





## Energetic Electron Flux Predictions in the Near-Earth Plasma Sheet From Solar Wind Driving

B. M. Swiger<sup>1</sup> , M. W. Liemohn<sup>1</sup> , N. Y. Ganushkina<sup>1,2</sup> , and S. V. Dubyagin<sup>2</sup> 

<sup>1</sup>Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA, <sup>2</sup>Finnish Meteorological Institute, Helsinki, Finland

### Special Section:

Machine Learning in  
Heliophysics

### Key Points:

- New model predicts plasma sheet electron flux (0.08–93 keV) from solar wind (SW) within a factor of two of observed based on median symmetric accuracy metric
- Input analysis of new model supports that the most impactful drivers are SW speed, electric field, and north-south component of interplanetary magnetic field
- Short-term changes of electron flux (<1 hr) in the plasma sheet are not predicted from high resolution (5 min) SW inputs

### Supporting Information:

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

### Correspondence to:

B. M. Swiger,  
swigerbr@umich.edu

### Citation:

Swiger, B. M., Liemohn, M. W., Ganushkina, N. Y., & Dubyagin, S. V. (2022). Energetic electron flux predictions in the near-Earth plasma sheet from solar wind driving. *Space Weather*, 20, e2022SW003150. <https://doi.org/10.1029/2022SW003150>

Received 7 MAY 2022

Accepted 19 SEP 2022

**Abstract** Suprathermal electrons in the near-Earth plasma sheet are important for inner magnetosphere considerations. They are the source population for outer radiation belt electrons and they pose risks to geosynchronous satellites through their contribution to surface charging. We use empirical modeling to address relationships between solar driving parameters and plasma sheet electron flux. Using Time History of Events and Macroscale Interactions during Substorms, OMNI, and Flare Irradiance Spectral Model Version 2 data, we develop a neural network model to predict differential electron flux from 0.08 to 93 keV in the plasma sheet, at distances from 6 to 12  $R_E$ . Driving parameters include solar wind (SW) density and speed, interplanetary magnetic field (IMF)  $B_z$  and  $B_y$ , solar extreme ultraviolet flux, IMF  $B_z$  ultra-low frequency (ULF) wave power, SW-magnetosphere coupling functions  $P_{\alpha 1}$  and  $NX_{CF}$ , and the 4-hr time history of these parameters. Our model predicts overall plasma sheet electron flux variations with correlation coefficients of between 0.59 and 0.77, and median symmetric accuracy in the 41%–140% range (depending on energy). We find that short time-scale electron flux variations are not reproduced using short time-scale inputs. Using a recently published technique to extract information from neural networks, we determine the most important drivers impacting model prediction are  $V_{SW}$ ,  $VB_S$ , and IMF  $B_z$ . SW-magnetosphere coupling functions that include IMF clock angle, IMF  $B_z$  ULF wave power, and IMF  $B_y$  have little impact in our model of electron flux in the near-Earth plasma sheet. The new model, built directly on differential flux, outperforms an existing model that derives fluxes from plasma moments, with the performance improvement increasing with increasing energy.

**Plain Language Summary** In near-Earth space, electrons are energized and transported toward the Earth, where they pose hazards to spacecraft or travel to the upper atmosphere to generate aurora. Previous studies have shown that much of the variations are attributed to changes in the upstream solar wind (SW), the driving conditions from the Sun. Yet, it has been difficult to determine which variations in the SW are most predictive of changes in the near-Earth electrons. To help confront these difficulties, we have developed a machine learning neural network model that predicts the variations of electrons using inputs of the SW and sunlight levels. We report details of the development of this model and assess the model's performance. The model well reproduces global variations of near-Earth electrons, yet does not fully reproduce flux variations during all space weather events. By ranking the importance of the model inputs, we show that individual SW parameters are more important to predictions than using some previously defined and investigated coupling functions calculated using those same parameters.

## 1. Introduction

Separating the magnetotail into northern and southern lobes, the plasma sheet extends from the edge of the dominance of Earth's magnetic dipole field to several dozen Earth radii ( $R_E$ ;  $R_E$  is the Earth's radius, 6,371 km) downtail, and its azimuthal extent ranges several hours of magnetic local time (MLT; Hill, 1974). The near-Earth (6–12 $R_E$ , 06–18MLT) plasma sheet, which contains the transition region from stretched to dipolar magnetic field, is an important driver of inner magnetosphere dynamics (e.g., Artemyev et al., 2018; Ganushkina et al., 2015; Schield et al., 1969; Southwood & Wolf, 1978; Wolf et al., 2007). Suprathermal electrons (~1 to 100 keV) in the near-Earth plasma sheet have particular effectiveness at altering the inner magnetosphere (e.g., Garrett, 1979; Turner & Li, 2008; Yu et al., 2019), and they are the seed population for radiation belt MeV electrons (e.g., Allison et al., 2019; Chen et al., 2007; Fu et al., 2011; Horne et al., 2005; Jaynes et al., 2015; Kennel & Petschek, 1966; Kennel & Thorne, 1967; Li et al., 2008, 2012; Malykhin et al., 2021). They contribute to spacecraft surface charging environments (e.g., Davis et al., 2008; Garrett, 1981; Koons et al., 1999), posing a space weather hazard to satellites at geosynchronous orbit (GEO; e.g., Choi et al., 2011; Ganushkina et al., 2021; Matéo-Vélez

© 2022. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](https://creativecommons.org/licenses/by/4.0/), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

et al., 2018; Thomsen et al., 2013); and, they can penetrate deep inside GEO to medium Earth orbit (MEO; Motoba et al., 2021; Turner et al., 2015).

Variations in the near-Earth plasma sheet are dependent on solar wind (SW) conditions upstream of Earth's bow shock (e.g., Borovsky et al., 1998; Burin des Roziers, Li, Baker, Fritz, Friedel, et al., 2009; Cao et al., 2013; Nagata et al., 2008; Nishida & Lyon, 1972; Terasawa et al., 1997; Wing et al., 2005). Many previous studies have used empirical modeling to study the relationships between solar driving and plasma sheet electron flux. Using Geotail in the plasma sheet and ACE in the SW, Luo et al. (2011) investigated the distribution of fluxes of electrons with energy >38 keV, dependent upon interplanetary conditions. They found that higher SW speed and southward interplanetary magnetic field (IMF) resulted in higher fluxes. They developed an empirical 2-D model predicting >38 keV electron fluxes in the plasma sheet as a function of SW velocity and density, and IMF magnitude and north-south components. Although the modeled fluxes had a relatively high correlation coefficient ( $R$ ; defined in Appendix B) with the observed ( $R = 0.86$ ), it was limited by using integrated flux with a lower energy boundary at 38 keV. Wang et al. (2013) used both Geotail and Time History of Events and Macroscale Interactions during Substorms (THEMIS) in the plasma sheet to develop a model predicting total plasma sheet pressure (with contribution from both ions and electrons) as a function of fixed SW dynamic pressure and planetary  $K$  ( $K_p$ ) index values. Yue et al. (2015) also used Geotail and THEMIS to model plasma sheet pressure, yet only during substorm growth phases. They found that the sunspot number was a controlling factor of plasma sheet pressure, in addition to SW dynamic pressure, combinations of IMF direction and magnitude (used to estimate cross polar cap potential) and the auroral electrojet (AE) index. For modeling geomagnetic storm-time electron plasma moments, Dubyagin et al. (2016) used THEMIS for the plasma sheet and OMNI for interplanetary parameters. They constructed a model based on dependencies of storm-time plasma sheet electron temperature to SW flow speed and plasma sheet electron density to SW proton density and southward IMF. Wang et al. (2017) used THEMIS for the plasma sheet and ACE, WIND, and Geotail spacecraft for interplanetary parameters combined with the machine learning method of support vector machine to predict electron temperature in the plasma sheet (with a limited domain of  $6 < R < 8R_E$  and  $21 < \text{MLT} < 3$ ). Investigating the relationship between SW and plasma sheet parameters, they found that electron temperature in this region of the plasma sheet can increase to beyond its 75th percentile (>3 keV) when the IMF north-south magnitude is weak but IMF ultra-low frequency (ULF) wave power remains high for several hours. Zou et al. (2020) trained a neural network to predict electron flux in the near-Earth plasma sheet from 7 to  $12R_E$ . Their model inputs included the AE and Disturbance Storm Time geomagnetic indices, F10.7, and several SW parameters with time lags of up to 48 hr. They achieved high correlation coefficient values ( $R = 0.92\text{--}0.95$ , depending on energy). Stepanov et al. (2021) built a regression model for electron fluxes in the near-Earth plasma sheet at three energies—10, 31, and 93 keV—using THEMIS and OMNI for the model output and input, respectively, looking for long-term (1–48 hr) correlations. They subdivided the plasma sheet into nine bins (three azimuthal and three radial), created a regression model for each bin, and analyzed the interplanetary time delays in evaluating the best correlation coefficients between modeled and observed electron flux. They concluded that the two best interplanetary predictors for plasma sheet electron flux are the SW speed and convective electric field modulated by the sine of the IMF clock angle.

The preceding models have investigated many aspects of solar driving of plasma sheet electron flux variations. However, several relationships between energetic electron flux in the plasma sheet and solar driving have yet to be fully investigated. This paper focuses on three such relationships.

1) *Solar driving effects on short time-scale electron flux variations.* Studies between upstream SW and plasma sheet electron flux to date have not considered small-scale variations (on time scales of minutes to an hour) in the plasma sheet electron flux. Yet, it is common for suprathermal electron flux in the plasma sheet to vary beyond an order of magnitude within minutes (e.g., Bame et al., 1966; Burin des Roziers, Li, Baker, Fritz, McPherron, et al., 2009). Since local processes within the plasma sheet contribute to short-term variations (e.g., Angelopoulos et al., 1992; Duan et al., 2014; Fu et al., 2011), any dependence of short-term variations on solar drivers has been difficult to infer.

2) *Influence of the ionosphere on plasma sheet electron flux.* The ionosphere sourcing ions to the plasma sheet is well established (e.g., Artemyev et al., 2020; Chappell et al., 1987; Kistler et al., 2019; Kronberg et al., 2012; Liemohn et al., 2005; Ohtani et al., 2011). Electrons of ionospheric origin in the plasma sheet has been less studied than ions, however there is some evidence that the ionosphere can play a similar role for the electron plasma

sheet. It has been suggested that energetic electron field-aligned currents observed in the plasma sheet originated in the auroral ionosphere (Zheng et al., 2012). Also, with increasing ionospheric activity, which is dependent on solar activity that varies with the 11-year solar cycle, the field-aligned, plasma sheet connected Region 1 and Region 2 currents are enhanced (Ohtani et al., 2014).

3) *The role of SW-magnetosphere coupling functions in driving electron flux variations.* Many SW-magnetosphere coupling functions have been proposed and shown to have strong correlations with geomagnetic indices through many decades of previous research (e.g., Balikhin et al., 2010; Boynton et al., 2011; Burton et al., 1975; Kan & Lee, 1979; Lockwood et al., 2019; Newell et al., 2007; Perreault & Akasofu, 1978; Vassiliadis et al., 1995). Many of these coupling functions are physically motivated, and they are expected to correlate well with directly driven magnetospheric phenomena. Relevant to our study, it has been suggested that SW-magnetosphere coupling functions can have a role as a driver for plasma sheet suprathermal electron flux variations and their transport to GEO (e.g., Denton et al., 2019; Kaufmann, 2012; V. A. Sergeev et al., 2014).

Empirical modeling is one approach to investigate the preceding three associations between solar driving and the electron fluxes in the plasma sheet: small-scale variations, ionospheric control, and role of SW-magnetosphere coupling functions. Additionally, there are several desirable attributes for an electron plasma sheet model to have to further our understanding of the dependence on SW driving as well as provide forecasting utility.

1. A model that is built to reproduce electron fluxes, not moments, is recommended to more accurately predict fluxes (Dubyagin et al., 2019).
2. We would like the model output to be specifiable by exact location and by energy, with sufficient energy range and channels to meaningfully fit the modeled spectra to a distribution function and obtain plasma moments.
3. The spatial domain should span the night side from 18 to 06MLT and include the transition region between stretched magnetic field and dipolar magnetic field.
4. The temporal domain should include all phases of the solar cycle and all levels of geomagnetic activity.
5. For forecasting capability, model inputs must not include driving parameters that are internal to the magnetosphere, such as geomagnetic indices.

The empirical models highlighted in the previous paragraph each include one or more of these model characteristics, yet none encompass all of them.

In this study, we are addressing the aforementioned SW and plasma sheet electron flux relationships by adding characteristics to a preliminary model (Swiger et al., 2020) that will also incorporate the desirable attributes introduced in the preceding paragraph. We now have THEMIS observations of the plasma sheet which span a period longer than an  $\sim 11$ -year sunspot cycle; and, we have consistent SW and IMF data from OMNI for the same period. Using the abundance of these data, we applied a neural network algorithm to build a predictive model of plasma sheet electron flux. We include external drivers of SW, IMF, SW-magnetosphere coupling functions, and solar extreme ultraviolet (EUV) photon flux—all at high time resolution. The spatial domain spans 6–12 $R_E$  and 18–06 MLT. We can specify electron flux at 31 different energy channels between 82.5 eV and 93 keV at any point within the spatial domain.

We describe the development of the model in Section 2. Model performance is assessed in Section 3, including comparison with a similar, existing model (Section 3.3). The driving parameters of near-Earth plasma sheet electron flux are analyzed in Section 3.4. Implications of the model output to relationships between solar driving and plasma sheet electron flux are discussed in Section 4. Further model use and a summary of our main findings are stated in Section 5.

## 2. Model Development

The THEMIS spacecraft (Angelopoulos, 2008) have been observing suprathermal electron flux in the near-Earth plasma sheet since late 2007. There are accompanying observations of the SW and IMF from OMNI (King & Papitashvili, 2005), and solar EUV photon flux data are provided by the Flare Irradiance Spectral Model Version 2 (FISM-2) (Chamberlin et al., 2020). These data constitute millions of samples over a period that ranges through all phases of a  $\sim 11$ -year sunspot cycle. They are robust to climatological considerations and propitious to applications of machine learning methods. Our machine learned model is a feed-forward neural network that connects

solar drivers and the near-Earth electron plasma sheet response. The neural network takes multiple variables as inputs and multiple variables as outputs.

