

# PRIME-SH: A Data-Driven Probabilistic Model of Earth's Magnetosheath

C. O'Brien<sup>1</sup>, B. M. Walsh<sup>1,2</sup>, Y. Zou<sup>3</sup>, R. Qudsi<sup>1</sup>, S. Tasnim<sup>4</sup>, H. Zhang<sup>5</sup>, D.G. Sibeck<sup>6</sup>

<sup>1</sup>Center for Space Physics, Boston University, Boston, MA, USA

<sup>2</sup>Department of Mechanical Engineering, Boston University, Boston, MA, USA

<sup>3</sup>Johns Hopkins University Applied Physics Lab, Laurel, MD, USA

<sup>4</sup>Institute for Solar-Terrestrial Physics, German Aerospace Center (DLR), Neustrelitz, Germany

<sup>5</sup>Computer Science Department, University of Alabama in Huntsville, Huntsville, AL, USA

<sup>6</sup>Heliophysics Science Division, NASA/GSFC, Greenbelt, MD, USA

## Key Points:

- PRIME-SH is an algorithm that predicts plasma and magnetic field in Earth's magnetosheath using inputs from in-situ monitors at L1.
- PRIME-SH accurately predicts the magnetosheath conditions in a statistical sense and its predictions obey mass, momentum, and energy conservation at the shock.
- PRIME-SH can be used to easily assemble continuous maps of the magnetosheath, addressing spatial limitations of in-situ data.

---

Corresponding author: Connor O'Brien, [obrienco@bu.edu](mailto:obrienco@bu.edu)

18 **Abstract**

19 A data-driven model of Earth's magnetosheath is developed by training a Bayesian  
 20 recurrent neural network to reproduce Magnetospheric MultiScale (MMS) measurements  
 21 of the magnetosheath plasma and magnetic field using measurements from the Wind space-  
 22 craft upstream of Earth at the first Earth-Sun Lagrange point (L1). This model, called  
 23 PRIME-SH in reference to its progenitor algorithm PRIME (Probabilistic Regressor for  
 24 Input to the Magnetosphere Estimation), is shown to predict spacecraft observations of  
 25 magnetosheath conditions accurately in a statistical sense with a continuous rank prob-  
 26 ability score (CRPS) of  $0.227\sigma$  and more accurately than current analytical models of  
 27 the magnetosheath. Furthermore, PRIME-SH is shown to reproduce physics not explic-  
 28 itly enforced during training, such as field line draping, the dayside plasma depletion layer,  
 29 the magnetosheath flow stagnation point, and the Rankine-Hugoniot MHD shock jump  
 30 conditions. PRIME-SH has the additional benefits of being computationally inexpen-  
 31 sive relative to global MHD simulations, being capable of reproducing kinetic physics such  
 32 as temperature anisotropy, and being capable of reliably estimating its own uncertainty  
 33 to within 3.5%.

34 **Plain Language Summary**

35 As the solar wind encounters Earth's magnetosphere and diverts around it, a shock  
 36 is formed that heats and compresses the plasma and warps the magnetic field frozen into  
 37 it. This shocked plasma and magnetic field, known as the magnetosheath, is what drives  
 38 energy transfer at the magnetopause. Due to orbital constraints there is no continuous  
 39 in-situ monitor of magnetosheath conditions. Studies of solar wind magnetosphere in-  
 40 teraction typically rely on solar wind conditions measured at L1 propagated to Earth  
 41 by some algorithm, which are then either used directly or used to drive some model of  
 42 the magnetosheath. This process has numerous points of uncertainty, from the choice  
 43 of propagation algorithm to the choice of magnetosheath model (or lack thereof). To ad-  
 44 dress these concerns with the traditional approach, this study develops a data-driven model  
 45 of the magnetosheath that uses data from L1 as its input. This new model, called PRIME-  
 46 SH, adapts a Bayesian recurrent neural network architecture that is capable of estimat-  
 47 ing uncertainties for its predictions. This new model is verified to be accurate in a sta-  
 48 tistical sense, and is also capable of representing physics that is not explicitly incorpo-  
 49 rated in the model during training.

50 **1 Introduction**

51 The region of turbulent, shocked solar wind plasma downstream of Earth's bow shock  
 52 is known as the magnetosheath. The magnetosheath plasma and magnetic field trans-  
 53 fer energy to Earth's magnetosphere via magnetic reconnection and viscous interaction  
 54 (Dungey, 1961; Axford, 1964). Despite this, the solar wind conditions upstream of the  
 55 bow shock are frequently taken as the input to the system in studies of solar wind-magnetosphere  
 56 interaction. This is because there is no continuous in-situ magnetosheath monitor due  
 57 to orbital constraints. In order to know the conditions in the magnetosheath during pe-  
 58 riods of time or locations in space where there is no in-situ monitor in the magnetosheath  
 59 available, it is therefore necessary to model the magnetosheath.

60 Early models of the magnetosheath were gas dynamic in nature, incorporating some  
 61 physical assumptions and including limited consideration of the magnetic field outside  
 62 the magnetopause (Spreiter et al., 1966; Spreiter & Alksne, 1969). These models have  
 63 matured through the inclusion of additional physics into modern MHD codes (e.g. Powell  
 64 et al. (1999); Lyon et al. (2004)), which offer spatially and temporally complete model  
 65 magnetosheaths at the cost of some physical assumptions and increased computational  
 66 expense. In situations where the computational expense of MHD modeling is prohibitive,

some magnetosheath modeling efforts fit analytical expressions derived from physical assumptions to spacecraft measurements of the magnetosheath (Kobel & Flückiger, 1994; Soucek & Escoubet, 2012; Tsyganenko et al., 2023). Others, such as the recent Mshpy23 model (Jung et al., 2024), parameterize the outputs of MHD models to reduce their computational cost but retain some of their accuracy. A shared feature of these approaches is that they all include physical assumptions. While they may often be valid, there remains differences between their outputs and the actual magnetosheath that can limit their representational power. This issue could be addressed by reducing the number of assumptions used to construct the model; hybrid-Vlasov codes capable of simulating the entire magnetosheath have recently come online (Von Alfthan et al., 2014; Hoilijoki et al., 2016) but come with an even higher computational cost than MHD codes and still simplify some physics.

One possible way of addressing this limitation is the use of neural network codes that do not assume a functional form or simplified physics. Neural networks have been used to assemble models of geophysical quantities for the past few decades since the early relativistic electron flux model of Koons and Gorney (1991), and have continued to be regularly utilized for space physics tasks. These algorithms do not require physical assumptions to construct tractable or analytical descriptions of the magnetosheath plasma and magnetic field, and are also computationally lightweight. In particular, new Bayesian recurrent neural network architectures have shown good performance in spatio-temporal inversion tasks such as electron density in the inner magnetosphere (Huang et al., 2022).

A crucial aspect of any prediction algorithm that is typically lacking in magnetospheric physics (and that is addressed by Bayesian neural networks) is uncertainty quantification (Borovsky, 2021). There is growing evidence that uncertainty in solar wind data affects correlation studies of the cross polar cap potential (Sivadas et al., 2022), development of solar wind-magnetosphere coupling functions (Lockwood et al., 2019), and global MHD simulation outputs (Al Shidi et al., 2023); feeding solar wind data with some uncertainty into a magnetosheath model with some inherent uncertainty results in a compounding effect. Since it is the shocked solar wind at the magnetopause rather than the solar wind upstream of the bow shock that interacts with the magnetosphere, this uncertainty has the potential to affect any study that tries to associate solar wind conditions with magnetospheric response in a way that is difficult to account for without a magnetosheath model that estimates this uncertainty.

Another challenge with traditional models aside from their physical assumptions is the fact that they typically use solar wind data that has been propagated from in-situ monitors far from Earth as input. Much like the magnetosheath, there is no continuous in-situ monitor of the solar wind near Earth due to orbital constraints. In order to obtain inputs for each of the previously mentioned models, data from monitors at the L1 position  $235R_E$  (1,500,000 km) ahead of Earth need to be propagated to Earth to account for the travel time of the solar wind plasma and interplanetary magnetic field (generally 30–60 minutes). This propagation task is made difficult by the structure and dynamics of the solar wind (Borovsky, 2018), and a great deal of varied algorithms have been developed in order to propagate between L1 and Earth accurately. One such algorithm, the Probabilistic Regressor for Input to the Magnetosphere Estimation (PRIME) (O’Brien et al., 2023) was recently developed to address some of these difficulties with traditional propagation algorithms, and its Bayesian recurrent neural network architecture is well suited to be adapted to the problem of magnetosheath prediction from L1 inputs (since the physics of solar wind propagation is the first “step” of that task).

In this study, a new algorithm capable of predicting magnetosheath plasma and magnetic field conditions given measurements made by an in-situ monitor at L1 is developed. This algorithm, named PRIME-SH after its progenitor algorithm PRIME (O’Brien et al., 2023), requires a dataset of in-situ magnetosheath measurements and associated solar wind inputs at L1 (Section 2), a network architecture adapted from PRIME and

120 optimized for the magnetosheath (Section 3). Outputs from PRIME-SH are validated  
 121 in a statistical sense on a holdout dataset. PRIME-SH is subjected to additional vali-  
 122 dation verifying that it reproduces some expected physics (Section 4), after which the  
 123 results can be summarized and discussed (Section 5).

## 124 2 Data

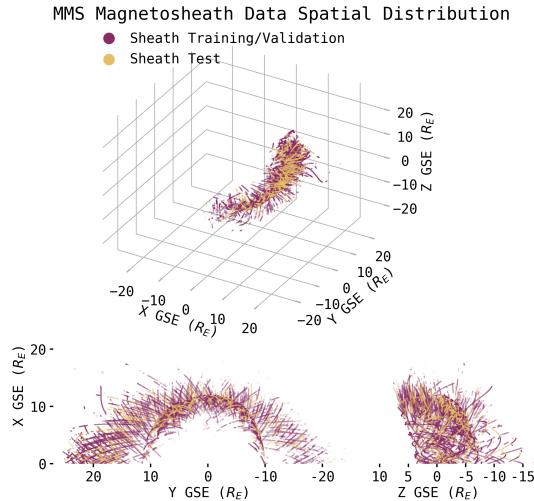
### 125 2.1 MMS Target Dataset

126 Plasma and magnetic field data from the Magnetospheric Multi Scale 1 (MMS-1)  
 127 spacecraft's (Burch et al., 2016) Fast Plasma Investigation (FPI) (Pollock et al., 2016)  
 128 and Fluxgate Magnetometer (FGM) (Russell et al., 2016) instruments are utilized as tar-  
 129 gets for the algorithm to be optimized against. MMS is a constellation of four spacecraft  
 130 designed to study magnetic reconnection at Earth's magnetopause and neutral sheet. It  
 131 therefore spends considerable time in Earth's magnetosheath and carries instruments par-  
 132 ticularly designed to measure the plasma and magnetic field there, making data it col-  
 133 lects highly suitable for use as targets to optimize PRIME-SH. The large volume of data  
 134 produced by MMS-1's instruments have motivated the development of automated clas-  
 135 sification, identification, and segmentation tools for MMS data that allow rapid selec-  
 136 tion of large amounts of data with particular features or from particular plasma regimes.

137 To assemble a solar wind dataset using MMS, an automatic tool developed by Olshevsky  
 138 et al. (2021) is used to classify all MMS-1 FPI 3D ion distributions from September 2nd  
 139 to January 1st 2023. The classifier is capable of discriminating between magneto-  
 140 spheric, magnetosheath, non-foreshock solar wind, and foreshock plasma through the shape  
 141 of the ion distribution function, and outputs a normalized probability that a given dis-  
 142 tribution belongs to each class. Periods of time where MMS-1 is in the magnetosheath  
 143 with probability  $p > 0.7$  are found using the classifier; all other time periods are removed  
 144 thereby removing the magnetosphere, solar wind, foreshock, and ambiguous classifica-  
 145 tions from the dataset. Remaining FGM magnetic field and FPI ion moments are av-  
 146 eraged in 100 second bins. Since the classifier is trained only on data from dayside or-  
 147 bits, any data on the nightside (GSE  $X < 0$ ) are removed. The full spatial distribu-  
 148 tion of the magnetosheath data are shown in Figure 1.

