail data required 200 training epochs to converge without overfitting, while models trained on MMS data required 500 training epochs. We pick the input timeseries size of PRIME-PS to be 5.5 hr (200 entries at 100 s cadence) with no lead time ("stride"). This was chosen because the magnetotail's dynamics have been shown to depend on the prior 4–5 hr of the solar wind time history (Borovsky et al., 1998; Denton & Borovsky, 2009; Wing et al., 2006). We add 30 min on top of the high end of that range to account for the solar wind travel time from L1 to Earth. Model hyperparameters (layer number and size, dropout rate) were chosen by grid search (1–4 layers, 32–320 nodes per layer in 32-node steps, 0%/20%/50% dropout rate). Each configuration was trained on the Geotail data set and scored on the 2005–2010 test set (Section 3.7). We then kept the hyperparameters fixed for all other data sets and train/test splits, to isolate the effects of changing the data set composition.

3.6. Evaluation Metrics

We employ four metrics to assess model performance. The first three, MAE, Coefficient of Determination $\left(R^2\right)$, and Pearson Correlation Coefficient (r), are used for all methods; for PRIME-PS, these metrics are computed using the predicted mean (μ) of the output Gaussian distribution. Unless otherwise noted, all PRIME-PS predictions shown in this work correspond to μ, while σ defines the uncertainty bands in Figures 10 and 11. The fourth, CRPS, is applied exclusively to the probabilistic PRIME-PS model.

The Mean Absolute Error measures the average magnitude of prediction errors:

10.1029/2026JH001589

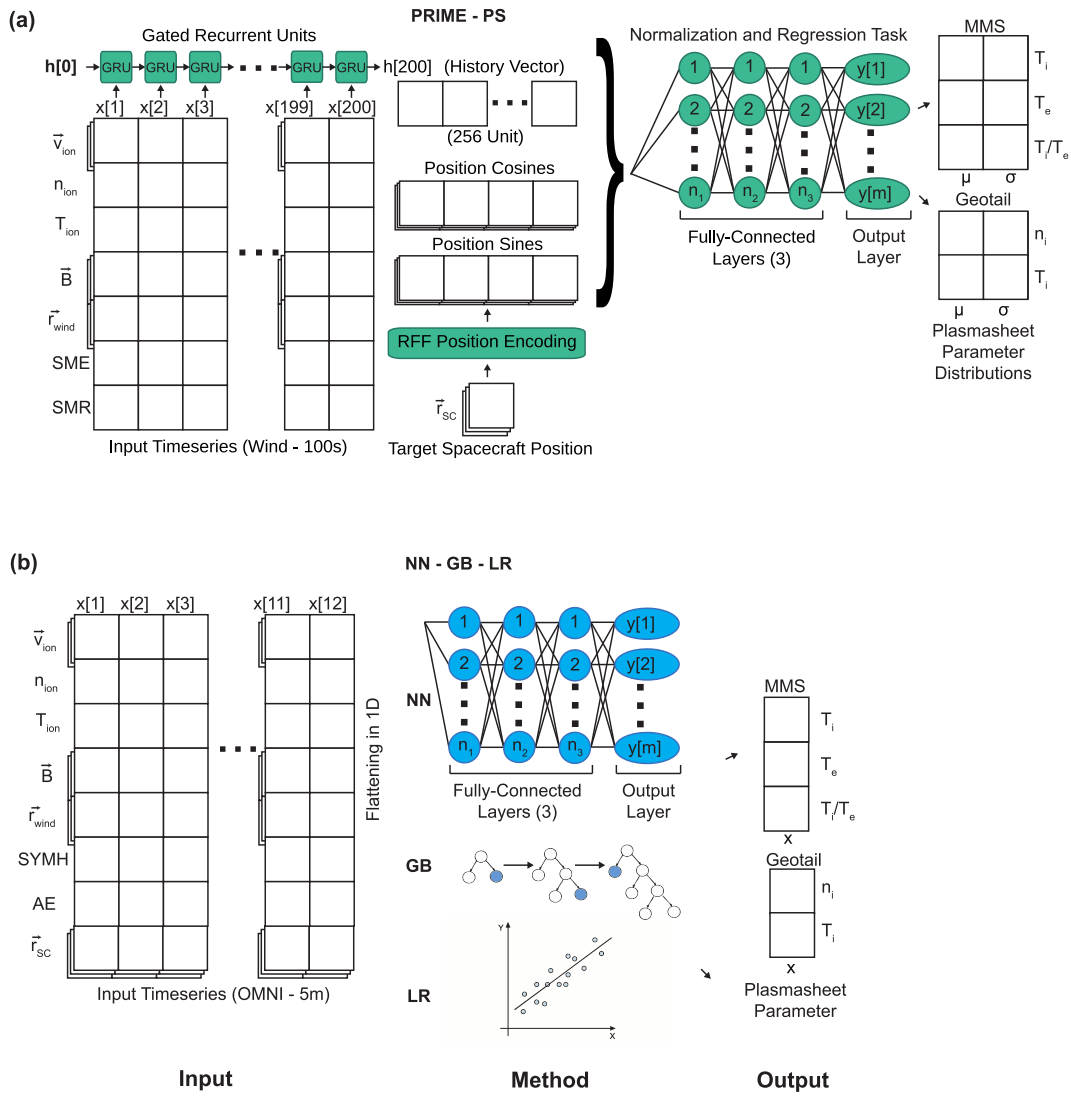


Figure 2. Schematic of the methodology used for the machine learning methods presented in this work. Panel (a) describes PRIME-PS methodology, highlighting the Wind spacecraft input, the presence of recurrent units, and the output producing a distribution rather than a unique value. Panel (b) describes the methodology taken for the shallow neural network, the gradient boosting, and linear regression in which the input is taken from OMNI database in 5 min cadence, and flattened into 1D before being introduced to the regression algorithm. In both panels, $\vec{r}_{SC}$ denotes the target spacecraft position in the plasma sheet, while $\vec{r}_{wind}$ denotes the location of the upstream solar wind monitor near L1, included as an input.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right|,$$

where y_i is the observed value, $\hat{y}_i$ is the predicted value, and N is the number of samples.

The Coefficient of Determination quantifies the proportion of variance explained by the model:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2},$$

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10.1029/2026JH001589

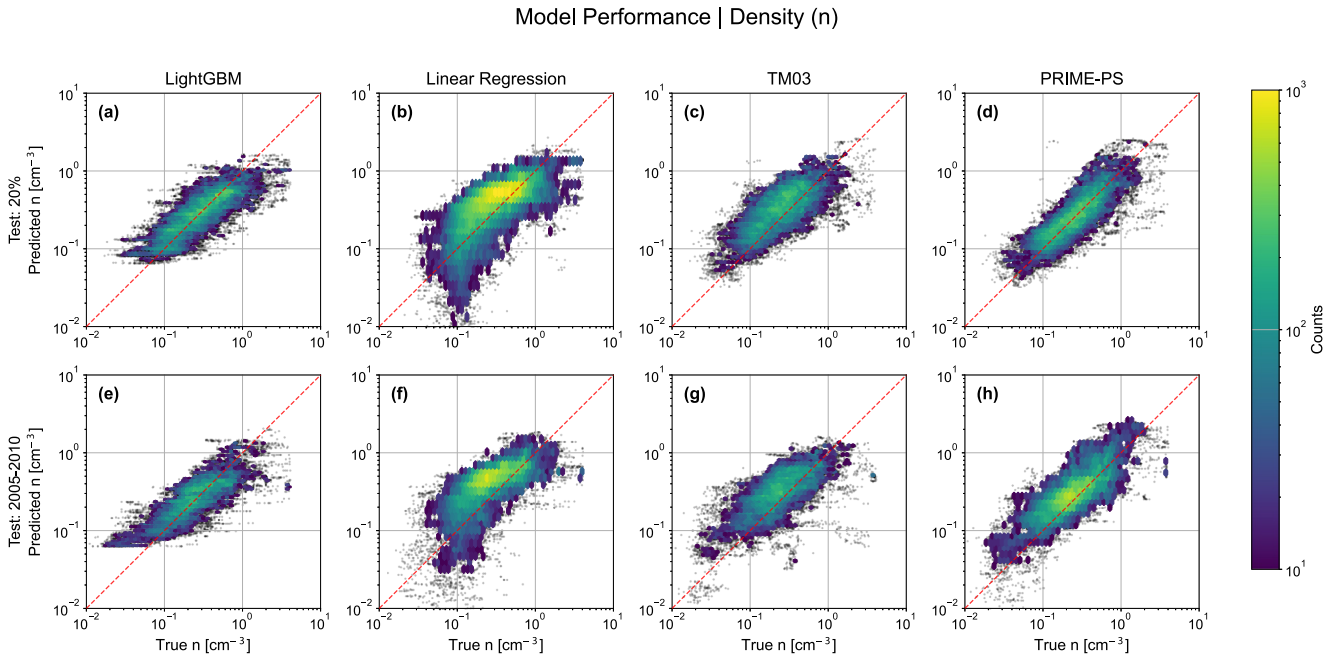


Figure 3. Comparison of four different models for predicting plasma number density (n) using Geotail in log-log scale. Each panel shows the predicted density versus the true observed density for the "strict" data set. The columns correspond to the four models: LightGBM, Linear Regression, TM03, and PRIME-PS. The rows represent two different training and testing methodologies: a chronological split where the final 20% of the data serves as the test set (top row, (a–d)), and a fixed-date split where the period 2005–2010 is used for testing (bottom row, (e–h)). The color of the hexagonal bins indicates the count of data points, with sparse outliers (counts <10) shown as individual points. The dashed red line represents a perfect one-to-one correspondence.

where $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ is the mean of observed values. An $R^2 = 0$ indicates a model with performance identical to the average value of the output, while $R^2 < 0$ describes a model with predictability worse than the mean value, and an $R^2 = 1$ a perfect one.

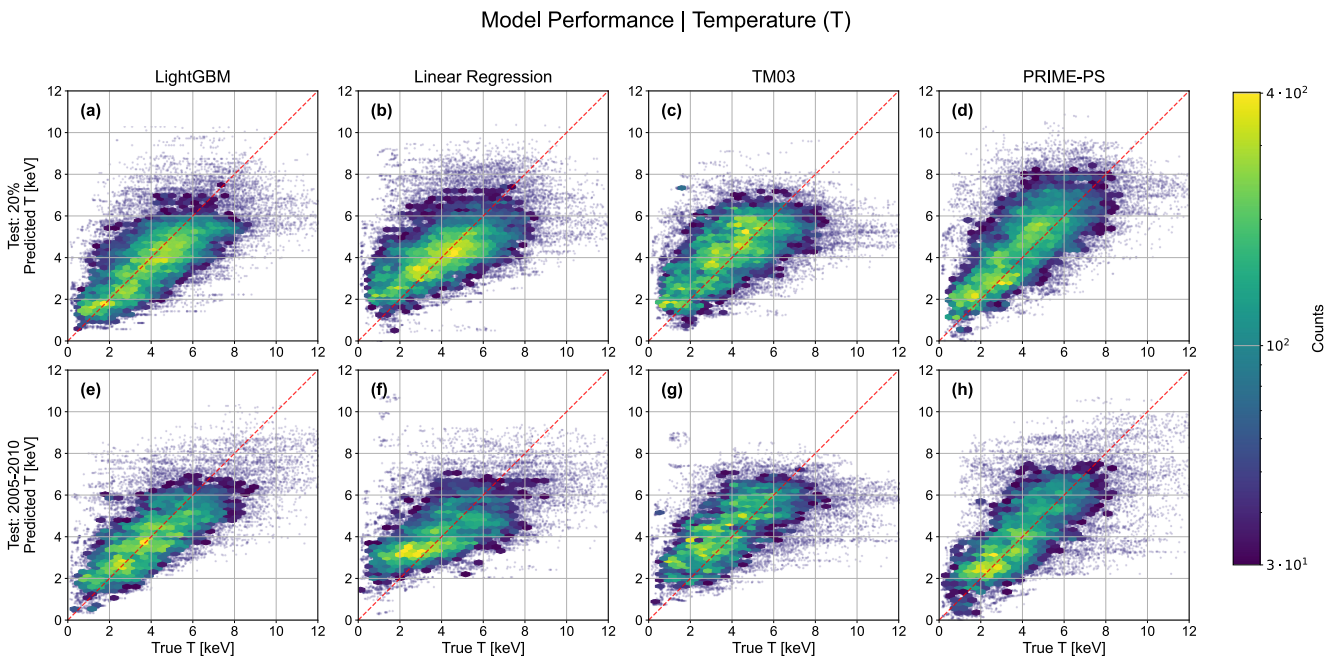


Figure 4. Comparison of four different models for predicting ion temperature (T_i) using Geotail. Each panel shows the predicted temperature versus the true observed temperature for the "strict" data set. The format of the plots is identical to Figure 3.