## 2.1. Data

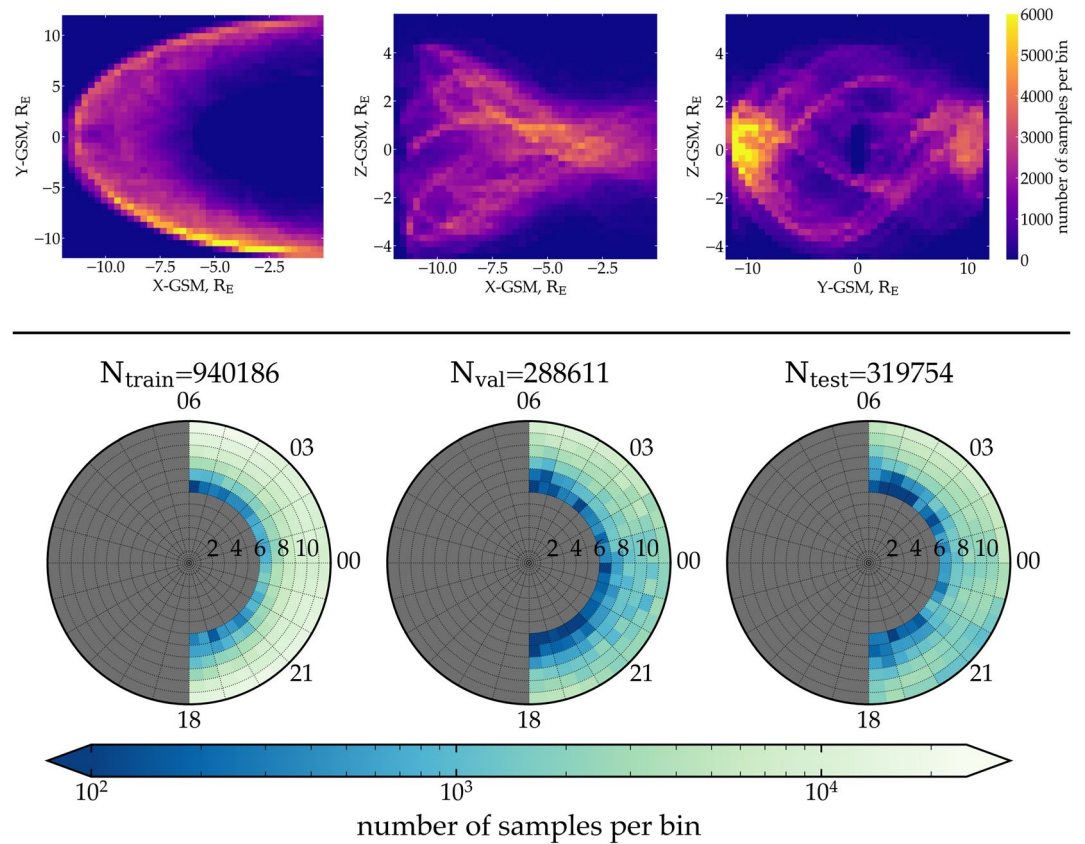
Electron flux data in the near-Earth plasma sheet come from three THEMIS probes—Th-A, Th-D, and Th-E. These spacecraft have been orbiting Earth continuously from 2007-present with a nominal spin rate of 3 s in near-equatorial orbits with perigee  $\sim 1.5R_E$  and apogee  $\sim 12R_E$  (Harvey et al., 2008). Each probe carries an Electrostatic Analyzer (ESA; McFadden, Carlson, Larson, Ludlam, et al., 2008) and Solid State Telescope (SST; Angelopoulos et al., 2008). For this study, we used the combination of these data from both instruments (GMOM—level 2 ground moments data product), which provides electron flux in energy channels ranging from a few eV to hundreds of keV. For this study, we used 31 energy channels spanning 83 eV to 93 keV (see Supporting Information S1 for a full list of energies). An example of THEMIS electron flux observations can be seen in Section 3.2. In the lineplot of observed flux (second from bottom panel of Figure 3), one can see typical variations of the flux levels of 0.08–93 keV electrons in the plasma sheet.

Electron measurements taken by the ESA instrument are susceptible to several sources of contamination which can lead to spurious data (McFadden, Carlson, Larson, Bonnell, et al., 2008). These include contamination from photoelectrons and errors from spacecraft charging and they contribute to artificially high count rates at lower energies  $< 50$  eV (McFadden, Carlson, Larson, Bonnell, et al., 2008). Electrons that enter the ESA instrument can further scatter before impacting the detector, resulting in additional, non-plasma source electron counts for energies up to 200 eV. We are using energies  $> 83$  eV, and therefore the effects from photoelectrons and spacecraft charging are negligible. However, the observations from the lower energies of our data set—up to 200 eV—are possibly contaminated with these internally produced scattered electrons. Another source of error of ESA electron counts comes from the instrument becoming saturated with high levels of electron counts that occur in higher density regions of the magnetosphere. In selecting our data, we rejected data taken during times when the ESA instrument was saturated.

The SST instrument also has some possible measurement errors that we needed to consider. Other studies have identified contamination and saturation in electron channels above 100 keV (e.g., Dubyagin et al., 2019; Ni et al., 2011). In a recent study by Stepanov et al. (2021), it was found that even for SST data that was marked as saturated, the electron flux values continued to increase as the THEMIS probe moved to lower altitudes. Moreover, the Stepanov et al. (2021) model that included SST data marked as saturated proved to more accurately reproduce observations than the model that used only data marked as “good data.” In our study, we use SST energy channels at  $\leq 93$  keV only, and we follow the Stepanov et al. (2021) evidence that SST data that are marked as saturated can still be scientifically useful. Therefore, we included the data if only the SST (but not the ESA) was saturated. For all THEMIS data, we excluded periods that were marked as during eclipse or outside the magnetosphere. Descriptions of further paring of the THEMIS observations continues in the next section.

### 2.1.1. Observed Output Data Set

Samples of the electron flux in the near-Earth plasma sheet were obtained in the following steps. (a) We compiled all GMOM pterf (ESA Reduced, SST Full at approximately 3-s resolution) data from the THEMIS probes A, D, and E from January 2008 through December 2020. (b) We removed all data where the ESA instrument was saturated. Along with good data quality flags, if only the SST instrument was marked as saturated, we kept the data. This follows the discussions of Dubyagin et al. (2019) and Stepanov et al. (2021) that showed the data with SST saturation do not invalidate results and significantly improve the accuracy of reconstructed fluxes. (c) We down-sampled the data from 3-s resolution to 1-min resolution by taking the average of the flux at each minute. (d) We restricted our sampling region to the night side, and to the near-Earth plasma sheet. The spatial domain is from  $6R_E \leq XY_{GSM} \leq 12R_E$ , and the azimuthal extent is from 18MLT (dawn) to 06MLT (dusk) across midnight.  $XY_{GSM} = \sqrt{X_{GSM}^2 + Y_{GSM}^2}$ , and GSM is the Geocentric Solar Magnetospheric coordinate system. The  $Z_{GSM}$  direction is not constrained. Since the plasma sheet undergoes flapping in the  $Z_{GSM}$  direction



**Figure 1.** The 2D plane sampling locations of the Time History of Events and Macroscale Interactions during Substorms data used in the development of the model. Top row: sampling distributions in the GSM  $XY$ -,  $XZ$ -, and  $YZ$ -planes. Bottom row: number of observations shown in each bin is the number of samples used in the training (left), validation (center), and testing (right) data sets. Each bin in bottom row is  $1R_E$  by 0.5 hr of magnetic local time.

(e.g., Runov et al., 2009; Tsutomu & Teruki, 1976; V. Sergeev et al., 2003), we relied on the low inclination of the THEMIS probes' orbits and a plasma- $\beta \geq 1$  criterion (e.g., Ruan et al., 2005) to indicate that we were taking samples from the plasma sheet. In the  $Z_{GSM}$  direction, >98% of the samples were within  $\pm 4 R_E$ . See Figure 1 for sampling distributions in the GSM  $XY$ -,  $XZ$ -, and  $YZ$ -planes. (e) After applying steps (1–4) above, we compiled the electron flux at 31 energies (see Supporting Information S1 for list) from THEMIS-A, -D, and -E into a single data set comprising a total number of 1,722,949 1-min samples.

$$\vec{y}_i = [\log_{10}(j_{1,i}), \log_{10}(j_{2,i}), \dots, \log_{10}(j_{m,i})] \quad (1)$$

Then, we converted the differential energy flux units into  $\log_{10}$  differential number flux (with units of  $\log_{10}(1/\text{cm}^2 \text{ s sr keV})$ ). We put these observations in the form of Equation 1, where  $j_{k,i}$  is the flux of the  $k$ th energy channel for observation  $i$ .

The observations  $\vec{y}_i$  were separated into three groups, based on year, for the training, validation, and test sets. We trained our model on data from years 2008, 2009, 2011, 2012, 2014, 2017, and 2018. The validation set uses data from years 2013, 2016, and 2019, and it was used to optimize the free parameters of the neural network algorithm (described in Section 2.2.2). The test set contains data from years 2010, 2015, and 2020. None of the data in the test set was “seen” by the model during training nor optimizing, and therefore is considered as out-of-sample data that we used to test our model's performance. Although there may be some autocorrelation of electron flux within the plasma sheet, which would allow “bleed” of information between train, validation, and test sets at the yearly boundaries, we assert that any such cross-over is negligible compared to the large number of independent examples in each set. Our training, validation, and test split

strategy is an attempt to remove any bias that would be introduced by solar cycle phases. For example, we do not want to train on data taken solely during solar minimum. Moreover, our choice of years for the test set allows us to calculate more meaningful goodness-of-fit metrics, as we will be using data from multiple solar cycle phases. Yet, we wanted to keep sufficient data in the train and validation sets to maximize the potential for the model to learn. After this split, and after further reductions due to missing data described in Section 2.1.2, the train set contained 940,186 samples, the validation set contained 288,611 samples, and the test set contained 319,754 samples.

The top row of Figure 1 shows two-dimensional sampling distributions in the GSM  $XY$ -,  $XZ$ -, and  $YZ$ -planes of the full data set (train, validation, and test data combined). The bottom row of panels in Figure 1 are the sampling locations in the  $XY$ -plane separated into the train, validation, and test data. For the bottom panels, the distributions are organized by bins in radial distance and MLT. Each bin is 0.5 MLT by  $1R_E$  (we show gray grid lines at each hour of MLT). The color in each bin indicates the number of samples per bin, corresponding to the color bar shown below the panels. Gray color indicates areas outside of the model's spatial domain. The MLT values are labeled on the outside edge of the circle, and the radial distance, in  $R_E$ , is labeled outward from the center. There are more samples near the apogee of the spacecraft around 11–12 $R_E$  than at other radial distances, and generally the number of samples decreases with decreasing radial distance. These aspects of the sampling distribution are due to both the spacecraft spending more time near their apogees and the fact that it is less likely to satisfy the  $\beta \geq 1$  criterion the closer the observations are taken to Earth where the dipole magnetic field becomes more dominant.

For all three subsets—train, validation, and test—spatial bins greater than  $8R_E$  have more than  $10^3$  samples and many exceed  $10^4$  samples. At  $6$ – $7R_E$  there is much lower numbers of samples, often less than  $10^2$  and in the smallest case on order of  $10^1$ . Nonetheless, Figure 1 shows that we have sufficient overall numbers of samples for our train, validation, and test sets, and that the samples are well distributed in the model's spatial domain. Sorting the data into spatial bins is done only for consideration of the sampling distribution. We did not use these bins during the development of the model, nor in analyzing the model's performance.

## 2.1.2. Observed Model Inputs

For our inputs to the neural network, we use a combination of SW, IMF, SW-magnetosphere coupling functions, solar EUV photon flux, and sampling location. For each observation,  $\vec{y}_i$ , in the observed output data, each input,  $\vec{x}_i$ , contains the following. (a) OMNI parameters; (b) SW-magnetosphere coupling functions; (c) IMF ULF wave power; (d) FISM-2 model output; and (e) the sampling location at the time of each  $\vec{y}_i$ .

### 2.1.2.1. OMNI Parameters

We chose SW flow speed  $V_{SW}$ , number density  $N_{SW}$ , IMF  $B_z$ , and IMF  $B_y$  (the latter two in GSM-coordinates) from the 1-min High Resolution OMNI data set. For missing values, we applied a linear interpolation method, discarding periods where there were missing data for longer than 1 hr. For each observation timestamp in the plasma sheet observations, we constructed an input of OMNI parameters representing the state of the SW for the previous 4 hr. At this step, if any of the 4-hr periods contained missing data (after discarding data with gaps greater than 1 hr), we also removed the corresponding plasma sheet data. This reduced the total number of samples to 1,548,551. All of the OMNI data used in the development of the neural network were standardized by subtracting the climatic mean and dividing by the climatic standard deviation. That is,

$$p_{\text{norm},i} = \frac{p_i - \mu_p}{\sigma_p}, \quad (2)$$

where  $p_{\text{norm},i}$  is the standardized value of the  $i$ th value of parameter  $p$ ; and,  $\mu_p$  and  $\sigma_p$  are the mean and standard deviation, respectively, of all  $p$  from January 2008 through December 2020. In addition to making it easier to train the neural network, standardizing has the benefit of negating the effect that physical units have on the input values.

### 2.1.2.2. SW-Magnetosphere Coupling Functions

Using the OMNI data parameters from Section 2.1.2.1, we calculated several SW-magnetosphere coupling functions. The first SW-magnetosphere coupling function that we used is the rectified SW convective electric field, defined by

$$VB_S = \begin{cases} V_X |B_Z| & \text{if } B_Z \leq 0 \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where  $V_X$  is the SW speed in the GSM X-direction, making  $VB_S$  a negative quantity in our calculations.  $VB_S$  is well correlated with geomagnetic activity indices (e.g., Arnoldy, 1971; Burton et al., 1975; Foster et al., 1971; Kan & Lee, 1979; Kelley et al., 2003; Rostoker & Fälthammar, 1967; Shukhtina et al., 2005; Troshichev et al., 2011; Vassiliadis et al., 1995). We include it as an input to account for the effect of strong positive dawn-dusk electric field (strong southward IMF) in the SW that merges with the magnetosphere and results in a strong dawn-dusk electric field in the magnetotail, leading to increased  $E \times B$  drift of particles in the plasma sheet.

$$NX_{CF} = P_{SW}^{1/2} V_{SW}^{4/3} B_T \sin^6 \left( \frac{\theta}{2} \right) \quad (4)$$

The second SW-magnetosphere coupling function we used as an input to the model takes the form of Equation 4. The coupling function in Equation 4, referred to here as  $NX_{CF}$  was one of the coupling functions found by Boynton et al. (2011) using Nonlinear AutoRegressive Moving Average with exogenous inputs to have the best prediction of the Disturbance storm time index. We include it as an indicator for the model to predict the plasma sheet electron flux during geomagnetic storms. In Equation 4,  $P_{SW}$  is the SW dynamic pressure,  $V_{SW}$  is the SW flow speed,  $B_T = \sqrt{B_Y^2 + B_Z^2}$  is the tangential IMF magnitude, and  $\theta = \arctan(B_Y/B_Z)$  is the IMF clock angle.

The third SW-magnetosphere coupling function that we used quantifies solar energy input to the magnetosphere. Described in detail by Lockwood et al. (2019) and Lockwood (2019),  $P_{\alpha 1}$  accounts for both kinetic particle energy in the SW and Poynting flux energy impacting Earth's magnetosphere. The  $P_{\alpha 1}$  function, including the  $P_\alpha$  factor, takes the form

$$P_{\alpha 1} = P_\alpha \left[ 1 + 2f_s \cos^2(\phi) M_A^{(2\alpha-2)} \right] \quad (5)$$

$$P_\alpha = K N_{sw}^{(2/3-\alpha)} V_{sw}^{(7/3-2\alpha)} B^{2\alpha} \sin^4(\theta/2) \quad (6)$$

where  $K$  is a constant, if the magnetic moment of the Earth and the mass of the SW are approximated as constant.  $N_{sw}$ ,  $V_{sw}$ ,  $B$ , and  $\theta$  are the SW number density, SW flow speed, IMF magnitude, and IMF clock angle, respectively.  $P_\alpha$  accounts for the kinetic energy flux of SW particles.  $P_{\alpha 1}$  adds Poynting flux energy to the total SW energy entering the magnetosphere by including the Alfvén Mach number,  $M_A$ , and by considering the angle between the SW velocity and the IMF,  $\phi$ . We used the parameters  $\alpha = 0.38$  and  $f_s = 0.74$ , which were empirically fit by Lockwood (2019) to optimize correlation with the daily averaged *am* geomagnetic index (Mayaud, 1980)—an indicator of substorm activity.