### 149 2.2 Wind Input Dataset

150 The input solar wind data at L1 comes from the Magnetic Field Investigation (MFI)  
 151 (Lepping et al., 1995) and Solar Wind Experiment (SWE) (Ogilvie et al., 1995) aboard  
 152 the Wind spacecraft. Wind was selected for this study because it had the best coverage  
 153 over the time period of the MMS-1 dataset used here (September 2nd 2015 to January  
 154 1st 2023). Key parameter (KP) moments data are utilized, resulting in time series of plasma  
 155 flow velocity  $\vec{v}$  (GSE coordinates), ion density  $n_{ion}$ , ion thermal speed  $v_{\perp th}$ , and IMF  
 156  $\vec{B}$  (GSM coordinates) at a 100 second cadence. Due to the difficulty involved with space-  
 157 craft intercalibration data from other L1 monitors are not included in this study (King,  
 158 2005). To give PRIME-SH information about the spatial separation of the input and tar-  
 159 get spacecraft and the location in the sheath at which the prediction is being made, the  
 160 positions of Wind and MMS-1 in GSE coordinates are included in the input data. Miss-  
 161 ing data are linearly interpolated and flagged so they can be excluded if necessary. The  
 162 precise windows of time in the Wind dataset used as input to predict each MMS target  
 163 heavily influence the performance of the optimized algorithm; these and other param-  
 164 eters pertaining to the exact construction of the dataset therefore must be optimized through  
 165 hyperparameter search (see Section 3.2).



**Figure 1.** 3D spatial distribution of the 117,427 magnetosheath MMS-1 data points split into 80% training/validation (purple) and 20% test (yellow) subsets. Data consists of  $\vec{B}_{GSM}$ ,  $\vec{V}_{GSE}$ ,  $n_i$ ,  $T_{i\parallel}$ , and  $T_{i\perp}$  from September 2nd 2015 to January 1st 2023. Train/validation/test split is as used in the optimized dataset (see Section 3.2).

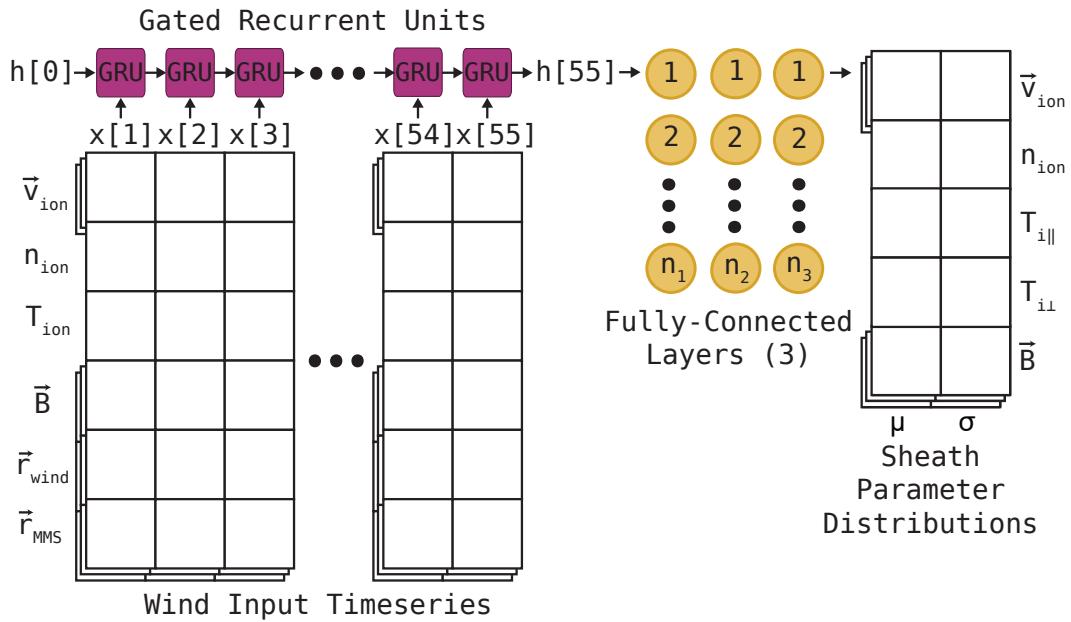
### 3 Algorithm Methodology

#### 3.1 Network Architecture

The overall architecture selected for the algorithm is similar to that utilized to construct PRIME (Probabilistic Regressor for Input to the Magnetosphere Estimation), an algorithm that predicts the solar wind near Earth using data from the Wind spacecraft at L1 (O'Brien et al., 2023). The Bayesian recurrent neural network architecture developed for PRIME is well suited to be adapted to the task of magnetosheath prediction for several reasons. First, it is capable of incorporating information about the time history of solar wind at L1 into its predictions which is important for predicting the solar wind and the evolution of the magnetosheath. Second, it is capable of assigning uncertainties to its predictions which is crucial in the frequently turbulent environment in the magnetosheath. Third, it has proven to be accurate when applied to the task of solar wind propagation, which is essentially the first step of the task undertaken by PRIME-SH.

The overall form of PRIME-SH is shown in Figure 2. Like PRIME, PRIME-SH utilizes a Gated Recurrent Unit (GRU) sequence (See Cho et al. (2014)) that is fed into fully connected neural network (FCNN) layers (See Bebis and Georgopoulos (1994)). The last layer of neurons are taken to be the mean and variance of a Gaussian probability distribution for each parameter rather than single scalar values (Nix & Weigend, 1994; Lakshminarayanan et al., 2017). The input feature size is 14, and the output feature size is 9. The algorithm is implemented in the Keras high-level API for tensorflow (<https://keras.io/api/>). Details of the architecture such as the length of the input time series and the size of each layer do not have optimal values that can be determined a priori. Instead, they are chosen via hyperparameter tuning (See Section 3.2).

The loss criterion used to optimize the algorithm during training is chosen to be the continuous rank probability score (CRPS) (Matheson & Winkler, 1976; Hersbach, 2000). The CRPS is a common scoring metric used to compare probabilistic forecasts for weather prediction (Zamo & Naveau, 2018). For a detailed description of the CRPS



**Figure 2.** Schematic of PRIME-SH’s neural network architecture, based on the architecture of PRIME (O’Brien et al., 2023). Note that the Gated Recurrent Unit (GRU) sequence feeds into a Fully Connected Neural Network (FCNN) in order to output a mean and variance for each desired parameter instead of a single value. Vector quantities such as magnetic field, flow velocity, and spacecraft position are stacked to show that they constitute three units in the input/output but describe one physical vector quantity. Exact layer size and additional regularization features (see Table 1) chosen via hyperparameter search.

see Section 2 of Camporeale and Carè (2021) or Section 3.1 of O'Brien et al. (2023). Briefly, the continuous rank probability score is given by

$$CRPS = \int_{-\infty}^{\infty} [F(y) - H(y - y_{obs})]^2 dy \quad (1)$$

where  $F(y)$  is the cumulative distribution function of a probabilistic prediction for some observation  $y_{obs}$  and  $H(y)$  is the Heaviside step function (Wilks, 2011). The continuous rank probability score is desirable as a loss function because it more symmetrically punishes over and under confident predictions than the negative log probability density (the most commonly used score for probabilistic predictions) (Camporeale & Carè, 2021). A side benefit is that the CRPS has the same unit as the variable of interest, making it more intuitively human-readable. In the case of Gaussian predictions with mean  $\mu$  and variance  $\sigma^2$  the CRPS is given by

$$CRPS(y_{obs}, \mu, \sigma) = \sigma \left[ \frac{y_{obs} - \mu}{\sigma} \operatorname{erf} \left( \frac{y_{obs} - \mu}{\sqrt{2}\sigma} \right) + \sqrt{\frac{2}{\pi}} e^{-\frac{(y_{obs} - \mu)^2}{2\sigma^2}} - \frac{1}{\sqrt{\pi}} \right] \quad (2)$$

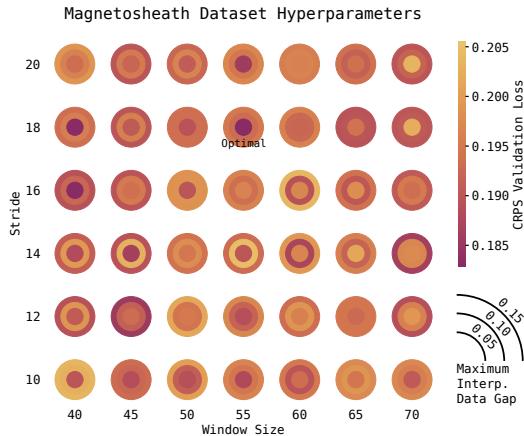
(Gneiting et al., 2005). Since PRIME-SH outputs Gaussian probability distributions, and since CRPS is negatively oriented, Equation 2 is used as a loss function during training. The 18 output units in PRIME-SH's last layer are taken to be the means ( $\mu_s$ ) and variances ( $\sigma_s$ ) defining a Gaussian probability distribution for each parameter. During training the CRPS over all nine parameters in the target dataset are averaged with equal weight assigned to all parameters.

The primary limitation of the CRPS as a loss function training probabilistic algorithms is the fact that it does not explicitly enforce reliability of the algorithm's predicted uncertainties (Camporeale et al., 2019). Reliability is measure of the degree to which a probabilistic forecast's uncertainties are statistically consistent with the observed probabilities of the events the forecast seeks to predict (Anderson, 1996). It has been shown that accuracy and reliability are competing metrics that must be balanced, and that simply minimizing the CRPS does not necessarily mean that the resulting model is reliable (Camporeale & Carè, 2021). Since reliability is not explicitly enforced, the reliability of PRIME-SH's uncertainties must be verified after training (See Section 4.1) (Tasistro-Hart et al., 2021).

### 3.2 Algorithm Optimization

Optimization of PRIME-SH follows a three step process. First, the optimal length, lead time, and composition of the input timeseries dataset is determined (the dataset hyperparameter search). Then the algorithm hyperparameters are systematically varied in order to find the optimal algorithm, then finally the optimal algorithm is instantiated and trained. This algorithm then becomes the canonical version of PRIME-SH.

Given a particular time when a prediction of the magnetosheath conditions is desired, it is difficult to say a priori what time period of Wind data from L1 contains the necessary information to make that prediction (especially given the flexible nature of neural network algorithms). Since the solar wind typically takes 30 to 60 minutes to get from L1 to Earth, there is likely only so much time history that can be incorporated before including more yields diminishing returns in terms of accuracy. Similarly, it is likely that including conditions at L1 right up until the time the sheath prediction is desired is not necessary, since the solar wind at that time has not had sufficient time to get to Earth. To find the optimal start and stop times of the timeseries used to make each prediction, a range of start and stop times are tested by optimizing a test version of PRIME-SH using different input time series lengths (windows) and lead times before each prediction (strides). It is also likely that large data gaps that are filled with interpolated data can affect the algorithm's performance, therefore a range of permissible data gap sizes are



**Figure 3.** Results from dataset optimization trials over timeseries window (length), stride (lead time), and permitted fraction of interpolated data. Units for window and stride are 100s (the Wind KP data cadence). The optimal set (window 55, stride 18, largest interp. fraction  $\leq 5\%$ ) is shown in darkest green and labelled “optimal”. Loss is given in dimensionless units of parameter interquartile range to ensure comparability of CRPS for each parameter.

also tested (expressed in terms of fractions of the window size). Whichever parameters produce a model that can achieve the best results on the validation dataset before overfitting are taken as optimal. When training these test models and for any time a model is trained, the input/target datasets are split into 60% training, 20% validation, and 20% test subsets. Since temporally adjacent entries in the input dataset are almost entirely overlapping, randomly assigning input/target pairs to each subset results in significant data leakage. To avoid this, the full dataset is split into independent blocks four times the length of the timeseries window used as input (i.e. for a window size of 55 measurements/  $\sim 1$  hour 32 minutes, the dataset is split into chunks of length 220 measurements/  $\sim 6$  hours 8 minutes) and those blocks are then assigned to each subset in order to achieve a 60%-20%-20% train-validation-test split. To ensure that no parameter dominates others due to their absolute relative values, each subset is rescaled to the interquartile range of the training set in order to account for outliers without leaking information about the validation/test sets during training. Results on the validation dataset from the search are shown in Figure 3.