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2
Performance Metrics for Each Method When Evaluating Plasma Density (n_i) in Units of [cm^{-3}]

	Strict				Flexible			
Method	MAE	R^2	r	CRPS	MAE	R^2	r	CRPS
Testing: Last 20% of data								
LightGBM	0.129(2)	0.373(9)	0.631(6)	–	0.330(14)	0.002(2)	**0.072(5)**	–
Neural Net	0.152(2)	0.325(10)	0.603(8)	–	0.353(12)	0.000(1)	0.055(4)	–
Linear Reg	0.173(2)	0.265(12)	0.620(9)	–	0.367(12)	0.001(1)	0.066(5)	–
PRIME-PS	**0.109(2)**	**0.473(18)**	**0.709(11)**	0.081(2)	**0.309(16)**	**0.010(1)**	0.071(6)	0.281(16)
TM03	0.163(2)	0.208(18)	0.570(12)	–	0.339(2)	0.000(22)	0.055(21)	–
Testing: Fixed dates 2005–2010								
LightGBM	0.151(2)	0.251(20)	0.627(16)	–	0.185(3)	0.032(15)	0.176(22)	–
Neural Net	0.169(2)	0.175(22)	0.518(17)	–	0.175(3)	0.115(16)	0.374(18)	–
Linear Reg	0.197(2)	0.162(13)	0.543(13)	–	0.256(5)	−0.051(52)	0.318(23)	–
PRIME-PS	**0.113(2)**	**0.465(25)**	**0.689(19)**	0.083(2)	**0.119(3)**	0.148(73)	**0.441(98)**	0.089(3)
TM03	0.160(2)	0.234(17)	0.523(17)	–	0.167(2)	**0.159(17)**	0.425(19)	–

Note. Values in parentheses denote the half-width of the 95% bootstrap confidence interval (1,000 resamples) applied to the preceding digit(s); for example, 0.109(2) denotes 0.109 ± 0.002. These intervals quantify test-sample variability only; a cross-validation scheme that reshuffles train/test splits would yield wider intervals. Bold values indicate the best performance for each metric within each testing data set. Lower MAE and higher R^2 and r indicate better performance.

The Pearson Correlation Coefficient measures the linear relationship between predictions and observations:

$$r = \frac{\sum_{i=1}^{N}(y_i - \bar{y})(\hat{y}_i - \overline{\hat{y}})}{\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{N}(\hat{y}_i - \overline{\hat{y}})^2}},$$

where $\overline{\hat{y}} = \frac{1}{N}\sum_{i=1}^{N}\hat{y}_i$ is the mean of predicted values.

Finally, CRPS quantifies the discrepancy between a predictive cumulative distribution function F and the cumulative distribution of the observation (Heaviside step function H at the observation y_{obs}):

$$\text{CRPS}(F, y_{\text{obs}}) = \int_{-\infty}^{\infty}\left[F(z) - H(z - y_{\text{obs}})\right]^2 \, dz.$$

For a Gaussian forecast with mean μ and standard deviation σ, an analytic expression is given by (Camporeale & Care, 2021; Gneiting et al., 2005; O'Brien et al., 2024):

$$\text{CRPS}(y_{\text{obs}}, \mu, \sigma) = \sigma\left[\frac{y_{\text{obs}} - \mu}{\sigma}\cdot \text{erf}\left(\frac{y_{\text{obs}} - \mu}{\sqrt{2}\,\sigma}\right) + \sqrt{\frac{2}{\pi}}\exp\left(-\frac{(y_{\text{obs}} - \mu)^2}{2\sigma^2}\right) - \frac{1}{\sqrt{\pi}}\right].$$

We use CRPS as the training loss and evaluation metric for the PRIME-PS model because it symmetrically penalizes both over- and under-confident probabilistic forecasts in the same units as the target variable.

To quantify the statistical uncertainty of these metrics, we employ bootstrap resampling of the test set. For each evaluation, we draw N observation–prediction pairs with replacement from the N available test samples, compute all metrics on the resampled set, and repeat the procedure 1,000 times. The 2.5th and 97.5th percentiles of the resulting distribution define the 95% confidence interval reported in Tables 2 and 3. We note that this procedure captures uncertainty from the finite test sample but not from training variability; a cross-validation approach that also reshuffles the training set would produce wider intervals.

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Savvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 3
Performance Metrics for Each Method When Evaluating Ion Temperature (T_i) in Units of [keV]

	Strict				Flexible			
Method	MAE	R^2	r	CRPS	MAE	R^2	r	CRPS
Testing: Last 20% of data								
LightGBM	1.21(1)	0.343(7)	0.667(4)	–	1.22(1)	0.366(5)	0.670(5)	–
Neural Net	1.25(1)	0.341(8)	0.635(5)	–	1.29(1)	0.367(6)	0.644(4)	–
Linear Reg	1.26(1)	0.344(9)	0.609(6)	–	1.29(2)	0.377(5)	0.635(4)	–
PRIME-PS	**1.19(1)**	**0.346(15)**	**0.704(6)**	0.833(7)	**1.11(1)**	**0.483(7)**	**0.739(3)**	0.779(4)
TM03	1.36(1)	0.271(9)	0.556(6)	–	1.37(1)	0.279(7)	0.576(4)	–
Testing: Fixed dates 2005–2010								
LightGBM	1.34(1)	0.394(8)	0.672(6)	–	1.35(2)	0.411(4)	0.685(4)	–
Neural Net	1.34(1)	0.395(10)	0.655(7)	–	1.34(2)	0.402(4)	0.668(5)	–
Linear Reg	1.36(1)	0.380(11)	0.619(8)	–	1.35(3)	0.425(3)	0.655(4)	–
PRIME-PS	**1.12(1)**	**0.555(8)**	**0.756(5)**	0.802(9)	**1.09(1)**	**0.586(5)**	**0.773(3)**	0.787(5)
TM03	1.43(1)	0.308(10)	0.563(8)	–	1.44(1)	0.340(7)	0.591(5)	–

Note. Values in parentheses denote the half-width of the 95% bootstrap confidence interval (1,000 resamples) applied to the preceding digit(s); for example, 1.19(1) denotes 1.19 ± 0.01. These intervals quantify test-sample variability only; reshuffling the train/test partition would yield wider intervals. Bold values indicate the best performance for each metric within each testing data set. Lower MAE and higher R^2 and r indicate better performance.

3.7. Testing Data Sets

To evaluate our Geotail-derived models and assess their generalization to unseen data, we perform three different tests. First, we use a standard holdout approach with the final 20% of data chronologically as the test set, yielding approximately 55,000 data points with sufficient spatial distribution (Figure 1). We select the final 20% chronologically rather than randomly to avoid contamination and data leakage. Second, we reserve a separate 2005–2010 test set (approximately 35,000 points), excluded from training. It serves two purposes: it tests whether performance depends on the epoch of the data, and, by being from 2005 onward, it is also unseen by the empirical baseline TM03, which was trained only on pre-2005 data. This lets us compare our model and TM03 on common, out-of-sample ground. Third, we test on observations from the THEMIS spacecraft, which provides truly independent data from different instrumentation during a geomagnetic storm (discussed below). For MMS-related modeling, we use only the 20% holdout test set and the THEMIS interval from the geomagnetic storm event discussed below. We note that the 20% holdout set was used to confirm training convergence (i.e., to verify that both training and test loss had saturated), which introduces a mild selection effect. The 2005–2010 block was not used for training or convergence monitoring, but was used to select the PRIME-PS hyperparameters. The THEMIS storm interval was not used at any stage and therefore provides a strictly independent evaluation.

4. Results

4.1. Plasma Sheet Density and Ion Temperature Using Geotail

Figure 3 presents scatter plots of predicted values versus ground truth for the different methodological combinations using the "strict" data set in log-log scale. A linear-scale version is provided in Figure S1 in Supporting Information S1. As quantified in Table 2, the methods exhibit comparable performance on the "strict" data set, whereas the "flexible" data set yields significantly lower accuracy. This discrepancy highlights the difficulty of distinguishing valid plasma sheet intervals from adjacent regions, such as the magnetosheath and boundary layers. Including more marginal intervals increases the statistical sample, but it can also introduce classification false positives, thereby degrading algorithmic performance.

The discrepancy between the two test sets is also informative. The flexible data set degrades density performance severely on the chronological test (MAE increases from 0.109 to 0.309 cm^{-3}) but only modestly on the 2005–2010 test set (0.113–0.119 cm^{-3}), although the ratio of flexible to strict points is similar in both (~2.8). The

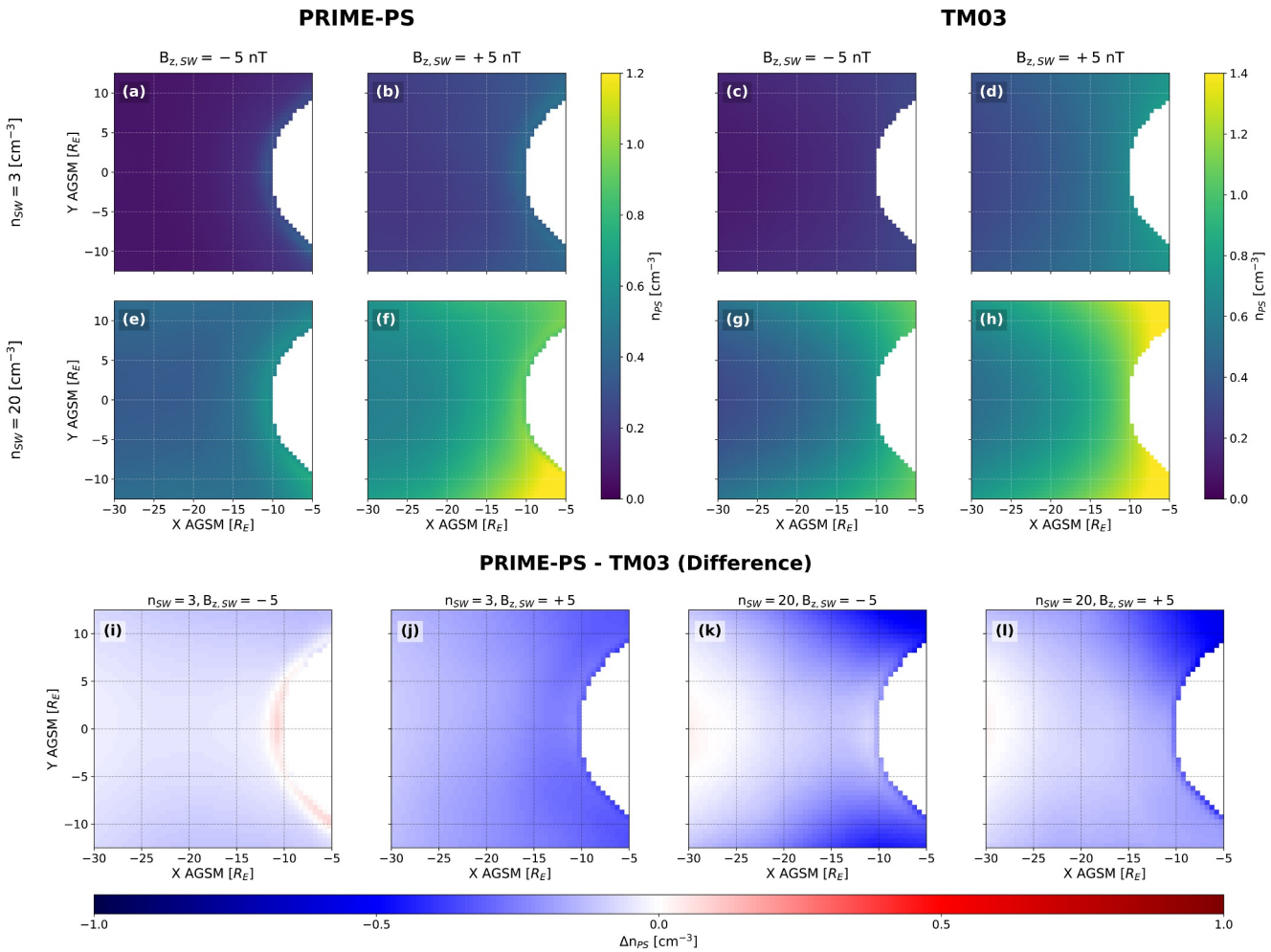


Figure 5. Comparison of equatorial plasma sheet density (n_{PS}) predicted by the PRIME-PS and TM03 models in the X-Y aberrated Geocentric Solar Magnetospheric plane. Panels (a)–(h) display density distributions for low ($n_{SW} = 3$ cm^{-3}) and high ($n_{SW} = 20$ cm^{-3}) solar wind densities under southward ($B_{z,SW} = -5$ nT) and northward ($B_{z,SW} = +5$ nT) interplanetary magnetic field conditions. The left group represents PRIME-PS, while the right group represents TM03. The bottom panels (i)–(l) show the density difference (Δn_{PS}), calculated as PRIME-PS minus TM03, highlighting spatial discrepancies and asymmetries between the two models under identical solar wind driving conditions. The exact model-driving parameters, excluding the variables varied, are provided in Supporting Information S1. All solar wind input parameters are held constant over the full 5.5-hr look-back window, representing a steady-state driving scenario.

difference is therefore not driven by sample size. A likely cause is that the flexible criteria admit more magnetosheath and boundary layer points in the chronological test (approximately 2016–2022) than in the 2005–2010 test set; because such points differ strongly in density from the CPS, even a small number of them can dominate the error, consistent with the near-zero R^2 and r of all methods, including TM03, on this subset. Changes in the Geotail orbit and the aging of its instrumentation in the final years of the mission may also have increased the fraction of misclassified or lower-quality measurements. The performance of the flexible classification is therefore not uniform in time and depends on the sampling and data quality of the test interval. The residual distributions support this interpretation. On the strict test sets, density residuals are negatively skewed (skewness -3 to -6) and heavy-tailed (excess kurtosis 25–80) for all methods, including TM03, indicating that high densities are systematically underestimated regardless of the algorithm. On the flexible sets, the excess kurtosis of the density residuals rises to several hundred, showing that a small number of very large errors dominates the metrics. Temperature residuals, by contrast, remain closer to Gaussian (|skewness| $\leq$ 1, excess kurtosis 2–5) in all cases.