Using the interpolated OMNI parameters from Section 2.1.2.1, we calculated these three coupling functions— $VB_S$ ,  $NX_{CF}$ , and  $P_{\alpha 1}$ —for all available OMNI (excluding periods with missing data greater than 1 hr) from January 2008 through December 2020. Similarly to standardizing the individual OMNI parameters themselves, we standardized the calculated coupling functions using Equation 2, with an exception for  $VB_S$ . We wanted to keep  $VB_S$  as a rectified input (i.e., no positive values); therefore, in standardizing  $VB_S$  we only divided by the climatic standard deviation and did not subtract the mean.

### 2.1.2.3. IMF $B_z$ ULF Wave Power

Lyons et al. (2009) and Kim et al. (2009) investigated the effect that ULF wave power in the IMF  $B_z$  (henceforth called  $B_{z,ULF}$ ) has on the coupled ionosphere-magnetosphere system. They found evidence that high  $B_{z,ULF}$  leads to enhanced magnetospheric and substorm activity. Later, Wang et al. (2017) showed that during weak or northward IMF  $B_z$  and high  $B_{z,ULF}$ , there was significantly increased electron temperature in the plasma sheet. The physical rationale for this finding, and for including  $B_{z,ULF}$  as an input to our model, is that higher ULF wave power impacting the magnetosphere may result in enhanced radial diffusion, bringing higher energy electrons in the plasma sheet farther earthward.

We used OMNI IMF  $B_z$  in GSM coordinates to calculate  $B_{z,ULF}$  following a procedure similar to one used by Wang et al. (2017), albeit with modifications. Details of the calculation of  $B_{z,ULF}$  can be found in Appendix A. Like the other parameters derived from OMNI data, we standardized the  $B_{z,ULF}$  by applying Equation 2, with the climatic mean and standard deviation taken from the period 2008–2020.

### 2.1.2.4. FISM-2 Output

We include photons that are impacting the dayside ionosphere in a night-side plasma sheet model to account for ionospheric influences on electrons in the plasma sheet (e.g., ionospheric outflow or increased ionospheric currents connected to the near-Earth plasma sheet). FISM-2 is an empirical model whose output includes EUV photon flux in spectral bands from 0.1 to 190 nm, with a time resolution of either one or 5 min. The time resolution of FISM-2 makes it more suitable to attempt to capture small scale plasma sheet variations than using an input with a much lower time resolution, for example, F10.7, for solar EUV flux. As an input to our model, we used a single band of FISM-2 output at  $\sim 20$  eV (54–65 nm), as this energy is sufficient to produce free electrons through photoionization of Oxygen in the upper atmosphere. FISM-2 data were taken from the LASP Interactive Solar Irradiance Data Center hosted by the Laboratory for Atmospheric and Space Physics (LASP). We matched the closest time stamps of FISM-2 data with the time stamps of the  $\vec{y}_i$  observations. Similarly to the previously described inputs, we standardized FISM-2 data using the mean and standard deviation of the available FISM-2 data from 2008 to 2020.

### 2.1.2.5. Sampling Location

The sampling location within the plasma sheet was included using the radial distance from Earth and the azimuthal direction. Information about the radial distance was encoded using  $XY_{GSM}$ . The azimuthal direction was encoded with  $\sin \phi$  and  $\cos \phi$ , where  $\phi = \frac{\pi}{12}$  MLT. We standardized the radial distance by subtracting the mean and dividing by the standard deviation of all plasma sheet samples identified by the methods of Section 2.1.1.

### 2.1.2.6. Combining Inputs

For each  $\vec{y}_i$ , a single  $\vec{x}_i$  is constructed by combining all of the OMNI, coupling functions, FISM-2, and sampling location data into a single one-dimensional input. Each input takes the form

$$\vec{x}_i = \left[ \vec{f}_{t_i-\Delta t}, \vec{f}_{t_i-\Delta t-\delta t}, \vec{f}_{t_i-\Delta t-2\delta t}, \dots, \vec{f}_{t_i-\Delta T}, \sin \phi, \cos \phi, r_i \right], \quad (7)$$

where

$$\begin{aligned} \phi &= \frac{2\pi}{24} \text{MLT} \\ r_i &= \frac{R_i - \mu_R}{\sigma_R} \\ R_i &= \sqrt{X_{GSM,i}^2 + Y_{GSM,i}^2} \\ \mu_R &= \frac{1}{N} \sum_{i=1}^N R_i \\ \sigma_R &= \sqrt{\frac{1}{N} \sum_{i=1}^N (R_i - \mu_R)^2} \\ \vec{f}_t &= [B_{Z,t}, B_{Y,t}, V_{SW,t}, N_{SW,t}, P_{a1,t}, (NX_{CF})_t, V B_{s,t}, B_{Z,ULF,t}, \text{FISM2}_t] \end{aligned}$$



The  $t$  in  $\vec{f}_i$  represents the time of the observation of  $\vec{y}_i$ . We used a time delay of  $\Delta t = 10$  min and a time step of  $\delta t = 5$  min in constructing inputs with time history. Our longest time delay is  $\Delta T = 240$  min. Therefore, each  $\vec{x}_i$  contains time history from  $-10$  to  $-240$  min in 5 min steps for each of the nine input parameters in  $\vec{f}_i$ . Adding in the three position inputs without time history— $\sin \phi$ ,  $\cos \phi$ , and  $r$ —the total number of inputs for each  $\vec{x}_i$  becomes 426.

Combining the  $\vec{x}_i$  in each set—train, validation, and test—we construct full input arrays  $\mathbf{X}$  (train),  $\mathbf{X}'$  (validation), and  $\mathbf{X}''$  (test) as

$$\mathbf{X} = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_N \end{bmatrix} \quad \mathbf{X}' = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_M \end{bmatrix} \quad \mathbf{X}'' = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_L \end{bmatrix} \quad (8)$$

$$\mathbf{Y} = \begin{bmatrix} \vec{y}_1 \\ \vec{y}_2 \\ \vdots \\ \vec{y}_N \end{bmatrix} \quad \mathbf{Y}' = \begin{bmatrix} \vec{y}_1 \\ \vec{y}_2 \\ \vdots \\ \vec{y}_M \end{bmatrix} \quad \mathbf{Y}'' = \begin{bmatrix} \vec{y}_1 \\ \vec{y}_2 \\ \vdots \\ \vec{y}_L \end{bmatrix} \quad (9)$$

for  $N = 940,186$  train,  $M = 288,611$  validation, and  $L = 319,754$  test samples. The observed arrays,  $\mathbf{Y}$ ,  $\mathbf{Y}'$ , and  $\mathbf{Y}''$  for the train, validation, and test sets, respectively, are similarly constructed, shown in Equation 9.

## 2.2. Neural Network Description

The neural network finds a mapping between the inputs and outputs such that  $F(\tilde{\mathbf{X}}) = \hat{\mathbf{Y}}$ , where  $\hat{\mathbf{Y}}$  is the predicted output array after applying the trained neural network,  $F$ , to some input array  $\tilde{\mathbf{X}}$ . The function  $F$  takes the form of Equation 10.

$$F(\tilde{\mathbf{X}}) = \mathcal{A}(\tilde{\mathbf{X}}\mathbf{W}_0 + \mathbf{b})\mathbf{W}_1 = \hat{\mathbf{Y}} \quad (10)$$

In Equation 10,  $\mathcal{A}$  is an activation function,  $\mathbf{W}_0 \in \mathbb{R}^{426 \times 192}$  and  $\mathbf{W}_1 \in \mathbb{R}^{192 \times 31}$  are weight matrices, and  $\mathbf{b} \in \mathbb{R}^{192}$  is a bias vector. The goal is to find  $\mathbf{W}_0$ ,  $\mathbf{W}_1$ , and  $\mathbf{b}$  that minimize a loss metric between  $\hat{\mathbf{Y}}$  and  $\mathbf{Y}$ . We have intended our training set  $(\mathbf{X}, \mathbf{Y})$  to be a well-represented sample of the true population, so that applying the trained network to new observations, such as  $(\mathbf{X}'', \mathbf{Y}'')$ , the loss metric between  $\hat{\mathbf{Y}}''$  and  $\mathbf{Y}''$  will also be small.

### 2.2.1. Neural Network Training

To train the model, we updated the weights and biases— $\mathbf{W}_0$ ,  $\mathbf{W}_1$ , and  $\mathbf{b}$ —using the Adam optimization algorithm (Kingma & Ba, 2017) to minimize our loss metric, the Huber function,  $H(\vec{y}, \hat{y})$ , defined in Equation 11.

$$\text{Huber} \equiv H(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2 & \text{for } |y - \hat{y}| \leq \delta, \\ \delta|y - \hat{y}| - \frac{1}{2}\delta^2 & \text{otherwise.} \end{cases} \quad (11)$$

In Equation 11,  $y$  and  $\hat{y}$  represent the observed and predicted outputs of the test data, respectively, and  $\delta$  is a tunable free parameter, optimized as described in Section 2.2.2.

The advantage of using a Huber loss is that the function behaves like mean squared error (MSE) at lower absolute residual  $|y - \hat{y}|$  values and behaves like mean absolute error (MAE) at higher absolute residual values. MSE tends to bias losses at higher residuals and MAE tends to bias losses at lower residuals. The values of the internal

**Table 1**  
*Hyperparameter Values Optimized Using Bayesian Tree-Structured Parzen Estimator Optimization*

Hyperparameter	Optimized value	Search space
Nodes (hidden layer)	192	Choice(32, 64, ..., 992, 1,024)
Dropout rate	0.4495	Uniform (0, 1)
Huber $\delta$	0.2293	Uniform (0.1, 10.0)
Adam $\beta_1$	0.4501	Uniform (0.001, 0.999)
Adam $\beta_2$	0.9913	Uniform (0.001, 0.999)
Adam learn rate	0.001	Choice ( $10^{-5}$ , $10^{-4}$ , $10^{-3}$ , $10^{-2}$ , $10^{-1}$ )
Batch size	100	Choice (100, 500, 1,000, 5,000, 10,000)

parameters of the Adam method,  $\alpha$ ,  $\beta_1$ , and  $\beta_2$  and the Huber value of  $\delta$  were chosen during hyperparameter optimization (see Section 2.2.2). The network is designed with an input layer of nodes (426 inputs), a single hidden layer (192 nodes), and an output layer (31 outputs). We start with a fully connected network—all nodes connected to all other nodes in adjacent layers—then use a technique called dropout between the input layer and the hidden layer. Dropout randomly removes a fraction of the connections during training to help prevent overfitting of the model to the training data (Srivastava et al., 2014). Our activation function, applied to every node of the hidden layer, was the rectified linear unit (ReLU), where  $\mathcal{A} = \text{ReLU}(x) = \max(0, x)$ . We updated the weights and biases using random sample batches of size 100, without replacement, and we cycled through all  $\sim 10^6$  training data samples 20 times. Therefore, the weights and biases were updated by the Adam optimization algorithm approximately 200,000 times. All model training was implemented by TensorFlow (Abadi et al., 2016) and Keras (Chollet et al., 2015) software.

### 2.2.2. Hyperparameter Optimization

When designing our neural network, we had many design choices, including choosing values for several hyperparameters. The hyperparameters are free parameters, for example, number of hidden layers, number of nodes in each layer, the neural network optimization algorithm (and its internal parameters), the loss function, etc. A list of hyperparameters that we optimized in our model is shown in Table 1. The values were found using a type of Bayesian optimization called Tree-structured Parzen Estimator (TPE). The TPE algorithm was implemented using hyperopt software (Bergstra et al., 2013). The search space for each hyperparameter (right-most column of Table 1) was defined by either randomly choosing discrete options (“choice”) or by selecting from a uniform distribution within minimum and maximum limits (“uniform”). The TPE algorithm is free to choose any combination of the hyperparameter values in the search space; however, rather than an exhaustive or random search, TPE creates a probabilistic model of the function to be optimized and iteratively suggests new hyperparameters based on the history of hyperparameter configurations—quickly converging to an optimized set of hyperparameters. In our case, the function was the neural network and the quantity that we were trying to minimize was the median symmetric accuracy (MSA) metric (defined in Appendix B) between the modeled and observed validation data. The MSA score stopped improving at around 50 iterations (even after 100 iterations there was no improvement beyond that obtained from 50). The values suggested by TPE optimization after 50 iterations are in the middle column of Table 1.

## 3. Model Assessment

A robust set of model-observation comparison metrics is advantageous in obtaining a thorough examination of the model performance in predictive capability (Liemohn et al., 2021). We calculated metrics using individual energy channels, and using flux from energy channels over wider energy ranges. The combined flux across energy channels was calculated as  $J_{\Delta E}$ .

$$J_{\Delta E} = \frac{1}{\Delta E} \int_{E_{\min}}^{E_{\max}} J(E) dE \approx \frac{1}{\Delta E} \sum_{i=1}^n J_{e_i} \Delta e_i. \quad (12)$$

To combine energy channels spanning a range of energy  $\Delta E$  from  $E_{\min}$  to  $E_{\max}$  for metric calculations and visualization, we computed a weighted average,  $J_{\Delta E}$ , of the flux in each energy channel,  $J_{e_i}$ , weighted by the width of each channel,  $\Delta e_i$ .

We used the common logarithm ( $\log_{10}$ ) of the flux to evaluate association, skill, and accuracy. Association between the model output and the observed data was assessed using the Pearson correlation coefficient ( $R$ ). Model skill was assessed using the prediction efficiency (PE) metric. For accuracy, we include the model-observation metric of MSE.

Actual flux values are used to calculate three additional metrics. We also quantified model accuracy using the MSA metric, and we assessed model bias with symmetric signed percentage bias (SSPB). Both MSA and SSPB provide a more realistic interpretation of model performance than when evaluating flux values varying by orders of magnitude (Morley et al., 2018). The inclusion parameter (IP) provides the percentage of predicted points that are within a factor of the corresponding observed points, and was adapted from the exclusion parameter defined in Mukhopadhyay et al. (2021). All metrics used are defined in Appendix B. For IP, we used  $\delta = 2$ , which provides the percentage of predicted points that are within a factor of two.

### 3.1. Global Model Performance

With global model performance, we are evaluating how well the model performed using the entire test data set, which was not involved in training the model. After training our model, we create a predicted set of outputs,  $\hat{\mathbf{Y}}''$ , using the observed test inputs,  $\mathbf{X}''$ .