The optimal window size is 55 measurements ( $\sim 5,500$  seconds,  $\sim 1$  hour 32 minutes), the optimal stride/lead time is 18 measurements ( $\sim 1,800$  seconds,  $\sim 30$  minutes). That is to say, for an MMS measurement at time  $t$ , the input timeseries from Wind runs from time  $t-5,500s-1,800s \approx t-122min$  to time  $t-1,800s \approx t-30min$ . The largest data gap that can be interpolated over is 4.6 minutes ( $\leq 5\%$  of the input window).

Once the optimal dataset structure is found, the optimal model configuration can be determined via hyperparameter search. The nine hyperparameters that are optimized are listed in Table 1, along with the values used for determining the optimal dataset, the optimal values used for the canonical version of PRIME-SH, and the search range for each hyperparameter. The hyperparameter search is conducted using the Hyperband tournament bracket style algorithm (Li et al., 2018) implemented in the KerasTuner API (O’Malley et al., 2019). The meaning of each hyperparameter is described in the following paragraph. After the optimal model configuration is determined, the canonical version of PRIME-SH is optimized on the training dataset for 20 epochs (the maximum before the loss on the validation dataset starts to increase).

	Dataset	HP Test	Canonical Algorithm	HP Range
GRU Layer	192	416		128-640
FCNN Layer 1	352	352		128-640
FCNN Layer 2	48	32		16-128
FCNN Layer 3	N/A	64		16-128
Normalization	Last Layer	Last Layer		Any Combination
Dropout Location	Last Layer	Last Layer		Any Combination
Dropout Rate	20%	35%		20%-50%
Optimizer	Adamax	Adam		Adam, Adamax, Adagrad
Learning Rate	$10^{-4}$	$10^{-4}$		$10^{-3}, 10^{-4}, 10^{-5}$

**Table 1.** Detailed layer sizes and architecture parameters for the test version of PRIME-SH used to optimize the dataset parameters (left column), the canonical version of PRIME-SH determined by hyperparameter search (middle column), and the range of each parameter for which the hyperparameter search was conducted (right column).

The nine hyperparameters are as follows (see also Table 1). The first four are the node sizes of the GRU layer and the following three fully-connected layers. The fifth is where in the algorithm sequence to perform a layer normalization step, which stabilizes neural networks during optimization to reduce the time it takes to optimize them (Ba et al., 2016). Layer normalization normalizes a given layer’s output vector before passing it to the next layer, which speeds up the convergence of the algorithm used to optimize the weights and biases of the algorithm by reducing the extent to which the gradients with respect to the weights in one layer covary with the outputs of the previous layer. The sixth and seventh hyperparameters are the dropout locations and rate used during training. Dropout is a technique to mitigate overfitting that involves randomly removing some percentage of the units from the network every training epoch. This prevents units from co-adapting which can lead to overfitting (Srivastava et al., 2014). The eighth and ninth hyperparameters are the optimization algorithm used to update the weights and biases in the network and that algorithm’s learning rate. Included in the search are the adaptive gradient descent algorithms Adam, Adamax, and Adagrad. An adaptive gradient descent algorithm changes the step size it uses to update parameter weights during optimization to avoid getting stuck in local minima or skipping over minima. Adam updates parameters according to estimates of first order and second moments and has been shown to be suitable for optimizing large algorithms (Kingma & Ba, 2017), Adamax updates parameters according to first order moments and the infinity norm and has been shown to be suitable for recurrent networks (Kingma & Ba, 2017), and Adagrad updates its gradient descent step size per parameter based on the number of updates the parameter receives during training making it suitable for sparse gradients (Duchi et al., 2011). Since each of these conditions could apply to PRIME-SH and the dataset used to optimize it, these three algorithms were included.

## 4 Results

### 4.1 Statistical Performance

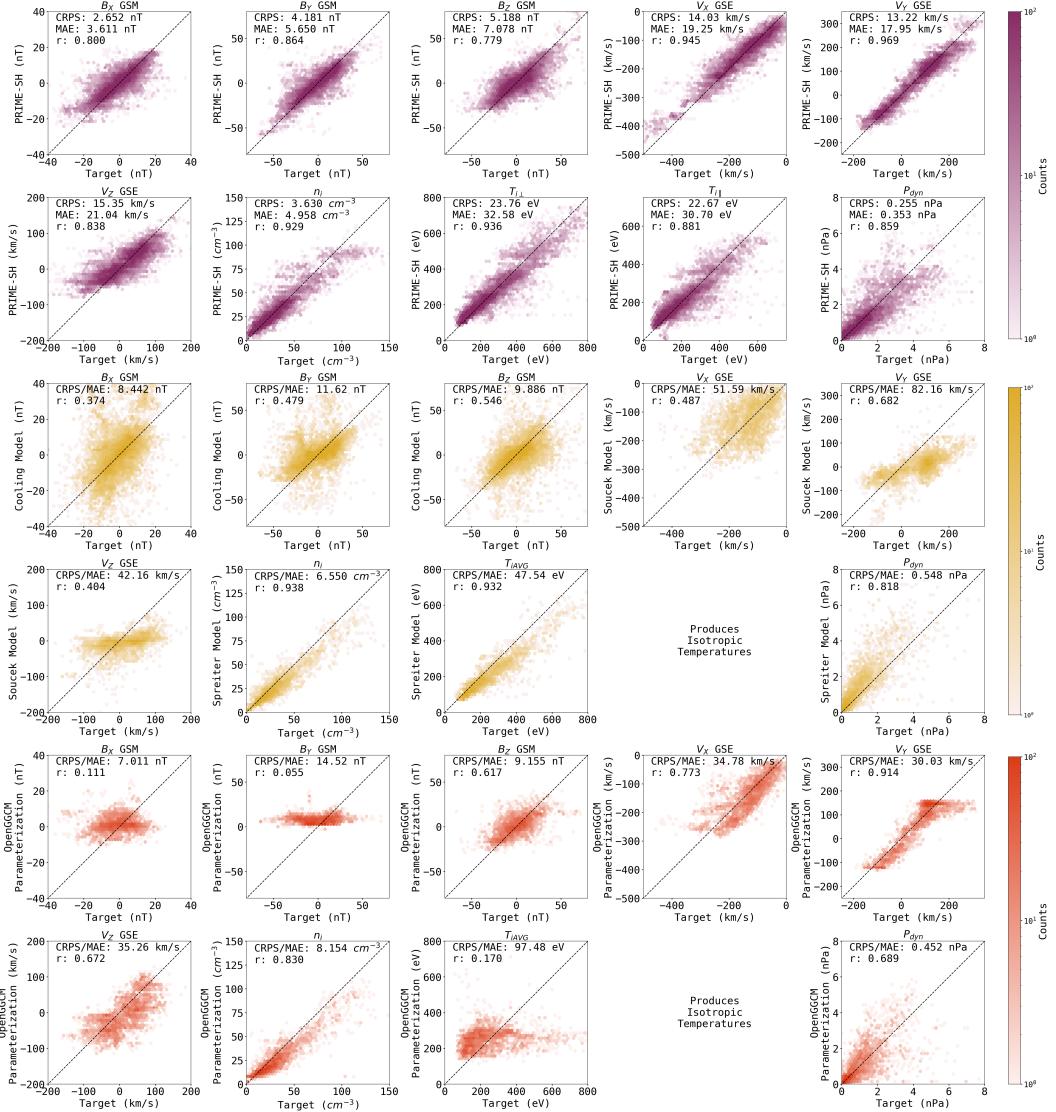
PRIME-SH’s performance is evaluated on the test dataset (not seen by the algorithm at any point during training) by calculating the CRPS between its predictions and the test dataset. Additionally, the mean absolute error (MAE) and Pearson’s r correlation coefficient are calculated between the means of PRIME-SH’s predicted probability distributions and the MMS test set thereby ignoring the uncertainty information. To gain a better sense of the accuracy of PRIME-SH’s predictions in a statistical sense, its

Parameter	PRIME-SH CRPS	PRIME-SH MAE	PRIME-SH r
$B_x$ GSM	2.65nT (0.296 $\sigma$ )	3.61nT (0.403 $\sigma$ )	0.800
$B_y$ GSM	4.18nT (0.245 $\sigma$ )	5.65nT (0.331 $\sigma$ )	0.864
$B_z$ GSM	5.19nT (0.323 $\sigma$ )	7.08nT (0.440 $\sigma$ )	0.779
$V_x$ GSE	14.03km/s (0.182 $\sigma$ )	19.25km/s (0.250 $\sigma$ )	0.945
$V_y$ GSE	13.22km/s (0.127 $\sigma$ )	17.95km/s (0.173 $\sigma$ )	0.969
$V_z$ GSE	15.35km/s (0.291 $\sigma$ )	21.04km/s (0.399 $\sigma$ )	0.838
$n_i$	3.63cm $^{-3}$ (0.169 $\sigma$ )	4.96cm $^{-3}$ (0.231 $\sigma$ )	0.929
$T_{i\perp}$	23.76eV (0.158 $\sigma$ )	32.58eV (0.216 $\sigma$ )	0.936
$T_{i\parallel}$	22.67eV (0.198 $\sigma$ )	30.70eV (0.268 $\sigma$ )	0.881
$P_{dyn}$	0.255nPa (0.224 $\sigma$ )	0.353nPa (0.311 $\sigma$ )	0.859

**Table 2.** Performance of PRIME-SH on the MMS test dataset across continuous rank probability score (CRPS, Equation 1), mean absolute error (MAE), and Pearson’s r correlation coefficient (also shown in Figure 4). CRPS is given in the units of each parameter as well as dimensionless units of standard deviations of each parameter in the MMS training dataset to facilitate comparison between each parameter.

outputs are compared to several analytical models and a parameterization of a popular MHD code for the same MMS-1 test dataset (Figure 4, Table 2). For magnetic field, the model derived in Cooling et al. (2001) is utilized. The Cooling et al. (2001) model essentially “drapes” the interplanetary magnetic field over the Shue et al. (1998) axisymmetric conic section magnetosheath model (based on Kobel and Flückiger (1994)). For magnetosheath flow, the model derived in Soucek and Escoubet (2012) is utilized. The Soucek and Escoubet (2012) model is partially based on Génot et al. (2011) and Kobel and Flückiger (1994), but extends those works to additional magnetopause and bow shock shapes. For density and temperature, the model derived in Spreiter et al. (1966) is utilized. The Spreiter et al. (1966) model is a gas dynamic model that assumes a nondissipative, ideal, compressible, steady flow. Additionally, PRIME-SH is compared to a parameterization of the OpenGGCM MHD code (Raeder et al., 2001, 2008) developed in Jung et al. (2024). This parameterization cannot capture small-scale structure in the MHD code’s outputs, but has been shown to be accurate when compared to observations and is importantly computationally lightweight enough to enable the statistical comparison in this study. The Soucek and Escoubet (2012) and Spreiter et al. (1966) models are implemented in the Mshpy23 package (Jung et al., 2024) and accept one minute resolution OMNI data as input (King & Papitashvili, 2020). The Spreiter et al. (1966) and OpenGGCM models produce isotropic temperatures, therefore their temperatures are compared to the average temperature measured by MMS  $T_{iAV} = (2T_{i\perp} + T_{i\parallel})/3$ . None of the models PRIME-SH is compared to have uncertainty information, therefore the MAE and CRPS reduce to the same form and number (Hersbach, 2000); both metrics are provided for PRIME-SH’s outputs so that all comparisons can be made.