Similarly, Table 3 and Figure 4 summarize the performance of these methods in modeling ion temperature within the plasma sheet. In contrast to the density results, performance is relatively uniform across both data sets and

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Savvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10.1029/2026JH001589

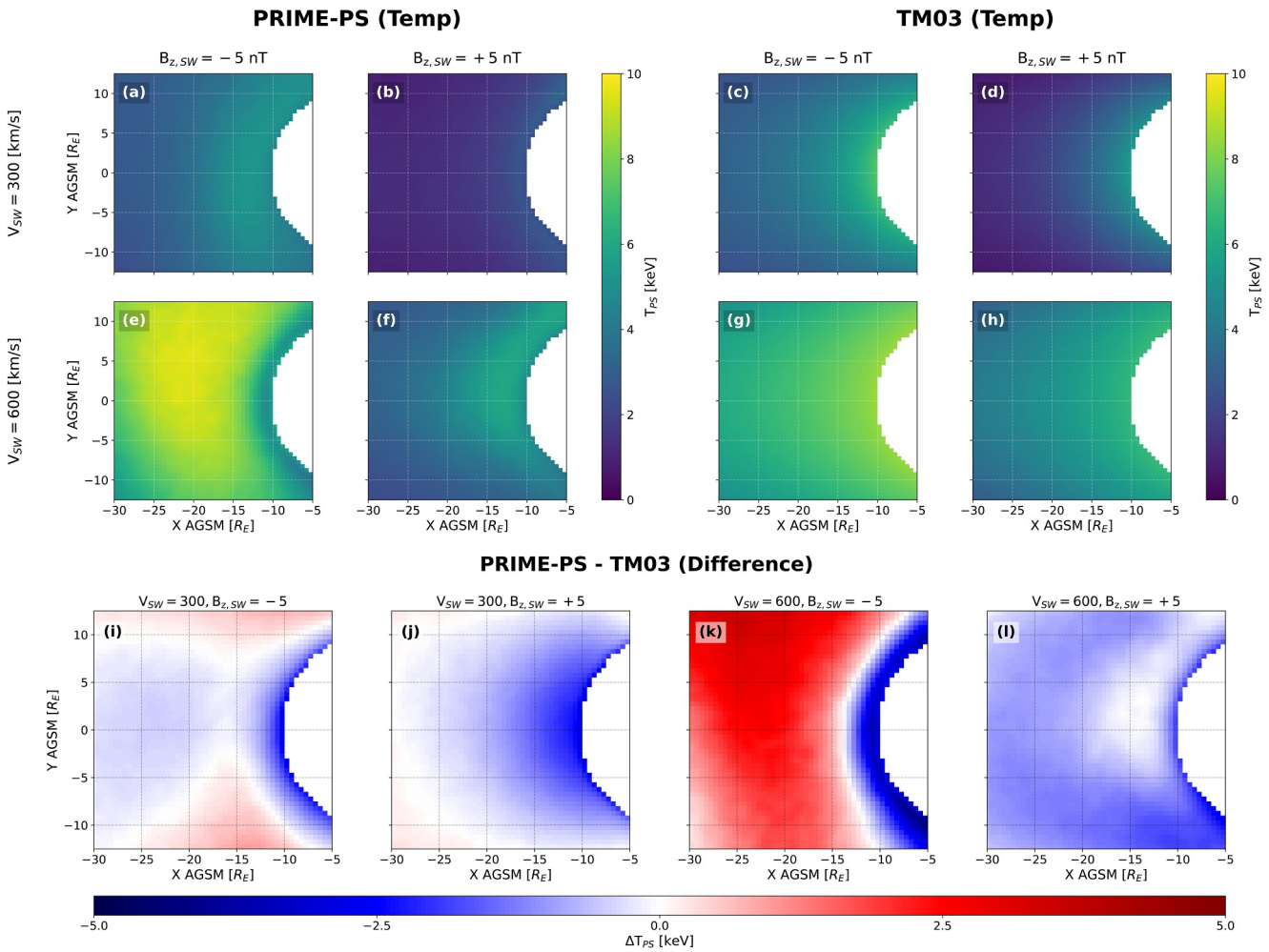


Figure 6. Comparison of equatorial plasma sheet ion temperature (T_{PS}) predicted by PRIME-PS (trained on Geotail data) and TM03 in the X-Y aberrated Geocentric Solar Magnetospheric plane. Panels (a)–(h) display temperature distributions for low (V_{SW} = 300 km/s) and high (V_{SW} = 600 km/s) solar wind speeds under southward ($B_{z,SW}$ = −5 nT) and northward ($B_{z,SW}$ = +5 nT) interplanetary magnetic field conditions. The left group represents PRIME-PS, while the right group represents TM03. The bottom panels (i)–(l) show the temperature difference (ΔT_{PS}), calculated as PRIME-PS minus TM03, highlighting spatial discrepancies and asymmetries between the two models under identical solar wind driving conditions. The exact model-driving parameters, excluding the variables varied, are provided in Supporting Information S1. All solar wind input parameters are held constant over the full 5.5-hr look-back window, representing a steady-state driving scenario.

models, with limited variance between methods. The notable exception is TM03, which clearly underperforms the ML methods, likely because its fixed functional form cannot represent the dawn-dusk asymmetries and nonlinear driver dependences discussed below.

Based on the performance metrics described above, we proceed primarily using PRIME-PS, which demonstrated marginally superior results. We have no theoretical basis to expect the other models to behave significantly differently in subsequent applications. While GB methods perform well with simpler inputs and minimal training, their limited ability to extrapolate beyond the training domain restricts their utility to feature analysis rather than predictive modeling in this context. Our analysis suggests that model complexity and input data set selection (e.g., choosing between Wind and OMNI data) play a limited role in statistical performance. The comparable results across different non-linear regression algorithms (GB and NN architectures) indicate that the problem is under-determined; a scarcity of observations or uncaptured complementary physical processes likely creates a performance bottleneck.

Consequently, we utilize the PRIME-PS model trained on the "strict" data set to generate a full map of plasma sheet quantities by projecting onto a grid of artificial satellite locations. Figure 5 compares PRIME-PS against TM03 under synthetic solar wind driving conditions. The NN generates an asymmetric density profile along the

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

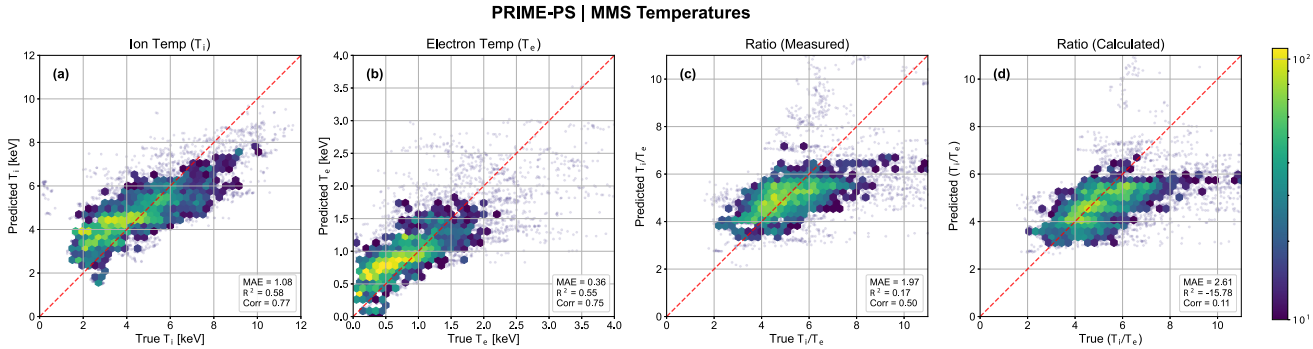


Figure 7. Predicted temperatures and temperature ratio versus the true observed variable for the Magnetospheric Multiscale data set using PRIME-PS. A chronological split where the final 20% of the data serves as the test set is used. Panel (a) corresponds to the ion temperature in keV, (b) shows the electron temperature in keV, (c) shows the directly predicted ion-to-electron temperature ratio, and (d) shows the ratio calculated from independently predicted T_i and T_e. Metrics for each quantity are given in the legend of each panel. The color of the hexagonal bins indicates the count of data points, with sparse outliers (counts <10) shown as individual points. The dashed red line represents a perfect one-to-one correspondence.

magnetotail, a result consistent with physical expectations and prior observations by other spacecraft (Wing & Newell, 1998; C.-P. Wang et al., 2007). It also captures the expected earthward increase in density and the additional density enhancement under northward IMF $(B_z > 0)$. The dawn-favoring density enhancement is evident for $B_z > 0$ but largely absent for $B_z < 0$. Under northward IMF, solar wind plasma enters the magnetosphere primarily along the dawn and dusk flanks through mechanisms such as Kelvin-Helmholtz instabilities and double-lobe reconnection (Sorathia et al., 2019; Wing & Newell, 1998; Wing et al., 2005). Entry appears more efficient on the dawn flank, and the duskward magnetic drift of hot ions leaves the cold, dense population concentrated on the dawnside, producing a persistent dawn-dusk density asymmetry (Wing et al., 2005; C.-P. Wang et al., 2007). That PRIME-PS recovers this pattern without any imposed spatial prior indicates that the model has learned the statistical imprint of these entry processes from the data. The asymmetry also depends on geomagnetic activity. Figure S3 in Supporting Information S1 shows the normalized dusk-dawn density difference, for $n_{SW} = 20\ \mathrm{cm}^{-3}$ and SMR of 0, −50, −100, and −200 nT, with all other inputs fixed at the values of Table S1 in Supporting Information S1. In both IMF orientations, the dawnside enhancement is clearly weaker at −200 nT than under quiet conditions, and the tailward plasma sheet ($X < -15\ R_E$) becomes nearly symmetric. The change is clearest under northward IMF, where most of the reduction already occurs between 0 and −50 nT, although the intermediate levels (−50 and −100 nT) do not order monotonically. This trend is consistent with Dubyagin et al. (2016), who found that the dawn-dusk asymmetry of the electron plasma sheet disappears for Dst <− 50 nT. However, SMR is physically correlated with other inputs, such as auroral indices and the solar wind driving, so varying it alone probes input combinations that are rare in the training data. We therefore regard Figure S3 in Supporting Information S1 as a qualitative sensitivity test; a model designed to account for input intercorrelations would allow a more rigorous attribution, which we leave to future work.

Figure 6 shows how the ion temperature differs between our ML approach and TM03. The asymmetric PRIME-PS pattern reflects the expected duskward enhancement of ion temperature (C. Wang et al., 2012; A. P. Walsh et al., 2014). This asymmetry can arise because ions undergoing gradient-curvature drift in the cross-tail magnetic field are preferentially transported toward dusk, gaining energy from the dawn-to-dusk convection electric field, whereas electrons drift dawnward (Wing et al., 2005). The model captures this drift signature without explicit physical encoding; Section 4.2 extends the comparison to both species using MMS. One caveat of the ML approach is that near-Earth observations under rare synthetic conditions (Figure 6, panel e) produce unexpected behavior, with ion temperature dropping significantly near the transition boundary from plasma sheet to inner magnetosphere. The same architecture trained with MMS observations does not exhibit this issue (Figure 8, discussed below), suggesting that spacecraft orbits and the resulting spatial coverage they produce can affect ML-driven modeling efforts.