#### 3.1.1. Qualitative Global Comparisons

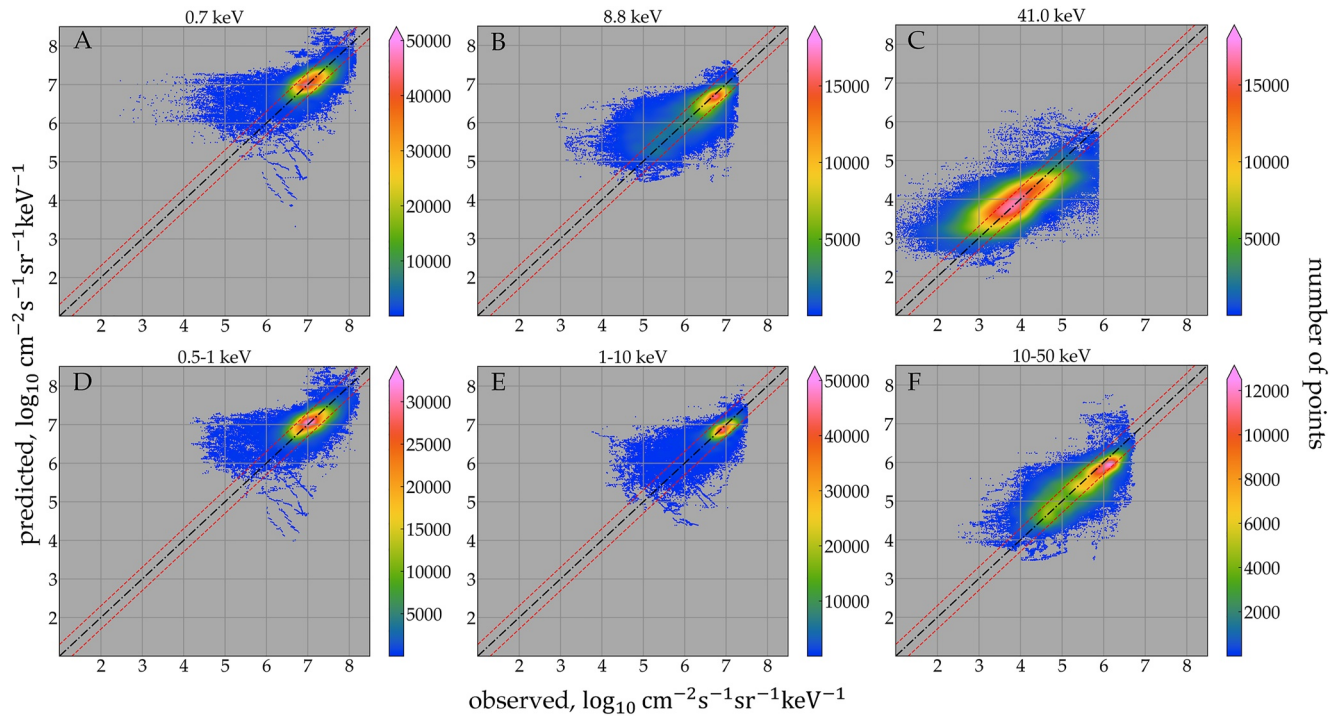
Shown in Figure 2 are the predicted versus observed  $\log_{10}$  flux values for several individual energy channels and combined energy ranges, as in Equation 12.

We chose to highlight the energy ranges of 1–10 and 10–50 keV because they are the energies most responsible for spacecraft surface charging at GEO (Matéo-Vélez et al., 2018; Thomsen et al., 2013); the range 0.5–1 keV was chosen as the population in the plasma sheet that is yet to be energized to these higher energies. Similar panels for all 31 energy channels predicted are shown in Supporting Information S1. The data span the entire test data set period—all observations and predictions from years 2010, 2015, and 2020. Observed flux values are shown on the horizontal axis and predicted flux values are shown on the vertical axis. In each panel a 1:1 dash-dot black line is drawn that indicates a hypothetical perfect prediction. The red dashed lines above and below the 1:1 line represent the extent with a factor of two for flux values ( $\pm \log_{10} 2$  on log scale). The color for each scatterplot varies according to the number of points in a two-dimensional histogram with square bins of dimension 0.34 by 0.34 log flux units ( $\log_{10}(1/[\text{cm}^2 \text{ s sr keV}])$ ). The values of the histogram are then used to interpolate the number of points at each location in the observation-prediction space. The saturation of the number of points beyond the maximum value of each color bar is shown as pink. Note that in panel C (41 keV) there are 10s of observations even lower than 1  $\log_{10}(1/[\text{cm}^2 \text{ s sr keV}])$  that we do not show. For all energies shown in Figure 2 (and in other energies not shown), the majority of observation-prediction paired points are within a factor of two, indicated by the red lines.

#### 3.1.2. Global Metrics

We calculated global metrics using the observed-modeled arrays  $(\mathbf{Y}'', \hat{\mathbf{Y}}'')$ . We present metric values for several energy channels and combined energy ranges in Table 2. The columns are the metrics calculated for either specific energy channels (in keV) or energy channels combined by Equation 12. The bottom two rows show the variance in  $\log_{10}$  of both the observed sample,  $y$ , and the predicted sample,  $\hat{y}$ .

Examining the metrics in Table 2 (and in Supporting Information S1, where all 31 energy channels are listed), we can see that the model has predictive capability. MSA metric values of 100% mean that the median modeled



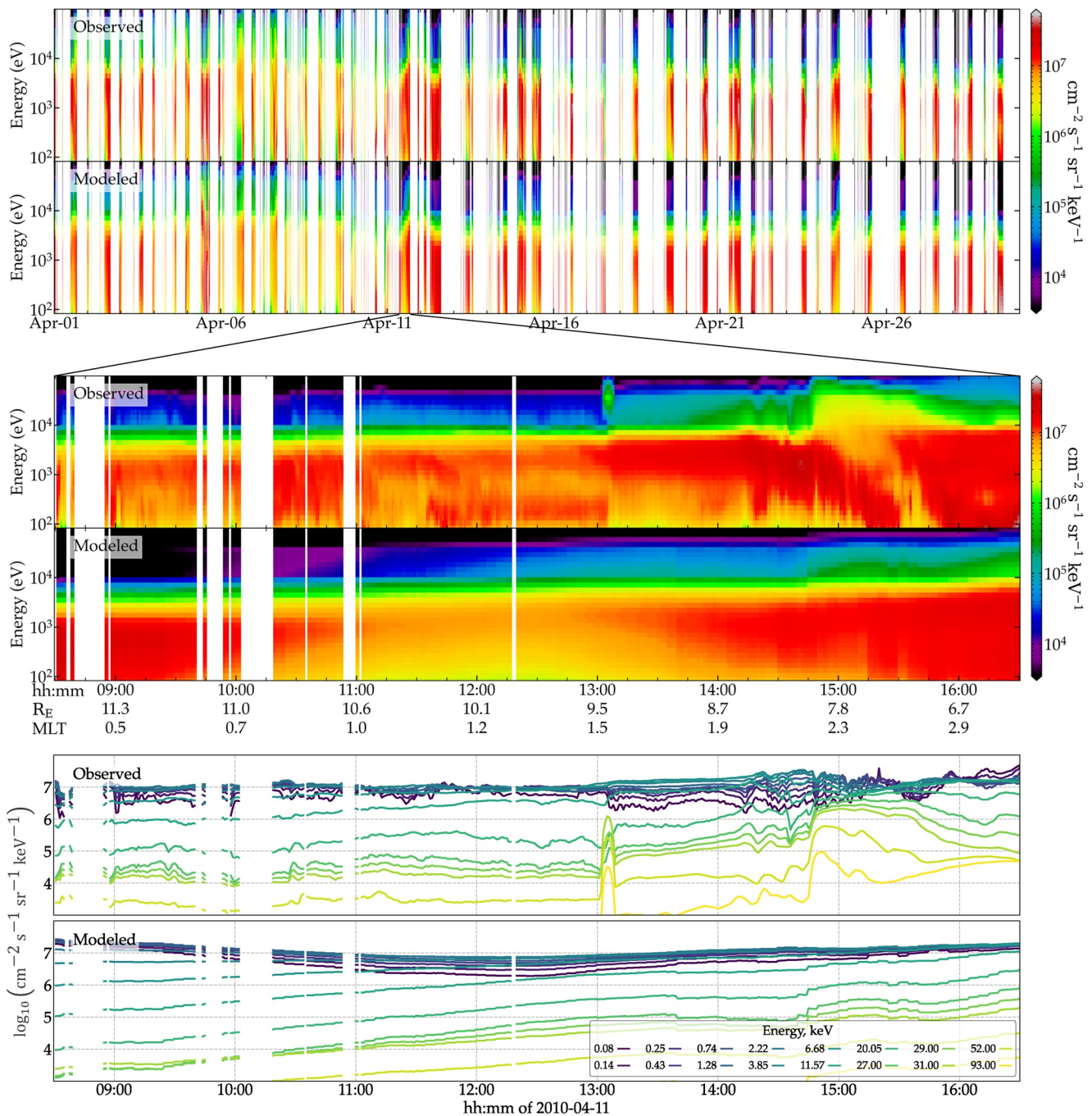
**Figure 2.** In each panel, observed flux is on the horizontal and predicted flux is on the vertical. A black dashed line shows hypothetical perfect prediction ( $\log_{10}\hat{y} = \log_{10}y$ ). Red dash-dot lines show lines of  $\log_{10}\hat{y} = \log_{10}y \pm \log_{10}2$ . Points inside the red lines represent predicted flux that is within a factor of two of the observed flux. The number of overlapping points is shown using the color scale. Saturation in the number of points beyond the maximum of each color bar is shown as pink. Note that the color bar range is not the same for each panel.

flux values are generally at a factor of two within the observed flux values. Most of the MSA metric values in Table 2, and 21 of the total 31 channels are  $\leq 100\%$ . The highest MSA value of 140% occurs for energy 26.4 keV. The eight channels between 0.74 and 5.1 keV (inclusive) have MSA values of  $\leq 50\%$ , whereas the 12 channels between 15.2 and 93 keV (inclusive) have MSA values of  $\geq 85\%$ . Another metric calculated with actual flux values—SSPB—shows the bias of the modeled fluxes compared to the observed fluxes. A model with zero bias would have an SSPB score of 0%. Deviations from 0 indicate either a systematic over-prediction (positive values) or a systematic under-prediction (negative values). For individual energy channels not listed in Table 2, SSPB shows that the model under-predicts all energies between 1 and 31 keV and has up to an 27% under-prediction

**Table 2**  
Selected Observed-Modeled Metrics Used in the Assessment of the Neural Network

Energy (in keV)	0.74	1.3	8.8	31	52	93	1–10	10–50
$R^*$	0.63	0.65	0.76	0.74	0.72	0.75	0.76	0.79
$PE^*$	0.39	0.42	0.57	0.55	0.49	0.54	0.58	0.62
$MSA^\dagger$	50%	46%	69%	87%	130%	130%	33%	77%
$MSE^*$	0.14	0.15	0.20	0.20	0.36	0.39	0.09	0.17
$SSPB^\dagger$	1.2%	−0.5%	−6.8%	−5.8%	20%	16%	−11%	−18%
$IP^\dagger$	73%	77%	61%	54%	44%	43%	84%	59%
$Var(y)^*$	0.23	0.25	0.46	0.45	0.71	0.83	0.21	0.44
$Var(\hat{y})^*$	0.10	0.11	0.25	0.26	0.30	0.38	0.11	0.29

*Note.* Columns show the calculated value for several different energy channels (energies in keV). Metrics calculated using  $\log_{10}$  flux values are marked with \*. Metrics calculated using actual flux values are marked with †. All metrics are defined in Appendix B.



**Figure 3.** Energy spectrograms of observed and modeled particle number flux of the THEMIS-E spacecraft during April 2010 for periods that fit plasma sheet criteria. The middle panel shows an 8 hr period on 11 April. The bottom panel shows line plots for selected energies for observed and modeled flux during the same 8 hr period. THEMIS-E = Time History of Events and Macroscale Interactions during Substorms probe E.

(at 26.4 keV). For energy channels <1 and >31 keV, the model over-predicts and has up to a 23% over-prediction (at 65.5 keV). The IP metric indicates the percentage of predicted points are within a factor of  $\delta$  from the observed points. For our calculations, we used  $\delta = 2$ . Overall, 59% of predicted points are within a factor of two of observed points. The same 21 channels that have MSA  $\leq 100\%$  also have IP  $\geq 50\%$ . The 26 keV energy channel has the lowest IP value of 40%. The predicted variance is systematically lower than the observed variance for all energy channels. In general, as the observed variance increases, the IP decreases. This indicates that the model is not able to capture large variations occurring on short time-scales.

The possible values for PE are in  $(-\infty, 1]$ , with unity indicating perfect prediction and anything less than zero indicating that the prediction is worse than simply using the observed mean for every model value. All PE scores are well above zero, indicating that the model does have some skill in predicting the full test set. The MSE of all channels is low compared to the variance of either the observed or the modeled values, which indicates that the data-model errors are less than the spread of either set against their mean value.

### 3.2. Local Model Performance

Aside from the global view of the model, we also probed whether the model has capability to capture small time-scale (sub-hourly) variations of plasma sheet electron flux. For this purpose, we selected a month of observations from a single spacecraft as it traversed into and out of the plasma sheet. We then zoomed in on 8 hr of single orbit and assessed the model performance using  $(\bar{y}, \hat{y})$  sample data from the month-long period and the 8 hr sub-period.

#### 3.2.1. Qualitative Local Comparisons

Figure 3 is showing the energy spectrograms from both observed and modeled electron flux during 1–29 April 2010, with the observed data from the THEMIS-E probe. The energy range is shown on the vertical axis, and spans the energy channels used for model development, from 82.5 eV to 93 keV. The color represents the electron differential number flux, with units of  $1/[\text{cm}^2 \text{ s sr keV}]$ . Flux  $\geq 5.5 \times 10^7$  is shown as gray and flux  $\leq 3.5 \times 10^3$  is shown as black. White gaps between the flux data are times when the THEMIS-E probe was outside of our model domain and plasma sheet criterion. The sub-period (middle and bottom pairs of panels of Figure 3) ranges from 08:30 to 16:30 on April 11. For this 8 hr period, the horizontal axis is labeled with the time, radial distance, and MLT of the observations. The middle two panels is showing electron flux spectrograms, and the bottom two panels are showing line plots for every other energy channel of observed and modeled flux.

For the month-long period of April 2010, the modeled data are following the trend of the observed data well. In the 8 hr sub-period starting 08:30 11 April 2010, there are some noteworthy similarities and differences between the modeled and observed flux. As the magnitude of the observed flux increases, so too does the modeled flux. This is seen, for example, at 13:00, when there is a sharp rise in flux at all energies in observed data. The model also has an increase in flux, although it is not as sharp, but more gradually and the increase starts earlier, at 12:30. Also, as the spacecraft moves closer toward Earth, the flux values at all energies gradually rise, and this trend is matched by the modeled flux. Noticeably, the model does not resolve the variations that are occurring on minute timescales.

In Figure 4, we show the  $\log_{10}(\text{observed}/\text{modeled})$  flux for the same periods as Figure 3. For both periods, it is easier to see when and at which energies the model over- or under-predicts. The month-long comparison (top panel) of Figure 4 shows that there are some orbits where ranges of flux were either highly under-predicted (blue/gray color) or highly over-predicted (red/black color). Yet for the majority of the periods through the plasma sheet, the modeled flux is within 0.5 orders of magnitude of the observed ( $\pm 0.5$  on color scale). Similarly for the 8 hr period on April 11, there are ranges of energy, and for small amounts of time where the model is under-predicting the observations by more than an order of magnitude. These are most noticeable for  $>10$  keV before 10:00, at 13:00, and 15:00–16:00.