On average, PRIME-SH predicts plasma parameters ( $\vec{v}$ ,  $n_i$ ,  $T_{i\perp}$ , and  $T_{i\parallel}$ ) slightly more accurately than magnetic field parameters. This is possibly due to the fact that fluctuations in magnetic field happen at a shorter timescale than those in the plasma, and neural networks tend to have more difficulty representing smaller scale variations than larger scale ones whether temporal or spatial in nature. PRIME-SH has a Pearson’s r higher than 0.75 for every parameter. There are no strong biases or systematic errors visible in Figure 4, only some amount of regression to the mean in the most extreme values of  $V_x$  and  $n_i$  (and therefore in  $P_{dyn}$  as well). Interestingly, PRIME-SH predicts magnetosheath conditions almost as accurately as its progenitor algorithm PRIME predicts solar wind conditions given the same type of input data from L1 (PRIME-SH’s average



**Figure 4.** Joint distributions of MMS-1 data (x axis) with predicted parameters from PRIME-SH (purple, top), three analytical magnetosheath models (yellow, middle), and a parameterization of the OpenGGCM MHD code (orange, bottom). CRPS, the mean absolute error (MAE), and Pearson's  $r$  correlation coefficient for each parameter shown in the top left of each distribution. The MAE is calculated between the peaks of PRIME-SH's predicted distributions and each MMS observation (thereby throwing away uncertainty information). A perfect prediction corresponds to the line  $y = x$ , plotted overtop of each distribution for convenience.

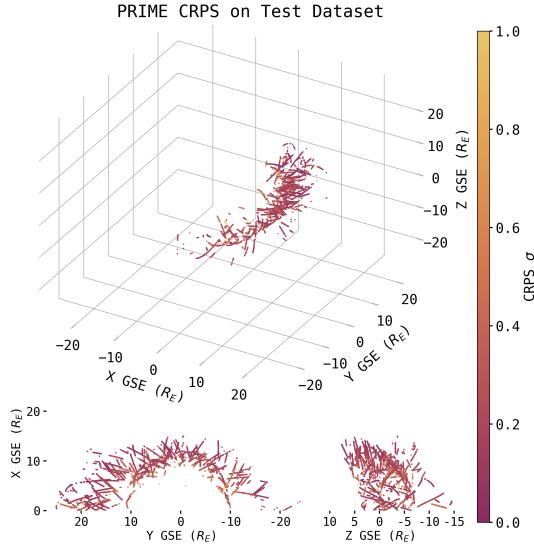
335 CRPS of  $0.221\sigma$  and PRIME's average CRPS of  $0.214\sigma$ ), despite that it has to represent  
 336 not only the physics of the solar wind's propagation from L1 to Earth but the physics  
 337 of the bow shock as well.

338 There is no parameter for which PRIME-SH is outperformed by any of the analytical  
 339 models considered here with respect to MAE or CRPS. For each component in  
 340 the magnetic field, PRIME-SH predicts MMS-1 observations more accurately than the  
 341 Cooling et al. (2001) model. Specifically, PRIME-SH's CRPS and MAE are both lower  
 342 than the Cooling et al. (2001) model's MAE, and PRIME-SH's Pearson's r is higher than  
 343 the Cooling et al. (2001) model's Pearson's r. There appears to be some systematic over-  
 344 prediction in the Cooling et al. (2001) model's outputs for  $B_X$ . This means that PRIME-  
 345 SH reproduces the actual magnetic field in the magnetosheath given upstream conditions  
 346 more accurately than the Cooling et al. (2001) model, but whether it produces a phys-  
 347 ically accurate draped field must be separately validated in Section 4.2.1. The Soucek  
 348 and Escoubet (2012) model has a large variance in  $V_X$  and does not reproduce fast flows  
 349 ( $> 300 \text{ km/s}$ ) as accurately as PRIME-SH does. It also underpredicts  $V_Y$  and  $V_Z$ , all  
 350 of which could be regression effects due to model outputs being too "smooth". The Spreiter  
 351 et al. (1966) comes the closest to outperforming PRIME-SH of any model considered here,  
 352 but still does not predict  $n_i$  or  $T_{iAVG}$  more accurately than PRIME-SH.

353 Compared to the parameterized MHD model, PRIME-SH has higher representa-  
 354 tional power and therefore higher accuracy across each parameter. For  $B_X$  and  $B_Y$  the  
 355 parameterized MHD model does not vary much at all (both have Pearson's  $r < 0.12$ ),  
 356 which could be consistent with the results presented in Jung et al. (2024) Figures 2, 3,  
 357 and 4. For plasma flow velocity, the parameterized MHD model clearly reaches the bounds  
 358 of its parameterization (most visible for  $V_Y < -120 \text{ km/s}$  and  $V_Y > 160 \text{ km/s}$ ). The  
 359 odd shape of the distribution for  $T_{iAVG}$  is also consistent with results presented in Jung  
 360 et al. (2024) Figure 2. The MHD model is more accurate than the associated analyti-  
 361 cal model for all parameters except  $B_Y$ ,  $n_i$ ,  $T_{iAVG}$ , and  $P_{dyn}$ , but is not more accurate  
 362 than PRIME-SH for any of the parameters it is capable of predicting.

363 PRIME-SH is a 3D model, and its outputs are valid over any regions covered by  
 364 MMS-1's orbit on the dayside (GSE  $X > 0R_E$ , GSE  $|Y| < 5R_E$ ). Since the magne-  
 365 tosheath has a great deal of variation across its extent, PRIME-SH's accuracy evaluated  
 366 against the test set is displayed in GSE coordinates in Figure 5. In general, PRIME-SH's  
 367 outputs are generally less accurate on predictions closer to the Earth than on those fur-  
 368 ther from the Earth. This suggests that PRIME-SH is less accurate during periods where  
 369 the magnetosheath is highly compressed or when it makes predictions close to the mag-  
 370 netopause. These periods are rare relative to nominal conditions in the training dataset,  
 371 so PRIME-SH being somewhat less accurate under these conditions is expected and should  
 372 be taken into account when using PRIME-SH. Also to be taken into account by poten-  
 373 tial users is that PRIME-SH has not been trained outside of the areas shown in Figure  
 374 5 and using it to predict magnetosheath conditions outside those areas is not recommended  
 375 by the authors.

376 Since reliability is not enforced by the CRPS loss function during training, PRIME-  
 377 SH's output uncertainties must be validated quantitatively through the use of a reli-  
 378 ability diagram (Hamill, 1997, 2001). Following the procedure in Camporeale et al. (2019)  
 379 and Camporeale and Carè (2021), the standardized errors associated with prediction  $\mu_i, \sigma_i$   
 380 with  $i = 1, \dots, N$  are defined as  $\eta_i = (y_{obs,i} - \mu_i)/(\sqrt{2}\sigma_i)$ . The probability density of  
 381 a given Gaussian forecast is therefore  $\Phi_i = \frac{1}{2}[\text{erf}(\eta_i) + 1]$ , allowing the reliability dia-  
 382 gram to be constructed from the empirical cumulative distribution of  $\Phi_i$  given by  $C(\phi) =$   
 383  $\frac{1}{N} \sum_{i=1}^N H(\phi - \Phi_i)$  (with  $H$  being the Heaviside step function).  $C(\phi)$  is the observed  
 384 frequency as a function of the predicted frequency, the same as reliability diagrams of  
 385 forecasts of discrete events (e.g. those in Hamill (1997)). This method has the benefit  
 386 of not requiring binning, which has been shown to affect the results of reliability diagrams



**Figure 5.** PRIME-SH’s accuracy on the test dataset averaged across all nine target parameters in dimensionless standard deviation units ( $\sigma$ ). Targets arranged spatially in 3D (top), the GSE X-Y plane (bottom left), and the GSE X-Z plane (bottom right).

of discrete events (Bröcker & Smith, 2007).  $C(\phi)$  is calculated for all observations in the test dataset for each parameter and presented in Figure 6.

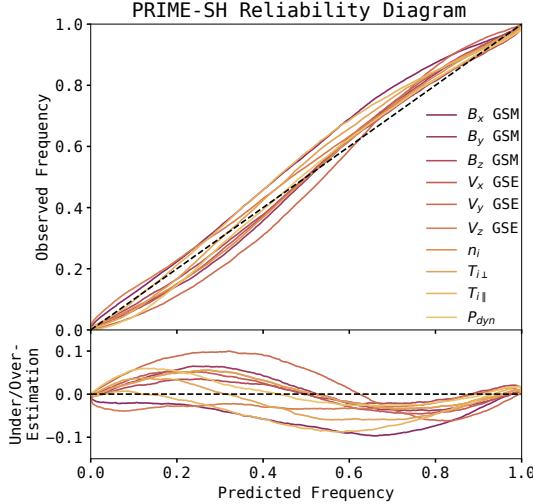
PRIME-SH is not perfectly reliable (its reliability diagram does not exactly follow the dashed line in Figure 6); it generally tends to overestimate the likelihood of unlikely events, and underestimate the likelihood of likely events. With the exception of  $V_Z$ ,  $B_X$ , and  $T_{||}$ , PRIME-SH tends to be conservative. This is not unexpected, as even models perfectly calibrated on training data can suffer calibration loss on the test dataset (Kull & Flach, 2015). The largest departures from perfect calibration are observed in  $V_Y$  GSE (predicts events that occur with  $p = 0.221$  as occurring with  $p = 0.320$ ),  $B_X$  GSM (predicts events that occur with  $p = 0.754$  as occurring with  $p = 0.657$ ), and  $T_{||}$  (predicts events that occur with  $p = 0.674$  as occurring with  $p = 0.586$ ). On average PRIME-SH is reliable to within 3.5% with a maximum difference 10% (calculated  $p_{obs} - p_{pred}$ ). This is roughly as reliable as its progenitor algorithm PRIME and other probabilistic prediction algorithms for space weather tasks (e.g. Tasistro-Hart et al. (2021)), but less reliable than those that use loss functions that enforce reliability explicitly (e.g. Hu et al. (2022)).

#### 4.2 Physical Validation

While a model’s accuracy and reliability are important to quantify in a statistical sense, it is also important to verify that a model can reproduce expected physics. This is especially important for neural network models, which can relatively easily overfit and reproduce a dataset’s noise rather than the underlying data representation or physics. In the following sections PRIME-SH’s outputs for synthetic data are investigated to ensure that it can reproduce magnetic field and plasma physics in the magnetosheath.

##### 4.2.1 Field Line Draping and Uncertainty

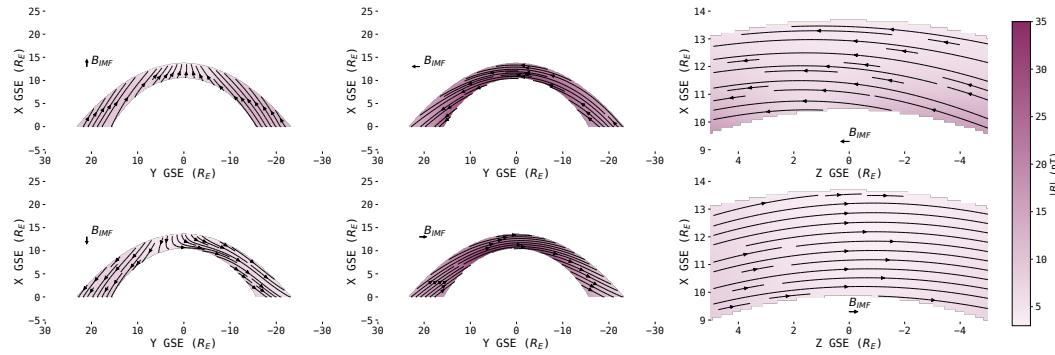
Since the interplanetary magnetic field is frozen into the solar wind plasma, as the plasma is shocked and diverts around the magnetopause the magnetic field “drapes” over



**Figure 6.** Reliability diagram constructed from PRIME-SH’s outputs on the test dataset for each parameter. Shown versus the predicted frequency of the observation from PRIME-SH are the value of the observed frequency (top) and the deviation from perfect reliability (bottom). For the bottom plot, a given parameter being over (under) the line by an amount corresponds to PRIME-SH over (under) predicting the frequency by that amount.

the obstacle forming a tangential discontinuity at the magnetopause (Crooker et al., 1985). In order to verify that PRIME-SH captures this feature of the magnetosheath, outputs are generated on a grid of points for the same input data. The grid is chosen to lie in the GSE X-Y or GSE X-Z plane (depending on IMF orientation) with a grid scale of  $0.1R_E$ . All grid cells inside the Shue et al. (1998) or outside the Jelínek et al. (2012) bow shock (calculated using the conditions at L1 used as inputs for PRIME-SH) are left unused. Only grid cells in regions well sampled by the MMS training data are included, hence the Z extent is restricted to  $\pm 5R_E$  away from the ecliptic and the nightside is not included (see Figure 1). The input data are chosen to be a  $400\text{km/s}$  solar wind only in the GSE X direction with otherwise average solar wind conditions from the Wing L1 dataset:  $|B| = 5.34nT$ ,  $V_X = -400\text{km/s}$ ,  $V_Y = 0\text{km/s}$ ,  $V_Z = 0\text{km/s}$ ,  $n_i = 7.12\text{cm}^{-3}$ , and  $v_{th} = 34.9\text{km/s}$ . In order to investigate whether PRIME-SH is capable of draping, conditions on the grid are calculated for six different IMF orientations: one radial toward Earth (cone angle  $0^\circ$ ), one downward (cone angle  $-90^\circ$ ), one duskward (cone angle  $+90^\circ$ ) one radial away from Earth (cone angle  $180^\circ$ ), one purely northward (clock angle  $0^\circ$ ), and one purely southward (clock angle  $180^\circ$ ). Shown in Figure 7 are these six grids, with the sheath magnetic field streamlines plotted in black arrows and the magnitude of B in each cell in color.