In the PRIME-PS model, the temperature and density profiles for high solar wind density suggest that the plasma sheet density is higher and temperature lower on the dawnside near the inner boundary of the plasma sheet (Figures 5f and 6f). This feature is likely related to a combination of solar wind plasma entry along the dawn flank and a stagnation point where $\mathbf{E} \times \mathbf{B}$ drift and corotation are approximately balanced by curvature and gradient

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10.1029/2026JH001589

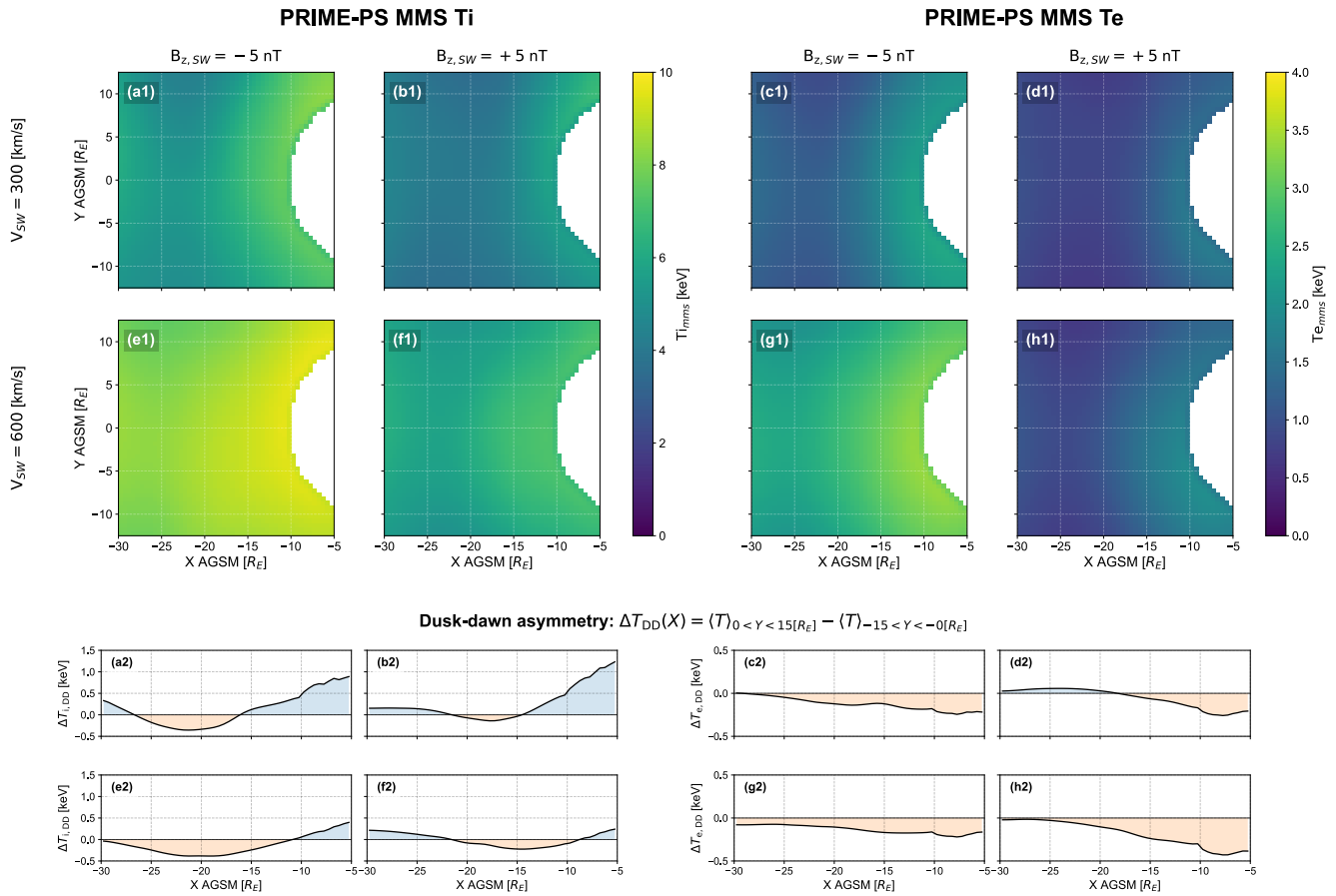


Figure 8. Equatorial plasma sheet ion (T_i) and electron (T_e) temperatures predicted by the PRIME-PS model in the X-Y aberrated Geocentric Solar Magnetospheric plane. Panels (a1)–(h1) display ion and electron temperature distributions for low (V_{SW} = 300 km/s) and high (V_{SW} = 600 km/s) solar wind speeds under southward ($B_{z,SW}$ = −5 nT) and northward ($B_{z,SW}$ = +5 nT) interplanetary magnetic field conditions. The left group maps ion temperatures, while the right group maps electron temperatures. The bottom panels (a2)–(h2) show the dusk-dawn temperature asymmetry (ΔT_{DD}), highlighting spatial variations across the dawn and dusk flanks under identical solar wind driving. As expected from charge-dependent magnetic drifts, ions are hotter on the dusk side and electrons on the dawn side. The exact model-driving parameters, excluding the variables varied, are provided in Supporting Information S1. All solar wind input parameters are held constant over the full 5.5-hr look-back window, representing a steady-state driving scenario.

drifts (Friedel et al., 2001). During active times, the inner boundary of the plasma sheet and the stagnation point may get pushed further inward to $r < 10\ R_E$ or even to geosynchronous orbit (Denton et al., 2005; Korth et al., 1999; Wing & Johnson, 2009).

4.2. Plasma Sheet Ion-to-Electron Temperature Ratio Using MMS

Next, we use MMS to quantify a more complex variable: the ratio between ion and electron temperature (T_i/T_e), and its components, ion and electron temperatures. T_i/T_e as a quantity is highly useful for informing global simulations (see, e.g., Sadeghzadeh et al., 2021), particularly when initializing particle distributions near the inner magnetosphere, where typical assumptions (e.g., magnetohydrodynamics, MHD) are no longer valid (Merkin et al., 2019; Sorathia et al., 2020). The ratio T_i/T_e is expected to vary significantly throughout the magnetotail, driven by large-scale convection patterns (Shustov et al., 2021) as well as the localized effects of transient processes, such as high-speed flows (Runov et al., 2015).

Consistent with our previous investigation (i.e., density and ion temperature using Geotail), we evaluated each methodology and determined that PRIME-PS performs slightly better. Therefore, for the sake of brevity, our comparison focuses on PRIME-PS, empirical baselines, and instrumental effects. In particular, we derived partial moments (Section 2), which differ at a statistically significant level from the standard L2 MMS moments. Full moments give an average plasma sheet T_i/T_e near 5.5, while partial moments give values of about 4, a difference

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

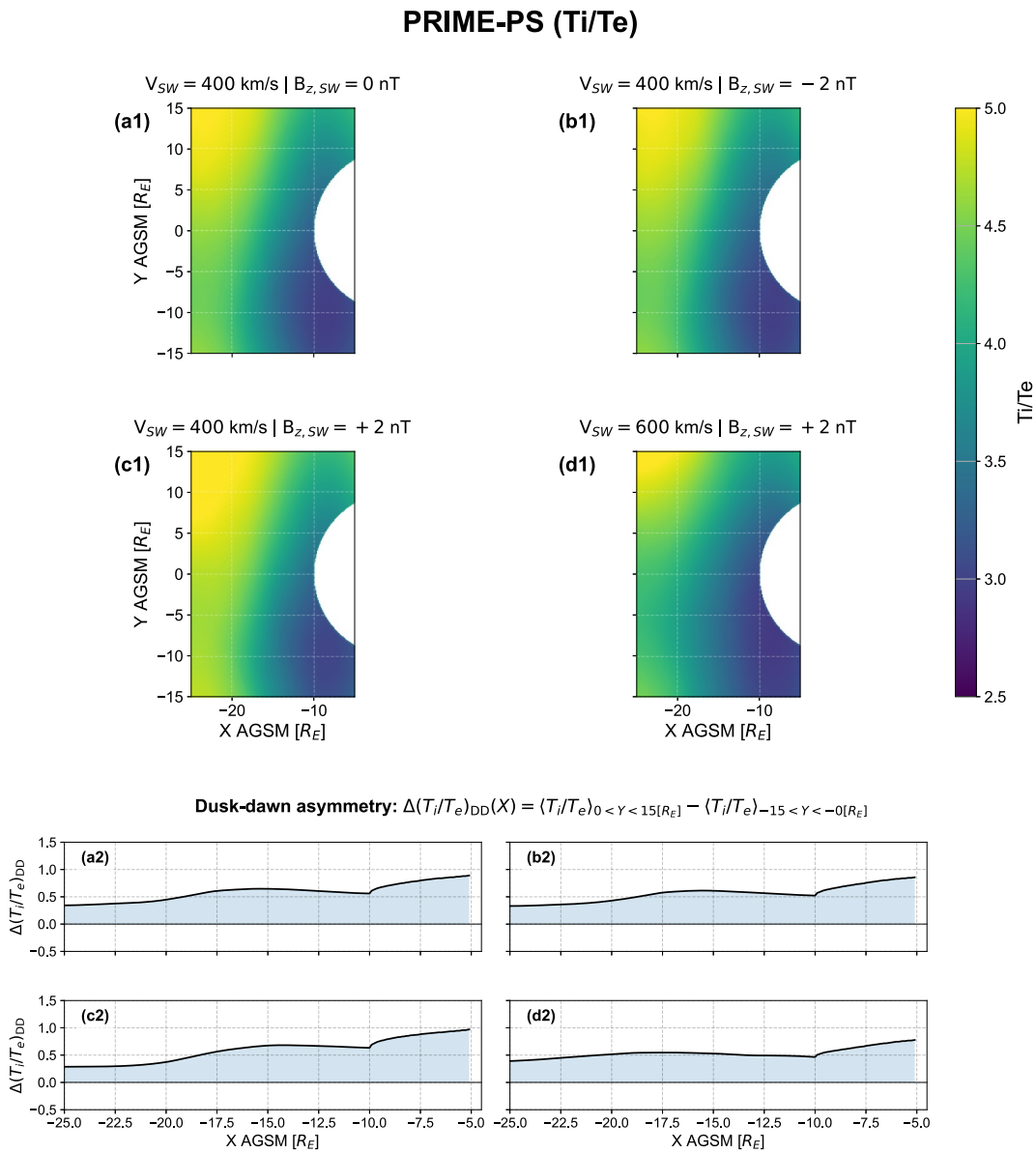


Figure 9. Equatorial plasma sheet ion-to-electron temperature ratio (T_i/T_e) predicted by the Magnetospheric Multiscale PRIME-PS model in the X-Y aberrated Geocentric Solar Magnetospheric plane. Panels (a1)–(d1) show the modeled T_i/T_e distributions for selected solar wind driving conditions: V_{SW} = 400 km/s with $B_{z,SW}$ = 0, −2, and +2 nT, and V_{SW} = 600 km/s with $B_{z,SW}$ = +2 nT. Panels (a2)–(d2) show the corresponding dusk-dawn asymmetry profiles, $\Delta(T_i/T_e)_{DD}$, computed as the difference between the dusk-side and dawn-side averages at each distance. The exact model-driving parameters, excluding the variables varied, are provided in Supporting Information S1. All solar wind input parameters are held constant over the full 5.5-hr look-back window, representing a steady-state driving scenario.

of roughly 30%. These are not two definitions of the same quantity. The partial moments exclude contaminating populations that bias the full moments, and are therefore the more accurate estimate. For electrons, these are low-energy spacecraft photoelectrons; for ions, they are background counts from penetrating radiation (Gershman et al., 2019). The same reasoning constrains our cross-mission validation: missions such as THEMIS carry their own instrumental systematics, distinct from those of MMS, which a simple offset cross-calibration cannot fully remove.