#### 3.2.2. Local Metrics

We calculated the same metrics as in Section 3.1.2 using data for only the April 2010 period and 8 hr sub-period (see Table 3). For both the period and sub-period, the higher energies ( $>8$  keV) are wholly under-predicted by the model. Yet, the trend of the observed flux for these energies is being followed, indicated by the higher  $R$  values. The higher variance of these energies is not able to be captured by the model, leading to low IP and high MSA values.

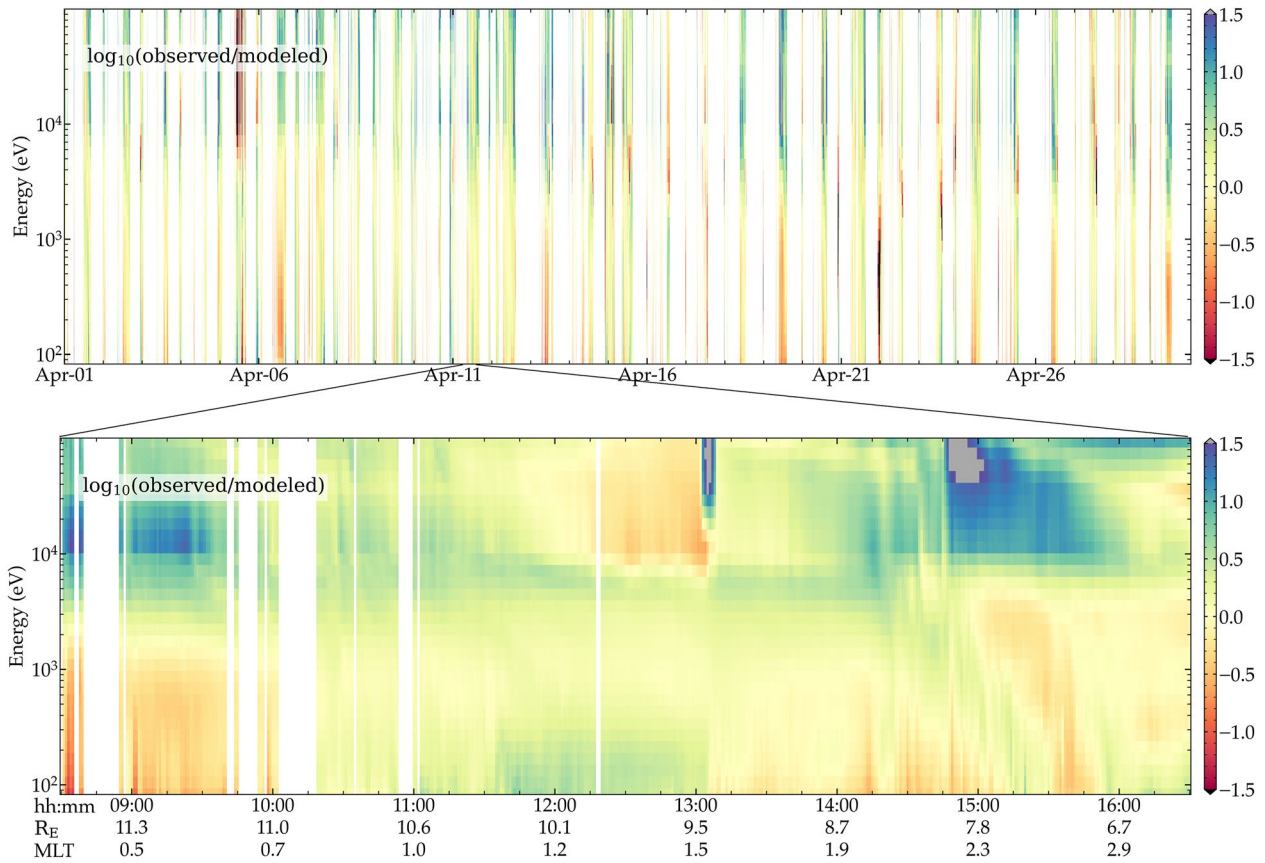


Figure 4. The  $\log_{10}$  (observed/modeled) for the same periods shown in Figure 3.

However, for thermal energies (around 1 keV) where the bulk of the electrons are, the prediction is much better. These energies also have the least variance, and, as mentioned in Section 3.1.2, the model is not able to capture large variations occurring on small time-scales. In the 8 hr sub-period, the IP for these energies are above 90%, and the MSA values are below 30%, indicating a very high accuracy.

We also calculated metrics for six spatial bins, shown schematically in Figure 5. The bins are bounded by radial distance (6, 9, and  $12R_E$ ) and MLT (18, 22, 02, 06). Metrics for two energy ranges 1–10 and 10–50 keV, calculated for these bins, are shown in Table 4. We show only metrics for energy ranges in Table 4 for conciseness and because the metrics for energy channels within each energy range are similar to the metrics for that range. As with previous metrics, the model has better accuracy at lower energies (1–10 keV) than at higher energies. Nearly all of the metrics for both the 1–10 and 10–50 keV energy ranges are better for distances  $>9R_E$  than for distances  $<9R_E$ . This is likely due to the sampling distribution (shown in Figure 1) having much higher samples at distances  $>9R_E$ . Similarly for both energy ranges, the metrics show that the model performs better in the pre-dawn side (02–06 MLT, regions A and D) than in the midnight region (22–02 MLT, regions B and E); and, better in the midnight region than the post-dusk side (18–22 MLT, regions C and F). The improvement in model performance as MLT approaches dawn can be attributed to the drastic decrease in the observed variance for the 1–10 keV energies for the pre-dawn sector versus the post-dusk sector. As shown in Figures 3 and 4 and Table 3, the model is not able to reproduce short time-scale variations. Therefore, in the pre-dawn region, when the observed fluxes do not vary as much, the model has better performance.

### 3.3. Model Comparison

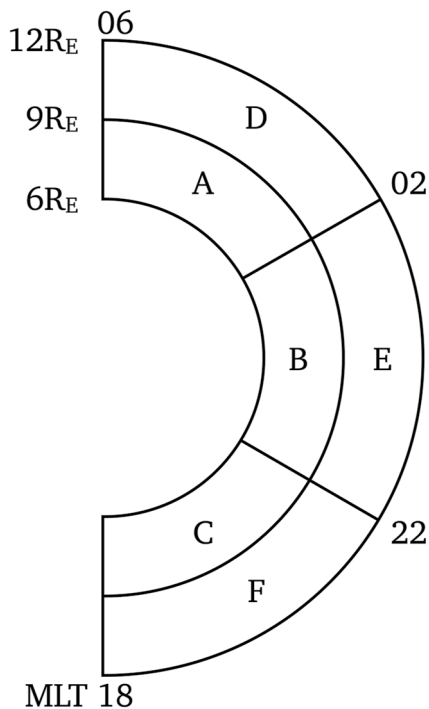
The Dubyagin et al. (2016) study modeled plasma sheet electron density and temperature using SW and IMF input parameters. The model, which we call Dubyagin2016, was developed using THEMIS data from geomagnetic

**Table 3**  
Observed-Modeled Metrics Calculated for the Two Periods Shown in Figure 3 Only

	Energy (in keV)	0.74	1.3	8.8	31	52	93
1–29 April 2010	$R$	0.54	0.50	0.65	0.73	0.75	0.77
	PE	0.17	0.24	0.37	0.29	0.28	0.32
	MSA	59%	50%	85%	180%	220%	290%
	MSE	0.14	0.12	0.21	0.41	0.48	0.56
	SSPB	0.25%	−3.9%	−44%	−140%	−190%	−250%
	IP	66%	74%	55%	34%	29%	25%
	Var( $y$ )	0.16	0.15	0.34	0.58	0.66	0.82
	Var( $\hat{y}$ )	0.06	0.05	0.21	0.37	0.39	0.46
08:30–16:30, 11 April 2010	$R$	0.42	0.51	0.92	0.78	0.83	0.88
	PE	−0.84	−0.11	−0.54	0.29	0.30	0.39
	MSA	25%	19%	140%	120%	140%	160%
	MSE	0.03	0.02	0.15	0.35	0.36	0.41
	SSPB	−8.5%	−7.8%	−140%	−110%	−140%	−160%
	IP	91%	98%	30%	45%	33%	37%
	Var( $y$ )	0.02	0.02	0.10	0.50	0.52	0.66
	Var( $\hat{y}$ )	0.03	0.02	0.18	0.24	0.25	0.31

Note. The format is similar to Table 2. Var( $y$ ) and Var( $\hat{y}$ ) are the observed and modeled sample variance, respectively.

storm periods in the plasma sheet from 6 to 11  $R_E$  and 18–06 MLT (across midnight). We convert Dubyagin2016 electron density and temperature into differential flux assuming either Maxwellian or Kappa distributions. When compared to THEMIS differential flux observations, Dubyagin et al. (2019) found that their derived flux differed by over an order of magnitude from observed flux for energies  $\geq 40$  keV. For thermal energies,  $\sim 1$  to 10 keV, the derived fluxes were within a factor of two of observed fluxes.



**Figure 5.** Spatial regions used for assessing the model's performance at different radial distances and local times.

Here, we present the results of calculating the differential number flux at four reference energies for non-storm periods using the Dubyagin2016 model, and their comparison with output from the model presented in this paper, which we refer to as SWPSNN (Solar Wind Plasma Sheet Neural Network). We use the orbit path of THEMIS-E during April 2010 to build input data sets for both the Dubyagin2016 and SWPSNN models. The observations are a subset of the SWPSNN test data set. They include times when the observations were valid for the SWPSNN and when there were no missing OMNI data. Due to excluding missing OMNI data entirely, rather than interpolating as was done for the data in Section 2.1.2.1, there are fewer data points for the period in this section than in Section 3.2. In Section 3.2, the number of points in April 2010 is 10,280, and in this section, the number of points for April 2010 is 8,048.

Figure 6 shows observed, SWPSNN, and Dubyagin2016 reconstructed electron flux at the four reference energies of 1, 9, 31, and 52 keV. Three different distributions—Maxwellian,  $\kappa = 4$ , and  $\kappa = 2$ —were used to calculate the reconstructed flux from Dubyagin2016 electron density and temperature. Each panel shows the same 6-hr period, starting on 21 April 2010 1200 UT. The period chosen to show is representative of any such 6-hr period from the entire month of April 2010. For the thermal energies at 1 and 9 keV,  $\kappa = 2$  and 4 both are reasonable reconstructions of the observed flux. However,



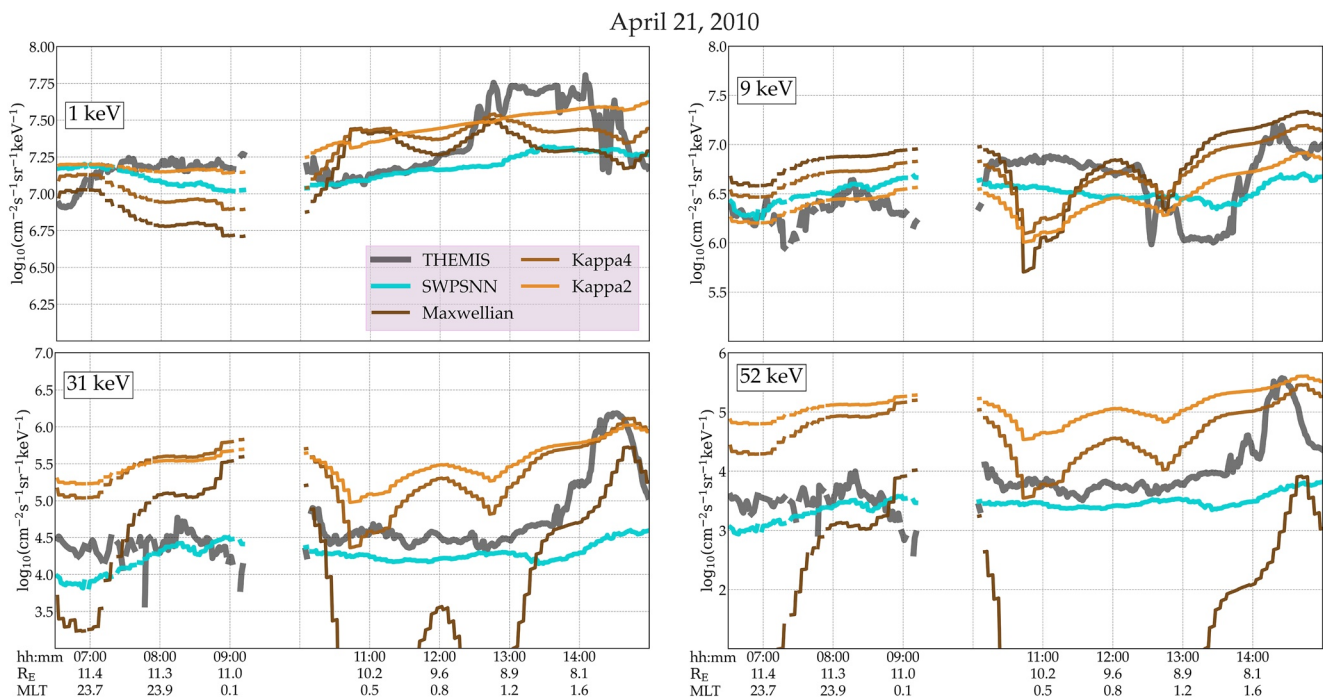
**Table 4**  
Selected Metrics for Spatial Subregions A to F Shown in Figure 5

	A		B		C		D		E		F	
Energy (in keV)	1–10	10–50	1–10	10–50	1–10	10–50	1–10	10–50	1–10	10–50	1–10	10–50
<i>R</i>	0.67	0.73	0.64	0.74	0.60	0.68	0.63	0.75	0.65	0.76	0.67	0.69
PE	0.44	0.42	0.37	0.50	0.33	0.45	0.38	0.56	0.43	0.57	0.44	0.48
MSA	31%	67%	58%	110%	150%	120%	25%	68%	27%	76%	67%	86%
SSPB	−12%	−47%	−36%	−46%	−23%	4.1%	−12%	−19%	−4.1%	−10%	−13%	−5.8%
IP	93%	67%	71%	46%	39%	45%	96%	64%	92%	59%	61%	55%
Var(y)	0.07	0.21	0.18	0.47	0.64	0.49	0.04	0.31	0.07	0.40	0.41	0.32

consistent with the Dubyagin et al. (2019) study, at the superthermal energies—31 and 52 keV, the Dubyagin2016 reconstructed fluxes differ by more than an order of magnitude from observed fluxes.

In contrast, the SWPSNN model, built directly from fluxes, matches reasonably well with the observed fluxes at all of the energies shown in Figure 6.