As can be seen in Figure 7, PRIME-SH reproduces the draping of the magnetic field in the magnetosheath well despite the frozen in condition not being enforced during training. For cone angles of  $\pm 90^\circ$  the magnetic field piles up at the nose of the magnetopause, much more than it does for radial IMF. This can be seen in the magnitude of the magnetic field, which is higher at the nose than the flanks for cone angles of  $\pm 90^\circ$ . For cone angles of  $0^\circ$  or  $180^\circ$ , the flanks have a relatively higher magnetic field than the nose (though it is not as strong as the field at the nose in the cone angle  $\pm 90^\circ$  case). For northward IMF, somewhat more magnetic field pileup is observed at the northern and southern flanks than for the southern IMF case. The magnetic field magnitude is also slightly higher overall in the northward IMF case than in the southward IMF case. These maps



**Figure 7.** Magnetosheath conditions output by PRIME-SH using synthetic data for six different IMF orientations (Shown with arrows in top left or bottom). Plasma conditions are average conditions from the input dataset, magnetic field magnitude is  $5.34nT$  (the average magnitude from the input dataset). Shown in color is the magnitude of  $B$ , and the arrows are  $B_X$  and  $B_Y$  GSM field lines (for the left four plots) or the  $B_X$  and  $B_Z$  GSM field lines (for the right two plots).

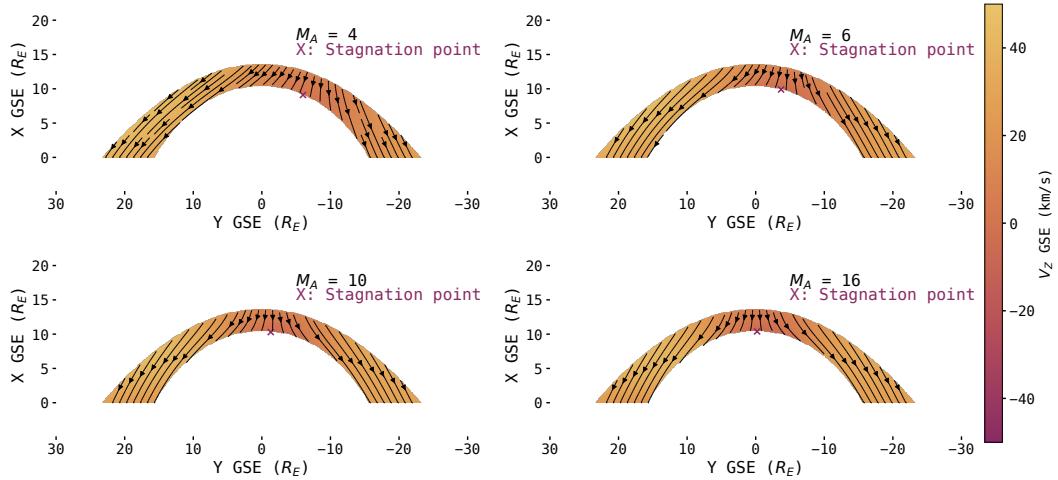
suggest that a lower reconnection rate at the nose causes magnetic field pileup and rearrangement in the sheath as many studies have predicted.

#### 4.2.2 Stagnation Point

As the solar wind plasma diverts and is slowed around the magnetopause, a region known as the stagnation point develops where there is very little to no plasma flow (Spreiter et al., 1966). For radial flow and typical Parker spiral magnetic field orientation, this point is thought to be roughly located at the nose of the magnetopause (with slight aberration from Earth's  $\approx 30\text{km/s}$  motion in the negative GSE Y direction). MHD theory predicts that for a Parker spiral IMF, the stagnation point should deflect downward for solar wind flows with low Alfvén Mach numbers (Russell et al., 1981). Here PRIME-SH is used to assemble predictions on more  $0.1R_E$  grids of the same configuration as Section 4.2.1, however this time the Alfvén Mach number of the synthetic dataset is varied from  $M_A = 4$  to  $M_A = 16$  (the solar wind typically has  $M_A \approx 10$ ). The density and velocity are held the same ( $n_i = 7.12$ ,  $V_X = -400\text{km/s}$ ) and the magnetic field is kept at a  $45^\circ$  Parker spiral as its magnitude is decreased in steps from  $12nT$  to  $2.4nT$  to yield the four Alfvén Mach numbers. Shown in Figure 8 are these four grids, with the X and Y GSE plasma flow velocity depicted with black arrows and the Z GSE flow velocity in color. Also depicted is the stagnation point, marked with a purple X.

As can be seen in Figure 7, PRIME-SH produces continuous flow maps that divert around the magnetopause for all four Alfvén mach numbers. Additionally, as the Alfvén Mach number decreases the stagnation point is observed to move downward as predicted by MHD theory and simulations. This feature is hard to observe using in-situ instruments, but here through what is essentially a spatio-temporal inversion the feature is shown to occur in reality.

One interesting feature is that there appears to be some weak dawn-dusk asymmetry in the flow velocity maps produced by PRIME-SH. This could be due to biases in MMS-1's orbit showing up in PRIME-SH's outputs, as the asymmetry does not appear in MHD simulations of the magnetosheath. However, other experimental work has also found dawn-dusk asymmetries in the magnetosheath properties (Walsh et al., 2012; Dimmock & Nykyri, 2013).



**Figure 8.** Magnetosheath conditions output by PRIME-SH using synthetic data at four different Alfvén Mach numbers ( $M_A = 4, 6, 10, 16$ ). Average plasma conditions from the input dataset are used, magnetic field is a Parker spiral orientation whose magnitude is varied for each case to result in the four Alfvén Mach numbers. Shown in color is the Z GSE velocity, and the arrows are the X and Y GSE velocity. The point of minimum velocity in the sheath (the stagnation point) is marked with the purple X.

#### 4.2.3 Shock Jump Conditions

Shocks, whether they are collisional or collisionless, conserve mass, momentum and energy. The Rankine-Hugoniot shock jump conditions are formulations of each of these conservation laws in terms of the conditions upstream and downstream of the shock. For an MHD shock, define the shock normal direction to be  $\hat{n}$ , the plasma flow velocity to be  $\vec{v}$ , the plasma mass density to be  $\rho$ , the thermal pressure to be  $P$ , the specific heat ratio to be  $\gamma$ , and the magnetic field to be  $\vec{B}$ . For some quantity  $\vec{X}$  upstream and downstream of the shock, define the notation  $\vec{X}_{up} - \vec{X}_{down} = [\vec{X}]$ . Mass conservation upstream and downstream of the shock can then be written:

$$[\rho \vec{u} \cdot \hat{n}] = 0 \quad (3)$$

Momentum conservation (with magnetic pressure included) can be written:

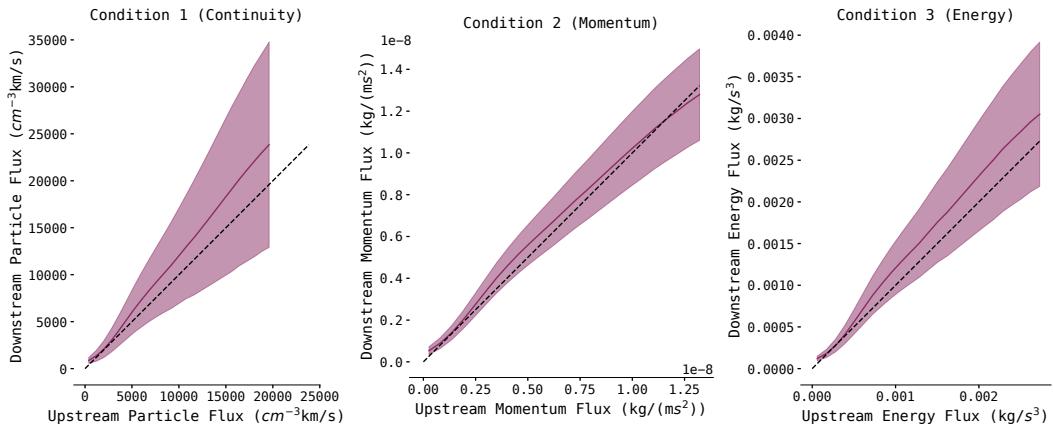
$$[\rho \vec{u} \cdot (\vec{u} \cdot \hat{n}) + (P + \frac{\vec{B}^2}{2\mu_0}) \hat{n} - \frac{(\vec{B} \cdot \hat{n}) \vec{B}}{\mu_0}] = 0 \quad (4)$$

Energy conservation can be written:

$$[\vec{u} \cdot \hat{n} \left( \frac{\rho \vec{u}^2}{2} + \frac{\gamma}{\gamma - 1} P + \frac{\vec{B}^2}{\mu_0} \right) - \frac{(\vec{B} \cdot \hat{n})(\vec{B} \cdot \vec{u})}{\mu_0}] = 0 \quad (5)$$

(Kallenrode, 2010).

None of these conditions are explicitly enforced during training, but they are part of the underlying physics PRIME-SH should be representing. To validate that PRIME-SH reproduces these conservation laws, a range of synthetic solar wind conditions with densities ranging from  $1\text{cm}^{-3}$  to  $50\text{cm}^{-3}$  with  $V_{\text{GSE}} = -400\text{km/s}$ ,  $\vec{B} = (-4nT)\hat{x} + (-4nT)\hat{y}$ , and  $v_{th} = 30\text{km/s}$  are initialized and used to generate predictions just behind the Jelínek et al. (2012) bow shock nose. This range was chosen to reflect the full



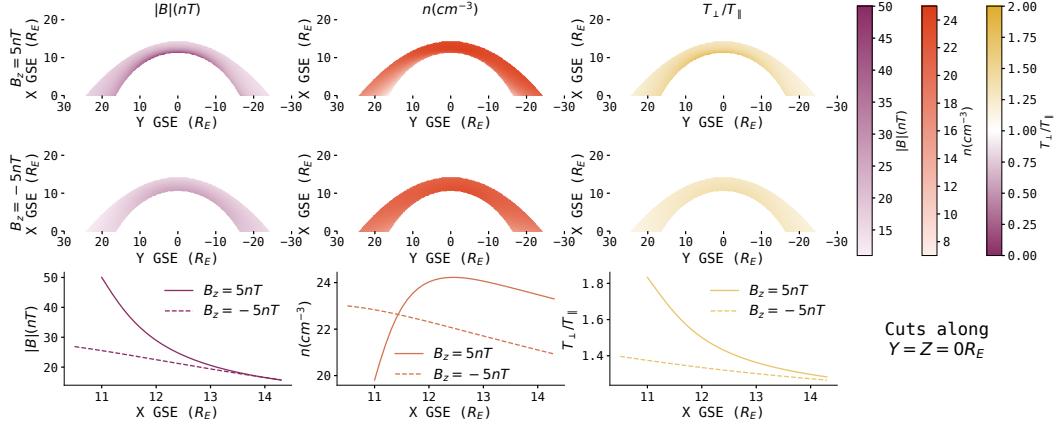
**Figure 9.** Particle, momentum, and energy fluxes calculated across a range of synthetic input conditions roughly corresponding to the range of the training dataset. Fluxes are calculated just upstream of the bow shock nose (using the input data) and just downstream (using PRIME-SH’s outputs), and uncertainties are calculated by propagating PRIME-SH’s predicted uncertainties through the MHD shock jump condition equations. Within PRIME-SH’s predicted uncertainties the three Rankine-Hugoniot MHD jump conditions are obeyed.

range of densities from the input dataset, which results in better coverage of the range of the three upstream fluxes observed than varying other conditions such as velocity. Equations 3, 4, and 5 are used to calculate the particle, momentum, and energy flux from the synthetic input data (upstream) and from PRIME-SH’s outputs (downstream). The uncertainties predicted by PRIME-SH can be propagated through Equations 3, 4, and 5 to obtain uncertainties for the downstream fluxes as well. The downstream fluxes are plotted as a function of upstream fluxes in Figure 9.