Figure 7 summarizes the performance of PRIME-PS in modeling ion temperature, electron temperature, and ion-to-electron temperature ratios within the plasma sheet. Compared to Table 3, the performance of ion temperature

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

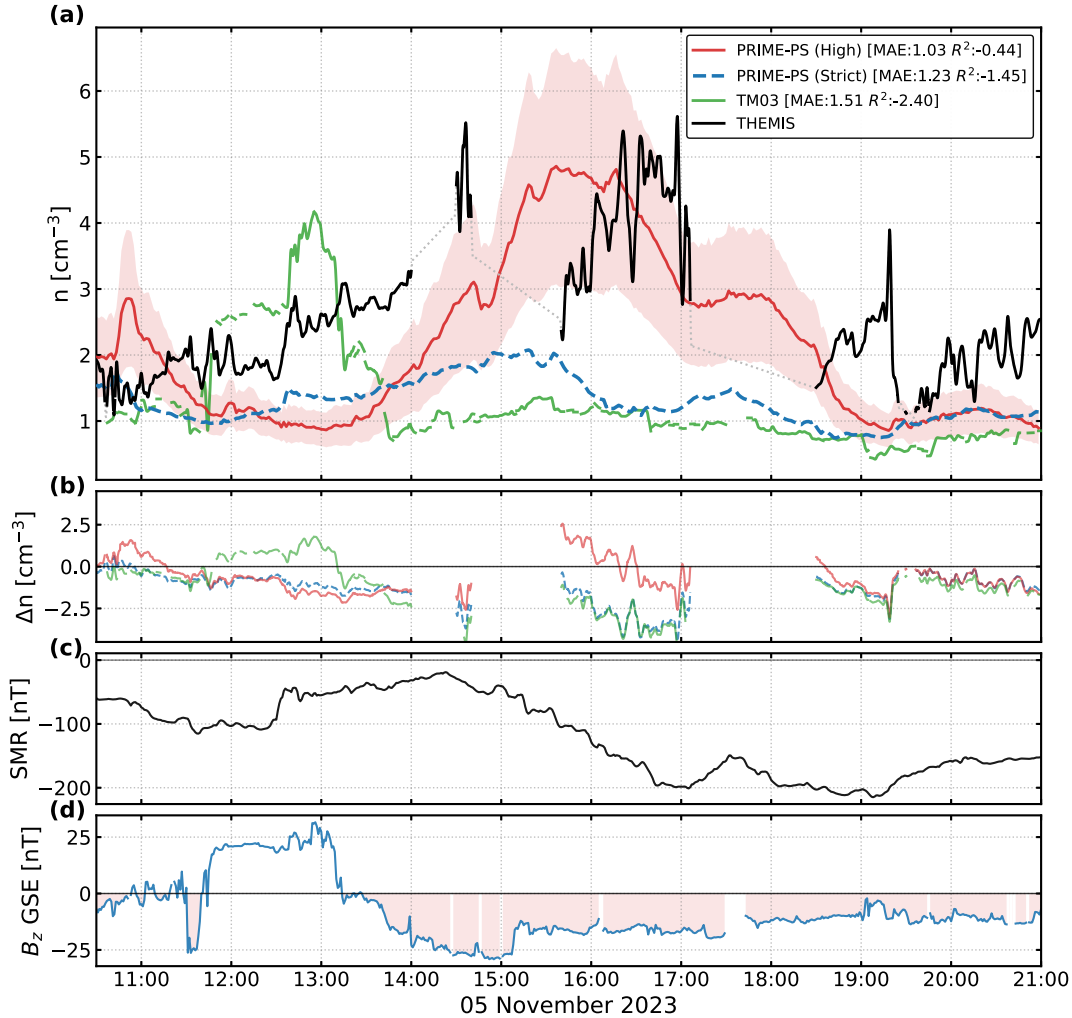


Figure 10. Time series comparison of plasma sheet ion density (n) predictions versus Time History of Events and Macroscale Interactions During Substorms (THEMIS) observations during the geomagnetic storm event on 05 November 2023. Panel (a) displays the observed density from THEMIS (black solid line for measured data; light grey dotted line for interpolated segments) alongside predictions from the TM03 empirical model (green solid line), PRIME-PS (High) (red solid line), and PRIME-PS (Strict) (blue dashed line) machine learning models. The red shaded region indicates the uncertainty bounds ($\pm 1\sigma$) for the PRIME-PS (High) model. Performance metrics, including Mean Absolute Error and R^2, are provided in the legend for each model. Panel (b) shows the density residuals (Δn), calculated as the model prediction minus the observed THEMIS value, highlighting temporal variations in model overestimation ($\Delta n > 0$) and underestimation ($\Delta n < 0$). Panel (c) shows the SMR index as an indicator of storm-time geomagnetic activity, while panel (d) shows the interplanetary magnetic field B_z component in GSE coordinates from the 1-min OMNIWeb database, providing context for the driving and geomagnetic conditions. For illustrative purposes, data are smoothed using a 100 s moving average window.

across all metrics is similar, showing that while spacecraft variations exist due to different instrumentation, the predictability of ion temperature is similarly high between Geotail and MMS. Electron temperature performance is similar, with a lower MAE due to ions being typically hotter than electrons throughout the whole magnetotail. Finally, PRIME-PS predicts T_i/T_e well when trained on it directly, but the ratio constructed from independently predicted T_i and T_e underperforms.

To visualize the spatial structure of ion and electron temperatures separately, we project the MMS PRIME-PS model onto a grid of synthetic spacecraft locations, analogous to the Geotail analysis above. Figure 8 shows the resulting equatorial maps of T_i and T_e under varying solar wind speed and IMF B_z conditions. As expected, ion temperature increases with solar wind speed and exhibits a clear duskward enhancement, consistent with charge-dependent gradient-curvature drift (C. Wang et al., 2012; A. P. Walsh et al., 2014). Interestingly, a dawnside

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

increase of ion temperature appears farther from Earth, at roughly −25 to −15 R_E. Because MMS observations are sparse at those distances, it is unclear whether this feature is physical or an artifact of uneven spatial sampling. The latter interpretation is supported by the Geotail-trained model (Figure 6), which, drawing on far more extensive tailward coverage, shows no comparable dawnside enhancement at those distances. Electron temperature, by contrast, shows a weaker but opposite asymmetry, with slightly elevated values on the dawn side. The bottom panels of Figure 8 quantify this dusk-dawn asymmetry directly: ΔT_{DD} is positive for ions (duskward enhancement) and negative for electrons (dawnward enhancement) across all driving conditions, reproducing the expected separation of species-dependent drift effects. Notably, this separation persists under northward IMF as well.

Using Tsyganenko and Mukai (2003) for ion temperature and Dubyagin et al. (2016) for electron temperature, one can obtain an empirical model of T_i/T_e driven by solar wind conditions. However, because these models were derived from different spacecraft (Geotail and THEMIS, respectively), their spatial domains do not overlap, and the combined output produces extreme values under certain conditions. We therefore used the MMS-trained PRIME-PS model and the driving conditions from Figure 10 of Dubyagin et al. (2016) to evaluate T_i/T_e in synthetic data and examine any asymmetries. Figure 9 shows the resulting T_i/T_e maps for four representative solar wind conditions. The ratio ranges from approximately 3 on the dawn side to around 5 on the dusk side, with the duskward enhancement being slightly elevated during northward IMF. The corresponding dusk-dawn asymmetry profiles (panels a2–d2) confirm a persistent positive $\Delta(T_i/T_e)_{\mathrm{DD}}$, indicating that the dusk side consistently hosts a higher T_i/T_e than the dawn side. The asymmetry also appears to depend on driving strength. For northward IMF it weakens as V_{SW} increases from 400 to 600 km/s; however, given the scarcity of fast-wind, northward-IMF intervals in the training set, we cannot exclude that this trend partly reflects uneven sampling. Overall, these results are consistent with the individual temperature asymmetries shown in Figure 8, and demonstrate that PRIME-PS can produce reasonable T_i/T_e spatial distributions.

4.3. Comparison With THEMIS Observations During the November 2023 Geomagnetic Storm

Finally, to more rigorously evaluate the model's generalization capabilities, we assessed its performance during the geomagnetic storm of November 04–08, 2023. This event represents a particularly challenging test, as storm observations constitute statistical outliers (<5% of the training distribution). Furthermore, the absence of concurrent plasma sheet observations from Geotail or MMS during this interval ensures a strictly out-of-sample validation. The THEMIS-D probe, located at approximately (−12.65, 3.19, −2.11) R_E in GSM coordinates during this interval, provided the necessary in situ ground truth for this analysis. We focus here on ion density, as density moments are generally less susceptible to cross-instrument calibration issues than temperature. Consistent with our training methodology, the TM03 and PRIME-PS models were driven using upstream solar wind parameters and the specific THEMIS spacecraft location.

Figure 10 presents the comparison between THEMIS observations and model predictions. The model trained on the strict data set (PRIME-PS "Strict") performs only modestly better than TM03 during the storm. Supplementing the strict data set with the high-density intervals (PRIME-PS "High") markedly improves the agreement with observations, reducing the MAE by approximately 30% relative to TM03 (and by about 16% relative to PRIME-PS "Strict") and capturing the density enhancements characteristic of the storm main phase. The temporal structure of the observed density variations is well reproduced, with the primary discrepancy being a time delay rather than a failure to capture the amplitude. For this event, retaining high-density intervals that standard classification schemes would discard clearly improves the storm-time prediction. While the residual errors indicate that further improvement is needed to fully resolve the most extreme density peaks, PRIME-PS already represents a clear improvement over the empirical baseline.

Figure 11 shifts the focus to the MMS-trained ion-to-electron temperature ratio model and provides a complementary view of the model performance during the same storm interval. The model using direct predictions and the one constructed from the ratio of independently predicted ion and electron temperatures follow the observed temporal evolution, although with offsets that likely reflect cross-instrument calibration differences between MMS-derived training data and THEMIS observations.

4.4. Input Feature Importance

As a first-order view of input importance, we provide a SHAP analysis of the GB model (Figure S2 in Supporting Information S1) and summarize the key findings here. We retain this analysis in Supporting Information S1 rather

10.1029/2026JH001589

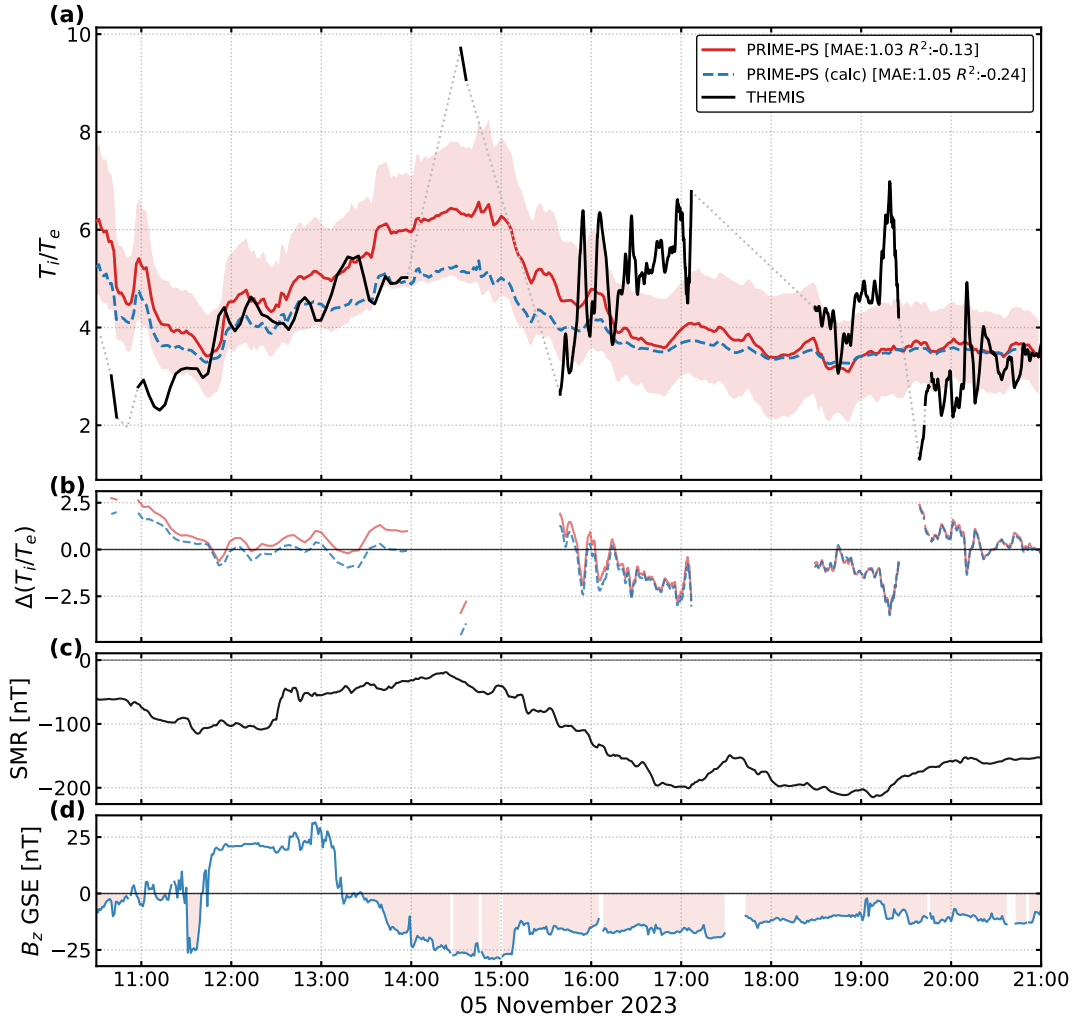


Figure 11. Time series comparison of plasma sheet ion-to-electron temperature ratio (T_i/T_e) predictions versus Time History of Events and Macroscale Interactions During Substorms (THEMIS) observations during the geomagnetic storm event on 05 November 2023. Panel (a) displays the observations from THEMIS (black solid line for measured data; light grey dotted line for interpolated segments) alongside predictions from PRIME-PS using direct predictions (red solid line) and PRIME-PS (calc; blue dashed line), obtained by dividing the predicted ion temperature by the predicted electron temperature. The red shaded region indicates the uncertainty bounds ($\pm 1\sigma$) for the PRIME-PS model. Performance metrics, including Mean Absolute Error and R^2, are provided in the legend for each model. Panel (b) shows the residuals ($\Delta(T_i/T_e)$), calculated as the model prediction minus the observed THEMIS value, highlighting temporal variations in model overestimation ($\Delta(T_i/T_e) > 0$) and underestimation ($\Delta(T_i/T_e) < 0$). Panel (c) shows the SMR index as an indicator of storm-time geomagnetic activity, while panel (d) shows the interplanetary magnetic field B_z component in GSE coordinates from the 1-min OMNIWeb database, providing context for the driving and geomagnetic conditions. For illustrative purposes, data are smoothed using a 100 s moving average window.