Using the Dubyagin2016 electron density and temperature, we reconstructed flux for five energies—1, 9, 31, 52, and 93 keV—for all 8,048 points in the April 2010 period. The metrics for  $\kappa = 2$  are shown in Table 5. Metrics for  $\kappa = 4$  and Maxwellian Dubyagin2016 versions are available in Supporting Information S1. Listed in Table 5 are metrics for five Dubyagin2016 modeled reference energies when compared to the THEMIS observations. The  $SS_{MSE}$  (Skill Score using a reference Mean Square Error) row, shown in bold, compares the Dubyagin2016 model output with the SWPSNN by considering the ratio of their MSE values (see Appendix B). For these calculations, the Dubyagin2016 model was used as the reference model; therefore, any values in the  $SS_{MSE}$  row greater than zero can be taken to mean that the SWPSNN has better skill than the Dubyagin2016 model, with comparative skill increasing as the  $SS_{MSE}$  metric goes closer to one. Considering all metrics in Table 5 shows that with increasing energy



**Figure 6.** Comparing Dubyagin2016 flux from Maxwellian,  $\kappa = 4$ , and  $\kappa = 2$  distributions with the Solar Wind Plasma Sheet Neural Network and with Time History of Events and Macroscale Interactions during Substorms observations.

**Table 5**  
*Metrics Calculated for Available DUBYAGIN2016 ( $\kappa = 2$ ) Model Output for April 2010*

Energy (keV)	1	9	31	52	93
$SS_{MSE}$	<b>0.48</b>	<b>0.28</b>	<b>0.40</b>	<b>0.56</b>	<b>0.70</b>
$R$	0.19	0.54	0.71	0.70	0.68
MSA	88%	98%	290%	630%	1,400%
SSPB	55%	-20%	290%	630%	1,400%
IP	54%	51%	27%	17%	8%

*Note.* The  $SS_{MSE}$  metric (bold row) compares the DUBYAGIN2016 model to the Solar Wind Plasma Sheet Neural Network (SWPSNN) model directly: scores that are  $>0$  indicate that the SWPSNN model outperforms the DUBYAGIN2016 model.

above the thermal range (1–10 keV), the performance of the DUBYAGIN2016 model decreases substantially faster than does the performance of the SWPSNN model. The IP, similarly to Tables 2–4, shows the percentage of predicted flux that is within a factor of two of the observed. For  $<10$  keV, DUBYAGIN2016 predicts roughly half of the flux to be within a factor of 2. However, at 31 keV the IP value drops to around a quarter and less for higher energy. We also see vastly decreasing accuracy and bias for the 31, 51, and 93 keV DUBYAGIN2016 fluxes, with the 52 and 93 keV fluxes at 630% and 1,400% MSA errors, respectively. Meanwhile, the SWPSNN flux, while also decreasing accuracy for increasing energies, does not suffer as dramatic loss in MSA (i.e., see Table 3), and the SWPSNN for the higher energies shows much greater skill (0.56 and 0.70  $SS_{MSE}$  for 52 and 93 keV, respectively) than the DUBYAGIN2016 model.

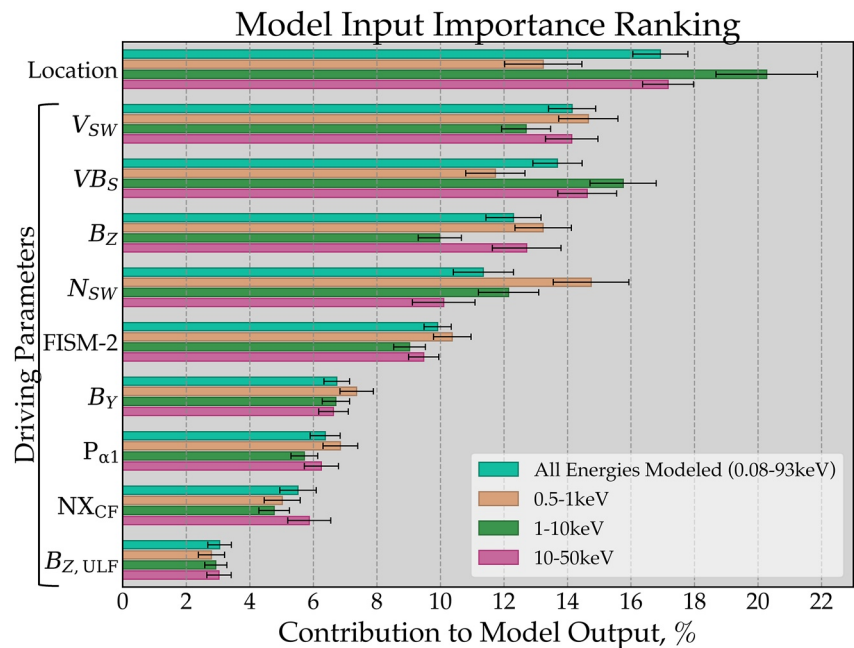
### 3.4. Driving Parameter Importance

Extracting information from black-box type machine learning models, such as the one used by us, is important if we are trying to understand the reasoning why the model performed either well or poorly. We additionally seek to use the model as a tool to infer relationships between input and output parameters of the modeled system. For neural networks, there are now available several techniques to extract and rank the importance of individual inputs that were used by the model to make its predictions. Ranking inputs in this way, while not providing insight into how a particular input is related to a particular prediction, will nonetheless provide useful information about which inputs warrant further study and analysis or which may be inconsequential to the system which we are trying to predict.

#### 3.4.1. Driving Parameter Ranking With DeepSHAP

Of the 426 individual inputs, there are 9 input types, listed in Figure 7, that we consider to be driving parameters of electron flux in the plasma sheet. For determining the importance of each driving parameter to our model output, we use an algorithm called DeepSHAP, which is a combination of SHapley Additive exPlanations (SHAP) (Lundberg & Lee, 2017) and Deep Learning Important FeaTures (Shrikumar et al., 2017) techniques. The DeepSHAP algorithm uses Shapley values (Shapley, 1953) to estimate the relative contribution from each input (one of 426 components of  $\vec{x}$ ) that went into producing each output (one of 31  $\log_{10}$  flux values in  $\vec{y}$ ). We refer the reader to Lundberg and Lee (2017) and Shrikumar et al. (2017) for detailed descriptions of how DeepSHAP works.

We used the DeepSHAP algorithm to calculate SHAP values using 200 randomly sampled inputs from the test set. We calculated the importance of model output with data from the test set since these were not used during the training of the model. The DeepSHAP algorithm provided SHAP values for each sample example (200), for each input (426), for each output (31). The next steps describe how we parsed these 2,641,200 SHAP values in the  $200 \times 426 \times 31$  array. (a) We start by taking the absolute value of all SHAP values, |SHAP|. The SHAP values could be positive or negative depending on how each input was weighted; however, we are interested only in the absolute contribution of each input. (b) We add the |SHAP| values across energy channels,  $\sum_{E_i} |SHAP|$ , resulting in a  $200 \times 426$  array. (c) For the 200 samples, we calculate the sample mean, standard deviation, and uncertainty of the  $\sum_{E_i} |SHAP|$  values. The uncertainty is defined as the standard deviation of the mean ( $SDOM = \frac{\sigma_s}{\sqrt{N}}$ , where  $\sigma_s$  is the sample standard deviation and  $N$  is the sample size). The values of the 200 samples from step 2 were replaced by their mean and uncertainty, reducing the array to  $2 \times 426$ . (d) We organized the inputs as driving parameters by grouping the time steps of similar parameters. For example, the driving parameter  $V_{SW}$  contains all 47 time steps of the SW speed (and similarly for all other inputs with time history). The Location parameter contains the radial distance,  $\cos \phi$ , and  $\sin \phi$  inputs; although, we do not consider the Location to be a “driving” parameter per se. (e) The contribution to model output of each parameter, as defined in step 4, is the sum of the sample means from step 3 within each parameter grouping. Likewise, the uncertainty in the contribution of each parameter is the sum of the uncertainties calculated in step 3. We added the means (as opposed to taking the maximum, e.g.,) to equally account for the contribution from all time steps for each parameter with time history. This reduced the array to a  $2 \times 10$  (sum of means and sum of uncertainties by the 10 input parameters). (f) We repeated steps 2–5 for each desired energy group: all energies modeled (0.08–93 keV; all energy channels), low energy (0.5–1 keV), medium energy (1–10 keV), and high energy (10–50 keV).

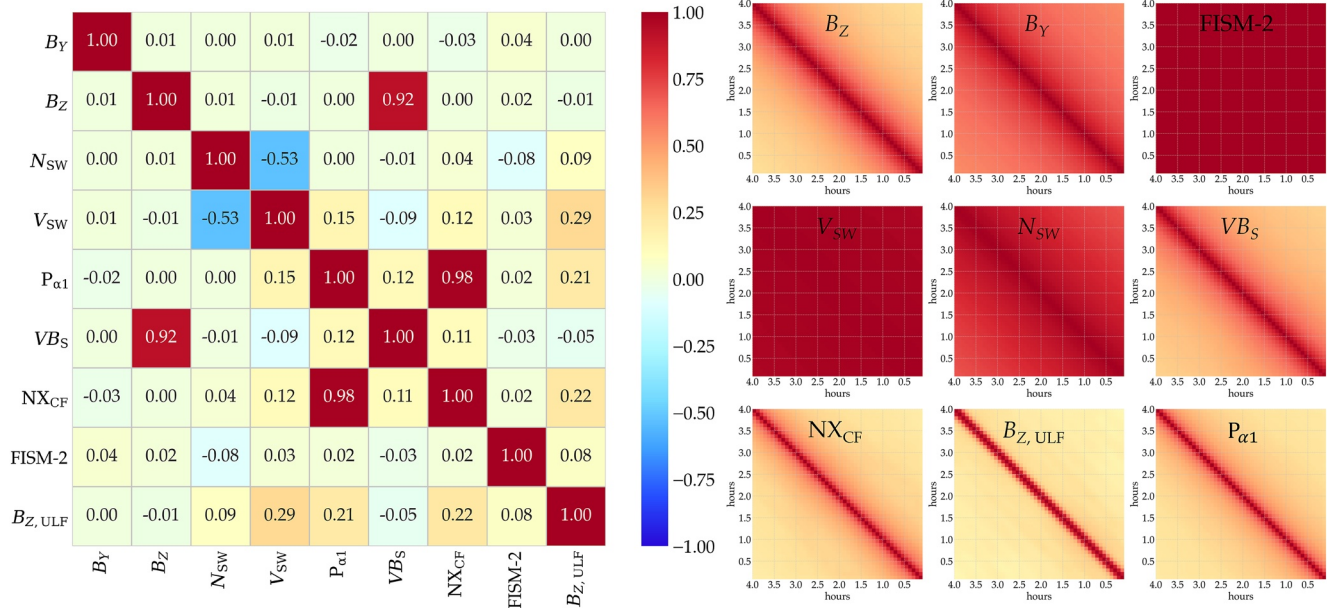


**Figure 7.** Input rankings by type of parameter. The parameters are listed by rank order according to their contribution for the full model output. Input contribution for subsets of model output, denoted by energy, are shown using different colored bars. The uncertainty of each bar represents the standard deviation of the mean. See text for further description.

The contributions of each of the driving parameters to model outputs are shown in Figure 7. We have 10 parameters, listed on the vertical axis. The “Location” parameter includes the contribution of the radial distance,  $\cos \phi$ , and  $\sin \phi$  inputs. The Location is included in this analysis as it is an input to the model; however, we do not consider the sampling location to be an external driving parameter of plasma sheet electron flux. The driving parameters— $V_{SW}$  through  $B_{Z,ULF}$  in Figure 7—include the contribution of that input at each time step from  $-10$  to  $-240$  min in 5 min increments. The inputs  $V_{SW}$  through  $B_{Z,ULF}$  are all of the parameters external to the magnetosphere, and we therefore consider them as driving parameters to plasma sheet electron flux. For each parameter, we calculated four different contributions, shown using four different colored bars. The contribution of each parameter to the full model output, comprising all 31 energy channels modeled, from 0.083 to 93 keV, is shown as cyan. The tan bars show the relative contribution of each parameter to only those channels in the energy range 0.5–1 keV. Similarly, the green bars are for modeled energy channels between 1 and 10 keV, and the magenta bars are for modeled energy channels between 10 and 50 keV. The parameters on the vertical axis are listed in descending order of contribution to the full model output (all energy channels, cyan bars). Uncertainties for each contribution are shown with black horizontal lines and represent the  $\pm$ SDOM.

Figure 7 shows that the least important driving parameters to model output are the coupling functions  $P_{\alpha 1}$  and  $NX_{CF}$ , the IMF  $B_z$  ULF wave power, and IMF  $B_y$ . For the “all energies modeled” group (full model output), the most important driving parameters are the SW speed ( $V_{SW}$ ), the rectified SW electric field ( $VB_S$ ), and the IMF  $B_z$ . For lower energies (0.5–1 keV), SW proton number density,  $N_{SW}$ , ranks as most impactful driving parameter, and for the 1–10 keV energy range,  $VB_S$  is the most impactful driving parameter.

We did not investigate the importance of time history of the inputs because many of them are highly autocorrelated. For example,  $V_{SW}$  has a 4-hr Pearson product-moment autocorrelation of 0.96 (see also Figure 8). On a global scale, whichever  $V_{SW}$  time step ranked the highest would be irrelevant to its actual importance, because the autocorrelation value implies that all other time steps would be almost equally as important. Nonetheless, investigating the time history importance of the driving parameters is a consideration for further investigation. Any such investigation is inherently complicated due to the intercorrelated nature of SW parameters and their time history (e.g., Borovsky, 2018). Furthermore, the interpretation of DeepSHAP output is difficult when assessing input variables that have correlation with each other. As shown in the next section, most of the input variables are not correlated. We further discuss the limitation of the DeepSHAP method for input variables  $P_{\alpha 1}$  and  $NX_{CF}$ , which have high correlation, in Section 4.



**Figure 8.** The Spearman rank correlation coefficients of the driving parameters (left panel). The auto-correlation of each driving parameter with itself is shown in the nine rightmost panels. The center color scale applies to all panels.

### 3.4.2. Driving Parameter Correlations

In Figure 8, we show the Spearman's rank correlation coefficients of driving parameters with one another, along with the Pearson product-moment autocorrelation of each parameter. We excluded the position parameter from this part of the analysis because we are only concerned with inputs which have time history. The left panel of Figure 8 shows Spearman's rank correlation of driving parameters for all samples from the combined train, validation, and test data. The Spearman's rank correlation assesses the monotonic relationship between variables regardless of the linearity of such relationships. It is an appropriate measure of correlation in our case, since the neural network is a non-linear function approximator. Mostly, the parameters are not correlated with each other. The largest correlation between two parameters are the coupling functions  $P_{\alpha 1}$  and  $NX_{CF}$ , which is expected as these two coupling functions share common factors in their calculations. Similarly, IMF  $B_Z$  and  $V_{B_S}$  have positive correlation.  $N_{SW}$  and  $V_{SW}$  have negative correlation, as expected (e.g., Le Chat et al., 2012). Aside from these and other minor correlations between the remaining parameters, there is generally a lack of rank correlation. In contrast, the autocorrelation of each driving parameter with its own time history is markedly much higher. In the nine right panels of Figure 8, we show the autocorrelation of each parameter using the combined train and test sets as the observed sample. In each panel, the correlation of input  $I$  at time  $t_a$  is correlated with  $I$  at time  $t_b$  and  $|t_a - t_b|$  ranges from 0 to 230 min. The parameter with the highest autocorrelation is FISM-2, having a 230 min correlation of  $>0.99$ . The parameter with the least amount of auto-correlation— $B_{Z, ULF}$ —has a 230 min correlation coefficient of 0.14.