Perfect conservation of each flux is represented by the dashed lines in Figure 9. As can be seen, while the quantities predicted by PRIME-SH do not perfectly conserve mass/particles, momentum, and energy, it does conserve them within the the  $1\sigma$  uncertainty bounds for each quantity. One contribution to this uncertainty is an experimental one. Although the instruments on Wind and MMS have been carefully calibrated, they were not calibrated together. Previous studies have found mismatches when comparing plasma and magnetic field parameters from different missions, even those with very similar instruments (King, 2005; Roberts et al., 2021). The points of largest *fractional* difference between upstream and downstream fluxes occur for the smallest fluxes (when  $n_{up} = 1\text{ cm}^{-3}$ ), which happens relatively infrequently in the input dataset. Despite the fact that mass/particle conservation, momentum conservation, and energy conservation were not explicitly enforced during training, PRIME-SH has been optimized such that it successfully represents the underlying physics to a degree that the three quantities are conserved.

#### 4.2.4 Plasma Depletion Layer

The plasma depletion layer is a transient region of the subsolar magnetosheath characterized by decreased density and increased magnetic field strength. This layer exists when the reconnection rate at the magnetopause is insufficient to prevent “pile-up” of magnetic flux, and as such is typically observed during periods of northward IMF (although it can sometimes be observed during periods of southward IMF). This “pile-up” can modify the local reconnection rate, and could even enable reconnection at the subsolar magnetopause for northward IMF (Anderson, 1996). It has also been shown that



**Figure 10.** Magnetosheath conditions output by PRIME-SH for synthetic input conditions with IMF purely northward ( $B_z = 5nT$ ) and purely southward ( $B_z = -5nT$ ). Average plasma conditions from the input dataset are used. Top row shows  $|B|$ ,  $n_i$ , and  $T_{\perp}/T_{\parallel}$  for  $B_z = 5nT$ , middle row shows the same for  $B_z = -5nT$ , and bottom row shows cuts along  $Y = Z = 0R_E$  for both magnetic field orientations for ease of comparison.

the plasma depletion layer has stronger temperature anisotropy than the rest of the magnetosheath, although it is currently unclear whether this is a formation mechanism of the region or simply a consequence of the flux pile-up (Phan & Paschmann, 1996). Despite the fact that the plasma depletion layer has been observed by in-situ spacecraft for many years (Cummings & Coleman, 1968), the dynamics and global geometry of the region is difficult to determine from observations due to their spatio-temporal ambiguity (Wang et al., 2004).

Both to verify PRIME-SH has been properly trained to replicate solar wind flow around the magnetosphere and to overcome the spatio-temporal ambiguity of in-situ observations, PRIME-SH is used to assemble predictions on more grids of the same configuration as Section 4.2.1 for northward ( $\vec{B} = 5nT\hat{z}$ ) and southward ( $\vec{B} = -5nT\hat{z}$ ) IMF. Plasma conditions are the same between each run ( $V_{GSE} = -400km/s$ ,  $n = 5cm^{-3}$ , and  $v_{th} = 30km/s$ , Alfvén Mach number 8). The magnetic field magnitude, density, and temperature anisotropy ( $T_{\perp}/T_{\parallel}$ ) are shown for each configuration in Figure 10 in the ecliptic and in cuts along the Sun-Earth line.

The plasma depletion layer can be readily identified in Figure 10 as the region of high  $|B|$ ,  $T_{\perp}/T_{\parallel}$  and low  $n$  close to the subsolar point in the northward IMF case that is not apparent in the southward IMF case. In the cuts along the Sun-Earth line, the density can be more readily observed to begin falling off about  $1R_E$  from the magnetopause, while at the same time  $|B|$  and  $T_{\perp}/T_{\parallel}$  begin to increase. This is contrasted with the southward case, in which all three parameters increase across the sheath somewhat linearly. This thickness is consistent with reported thicknesses from the literature which range from  $0.3R_E$  to  $1R_E$  for  $M_A = 8$ , depending on identification criteria (Wang et al., 2004). This validates that PRIME-SH has been trained to reproduce magnetic flux pile-up and its effects in the magnetosheath. Unlike numerical simulations, this is a spatial map of the plasma depletion layer based directly on observations rather than physical assumptions, which can cause deviation between predicted and observed global depletion layer configurations (Zwan & Wolf, 1976; Southwood & Kivelson, 1995).

545    **5 Conclusions**

546    A Bayesian recurrent neural network is trained to predict MMS-1 observations of  
 547    Earth's magnetosheath given timeseries input measured by the Wind spacecraft at L1.  
 548    This algorithm, called PRIME-SH in reference to its progenitor algorithm PRIME,  
 549    incorporates the time history of the solar wind at L1 to generate probability distributions  
 550    for magnetosheath plasma and magnetic field parameters. These probability distribu-  
 551    tions can be used to determine the uncertainty associated with PRIME-SH's predictions.

552    PRIME-SH is shown to have good performance in a statistical sense across a test  
 553    dataset of MMS-1 data not used during training (Average CRPS  $0.221\sigma$ ). The uncer-  
 554    tainties predicted by PRIME-SH are shown to be reliable to within 3.5% with a max-  
 555    imum difference 10% through a comparison to the test dataset. Additionally, PRIME-  
 556    SH predicts magnetosheath conditions more accurately than several popular analytical  
 557    models (Spreiter et al., 1966; Kobel & Flückiger, 1994; Cooling et al., 2001; Soucek &  
 558    Escoubet, 2012) and a parameterization of the OpenGGCM MHD code (Jung et al., 2024).  
 559    While statistical validation is important, it is also important to validate that a model  
 560    is indeed producing physical results. It is verified that the magnetic field values produced  
 561    by PRIME-SH across a grid of points in the magnetosheath "drape" across the magne-  
 562    topause in 3D for several different orientations of the upstream magnetic field. Plasma  
 563    flow velocities output by PRIME-SH across a grid of magnetosheath points divert around  
 564    the magnetopause as expected, and the point at which the flow stagnates moves dawn-  
 565    ward with decreasing Alfvén Mach number as predicted by MHD theory (Russell et al.,  
 566    1981). PRIME-SH is shown to conserve particle/mass flux, momentum flux, and energy  
 567    flux within  $1\sigma$  uncertainty across the bow shock for the range of input parameters it is  
 568    trained on. PRIME-SH is also capable of reproducing the plasma depletion layer given  
 569    input conditions for which the depletion layer is expected to form. From this it may be  
 570    concluded that PRIME-SH has indeed been optimized to represent the physics of solar  
 571    wind flow from L1, through the bow shock, and around the magnetopause.

572    PRIME-SH is not only more accurate in a statistical sense than current analyti-  
 573    cal models and MHD simulation parameterizations, but it also has additional function-  
 574    ality these other models do not. First, PRIME-SH outputs  $T_{\perp}$  and  $T_{\parallel}$  separately, kinetic  
 575    physics that MHD and analytical models currently lack. Additionally, PRIME-SH out-  
 576    puts uncertainties for with its outputs. These uncertainties were used in this study to  
 577    assign confidence intervals to fluxes calculated to verify that PRIME-SH conserves par-  
 578    ticles, mass, and energy. They could additionally be used to in more advanced techniques  
 579    such as regression recalibration or ensemble modeling. For example, it has been argued  
 580    using 3 month cadence averaged THEMIS data that the solar wind electric field and Poynt-  
 581    ing flux that can be driven across the bow shock into the magnetosheath saturate at their  
 582    highest values (Pulkkinen et al., 2016), which could be due to physics occurring at the  
 583    bow shock or due to statistical effects to due with uncertainty in assigning solar wind  
 584    conditions to magnetosheath observations. PRIME-SH's outputs can be used to expand  
 585    the dataset, improve its temporal resolution, and correct for the uncertainty in the so-  
 586    lar wind and magnetosheath conditions. It has also been argued that the solar wind alone  
 587    controls the rate of energy transfer across the magnetopause due to rearrangements of  
 588    the magnetosheath flow (Lopez, 2016), which can be investigated using PRIME-SH as  
 589    a kind of spatio-temporal inversion algorithm. In short, PRIME-SH is an accurate and  
 590    lightweight magnetosheath prediction algorithm that offers functionality no other mag-  
 591    netosheath prediction algorithm does, and enables new statistical and event-based stud-  
 592    ies of the magnetosheath.

593    **Acknowledgments**

594    Magnetospheric Multiscale, Wind, and OMNI data are available through the Coordinated  
 595    Data Analysis Web (CDAWeb) online portal at [https://cdaweb.gsfc.nasa.gov/istp\\_public/](https://cdaweb.gsfc.nasa.gov/istp_public/). Codes for dataset preparation, algorithm development, and analysis presented

in this paper are available at <https://github.com/connor-obrien888/primesh>. Authors CO, BMW, and YZ would like to acknowledge support from NASA grants 80NSSC21K0026 and 80NSSC20K1710. Author ST acknowledges support from the German Aerospace Center (DLR). DGS was supported by NASA's MMS Theory and Modeling program. The authors acknowledge the instrument teams for FPI, FGM, SWE, and MFI, as well as the other MMS and Wind instrument teams whose labor made this study possible.

## 603 References

- 604 Al Shidi, Q., Pulkkinen, T., & Welling, D. (2023, February). *Uncertainty Quantification in Global Simulations: Solar Wind Propagation Effects* (other). oral.  
 605 Retrieved 2023-02-28, from <https://meetingorganizer.copernicus.org/EGU23/EGU23-11764.html> doi: 10.5194/egusphere-egu23-11764
- 606 Anderson, J. L. (1996, July). A Method for Producing and Evaluating Probabilistic Forecasts from Ensemble Model Integrations. *Journal of Climate*, 9(7),  
 607 1518–1530. Retrieved 2023-07-21, from [http://journals.ametsoc.org/doi/10.1175/1520-0442\(1996\)009<1518:AMFPAE>2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0442(1996)009<1518:AMFPAE>2.0.CO;2) doi:  
 608 10.1175/1520-0442(1996)009<1518:AMFPAE>2.0.CO;2
- 609 Axford, W. (1964, January). Viscous interaction between the solar wind and the earth's magnetosphere. *Planetary and Space Science*, 12(1), 45–53. Retrieved 2021-11-28, from <https://linkinghub.elsevier.com/retrieve/pii/0032063364900674> doi: 10.1016/0032-0633(64)90067-4
- 610 Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016, July). *Layer Normalization*. arXiv. Retrieved 2023-07-18, from <http://arxiv.org/abs/1607.06450> (arXiv:1607.06450 [cs, stat])
- 611 Bebis, G., & Georgopoulos, M. (1994, October). Feed-forward neural networks. *IEEE Potentials*, 13(4), 27–31. Retrieved 2023-07-19, from <http://ieeexplore.ieee.org/document/329294/> doi: 10.1109/45.329294
- 612 Borovsky, J. E. (2018, October). The spatial structure of the oncoming solar wind at Earth and the shortcomings of a solar-wind monitor at L1. *Journal of Atmospheric and Solar-Terrestrial Physics*, 177, 2–11. Retrieved 2023-03-13, from <https://linkinghub.elsevier.com/retrieve/pii/S1364682617300159> doi: 10.1016/j.jastp.2017.03.014
- 613 Borovsky, J. E. (2021, February). Is Our Understanding of Solar-Wind/Magnetosphere Coupling Satisfactory? *Frontiers in Astronomy and Space Sciences*, 8, 634073. Retrieved 2023-04-13, from <https://www.frontiersin.org/articles/10.3389/fspas.2021.634073/full> doi: 10.3389/fspas.2021.634073
- 614 Bröcker, J., & Smith, L. A. (2007, June). Increasing the Reliability of Reliability Diagrams. *Weather and Forecasting*, 22(3), 651–661. Retrieved 2023-07-24, from <http://journals.ametsoc.org/doi/10.1175/WAF993.1> doi: 10.1175/WAF993.1
- 615 Burch, J. L., Moore, T. E., Torbert, R. B., & Giles, B. L. (2016, March). Magnetospheric Multiscale Overview and Science Objectives. *Space Science Reviews*, 199(1-4), 5–21. Retrieved 2021-12-01, from <http://link.springer.com/10.1007/s11214-015-0164-9> doi: 10.1007/s11214-015-0164-9
- 616 Camporeale, E., & Carè, A. (2021). ACCRUE: ACCURATE AND RELIABLE UNCERTAINTY ESTIMATE IN DETERMINISTIC MODELS. *International Journal for Uncertainty Quantification*, 11(4), 81–94. Retrieved 2023-01-05, from <http://www.dl.begellhouse.com/journals/52034eb04b657aea,3ec0b84376cff3d2,1801e97431c5911b.html> doi: 10.1615/Int.J.UncertaintyQuantification.2021034623
- 617 Camporeale, E., Chu, X., Agapitov, O. V., & Bortnik, J. (2019, March). On the Generation of Probabilistic Forecasts From Deterministic Models. *Space Weather*, 17(3), 455–475. Retrieved 2023-07-21, from <https://onlinelibrary>