than the main text because several of the input features are physically intercorrelated (e.g., solar wind velocity and geomagnetic indices), and the resulting SHAP attribution can redistribute importance among coupled inputs in ways that require more careful treatment than is within the scope of this work. Each bar gives a feature's normalized contribution, with the values summing to unity, and all time-history steps of a given input grouped into a single entry; resolving the importance of individual lags is left for future work. Consistent with empirical models such as TM03, spacecraft location dominates the density prediction, accounting for roughly 45% of the normalized importance in LightGBM. Ion temperature instead responds more to solar-wind velocity and geomagnetic indices, suggesting that density is set largely by position while temperature tracks external driving. Comparing the full and storm panels for each target (Figure S2 in Supporting Information S1), this balance shifts during the storm, with solar-wind driving overtaking location, as expected when the system is strongly driven.

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Savvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5. Discussion and Conclusions

In this work, we demonstrated a supervised ML application utilizing two different input data sets (Wind, OMNI) and two distinct missions containing multiple output targets (Geotail, MMS). Our findings can provide multiple insights into both modeling strategies and the physical processes governing the plasma sheet during quiet and disturbed conditions.

First, we showed that the definition and classification of the plasma sheet region significantly impact model training effectiveness. The "flexible" data set exhibited substantial performance degradation compared to the "strict" data set, highlighting the necessity of well-curated output targets. While classification currently relies on threshold-based criteria, unsupervised learning techniques, already applied to dayside observations (Nguyen et al., 2022; Olshevsky et al., 2021; Toy-Edens et al., 2024), could offer less biased alternatives for the complex nightside environment. Importantly, this performance degradation was primarily observed in density profiles (Tables 2 and 3), while temperature profiles remained relatively robust.

This distinction is physically consistent with the properties of the CPS and the PSBL. Statistical surveys indicate that while the PSBL and CPS often exhibit comparable bulk temperature moments, their densities differ more significantly, with the PSBL generally being more tenuous (Baumjohann et al., 1988, 1990; Vo et al., 2023). Consequently, the "flexible" data set introduces a higher number of lower-density PSBL points (effectively "false positives" for a CPS model) thereby skewing density predictions. In contrast, possibly due to the presence of high-speed ion beams of the PSBL (Birn et al., 2020), the effective temperatures are similar to the hot, isotropic CPS, making the temperature error less pronounced. It is crucial to note, however, that while the bulk T_i moments are similar, the distributions differ fundamentally: the CPS is thermalized and isotropic, while the PSBL ion distribution functions are highly anisotropic and consist of field-aligned beams. Furthermore, feature importance analysis suggests that spacecraft location primarily drives density patterns, whereas temperature might be more strongly associated with internal dynamics (i.e., geomagnetic indices as indicated in Figure S2 in Supporting Information S1). Thus, including boundary layer data disproportionately degrades the spatially dependent density model while having a limited effect on the dynamics-driven temperature model.

The Geotail-trained and MMS-trained ion temperature maps (Figures 6 and 8) also provide an indirect cross-mission consistency check. Under identical driving, both models independently recover the duskward temperature enhancement and the dawn-dusk asymmetry structure, which suggests that these features reflect physical patterns rather than mission-specific artifacts. A direct cross-mission test supports this picture and reveals mission-specific offsets (Figure S4 in Supporting Information S1). Applied to the MMS test set (final 20%), the Geotail-trained model retains a correlation of 0.67, and the MMS-trained model applied to the Geotail 2005–2010 strict test set retains 0.61, close to the within-mission values of 0.70–0.77. The Geotail-trained model, however, systematically underestimates the MMS temperatures, and the MMS-trained model overestimates the Geotail temperatures, which drives R^2 to near zero in both cases. The direction of these offsets matches the instrumental differences described in Section 2 and the expectations from Figures 6 and 8, which is that the MMS-derived T_i is systematically higher. The two data sets thus agree on the relative spatial and driver dependence of T_i but differ in absolute calibration. The models' disagreement, however, reveals a more subtle challenge: the spatial coverage of any single mission imprints itself on the trained model. The dawnside ion temperature feature visible in the MMS model at -25 to -15 R_E but absent in the Geotail model (Figures 6 and 8) illustrates this point. Where one spacecraft's orbit provides dense sampling, the model learns spatial structure confidently. However, where coverage is limited, apparent features may reflect artifacts rather than physics. This orbit dependence carries a practical implication: users of these models should weight predictions more heavily in regions well sampled by the training mission and treat sparsely covered regions with caution. More broadly, it argues for a future cross-calibrated model that ingests observations from multiple missions (Geotail, MMS, THEMIS, Cluster) onto a common reference frame, so that each spacecraft's orbital strengths compensate for the others' gaps. Such an effort would require careful intercalibration of plasma instruments across missions, but the payoff would be a model whose spatial fidelity is less dependent on the spatial coverage of each mission.

Comparing modeling approaches, the ML methods generally outperform the available empirical models, regardless of whether OMNI or Wind solar wind inputs are used, with the largest gains for ion temperature. Empirical models offer compact functional forms but they cannot capture the spatial structure that the ML data-driven models recover. Beyond statistical metrics, the primary advantage of our data-driven approach is the

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

reproduction of physically driven asymmetries (Figures 5, 6, 8, and 9). These asymmetries are well-documented in observational literature: the dawnside density enhancement is attributed to solar wind entry processes (Sorathia et al., 2019; C.-P. Wang et al., 2007; Wing & Newell, 1998), while the duskward ion temperature and temperature ratio enhancement arise from energy-dependent magnetic drifts (A. P. Walsh et al., 2014; C. Wang et al., 2012; Wing et al., 2005). Near Earth, the MMS-based model further resolves species-dependent asymmetries, with ions exhibiting duskward temperature enhancement and electrons showing a weaker dawnward enhancement (Figure 8), consistent with charge-dependent gradient-curvature drift. The resulting T_i/T_e maps (Figure 9) display a persistent duskward enhancement across all driving conditions. The dawnside density enhancement, by contrast, decreases with increasing ring current activity (Figure S3 in Supporting Information S1), consistent with Dubyagin et al. (2016). The model's ability to capture these features gives additional confidence in its utility for quantitative applications and in its ability to produce a magnetotail picture consistent with observations.

Another key result of this work is the demonstration that data set curation choices directly affect storm-time prediction quality. By evaluating our models against a THEMIS passage during a severe geomagnetic storm (Figure 10), we showed that including high-density intervals, often treated as outliers in standard classification, reduced the MAE by approximately 16% relative to the strict-only model and 30% relative to TM03, and enabled the model to reproduce the temporal structure of storm-time density enhancements. For this event, the choice of which data to include mattered as much as the choice of algorithm. The T_i/T_e predictions during the same storm interval (Figure 11) track the observed evolution somewhat better, which may indicate that the temperature ratio is more tightly constrained by solar wind driving than the absolute density. We note, however, that this improvement is not general. On the Geotail density model, PRIME-PS "High" increases the MAE from 0.113 to 0.123 cm^{-3} relative to PRIME-PS "Strict," because the added high-density points shift the learned distribution away from the typical conditions that dominate the training and test data. Even for the moderately disturbed subset of this test set (SMR < -20 nT, ∼3,000 samples), PRIME-PS "High" performs comparably to PRIME-PS "Strict" (MAE of 0.093 vs. 0.089 cm^{-3}). The storm-time benefit therefore rests on capturing better the overall trend and peak, and on capturing the severity of the single event we demonstrate. Establishing whether and under which conditions it generalizes across a broader set of geomagnetic storms remains an important next step.

Furthermore, the accuracy of the moments used as ground truth strongly affects model reliability, especially for the ion-to-electron temperature ratio (T_i/T_e). The standard MMS L2 moments and the partial moments used here differ by roughly 30%. Correcting these contamination effects is not a matter of preference but a condition for a reliable model. Beyond moment accuracy, instrumental systematics and data sparsity also contribute, and the residual storm-time errors may reflect physical processes not captured by solar wind input alone, consistent with recent findings on internal plasmaspheric sources (Kang et al., 2025). Oxygen outflows during strong driving conditions may also influence the locally measured ion populations at the plasma sheet (Kistler et al., 2010; Regoli et al., 2024; Upadhyay et al., 2025), making data-driven modeling efforts during these times even more challenging.

More broadly, the framework presented here is not limited to reproducing the spatial structure of the plasma sheet. Because PRIME-PS maps upstream solar wind conditions onto plasma sheet particle quantities through a learned propagation function, it can serve as a data-driven boundary or closure model within larger simulation frameworks. For example, an inner-magnetosphere or ring current model that requires plasma sheet conditions at its outer boundary, or a coupled MHD–ring current framework that must partition pressure between ions and electrons, could query PRIME-PS at each time step rather than prescribing fixed or parameterized profiles, providing time-dependent, solar-wind-driven boundary conditions that reflect the statistical behavior of the driven system. The approach also extends naturally to other magnetospheric regions: the same architecture could be retrained for other environments, provided sufficiently large and diverse data sets are available.

However, embedding a data-driven model inside a physics-based simulation introduces the risk of producing states that violate conservation laws or known transport processes, particularly when the simulation queries conditions outside the training distribution. Ensuring consistency with constraints such as pressure balance, particle conservation, and thermodynamic limits would likely require either physics-informed loss functions during training or post hoc enforcement layers. These safeguards may become especially important during extreme driving, where the training data are sparsest and unphysical extrapolation is most likely.

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Looking forward, several promising avenues exist to further improve storm-time prediction. Augmenting disturbed-time data sets by fusing in situ observations with high-resolution global magnetospheric simulations, a "reanalysis" strategy successfully employed in meteorology (Hersbach et al., 2020; Poduval et al., 2023; Rasp et al., 2020), offers a natural next step. While multi-spacecraft integration (Geotail, Cluster, MMS, THEMIS) is essential, it requires rigorous instrument intercalibration. Alternative techniques, such as imbalanced learning from either an algorithm perspective (Chu et al., 2025) or optimization of the data set (Camporeale, 2025), offer potential pathways to weight extreme events appropriately and to balance quiet-time accuracy with storm-time coverage without forcing a choice between the two regimes. Furthermore, symbolic regression may also be useful in conjunction with the previously mentioned techniques to obtain models that are more interpretable and based on fundamental physical laws (La Cava et al., 2021; Markidis et al., 2025).

Finally, we caution against modeling solutions that rely solely on local in situ parameters (e.g., magnetic field) to infer plasma properties due to the "non-uniqueness" problem. In plasma environments and especially in the magnetotail, similar magnetic field topologies can host vastly different plasma states, ranging from cold/dense to hot/tenuous (Sergeev et al., 2025; Wing et al., 2006), making external (i.e., solar wind) driving a necessary, albeit potentially insufficient, approach. Moreover, spatial and temporal variability can arise from the long transport timescales of solar wind plasma as it travels from the flanks to the plasma sheet center and then to the inner edge of the plasma sheet (Borovsky et al., 1998; Wing et al., 2006, 2014). Ultimately, resolving all the above-mentioned ambiguities likely requires a move from single-point (or effectively single-point due to the electron scale tetrahedron in the case of MMS) measurements to constellation-class missions (e.g., Links; National Academies of Sciences, Engineering, and Medicine, 2025) or model-data fusion frameworks.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

Geotail data are available through https://darts.isas.jaxa.jp/stp/geotail/data.html (Geotail/MGF_LEP Team, 1994). MMS measurements can be found through https://lasp.colorado.edu/mms/sdc/public/about/browse-wrapper/ or the Graphical User Interface (GUI) located in https://lasp.colorado.edu/mms/sdc/public/search/ (NASA, 2025). The SuperMAG indices are available via https://supermag.jhuapl.edu. THEMIS ground moment data can be found at https://themis.ssl.berkeley.edu/data_products/index.php (Angelopoulos, 2008a), while the OMNIweb data are accessed through https://spdf.gsfc.nasa.gov/pub/data/omni (King & Papitashvili, 2005), and the Wind observations can be accessed through the Coordinated Data Analysis Web (CDAWeb) online portal at https://cdaweb.gsfc.nasa.gov/istp_public/ (Wilson et al., 2021). The PRIME-PS model is available through the primesw Python package (O'Brien, 2026).