## 4. Discussion

Our assessment of the model is that it captures the variations of electron flux in the near-Earth plasma sheet overall, which is supported by the model-observation goodness-of-fit metrics. In Table 2, our model accuracy metric, MSA, is between 46% and 130%. We see also in Figure 2 that the bulk of modeled flux values are within a factor of 2 (in between the red lines) for all energies and energy ranges shown. An overall bias does exist, where for energies 1–31 keV the model has an under-prediction of  $\leq 27\%$  and an over-prediction of  $\leq 23\%$  for  $<1$  and  $>31$  keV energies. The metric values of  $R$  and PE likewise depend on energy. For individual energy channels,  $R$  values range from 0.59 to 0.77 and PE values range from 0.34 to 0.59. The PE values are well above zero for each channel, showing the skill that this model has in predicting plasma sheet flux.

We compared the model from this paper to a reference model whose output is electron density and temperature. From these bulk plasma properties, and assuming a distribution function such as Maxwellian, we can obtain

differential electron flux at any energy. The results of a direct comparison between the Dubyagin2016 model and this paper's model, the SWPSNN, show that the SWPSNN model has better predictive skill at all energies, and this skill improves with increasing energy above the thermal energy range (above 1–10 keV). Overall, the SWPSNN model, built on fluxes as recommended by Dubyagin et al. (2019), is better at than the Dubyagin2016 model at reproducing suprathermal flux in the near-Earth plasma sheet.

Despite overall improvements, the model fails to capture sub-hourly variations. Although variations on these timescales in the plasma sheet are clearly occurring in the observed data (i.e., Figure 3), the model cannot predict them. A couple of explanations for this deficiency are plausible. First, the inputs to the model, although used at a 5-min resolution, do not themselves typically have much variance on these short time scales. In Figures 7 and 8, we see that the driving parameters that have the most influence on model prediction also have some of the highest amounts of autocorrelation. The  $V_{SW}$  input, for example, has a 4-hr autocorrelation of  $>0.96$ . The next two most impactful inputs,  $VB_S$  and IMF  $B_z$ , have 4-hr autocorrelations of  $>0.26$  and 2-hr autocorrelations of  $>0.37$ . Indeed, many previous studies have examined the hours-long autocorrelation of SW and IMF parameters at 1 AU (e.g., Gosling & Bame, 1972; King & Papitashvili, 2005; Richardson & Paularena, 2001; Wicks et al., 2010). Since the model's most impactful driving parameters have hours-long auto-correlation, we surmise that it was unlikely that this model would be able to consistently predict magnetospheric variations on timescales less than that by which the model inputs are appreciably varying. Denton et al. (2019) faced a similar problem with their model being unable to account for variations in the plasma sheet that are occurring on timescales shorter than that of their input parameters (for which they used the  $Kp$  index with a 3-hr cadence). Second, much of the short time-scale variation seen in the observed flux data is the result of local plasma processes that are occurring within the plasma sheet, albeit influenced by external drivers. That is, the plasma sheet in general is responding to the SW variations, such as increased cross-tail electric field due to increased SW convective electric field. But on local scales—both temporally and spatially—electrons in the plasma sheet are beholden to respond to the local forces acting on them.

The FISM-2 driving parameter provides a relatively high contribution to model output, at 9%–10% for all energy groups. FISM-2 represents the flux for photons at energies between 19 and 23 eV. The long-term trend in the flux of photons at these energies are highly dependent on the 11-year sunspot cycle and the  $\sim 28$ -day solar rotation period. As the sunspot cycle phases from solar minimum to solar maximum, background (non-flare) solar EUV flux varies from low to high. By including FISM-2 model output as an input to our model, we therefore included information about the overall solar activity congruous with the sunspot cycle and the solar rotation period. That levels of solar activity hold sway over electrons in the plasma sheet has been modeled by Yue et al. (2015) and Zou et al. (2020). In Yue et al. (2015) study, the sunspot number was used as an indicator for plasma sheet pressure. Zou et al. (2020) included F10.7 (as possibly a proxy for sunspot number and/or solar EUV flux) as an input to their model. A possible explanation for an effect of solar activity on the near-Earth electron plasma sheet is that increased solar EUV flux will create a higher dayside ionospheric conductance from increased photoionization. Increased dayside conductance drives stronger field-aligned Region 1 and 2 currents between the plasma sheet and nightside ionosphere (Ohtani et al., 2014). The result is increased ionospheric outflow to the plasma sheet, leading to background plasma sheet electron flux levels that are dependent on the solar cycle. An investigation into any direct links between solar flare impact events or solar rotation and plasma sheet electron flux dependence is a consideration for a future study.

Two of the three SW-magnetosphere coupling functions, similar in form to  $V B \sin^\alpha \left( \frac{\theta}{2} \right)$ , (where  $\alpha$  is optimized for correlation with some geomagnetic index and  $\theta$  is the IMF clock angle), had meager impact on model output. Functions that contain the sine of the clock angle not being an important contribution to modeled electron flux in the plasma sheet conflicts with one of the results from the Stepanov et al. (2021) study, which had identified  $E_{kl} = V B_{XY} \sin^2 \left( \frac{\theta}{2} \right)$  as a main driver of near-Earth plasma sheet 10–200 keV electron flux. The SW-magnetosphere coupling functions that we used,  $NX_{CF}$  and  $P_{\alpha 1}$ , and especially their  $\alpha$  parameter, were optimized to best correlate with a particular geomagnetic index, and not with the electron flux variations within the plasma sheet directly. Although our model did not weight these coupling functions very much, it remains inconclusive whether they could—after optimizing them for the plasma sheet—provide a larger contribution to the driving of plasma sheet electron flux. Moreover, these coupling functions are highly correlated with each other, which makes their ranking from the DeepSHAP method more difficult to interpret. It is unlikely that the DeepSHAP algorithm could distinguish between these inputs to the model when assigning importance weights. The full contribution of SW-magnetosphere coupling functions to plasma sheet electron flux variations is a consideration for further study.

The least important input to impact model output was the IMF  $B_z$  wave power. Although Wang et al. (2017) found some dependence of plasma sheet electron temperature to  $B_{z,ULF}$ , the dependence was limited compared to the larger domain of our model. In Wang et al. (2017), the  $B_{z,ULF}$  was only relevant to plasma sheet temperature during strong  $B_{z,ULF}$  and weak or northward IMF. Our model, not selecting for these conditions, failed to assign much weight to the  $B_{z,ULF}$  inputs. Moreover, we are predicting electron fluxes at individual energies, whereas the Wang et al. (2017) study was investigating  $B_{z,ULF}$  on electron temperature only.

The Zou et al. (2020) study, similar in scope to this paper, modeled near-Earth plasma sheet 0.06–293 keV differential electron flux with  $R$  values above 0.91 and MSE values less than 0.13 (for comparison, see Table 2). In contrast to this paper, the Zou et al. (2020) model used inputs that were internal to the magnetosphere (geomagnetic indices) and input time history of 48 hr, both of which could add sufficient information of near-Earth electron flux variations to obtain such high correlation coefficients and low MSE values. Indeed, Stepanov et al. (2021) found that 10–200 keV electron flux in the near-Earth plasma sheet has a response “memory” to  $E_{kl}$  of up to 24 hr. Increasing the time history of the SWPSNN model input parameters is a consideration for a future investigation to further quantify the response of electron flux in the plasma sheet to longer durations of variations in these parameters. Geomagnetic indices and their time history may also be added to the SWPSNN model in a future study to investigate the correlation of plasma sheet electron flux with geomagnetic disturbances.

Uncertainty within the model comes from uncertainty in the SW measurements and their assumed propagation from L1 to the Earth's bow shock (e.g., King & Papitashvili, 2005; Weimer & King, 2008). There is also uncertainty that arises in using data from a single mission in training the model. For example, comparisons between THEMIS electron flux and other spacecraft, such as the LANL geosynchronous satellites, have shown a consistent offset in observed fluxes (Ni et al., 2011). Uncertainty also comes from the lack of information about local electron acceleration and transport processes occurring within the plasma sheet. We did not select any specific periods of geomagnetic activity when selecting the data that was used for the development of this model. Therefore, the minute-averaged data were sampled from the prevailing distribution of geomagnetic activity. Over long periods, this is dominated by quiet conditions (see, e.g., bottom panels in Figure 1 of Ganushkina et al. (2021), which shows the occurrence frequencies of geomagnetic index values over 16 years). Consequently, the periods in our data set where a large variation in flux occurred on short timescales were very few compared to periods where flux remained relatively constant. It is most likely that the algorithm optimized for the over-abundance of quiet geomagnetic activity samples. If we were to retrain the neural network but include only times during high geomagnetic activity as samples for training, it is possible that we would obtain different results. This is a consideration for a future study.

## 5. Conclusions

Suprathermal electrons from the plasma sheet appreciably contribute to inner magnetosphere dynamics and severe space weather environments at GEO and MEO. Specifying the intensity, duration, and location of electron flux enhancement in the near-Earth plasma sheet is important for increased understanding, improved forecasting, and potentially obtaining more accurate boundary conditions for physics-based inner magnetosphere models. Using empirical modeling, we investigated three relationships between solar driving and electron flux in the plasma sheet: (a) the possibility of using high resolution SW input parameters to improve predictions of short-term plasma sheet electron flux variations; (b) the indirect influence of solar EUV flux; and (c) whether SW-magnetosphere coupling functions are an important driver of plasma sheet electron flux.

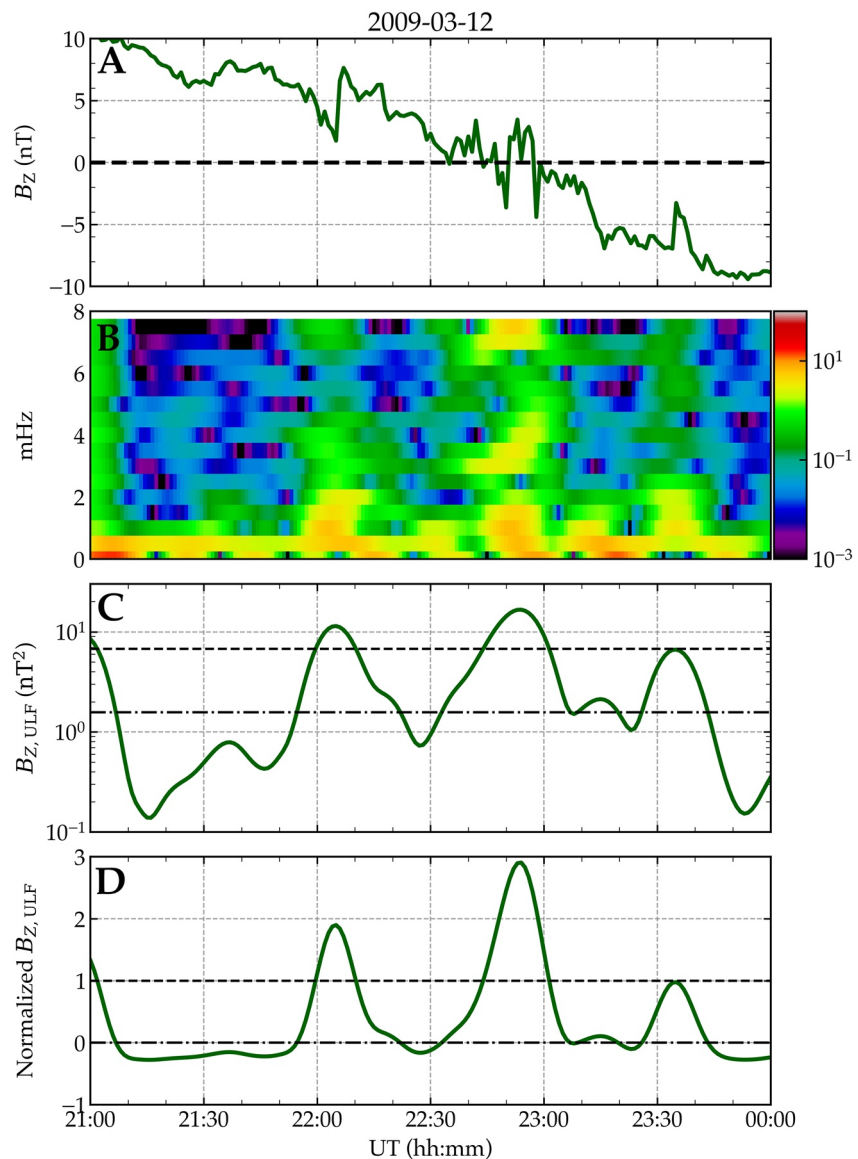
The model is a neural network that predicts 0.08–93 keV electron flux in the near-Earth plasma sheet from 6 to  $12R_E$  in radial distance and 18–06 MLT in azimuthal range. We trained the model using 12 years of data from the THEMIS mission to sample the plasma sheet, and all of the model's driving parameters are external to the magnetosphere: SW and IMF data from OMNI, and solar EUV flux from FISM-2. Our findings are the following:

1. Short-term variations— $\lesssim 1$  hr—of electron flux in the plasma sheet are not predicted from high resolution—5 min—SW input parameters.
2. The most important drivers of the modeled flux are  $IMFB_z$ ,  $V_{SW}$ , and  $VB_S$ , which is consistent with previous studies. Solar EUV flux provided a contribution to model output, yet was markedly not as influential as  $IMFB_z$ ,  $V_{SW}$ ,  $VB_S$ , or  $N_{SW}$ .
3. SW-magnetosphere coupling functions with a  $\sin^\alpha\left(\frac{\theta}{2}\right)$  term and  $B_{z,ULF}$  are amongst the least important external drivers of modeled plasma sheet electron flux.

The model has potential to be used as a representation of plasma sheet electron flux, given inputs of upstream SW, and solar EUV photon flux. It could be used as a boundary condition for first-principles based inner magnetosphere models or as part of a specification model for the electron radiation environment of near-Earth space.

### Appendix A: Calculation of $B_{Z,ULF}$

Figure A1 shows an example of the wave spectra and calculated wave power index,  $B_{Z,ULF}$ , for a 3 hr period on 12 March 2009. We show the interplanetary magnetic field (IMF)  $B_Z$  from OMNI in panel (a). Panel (b) shows the wave spectra from 0 to 8 mHz. Panel (c) is the calculated  $B_{Z,ULF}$  index found by adding all the spectra from 1.667 to 6.667 mHz at each time step. Panel (d) shows the standardized  $B_{Z,ULF}$  using Equation 2. The climatic mean and standard deviation were calculated with the period from January 2008 through December 2020.



**Figure A1.** A period of 3 hr of interplanetary magnetic field  $B_Z$  (a); the ultra-low frequency (ULF) wave spectra in the Pc5 band (b);  $B_{Z,ULF}$  index (c); and standardized  $B_{Z,ULF}$  (d). In panels (c and d), the climatic  $B_{Z,ULF}$  mean and standard deviation are shown with dash-dot and dashed lines, respectively.

$$w(n) = 0.5 - 0.5 \cos\left(\frac{2\pi n}{M-1}\right) \quad (\text{A1})$$

In Equation A1,  $0 \leq n \leq M - 1$ ,  $M$  is the number of points, and  $w(n)$  is the value of the Hann window at the  $n$ th point. A summary of our steps (modified from Wang et al. (2017)) is as follows:

1. Fill missing data with an Akima spline cubic interpolation (Akima, 1970).
2. Remove the background by subtracting a 60 min rolling mean from the original signal.
3. For a window of 30 min, shifted each minute, we calculated the wave spectra from the Fast Fourier Transform coefficients of the smoothed signal. In this step, the smoothed IMF  $B_z$  signal was first shaped by a Hann window (Equation A1).
4. The wave power index was calculated by adding the spectra in the Pc5 ultra-low frequency band (1.667–6.667 mHz).