- 650 .wiley.com/doi/abs/10.1029/2018SW002026 doi: 10.1029/2018SW002026
- 651 Cho, K., van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014, October). *On*  
652 *the Properties of Neural Machine Translation: Encoder-Decoder Approaches*.  
653 arXiv. Retrieved 2023-07-19, from <http://arxiv.org/abs/1409.1259>  
654 (arXiv:1409.1259 [cs, stat])
- 655 Cooling, B. M. A., Owen, C. J., & Schwartz, S. J. (2001, September). Role of the  
656 magnetosheath flow in determining the motion of open flux tubes. *Journal of*  
657 *Geophysical Research: Space Physics*, 106(A9), 18763–18775. Retrieved 2023-  
658 07-12, from <http://doi.wiley.com/10.1029/2000JA000455> doi: 10.1029/  
659 2000JA000455
- 660 Crooker, N. U., Luhmann, J. G., Russell, C. T., Smith, E. J., Spreiter, J. R., &  
661 Stahara, S. S. (1985, April). Magnetic field draping against the dayside magne-  
662 topause. *Journal of Geophysical Research: Space Physics*, 90(A4), 3505–3510.  
663 Retrieved 2024-01-05, from [https://agupubs.onlinelibrary.wiley.com/](https://agupubs.onlinelibrary.wiley.com/doi/10.1029/JA090iA04p03505)  
664 doi: 10.1029/JA090iA04p03505 doi: 10.1029/JA090iA04p03505
- 665 Cummings, W. D., & Coleman, P. J. (1968, September). Magnetic fields in  
666 the magnetopause and vicinity at synchronous altitude. *Journal of Geo-*  
667 *physical Research*, 73(17), 5699–5718. Retrieved 2024-03-12, from <http://doi.wiley.com/10.1029/JA073i017p05699> doi: 10.1029/JA073i017p05699
- 668 Dimmock, A. P., & Nykyri, K. (2013, August). The statistical mapping of mag-  
669 netosheath plasma properties based on THEMIS measurements in the mag-  
670 netosheath interplanetary medium reference frame: MAGNETOSHEATH  
671 STATISTICAL MAPPING. *Journal of Geophysical Research: Space Physics*,  
672 118(8), 4963–4976. Retrieved 2020-11-11, from <http://doi.wiley.com/10.1002/jgra.50465> doi: 10.1002/jgra.50465
- 673 Duchi, J., Hazan, E., & Singer, Y. (2011). Adaptive Subgradient Methods for Online  
674 Learning and Stochastic Optimization.
- 675 Dungey, J. W. (1961, January). Interplanetary Magnetic Field and the Au-  
676 roral Zones. *Physical Review Letters*, 6(2), 47–48. Retrieved 2021-11-  
677 28, from <https://link.aps.org/doi/10.1103/PhysRevLett.6.47> doi:  
678 10.1103/PhysRevLett.6.47
- 679 Gneiting, T., Raftery, A. E., Westveld, A. H., & Goldman, T. (2005, May). Cali-  
680 brated Probabilistic Forecasting Using Ensemble Model Output Statistics and  
681 Minimum CRPS Estimation. *Monthly Weather Review*, 133(5), 1098–1118.  
682 Retrieved 2023-05-22, from <http://journals.ametsoc.org/doi/10.1175/MWR2904.1> doi: 10.1175/MWR2904.1
- 683 Génot, V., Broussiou, L., Budnik, E., Hellinger, P., Trávníček, P. M., Lucek, E., &  
684 Dandouras, I. (2011, October). Timing mirror structures observed by Cluster  
685 with a magnetosheath flow model. *Annales Geophysicae*, 29(10), 1849–1860.  
686 Retrieved 2024-02-07, from <https://angeo.copernicus.org/articles/29/1849/2011/> doi: 10.5194/angeo-29-1849-2011
- 687 Hamill, T. M. (1997, December). Reliability Diagrams for Multicategory Probabilis-  
688 tic Forecasts. *Weather and Forecasting*, 12(4), 736–741. Retrieved 2023-07-  
689 24, from [http://journals.ametsoc.org/doi/10.1175/1520-0434\(1997\)012<0736:RDFMPF>2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0434(1997)012<0736:RDFMPF>2.0.CO;2) doi: 10.1175/1520-0434(1997)012(0736:  
690 RDFMPF)2.0.CO;2
- 691 Hamill, T. M. (2001, March). Interpretation of Rank Histograms for Verify-  
692 ing Ensemble Forecasts. *Monthly Weather Review*, 129(3), 550–560. Re-  
693 tried 2023-07-24, from [http://journals.ametsoc.org/doi/10.1175/1520-0493\(2001\)129<0550:IORHFV>2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2) doi: 10.1175/1520-0493(2001)  
694 129<0550:IORHFV>2.0.CO;2
- 695 Hersbach, H. (2000, October). Decomposition of the Continuous Ranked Prob-  
696 ability Score for Ensemble Prediction Systems. *Weather and Forecasting*,  
697 15(5), 559–570. Retrieved 2023-01-23, from [http://journals.ametsoc.org/doi/10.1175/1520-0434\(2000\)015<0559:DOTCRP>2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2) doi: 10.1175/  
698 0559:DOTCRP)2.0.CO;2
- 699 700

- 705 1520-0434(2000)015<0559:DOTCRP>2.0.CO;2
- 706 Hoilijoki, S., Palmroth, M., Walsh, B. M., Pfau-Kempf, Y., von Alfthan, S., Ganse,  
707 U., ... Vainio, R. (2016, May). Mirror modes in the Earth's magnetosheath:  
708 Results from a global hybrid-Vlasov simulation. *Journal of Geophysical  
709 Research: Space Physics*, 121(5), 4191–4204. Retrieved 2024-02-02, from  
710 <https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2015JA022026>  
711 doi: 10.1002/2015JA022026
- 712 Hu, A., Camporeale, E., & Swiger, B. (2022, September). *Multi-Hour Ahead Dst In-  
713 dex Prediction Using Multi-Fidelity Boosted Neural Networks*. arXiv. Retrieved  
714 2023-01-05, from <http://arxiv.org/abs/2209.12571> (arXiv:2209.12571  
715 [physics])
- 716 Huang, S., Li, W., Shen, X., Ma, Q., Chu, X., Ma, D., ... Goldstein, J. (2022,  
717 September). Application of Recurrent Neural Network to Modeling Earth's  
718 Global Electron Density. *Journal of Geophysical Research: Space Physics*,  
719 127(9). Retrieved 2022-10-25, from <https://onlinelibrary.wiley.com/doi/10.1029/2022JA030695> doi: 10.1029/2022JA030695
- 720 Jelínek, K., Němcéek, Z., & Šafránková, J. (2012, May). A new approach to mag-  
721 netopause and bow shock modeling based on automated region identification.  
722 *Journal of Geophysical Research: Space Physics*, 117(A5), 2011JA017252. Re-  
723 trieval 2024-01-05, from <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2011JA017252> doi: 10.1029/2011JA017252
- 724 Jung, J., K. Connor, H., P. Dimmock, A., Sembay, S., M. Read, A., Soucek, J., ...  
725 5Institute of Atmospheric Physics, Academy of Sciences of the Czech Repub-  
726 lic (2024). Mshpy23: a user-friendly, parameterized model of magnetosheath  
727 conditions. *Earth and Planetary Physics*, 8(1), 0–0. Retrieved 2023-09-27,  
728 from <https://www.eppcgs.org/en/article/doi/10.26464/epp2023065> doi:  
729 10.26464/epp2023065
- 730 Kallenrode, M.-B. (2010). *Space physics: an introduction to plasmas and particles in  
731 the heliosphere and magnetospheres ; with 12 tables, numerous excercises and  
732 problems* (3. ed., paperback ed ed.). Berlin Heidelberg: Springer.
- 733 King, J. H. (2005). Solar wind spatial scales in and comparisons of hourly Wind  
734 and ACE plasma and magnetic field data. *Journal of Geophysical Research*,  
735 110(A2), A02104. Retrieved 2020-05-26, from <http://doi.wiley.com/10.1029/2004JA010649> doi: 10.1029/2004JA010649
- 736 King, J. H., & Papitashvili, N. E. (2020). *OMNI 1-min Data Set*. NASA Space  
737 Physics Data Facility. Retrieved 2023-06-01, from <https://hpde.io/NASA/NumericalData/OMNI/HighResolutionObservations/Version1/PT1M> doi:  
738 10.48322/45BB-8792
- 739 Kingma, D. P., & Ba, J. (2017, January). *Adam: A Method for Stochastic Optimiza-  
740 tion*. arXiv. Retrieved 2023-07-18, from <http://arxiv.org/abs/1412.6980>  
741 (arXiv:1412.6980 [cs])
- 742 Kobel, E., & Flückiger, E. O. (1994, December). A model of the steady  
743 state magnetic field in the magnetosheath. *Journal of Geophysical Re-  
744 search: Space Physics*, 99(A12), 23617–23622. Retrieved 2024-01-05, from  
745 <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/94JA01778> doi:  
746 10.1029/94JA01778
- 747 Koons, H. C., & Gorney, D. J. (1991, April). A neural network model of the  
748 relativistic electron flux at geosynchronous orbit. *Journal of Geophysical  
749 Research: Space Physics*, 96(A4), 5549–5556. Retrieved 2024-01-16, from  
750 <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/90JA02380> doi:  
751 10.1029/90JA02380
- 752 Kull, M., & Flach, P. (2015). Novel Decompositions of Proper Scoring Rules for  
753 Classification: Score Adjustment as Precursor to Calibration. In A. Appice,  
754 P. P. Rodrigues, V. Santos Costa, C. Soares, J. Gama, & A. Jorge (Eds.),  
755 *Machine Learning and Knowledge Discovery in Databases* (Vol. 9284, pp. 68–