Acknowledgments
We acknowledge funding from the NASA DRIVE Science Center for Geospace Storms (CGS) under award 80NSSC22M0163. CO is supported by the National Science Foundation under Award No. 2524739. SW acknowledges the support of NASA Grants 80NSSC23K0904 and 80NSSC23K0899. KS was additionally supported by NASA Grants 80NSSC20K1833, 80NSSC23K0223, and 80NSSC24K1106.

References

Aminalragia-Giamini, S., Katsavrias, C., Papadimitriou, C., Daglis, I., Nasi, A., Brunet, A., et al. (2023). The EMERALD model for the estimation of the radial diffusion coefficients in the outer belt. *Space Weather*, *21*(1), e2022SW003283. https://doi.org/10.1029/2022sw003283

Aminalragia-Giamini, S., Raptis, S., Anastasiadis, A., Tsigkanos, A., Sandberg, I., Papaioannou, A., et al. (2021). Solar energetic particle event occurrence prediction using solar flare soft X-ray measurements and machine learning. *Journal of Space Weather and Space Climate*, *11*, 59. https://doi.org/10.1051/swsc/2021043

Angelopoulos, V. (2008a). THEMIS data. Retrieved from https://themis.ssl.berkeley.edu/themisdata/.UniversityofCalifornia,Berkeley/SpaceSciencesLaboratory

Angelopoulos, V. (2008b). The THEMIS mission. *Space Science Reviews*, *141*(1), 5–34. https://doi.org/10.1007/s11214-008-9336-1

Baumjohann, W., Paschmann, G., & Lühr, H. (1990). Pressure balance between lobe and plasma sheet. *Geophysical Research Letters*, *17*(1), 45–48. https://doi.org/10.1029/gl017i001p00045

Baumjohann, W., Paschmann, G., Sckopke, N., Cattell, C., & Carlson, C. (1988). Average ion moments in the plasma sheet boundary layer. *Journal of Geophysical Research*, *93*(A10), 11507–11520. https://doi.org/10.1029/ja093ia10p11507

Birn, J., Chandler, M., & Nakamura, R. (2020). Ion beams in the plasma sheet boundary layer: MMS observations and test particle simulations. *Journal of Geophysical Research: Space Physics*, *125*(2), e2019JA027113. https://doi.org/10.1029/2019ja027113

Borovsky, J. E., Thomsen, M. F., & Elphic, R. C. (1998). The driving of the plasma sheet by the solar wind. *Journal of Geophysical Research*, *103*(A8), 17617–17639. https://doi.org/10.1029/97ja02986

Burch, J., Moore, T., Torbert, R., & Giles, B. L. (2016). Magnetospheric multiscale overview and science objectives. *Space Science Reviews*, *199*(1–4), 5–21. https://doi.org/10.1007/s11214-015-0164-9

Camporeale, E. (2025). PARIS: Pruning algorithm via the representer theorem for imbalanced scenarios. *arXiv preprint arXiv:2512.06950*.

Camporeale, E., & Care, A. (2021). ACCRUE: Accurate and reliable uncertainty estimate in deterministic models. *International Journal for Uncertainty Quantification*, *11*(4), 81–94. https://doi.org/10.1615/int.j.uncertaintyquantification.2021034623

Cao, X., Halekas, J. S., Haaland, S., Ruhunusiri, S., & Glassmeier, K.-H. (2023). Using machine learning to characterize solar wind driving of convection in the terrestrial magnetotail lobes. *Frontiers in Astronomy and Space Sciences*, *10*, 1180410. https://doi.org/10.3389/fspas.2023.1180410

Chollet, F., et al. (2015). Keras. *GitHub*. Retrieved from https://github.com/fchollet/keras

Chu, X., Jia, L., McPherron, R. L., Li, X., & Bortnik, J. (2025). Imbalanced regression artificial neural network model for auroral electrojet indices (IRANNA): Can we predict strong events? *Space Weather*, *23*(5), e2024SW004236. https://doi.org/10.1029/2024sw004236

Chu, X., Ma, D., Bortnik, J., Tobiska, W. K., Cruz, A., Bouwer, S. D., et al. (2021). Relativistic electron model in the outer radiation belt using a neural network approach. *Space Weather*, *19*(12), e2021SW002808. https://doi.org/10.1029/2021sw002808

Delgado-Panadero, Á., Hernández-Lorca, B., García-Ordás, M. T., & Benítez-Andrades, J. A. (2022). Implementing local-explainability in gradient boosting trees: Feature contribution. *Information Sciences*, *589*, 199–212. https://doi.org/10.1016/j.ins.2021.12.111

Denton, M. H., & Borovsky, J. E. (2009). The superdense plasma sheet in the magnetosphere during high-speed-stream-driven storms: Plasma transport timescales. *Journal of Atmospheric and Solar-Terrestrial Physics*, *71*(10–11), 1045–1058. https://doi.org/10.1016/j.jastp.2008.04.023

Denton, M. H., Thomsen, M. F., Korth, H., Lynch, S., Zhang, J. C., & Liemohn, M. W. (2005). Bulk plasma properties at geosynchronous orbit. *Journal of Geophysical Research*, *110*(A7). https://doi.org/10.1029/2004JA010861

Devanandan, A. P., Keesee, A., Raptis, S., Ohtani, S., Merkin, V., & Gkioulidou, M. (2026). Statistical characteristics of stormtime bursty bulk flows. *Geophysical Research Letters*, *53*(3), e2025GL119632. https://doi.org/10.1029/2025GL119632

Dubyagin, S., Ganushkina, N. Y., Sillanpää, I., & Runov, A. (2016). Solar wind-driven variations of electron plasma sheet densities and temperatures beyond geostationary orbit during storm times. *Journal of Geophysical Research: Space Physics*, *121*(9), 8343–8360. https://doi.org/10.1002/2016ja022947

Friedel, R. H. W., Korth, H., Henderson, M. G., Thomsen, M. F., & Scudder, J. D. (2001). Plasma sheet access to the inner magnetosphere. *Journal of Geophysical Research*, *106*(A4), 5845–5858. https://doi.org/10.1029/2000JA003011

Geotail/MGF_LEP Team. (1994). *Geotail magnetic field (MGF) and low-energy plasma (LEP) data*. Institute of Space and Astronautical Science (ISAS). Retrieved from https://darts.isas.jaxa.jp/app/stp/geotail/ascii.html

Gershman, D. J., Dorelli, J. C., Avanov, L. A., Gliese, U., Barrie, A., Schiff, C., et al. (2019). Systematic uncertainties in plasma parameters reported by the fast plasma investigation on NASA's magnetospheric multiscale mission. *Journal of Geophysical Research: Space Physics*, *124*(12), 10345–10359. https://doi.org/10.1029/2019ja026980

Gneiting, T., Raftery, A. E., Westveld III, A. H., & Goldman, T. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Monthly Weather Review*, *133*(5), 1098–1118. https://doi.org/10.1175/mwr2904.1

Graham, D. B., Khotyaintsev, Y. V., & André, M. (2023). Langmuir and upper hybrid waves in Earth's magnetotail. *Journal of Geophysical Research: Space Physics*, *128*(10), e2023JA031900. https://doi.org/10.1029/2023ja031900

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *The Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hu, A., Shneider, C., Tiwari, A., & Camporeale, E. (2022). Probabilistic prediction of Dst storms one-day-ahead using full-disk SOHO images. *Space Weather*, *20*(8), e2022SW003064. https://doi.org/10.1029/2022sw003064

Hua, M., Tian, S., Bortnik, J., Ma, D., Wang, Y., Lyons, L., & Wygant, J. (2026). A neural network model of equatorial electric field structures in the inner magnetosphere. *Geophysical Research Letters*, *53*(3), e2025GL120163. https://doi.org/10.1029/2025gl120163

Huber, P. J. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo. *Annals of Statistics*, *1*(5), 799–821. https://doi.org/10.1214/aos/1176342503

Kang, S.-B., Fok, M.-C., Glocer, A., Ferradas, C., Huba, J., & Chen, L.-J. (2025). Plasmaspheric-origin ions as a major source of plasma sheet pressure during storm early main phase. *Geophysical Research Letters*, *52*(21), e2025GL117657. https://doi.org/10.1029/2025gl117657

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., et al. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, *30*.

King, J., & Papitashvili, N. (2005). Solar wind spatial scales in and comparisons of hourly Wind and ACE plasma and magnetic field data. *Journal of Geophysical Research*, *110*(A2). https://doi.org/10.1029/2004ja010649

Kingma, D. P., & Ba, J. (2017). Adam: A method for stochastic optimization. *arXiv:1412.6980 [cs]*. Retrieved from http://arxiv.org/abs/1412.6980

Kistler, L., Mouikis, C., Klecker, B., & Dandouras, I. (2010). Cusp as a source for oxygen in the plasma sheet during geomagnetic storms. *Journal of Geophysical Research*, *115*(A3). https://doi.org/10.1029/2009ja014838

Korth, H., Thomsen, M. F., Borovsky, J. E., & McComas, D. J. (1999). Plasma sheet access to geosynchronous orbit. *Journal of Geophysical Research*, *104*(A11), 25047–25061. https://doi.org/10.1029/1999JA900292

La Cava, W., Burlacu, B., Virgolin, M., Kommenda, M., Orzechowski, P., de França, F. O., et al. (2021). Contemporary symbolic regression methods and their relative performance. *Advances in Neural Information Processing Systems*.

Li, S., Kronberg, E., Mouikis, C., Luo, H., Ge, Y., & Du, A. (2023). Prediction of proton pressure in the outer part of the inner magnetosphere using machine learning. *Space Weather*, *21*(9), e2022SW003387. https://doi.org/10.1029/2022sw003387

Liu, H., Li, W., Tang, B., Norgren, C., Liu, K., Khotyaintsev, Y. V., et al. (2025). High-speed electron flows in the Earth magnetotail. *AGU Advances*, *6*(2), e2024AV001549. https://doi.org/10.1029/2024av001549

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*.

Ma, D., Chu, X., Bortnik, J., Claudepierre, S. G., Tobiska, W. K., Cruz, A., et al. (2022). Modeling the dynamic variability of sub-relativistic outer radiation belt electron fluxes using machine learning. *Space Weather*, *20*(8), e2022SW003079. https://doi.org/10.1029/2022sw003079

Markidis, S., Ekelund, J., Pennati, L., Hu, A., & Peng, I. (2025). Discovering governing equations of geomagnetic storm dynamics with symbolic regression. In *International conference on computational science* (pp. 122–136).

Martin, D., Hong, J., O'Brien, C., Filho, V. P. M., Kobayashi, J. R., Samara, E., & Gallego, J. (2025). CORONA-Fields: Leveraging foundation models for classification of solar wind phenomena. *arXiv*. https://doi.org/10.48550/arXiv.2511.09843

McElfresh, D., Khandagale, S., Valverde, J., Prasad C, V., Ramakrishnan, G., Goldblum, M., & White, C. (2023). When do neural nets outperform boosted trees on tabular data? *Advances in Neural Information Processing Systems*, *36*, 76336–76369.

McFadden, J., Carlson, C., Larson, D., Bonnell, J., Mozer, F., Angelopoulos, V., & Auster, U. (2008). THEMIS ESA first science results and performance issues. *Space Science Reviews*, *141*(1), 477–508. https://doi.org/10.1007/s11214-008-9440-2

Merkin, V. G., Panov, E. V., Sorathia, K., & Ukhorskiy, A. Y. (2019). Contribution of bursty bulk flows to the global dipolarization of the magnetotail during an isolated substorm. *Journal of Geophysical Research: Space Physics*, *124*(11), 8647–8668. https://doi.org/10.1029/2019ja026872

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Mukai, T., Machida, S., Saito, Y., Hirahara, M., Terasawa, T., Kaya, N., et al. (1994). The low energy particle (LEP) experiment onboard the Geotail satellite. *Journal of Geomagnetism and Geoelectricity*, *46*(8), 669–692. https://doi.org/10.5636/jgg.46.669

NASA. (2025). Magnetospheric multiscale (MMS) science data. Retrieved from https://lasp.colorado.edu/mms/sdc/public/.NASASpacePhysicsDataFacility(SPDF)/LASPMMSScienceDataCenter

National Academies of Sciences, Engineering, and Medicine. (2025). Solar and space physics for the nation: An overview of the 2024–2033 decadal survey. *The National Academies Press*. https://doi.org/10.17226/29150

Nguyen, G., Aunai, N., Michotte de Welle, B., Jeandet, A., Lavraud, B., & Fontaine, D. (2022). Massive multi-mission statistical study and analytical modeling of the Earth's magnetopause: 1. A gradient boosting based automatic detection of near-Earth regions. *Journal of Geophysical Research: Space Physics*, *127*(1), e2021JA029773. https://doi.org/10.1029/2021ja029773

Nishida, A. (1994). The geotail mission. *Geophysical Research Letters*, *21*(25), 2871–2873. https://doi.org/10.1029/94GL01223

O'Brien, C. (2026). connor-obrien888/prime: PRIME 2.0.0. *Zenodo*. https://doi.org/10.5281/zenodo.20802889

O'Brien, C., Walsh, B., Zou, Y., Qudsi, R., Tasnim, S., Zhang, H., & Sibeck, D. G. (2024). PRIME-SH: A data-driven probabilistic model of Earth's magnetosheath. *Journal of Geophysical Research: Machine Learning and Computation*, *1*(3), e2024JH000235. https://doi.org/10.1029/2024jh000235

O'Brien, C., Walsh, B. M., Zou, Y., Tasnim, S., Zhang, H., & Sibeck, D. G. (2023). PRIME: A probabilistic neural network approach to solar wind propagation from L1. *Frontiers in Astronomy and Space Sciences*, *10*, 1250779. https://doi.org/10.3389/fspas.2023.1250779

Ohtani, S., & Mukai, T. (2008). Statistical characteristics of the storm time plasma sheet. *Journal of Geophysical Research*, *113*(A1). https://doi.org/10.1029/2007ja012547

Olshevsky, V., Khotyaintsev, Y. V., Lalti, A., Divin, A., Delzanno, G. L., Anderzén, S., et al. (2021). Automated classification of plasma regions using 3D particle energy distributions. *Journal of Geophysical Research: Space Physics*, *126*(10), e2021JA029620. https://doi.org/10.1029/2021ja029620

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). PyTorch: An imperative style, high-performance deep learning library. In *Advances in neural information processing systems* (Vol. 32). Retrieved from https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf

Poduval, B., McPherron, R., Walker, R., Himes, M., Pitman, K., Azari, A., et al. (2023). AI-ready data in space science and solar physics: Problems, mitigation and action plan. *Frontiers in Astronomy and Space Sciences*, *10*, 1203598. https://doi.org/10.3389/fspas.2023.1203598

Pollock, C., Moore, T., Jacques, A., Burch, J., Gliese, U., Saito, Y., et al. (2016). Fast plasma investigation for magnetospheric multiscale. *Space Science Reviews*, *199*(1–4), 331–406. https://doi.org/10.1007/s11214-016-0245-4

Rahimi, A., & Recht, B. (2007). Random features for large-scale kernel machines. In *Proceedings of the 21st international conference on neural information processing systems* (pp. 1177–1184). Curran Associates Inc.

Raptis, S., Aminalragia-Giamini, S., Karlsson, T., & Lindberg, M. (2020). Classification of magnetosheath jets using neural networks and high resolution OMNI (HRO) data. *Frontiers in Astronomy and Space Sciences*, *7*, 24. https://doi.org/10.3389/fspas.2020.00024

Raptis, S., Merkin, V., Ohtani, S., Gkioulidou, M., & Regoli, L. H. (2024). Plasma sheet magnetic flux transport during geomagnetic storms. *Geophysical Research Letters*, *51*(18), e2024GL110839. https://doi.org/10.1029/2024gl110839

Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., & Thuerey, N. (2020). WeatherBench: A benchmark data set for data-driven weather forecasting. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002203. https://doi.org/10.1029/2020ms002203

Regoli, L., Gkioulidou, M., Ohtani, S., Raptis, S., Mouikis, C., Kistler, L., et al. (2024). Temporal evolution of O^+ population in the near-Earth plasma sheet during geomagnetic storms as observed by the magnetospheric multiscale mission. *Journal of Geophysical Research: Space Physics*, *129*(5), e2023JA032203. https://doi.org/10.1029/2023ja032203

Richard, L., Khotyaintsev, Y. V., Graham, D. B., & Russell, C. (2022). Are dipolarization fronts a typical feature of magnetotail plasma jets fronts? *Geophysical Research Letters*, *49*(22), e2022GL101693. https://doi.org/10.1029/2022gl101693

Runov, A., Angelopoulos, V., Gabrielse, C., Liu, J., Turner, D., & Zhou, X.-Z. (2015). Average thermodynamic and spectral properties of plasma in and around dipolarizing flux bundles. *Journal of Geophysical Research: Space Physics*, *120*(6), 4369–4383. https://doi.org/10.1002/2015ja021166

Sadeghzadeh, S., Yang, J., Wang, C.-P., Mousavi, A., Wang, W., Sun, W., et al. (2021). Effects of bubble injections on the plasma sheet configuration. *Journal of Geophysical Research: Space Physics*, *126*(6), e2021JA029127. https://doi.org/10.1029/2021ja029127

Sergeev, V., Beliuchenko, K., Divin, A., Merkin, V., & Sorathia, K. (2025). Distributions of cool/dense and hot/tenuous plasma regions in the magnetotail plasma sheet depending on substorm phases and Solar wind as follows from global magnetospheric simulations. *Journal of Geophysical Research: Space Physics*, *130*(8), e2025JA033889. https://doi.org/10.1029/2025ja033889

Shue, J.-H., Song, P., Russell, C., Steinberg, J., Chao, J., Zastenker, G., et al. (1998). Magnetopause location under extreme solar wind conditions. *Journal of Geophysical Research*, *103*(A8), 17691–17700. https://doi.org/10.1029/98ja01103

Shustov, P., Lukin, A. S., Zhang, X.-J., Artemyev, A. V., Petrukovich, A., & Angelopoulos, V. (2021). Evolution of thermal electron distributions in the magnetotail: Convective heating and scattering-induced losses. *Journal of Geophysical Research: Space Physics*, *126*(12), e2021JA029952. https://doi.org/10.1029/2021ja029952

Sitnov, M., Birn, J., Ferdousi, B., Gordeev, E., Khotyaintsev, Y., Merkin, V., et al. (2019). Explosive magnetotail activity. *Space Science Reviews*, *215*(4), 1–95. https://doi.org/10.1007/s11214-019-0599-5

Sitnov, M., Stephens, G., Motoba, T., & Swisdak, M. (2021). Data mining reconstruction of magnetotail reconnection and implications for its first-principle modeling. *Frontiers in Physics*, *9*, 644884. https://doi.org/10.3389/fphy.2021.644884

Smirnov, A., Berrendorf, M., Shprits, Y., Kronberg, E. A., Allison, H. J., Aseev, N. A., et al. (2020). Medium energy electron flux in Earth's outer radiation belt (MERLIN): A machine learning model. *Space Weather*, *18*(11), e2020SW002532. https://doi.org/10.1029/2020sw002532

Sorathia, K., Merkin, V., Panov, E., Zhang, B., Lyon, J., Garretson, J., et al. (2020). Ballooning-interchange instability in the near-Earth plasma sheet and auroral beads: Global magnetospheric modeling at the limit of the MHD approximation. *Geophysical Research Letters*, *47*(14), e2020GL088227. https://doi.org/10.1029/2020gl088227

Sorathia, K., Merkin, V., Ukhorskiy, A., Allen, R., Nykyri, K., & Wing, S. (2019). Solar wind ion entry into the magnetosphere during northward IMF. *Journal of Geophysical Research: Space Physics*, *124*(7), 5461–5481. https://doi.org/10.1029/2019ja026728

Stone, E. C., Frandsen, A., Mewaldt, R., Christian, E., Margolies, D., Ormes, J., & Snow, F. (1998). The advanced composition explorer. *Space Science Reviews*, *86*(1–4), 1–22. https://doi.org/10.1023/a:1005082526237

Toy-Edens, V., Mo, W., Raptis, S., & Turner, D. L. (2024). Classifying 8 years of MMS dayside plasma regions via unsupervised machine learning. *Journal of Geophysical Research: Space Physics*, *129*(6), e2024JA032431. https://doi.org/10.1029/2024ja032431

Tsyganenko, N., & Mukai, T. (2003). Tail plasma sheet models derived from geotail particle data. *Journal of Geophysical Research*, *108*(A3). https://doi.org/10.1029/2002ja009707

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Savvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Upadhyay, A., Gkioulidou, M., Ohtani, S.-i., & Cohen, I. J. (2025). Statistical analysis of storm-time O^+ outflows as observed by the Van Allen Probes. *Journal of Geophysical Research: Space Physics*, *130*(11), e2025JA034255. https://doi.org/10.1029/2025ja034255

Vo, T., Ergun, R., Usanova, M., & Chasapis, A. (2023). Mesoscale structure and properties of the terrestrial magnetotail plasma sheet from the magnetospheric multiscale mission. *Journal of Geophysical Research: Space Physics*, *128*(9), e2023JA031358. https://doi.org/10.1029/2023ja031358

Walsh, A. P., Haaland, S., Forsyth, C., Keesee, A. M., Kissinger, J., Li, K., et al. (2014). Dawn–dusk asymmetries in the coupled solar wind–magnetosphere–ionosphere system: A review. *Annales Geophysicae*, *32*(7), 705–737. https://doi.org/10.5194/angeo-32-705-2014

Walsh, B., Bhakyapaibul, T., & Zou, Y. (2019). Quantifying the uncertainty of using solar wind measurements for geospace inputs. *Journal of Geophysical Research: Space Physics*, *124*(5), 3291–3302. https://doi.org/10.1029/2019ja026507

Wang, C., Gkioulidou, M., Lyons, L. R., & Angelopoulos, V. (2012). Spatial distributions of the ion to electron temperature ratio in the magnetosheath and plasma sheet. *Journal of Geophysical Research*, *117*(A8). https://doi.org/10.1029/2012ja017658

Wang, C.-P., Lyons, L. R., Nagai, T., Weygand, J. M., & McEntire, R. W. (2007). Sources, transport, and distributions of plasma sheet ions and electrons and dependences on interplanetary parameters under northward interplanetary magnetic field. *Journal of Geophysical Research*, *112*(A10). https://doi.org/10.1029/2007ja012522

Wilson, L. B., III., Brosius, A. L., Gopalswamy, N., Nieves-Chinchilla, T., Szabo, A., Hurley, K., et al. (2021). A quarter century of wind spacecraft discoveries. *Reviews of Geophysics*, *59*, e2020RG000714. https://doi.org/10.1029/2020RG000714

Wing, S., Johnson, J., Chaston, C., Echim, M., Escoubet, C., Lavraud, B., et al. (2014). Review of solar wind entry into and transport within the plasma sheet. *Space Science Reviews*, *184*(1–4), 33–86. https://doi.org/10.1007/s11214-014-0108-9

Wing, S., & Johnson, J. R. (2009). Substorm entropies. *Journal of Geophysical Research*, *114*(A9). https://doi.org/10.1029/2008JA013989

Wing, S., Johnson, J. R., & Fujimoto, M. (2006). Timescale for the formation of the cold-dense plasma sheet: A case study. *Geophysical Research Letters*, *33*(23). https://doi.org/10.1029/2006gl027110

Wing, S., Johnson, J. R., Newell, P. T., & Meng, C.-I. (2005). Dawn-dusk asymmetries, ion spectra, and sources in the northward interplanetary magnetic field plasma sheet. *Journal of Geophysical Research*, *110*(A8). https://doi.org/10.1029/2005JA011086

Wing, S., & Newell, P. T. (1998). Central plasma sheet ion properties as inferred from ionospheric observations. *Journal of Geophysical Research*, *103*(A4), 6785–6800. https://doi.org/10.1029/97ja02994

Wing, S., Turner, D. L., Ukhorskiy, A. Y., Johnson, J. R., Sotirelis, T., Nikoukar, R., & Romeo, G. (2022). Modeling radiation belt electrons with information theory informed neural networks. *Space Weather*, *20*(8), e2022SW003090. https://doi.org/10.1029/2022SW003090

Xu, S., Huang, S., Yuan, Z., Deng, X., & Jiang, K. (2020). Prediction of the Dst index with bagging ensemble-learning algorithm. *The Astrophysical Journal Supplement Series*, *248*(1), 14. https://doi.org/10.3847/1538-4365/ab880e

29935210, 2026, 6, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JH001589 by Sarvas Raptis - Cochrane Greece , Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License