- 85). Cham: Springer International Publishing. Retrieved 2023-07-24, from  
[http://link.springer.com/10.1007/978-3-319-23528-8\\_5](http://link.springer.com/10.1007/978-3-319-23528-8_5) (Series Title:  
 Lecture Notes in Computer Science) doi: 10.1007/978-3-319-23528-8\_5
- Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017, November). *Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles*. arXiv. Retrieved 2023-01-05, from <http://arxiv.org/abs/1612.01474> (arXiv:1612.01474 [cs, stat])
- Lepping, R. P., Acuña, M. H., Burlaga, L. F., Farrell, W. M., Slavin, J. A., Schatten, K. H., ... Worley, E. M. (1995, February). The WIND magnetic field investigation. *Space Science Reviews*, 71(1-4), 207–229. Retrieved 2021-12-01, from <http://link.springer.com/10.1007/BF00751330> doi: 10.1007/BF00751330
- Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., & Talwalkar, A. (2018). Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. *Journal of Machine Learning Research*, 18(185), 1–52. Retrieved from <http://jmlr.org/papers/v18/16-558.html>
- Lockwood, M., Bentley, S. N., Owens, M. J., Barnard, L. A., Scott, C. J., Watt, C. E., & Allanson, O. (2019, January). The Development of a Space Climatology: 1. Solar Wind Magnetosphere Coupling as a Function of Timescale and the Effect of Data Gaps. *Space Weather*, 17(1), 133–156. Retrieved 2022-09-16, from <https://onlinelibrary.wiley.com/doi/abs/10.1029/2018SW001856> doi: 10.1029/2018SW001856
- Lopez, R. E. (2016, May). The integrated dayside merging rate is controlled primarily by the solar wind: THE SOLAR WIND CONTROLS DAYSIDE MERGING. *Journal of Geophysical Research: Space Physics*, 121(5), 4435–4445. Retrieved 2021-11-04, from <http://doi.wiley.com/10.1002/2016JA022556> doi: 10.1002/2016JA022556
- Lyon, J., Fedder, J., & Mobarry, C. (2004, October). The Lyon–Fedder–Mobarry (LFM) global MHD magnetospheric simulation code. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66(15–16), 1333–1350. Retrieved 2024-01-24, from <https://linkinghub.elsevier.com/retrieve/pii/S1364682604001439> doi: 10.1016/j.jastp.2004.03.020
- Matheson, J. E., & Winkler, R. L. (1976, June). Scoring Rules for Continuous Probability Distributions. *Management Science*, 22(10), 1087–1096. Retrieved 2023-07-18, from <https://pubsonline.informs.org/doi/10.1287/mnsc.22.10.1087> doi: 10.1287/mnsc.22.10.1087
- Nix, D., & Weigend, A. (1994). Estimating the mean and variance of the target probability distribution. In *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94)* (pp. 55–60 vol.1). Orlando, FL, USA: IEEE. Retrieved 2023-07-25, from <http://ieeexplore.ieee.org/document/374138/> doi: 10.1109/ICNN.1994.374138
- Ogilvie, K. W., Chornay, D. J., Fritzenreiter, R. J., Hunsaker, F., Keller, J., Lobell, J., ... Gergin, E. (1995, February). SWE, a comprehensive plasma instrument for the WIND spacecraft. *Space Science Reviews*, 71(1-4), 55–77. Retrieved 2021-12-01, from <http://link.springer.com/10.1007/BF00751326> doi: 10.1007/BF00751326
- Olshevsky, V., Khotyaintsev, Y. V., Lalti, A., Divin, A., Delzanno, G. L., Anderzen, S., ... Markidis, S. (2021, October). Automated classification of plasma regions using 3D particle energy distributions. *Journal of Geophysical Research: Space Physics*, 126(10). Retrieved 2022-03-15, from <http://arxiv.org/abs/1908.05715> (arXiv: 1908.05715) doi: 10.1029/2021JA029620
- O’Malley, T., Bursztein, E., Long, J., Chollet, F., Jin, H., Invernizzi, L., et al. (2019). Kerastuner. <https://github.com/keras-team/keras-tuner>.
- O’Brien, C., Walsh, B. M., Zou, Y., Tasnim, S., Zhang, H., & Sibeck, D. G. (2023, September). PRIME: a probabilistic neural network approach to solar

- wind propagation from L1. *Frontiers in Astronomy and Space Sciences*, 10, 1250779. Retrieved 2023-11-13, from <https://www.frontiersin.org/articles/10.3389/fspas.2023.1250779/full> doi: 10.3389/fspas.2023.1250779
- Phan, T. D., & Paschmann, G. (1996, April). Low-latitude dayside magnetopause and boundary layer for high magnetic shear: 1. Structure and motion. *Journal of Geophysical Research: Space Physics*, 101(A4), 7801–7815. Retrieved 2024-03-15, from <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/95JA03752> doi: 10.1029/95JA03752
- Pollock, C., Moore, T., Jacques, A., Burch, J., Gliese, U., Saito, Y., ... Zeuch, M. (2016, March). Fast Plasma Investigation for Magnetospheric Multi-scale. *Space Science Reviews*, 199(1-4), 331–406. Retrieved 2023-02-23, from <http://link.springer.com/10.1007/s11214-016-0245-4> doi: 10.1007/s11214-016-0245-4
- Powell, K. G., Roe, P. L., Linde, T. J., Gombosi, T. I., & De Zeeuw, D. L. (1999, September). A Solution-Adaptive Upwind Scheme for Ideal Magnetohydrodynamics. *Journal of Computational Physics*, 154(2), 284–309. Retrieved 2024-01-24, from <https://linkinghub.elsevier.com/retrieve/pii/S002199919996299X> doi: 10.1006/jcph.1999.6299
- Pulkkinen, T. I., Dimmock, A. P., Lakka, A., Osmane, A., Kilpua, E., Myllys, M., ... Viljanen, A. (2016, September). Magnetosheath control of solar wind-magnetosphere coupling efficiency. *Journal of Geophysical Research: Space Physics*, 121(9), 8728–8739. Retrieved 2021-10-29, from <https://onlinelibrary.wiley.com/doi/abs/10.1002/2016JA023011> doi: 10.1002/2016JA023011
- Raeder, J., Larson, D., Li, W., Kepko, E. L., & Fuller-Rowell, T. (2008, December). OpenGGCM Simulations for the THEMIS Mission. *Space Science Reviews*, 141(1-4), 535–555. Retrieved 2024-02-07, from <http://link.springer.com/10.1007/s11214-008-9421-5> doi: 10.1007/s11214-008-9421-5
- Raeder, J., McPherron, R. L., Frank, L. A., Kokubun, S., Lu, G., Mukai, T., ... Slavin, J. A. (2001, January). Global simulation of the Geospace Environment Modeling substorm challenge event. *Journal of Geophysical Research: Space Physics*, 106(A1), 381–395. Retrieved 2024-02-07, from <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2000JA000605> doi: 10.1029/2000JA000605
- Roberts, O. W., Nakamura, R., Coffey, V. N., Gershman, D. J., Volwerk, M., Varsani, A., ... Pollock, C. (2021, October). A Study of the Solar Wind Ion and Electron Measurements From the Magnetospheric Multiscale Mission's Fast Plasma Investigation. *Journal of Geophysical Research: Space Physics*, 126(10). Retrieved 2023-02-23, from <https://onlinelibrary.wiley.com/doi/10.1029/2021JA029784> doi: 10.1029/2021JA029784
- Russell, C. T., Anderson, B. J., Baumjohann, W., Bromund, K. R., Dearborn, D., Fischer, D., ... Richter, I. (2016, March). The Magnetospheric Multiscale Magnetometers. *Space Science Reviews*, 199(1-4), 189–256. Retrieved 2021-08-25, from <http://link.springer.com/10.1007/s11214-014-0057-3> doi: 10.1007/s11214-014-0057-3
- Russell, C. T., Zhuang, H., Walker, R. J., & Crooker, N. U. (1981, September). A note on the location of the stagnation point in the magnetosheath flow. *Geophysical Research Letters*, 8(9), 984–986. Retrieved 2023-10-19, from <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/GL008i009p00984> doi: 10.1029/GL008i009p00984
- Shue, J.-H., Song, P., Russell, C. T., Steinberg, J. T., Chao, J. K., Zastenker, G., ... Kawano, H. (1998, August). Magnetopause location under extreme solar wind conditions. *Journal of Geophysical Research: Space Physics*, 103(A8), 17691–17700. Retrieved 2020-05-11, from <http://doi.wiley.com/10.1029/10.1029/GL008i009p00984>

- 98JA01103 doi: 10.1029/98JA01103
- Sivadas, N., Sibeck, D., Subramanyan, V., Walach, M.-T., Murphy, K., & Halford, A. (2022, January). *Uncertainty in solar wind forcing explains polar cap potential saturation.* arXiv. Retrieved 2022-06-09, from <http://arxiv.org/abs/2201.02137> (Number: arXiv:2201.02137 arXiv:2201.02137 [astro-ph, physics:physics])
- Soucek, J., & Escoubet, C. P. (2012, June). Predictive model of magnetosheath plasma flow and its validation against Cluster and THEMIS data. *Annales Geophysicae*, 30(6), 973–982. Retrieved 2023-08-21, from <https://angeo.copernicus.org/articles/30/973/2012/> doi: 10.5194/angeo-30-973-2012
- Southwood, D. J., & Kivelson, M. G. (1995, December). Magnetosheath flow near the subsolar magnetopause: Zwan-Wolf and Southwood-Kivelson theories reconciled. *Geophysical Research Letters*, 22(23), 3275–3278. Retrieved 2024-03-12, from <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/95GL03131> doi: 10.1029/95GL03131
- Spreiter, J. R., & Alksne, A. Y. (1969). Plasma flow around the magnetosphere. *Reviews of Geophysics*, 7(1, 2), 11. Retrieved 2020-10-22, from <http://doi.wiley.com/10.1029/RG007i001p00011> doi: 10.1029/RG007i001p00011
- Spreiter, J. R., Summers, A. L., & Alksne, A. Y. (1966, March). Hydromagnetic flow around the magnetosphere. *Planetary and Space Science*, 14(3), 223–253. Retrieved 2020-11-09, from <https://linkinghub.elsevier.com/retrieve/pii/0032063366901243> doi: 10.1016/0032-0633(66)90124-3
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15(56), 1929–1958. Retrieved from <http://jmlr.org/papers/v15/srivastava14a.html>
- Tasistro-Hart, A., Grayver, A., & Kuvshinov, A. (2021, January). Probabilistic Geomagnetic Storm Forecasting via Deep Learning. *Journal of Geophysical Research: Space Physics*, 126(1), e2020JA028228. Retrieved 2023-07-21, from <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020JA028228> doi: 10.1029/2020JA028228
- Tsyganenko, N. A., Semenov, V. S., & Erkaev, N. V. (2023, November). Data-Based Modeling of the Magnetosheath Magnetic Field. *Journal of Geophysical Research: Space Physics*, 128(11), e2023JA031665. Retrieved 2024-01-03, from <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2023JA031665> doi: 10.1029/2023JA031665
- Von Alfthan, S., Pokhotelov, D., Kempf, Y., Hoilijoki, S., Honkonen, I., Sandroos, A., & Palmroth, M. (2014, December). Vlasiator: First global hybrid-Vlasov simulations of Earth's foreshock and magnetosheath. *Journal of Atmospheric and Solar-Terrestrial Physics*, 120, 24–35. Retrieved 2024-01-16, from <https://linkinghub.elsevier.com/retrieve/pii/S1364682614001916> doi: 10.1016/j.jastp.2014.08.012
- Walsh, B. M., Sibeck, D. G., Wang, Y., & Fairfield, D. H. (2012, December). Dawn-dusk asymmetries in the Earth's magnetosheath: MAGNETOSHEATH ASYMMETRIES. *Journal of Geophysical Research: Space Physics*, 117(A12), n/a–n/a. Retrieved 2020-10-22, from <http://doi.wiley.com/10.1029/2012JA018240> doi: 10.1029/2012JA018240
- Wang, Y. L., Raeder, J., & Russell, C. T. (2004, March). Plasma depletion layer: Magnetosheath flow structure and forces. *Annales Geophysicae*, 22(3), 1001–1017. Retrieved 2023-10-23, from <https://angeo.copernicus.org/articles/22/1001/2004/> doi: 10.5194/angeo-22-1001-2004
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (3rd ed ed.) (No. v. 100). Amsterdam ; Boston: Elsevier/Academic Press.
- Zamo, M., & Naveau, P. (2018, February). Estimation of the Continuous Ranked

- 925 Probability Score with Limited Information and Applications to Ensemble  
926 Weather Forecasts. *Mathematical Geosciences*, 50(2), 209–234. Retrieved  
927 2023-05-31, from <http://link.springer.com/10.1007/s11004-017-9709-7>  
928 doi: 10.1007/s11004-017-9709-7
- 929 Zwan, B. J., & Wolf, R. A. (1976, April). Depletion of solar wind plasma near a  
930 planetary boundary. *Journal of Geophysical Research*, 81(10), 1636–1648. Re-  
931 trieved 2024-03-12, from <http://doi.wiley.com/10.1029/JA081i010p01636>  
932 doi: 10.1029/JA081i010p01636