## Appendix B: Metric Definitions

The metrics used in our study are defined for reference in Table B1.

Metric	Common name	Formula	Range	Optimal
$R$	Pearson correlation coefficient	$\frac{\sum_{i=1}^n [y_i - \bar{y}][\hat{y}_i - \bar{\hat{y}}]}{\sqrt{\sum_{i=1}^n [y_i - \bar{y}]^2} \sqrt{\sum_{i=1}^n [\hat{y}_i - \bar{\hat{y}}]^2}}$	$[-1, 1]$	1
PE	Prediction efficiency	$1 - \frac{\text{MSE}(y, \hat{y})}{\text{Var}(y)}$	$(-\infty, 1]$	1
$\text{SS}_{\text{MSE}}$	Reference skill score	$1 - \frac{\text{MSE}(y, \hat{y})}{\text{MSE}_r(y, \hat{y}_r)}$	$(-\infty, 1]$	1
MSA	Median symmetric accuracy	$100 \left( e^{\left  \ln\left(\frac{\hat{y}}{y}\right) \right } - 1 \right)$	$[0, \infty)$	0
MSE	Mean square error	$\frac{1}{N-1} \sum_{i=1}^N (\hat{y}_i - y_i)^2$	$[0, \infty)$	0
SSPB	Symm. signed % bias	$100 \left( \text{sign} \left[ \text{M} \left( \ln \left( \frac{\hat{y}}{y} \right) \right) \right] \left( e^{\left  \ln \left( \frac{\hat{y}}{y} \right) \right } - 1 \right) \right)$	$(-\infty, \infty)$	0
IP	Inclusion parameter	$100 \left( \frac{1}{N} \sum_{i=1}^N \text{bool} \left[ \max \left( \frac{y_i}{\hat{y}_i}, \frac{\hat{y}_i}{y_i} \right) \leq \delta \right] \right)$	$[0, 100]$	100
Var	Sample variance	$\frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2$	n/a	n/a
$\bar{y}$	Sample mean	$\frac{1}{N} \sum_{i=1}^N y_i$	n/a	n/a

*Note.* In all formulas,  $y$  is the observed sample and  $\hat{y}$  is the modeled sample.  $\text{M}$  is the median function.  $\text{MSE}_r$  and  $\hat{y}_r$  are the mean squared error and modeled sample, respectively, of a reference model. The Range column shows the possible values for each metric. The Optimal column shows the value for each metric if the model has perfect data-model comparison.



## Data Availability Statement

Model implementation source code and examples are provided at <https://github.com/briswi/ESWPSNN>. The source code is also released with a DOI at <https://doi.org/10.5281/zenodo.7083352> (Swiger, 2022). Flare Irradiance Spectral Model Version 2 data are available from the LASP Interactive Solar Irradiance Data Center data center (<https://lasp.colorado.edu/lisird/>). OMNI and Time History of Events and Macroscale Interactions during Substorms data are available via CDAWeb (<https://cdaweb.gsfc.nasa.gov/>). The input and output data for the new model is deposited at the DeepBlue Data Repository (<https://deepblue.lib.umich.edu/data>), and can be accessed publicly at <https://doi.org/10.7302/araa-6f62> (Swiger et al., 2022). The authors thank developers of SpacePy, Matplotlib, pandas, TensorFlow/Keras, Hyperopt, and SHAP for making their software free and open source.

## Acknowledgments

BS performed the methodology and wrote the original draft. ML and NG provided resources, supervision and project administration. SD developed and implemented code for Dubyagin2016 fluxes. All authors participated in the investigation and formal analysis; and, reviewed and edited the manuscript. Work in this manuscript funded by the NASA FINESST Grant 80NSSC20K1504, and other NASA Grants NNX17A148G, 80NSSC20K0353, and NNX17AB87G. The contributions by SD and NG were also partly supported by the Academy of Finland (Grant 339329).

## References

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). Tensorflow: A system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation*. USENIX Association.
- Akima, H. (1970). A new method of interpolation and smooth curve fitting based on local procedures. *Journal of the Association for Computing Machinery*, *17*(4), 589–602. <https://doi.org/10.1145/321607.321609>
- Allison, H. J., Horne, R. B., Glauert, S. A., & Del Zanna, G. (2019). On the importance of gradients in the low-energy electron phase space density for relativistic electron acceleration. *Journal of Geophysical Research: Space Physics*, *2019JA026516*. <https://doi.org/10.1029/2019ja026516>
- Angelopoulos, V. (2008). The THEMIS mission. *Space Science Reviews*, *141*(1–4), 5–34. <https://doi.org/10.1007/s11214-008-9336-1>
- Angelopoulos, V., Baumjohann, W., Kennel, C. F., Coroniti, F. V., Kivelson, M. G., Pellat, R., et al. (1992). Bursty bulk flows in the inner central plasma sheet. *Journal of Geophysical Research*, *97*(A4), 4027. <https://doi.org/10.1029/91ja02701>
- Angelopoulos, V., Sibeck, D., Carlson, C. W., McFadden, J. P., Larson, D., Lin, R. P., et al. (2008). First results from the THEMIS mission. *Space Science Reviews*, *141*(1–4), 453–476. <https://doi.org/10.1007/s11214-008-9378-4>
- Arnoldy, R. L. (1971). Signature in the interplanetary medium for substorms. *Journal of Geophysical Research*, *76*(22), 5189–5201. <https://doi.org/10.1029/JA076i022p05189>
- Artemyev, A. V., Angelopoulos, V., Runov, A., & Zhang, X. J. (2020). Ionospheric outflow during the substorm growth phase: THEMIS observations of oxygen ions at the plasma sheet boundary. *Journal of Geophysical Research: Space Physics*, *125*(7), e2019JA027612. <https://doi.org/10.1029/2019JA027612>
- Artemyev, A. V., Zhang, X. J., Angelopoulos, V., Runov, A., Spence, H. E., & Larsen, B. A. (2018). Plasma anisotropies and currents in the near-Earth plasma sheet and inner magnetosphere. *Journal of Geophysical Research: Space Physics*, *123*(7), 5625–5639. <https://doi.org/10.1029/2018ja025232>
- Balikhin, M. A., Boynton, R. J., Billings, S. A., Gedalin, M., Ganushkina, N., Coca, D., & Wei, H. (2010). Data based quest for solar wind-magnetosphere coupling function. *Geophysical Research Letters*, *37*(24), L24107. <https://doi.org/10.1029/2010gl045733>
- Bame, S. J., Asbridge, J. R., Felthaus, H. E., Olson, R. A., & Strong, I. B. (1966). Electrons in the plasma sheet of the Earth's magnetic tail. *Physical Review Letters*, *16*(4), 138–142. <https://doi.org/10.1103/PhysRevLett.16.138>
- Bergstra, J., Yamins, D., & Cox, D. (2013). Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In S. Dasgupta & D. McAllester (Eds.), *Proceedings of the 30th International Conference on Machine Learning* (Vol. 28, pp. 115–123). PMLR. Retrieved from <http://proceedings.mlr.press/v28/bergstra13.html>
- Borovsky, J. E. (2018). On the origins of the intercorrelations between solar wind variables. *Journal of Geophysical Research: Space Physics*, *123*(1), 20–29. <https://doi.org/10.1002/2017ja024650>
- 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>
- Boynton, R. J., Balikhin, M. A., Billings, S. A., Wei, H. L., & Ganushkina, N. (2011). Using the NARMAX OLS-ERR algorithm to obtain the most influential coupling functions that affect the evolution of the magnetosphere. *Journal of Geophysical Research*, *116*(A5), A05218. <https://doi.org/10.1029/2010JA015505>
- Burin des Roziers, E., Li, X., Baker, D. N., Fritz, T. A., Friedel, R., Onsager, T. G., & Dandouras, I. (2009). Energetic plasma sheet electrons and their relationship with the solar wind: A Cluster and Geotail study. *Journal of Geophysical Research*, *114*(A2), A02220. <https://doi.org/10.1029/2008ja013696>
- Burin des Roziers, E., Li, X., Baker, D. N., Fritz, T. A., McPherron, R. L., & Dandouras, I. (2009). Cluster observations of energetic electron flux variations within the plasma sheet. *Journal of Geophysical Research*, *114*(A11), A11208. <https://doi.org/10.1029/2009ja014239>
- Burton, R. K., McPherron, R. L., & Russell, C. T. (1975). An empirical relationship between interplanetary conditions and Dst. *Journal of Geophysical Research*, *80*(31), 4204–4214. <https://doi.org/10.1029/JA080i031p04204>
- Cao, J., Duan, A., Reme, H., & Dandouras, I. (2013). Relations of the energetic proton fluxes in the central plasma sheet with solar wind and geomagnetic activities. *Journal of Geophysical Research: Space Physics*, *118*(11), 7226–7236. <https://doi.org/10.1002/2013ja019289>
- Chamberlin, P. C., Eparvier, F. G., Knoer, V., Leise, H., Pankratz, A., Snow, M., et al. (2020). The flare irradiance spectral model-version 2 (FISM2). *Space Weather*, *18*(12), e2020SW002588. <https://doi.org/10.1029/2020sw002588>
- Chappell, C. R., Moore, T. E., & Waite, J. H. (1987). The ionosphere as a fully adequate source of plasma for the Earth's magnetosphere. *Journal of Geophysical Research*, *92*(A6), 5896. <https://doi.org/10.1029/JA092iA06p05896>
- Chen, Y., Reeves, G. D., & Friedel, R. H. W. (2007). The energization of relativistic electrons in the outer Van Allen radiation belt. *Nature Physics*, *3*(9), 614–617. <https://doi.org/10.1038/nphys655>
- Choi, H.-S., Lee, J., Cho, K.-S., Kwak, Y.-S., Cho, I.-H., Park, Y.-D., et al. (2011). Analysis of geo spacecraft anomalies: Space weather relationships. *Space Weather*, *9*(6), S06001. <https://doi.org/10.1029/2010sw000597>
- Chollet, F. (2015). Keras. Retrieved from <https://keras.io>
- Davis, V. A., Mandell, M. J., & Thomsen, M. F. (2008). Representation of the measured geosynchronous plasma environment in spacecraft charging calculations. *Journal of Geophysical Research*, *113*(A10), A10204. <https://doi.org/10.1029/2008ja013116>
- Denton, M. H., Taylor, M. G. T., Rodriguez, J. V., & Henderson, M. G. (2019). Extension of an empirical electron flux model from 6 to 20 Earth radii using cluster/rapid observations. *Space Weather*, *17*(5), 778–792. <https://doi.org/10.1029/2018sw002121>

- Duan, A. Y., Cao, J. B., Dunlop, M., & Wang, Z. Q. (2014). Energetic electron bursts in the plasma sheet and their relation with BBFs. *Journal of Geophysical Research: Space Physics*, 119(11), 8902–8915. <https://doi.org/10.1002/2014ja020169>
- Dubyagin, S., Ganushkina, N., & Liemohn, M. (2019). On the accuracy of reconstructing plasma sheet electron fluxes from temperature and density models. *Space Weather*, 17(12), 1704–1719. <https://doi.org/10.1029/2019sw002285>
- 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>
- Foster, J. C., Fairfield, D. H., Ogilvie, K. W., & Rosenberg, T. J. (1971). Relationship of interplanetary parameters and occurrence of magnetospheric substorms. *Journal of Geophysical Research*, 76(28), 6971–6975. <https://doi.org/10.1029/JA076i028p06971>
- Fu, H. S., Khotyaintsev, Y. V., André, M., & Vaivads, A. (2011). Fermi and betatron acceleration of suprathermal electrons behind dipolarization fronts. *Geophysical Research Letters*, 38(16), L16104. <https://doi.org/10.1029/2011GL048528>
- Ganushkina, N. Y., Liemohn, M. W., Dubyagin, S., Daglis, I. A., Dandouras, I., De Zeeuw, D. L., et al. (2015). Defining and resolving current systems in geospace. *Annales Geophysicae*, 33(11), 1369–1402. <https://doi.org/10.5194/angeo-33-1369-2015>
- Ganushkina, N. Y., Swiger, B., Dubyagin, S., Matéo-Vélez, J. C., Liemohn, M. W., Sicard, A., & Payan, D. (2021). Worst-case severe environments for surface charging observed at LANL satellites as dependent on solar wind and geomagnetic conditions. *Space Weather*, 19(9), e2021SW002732. <https://doi.org/10.1029/2021sw002732>
- Garrett, H. B. (1979). Review of quantitative models of the 0- to 100-keV near-Earth plasma. *Reviews of Geophysics and Space Physics*, 17(3), 397–417. <https://doi.org/10.1029/RG017i003p00397>
- Garrett, H. B. (1981). The charging of spacecraft surfaces. *Reviews of Geophysics*, 19(4), 577–616. <https://doi.org/10.1029/RG019i004p00577>
- Gosling, J. T., & Bame, S. J. (1972). Solar-wind speed variations 1964–1967: An autocorrelation analysis. *Journal of Geophysical Research*, 77(1), 12–26. <https://doi.org/10.1029/JA077i001p00012>
- Harvey, P., Taylor, E., Sterling, R., & Cully, M. (2008). The THEMIS constellation. *Space Science Reviews*, 141(1–4), 117–152. <https://doi.org/10.1007/s11214-008-9416-2>
- Hill, T. W. (1974). Origin of the plasma sheet. *Reviews of Geophysics and Space Physics*, 12(3), 379–388. <https://doi.org/10.1029/RG012i003p00379>
- Horne, R. B., Thorne, R. M., Glauert, S. A., Albert, J. M., Meredith, N. P., & Anderson, R. R. (2005). Timescale for radiation belt electron acceleration by whistler mode chorus waves. *Journal of Geophysical Research*, 110(A3), A03225. <https://doi.org/10.1029/2004ja010811>
- Jaynes, A. N., Baker, D. N., Singer, H. J., Rodriguez, J. V., Loto'aniu, T. M., Ali, A. F., et al. (2015). Source and seed populations for relativistic electrons: Their roles in radiation belt changes. *Journal of Geophysical Research: Space Physics*, 120(9), 7240–7254. <https://doi.org/10.1002/2015ja021234>
- Kan, J. R., & Lee, L. C. (1979). Energy coupling function and solar wind-magnetosphere dynamo. *Geophysical Research Letters*, 6(7), 577–580. <https://doi.org/10.1029/GL006i007p00577>
- Kaufmann, R. L. (2012). Dependence of plasma sheet energy fluxes and currents on solar wind-magnetosphere coupling. *Journal of Geophysical Research*, 117(A10), A10230. <https://doi.org/10.1029/2012ja017995>
- Kelley, M. C., Makela, J. J., Chau, J. L., & Nicolls, M. J. (2003). Penetration of the solar wind electric field into the magnetosphere/ionosphere system. *Geophysical Research Letters*, 30(4), 2002GL016321. <https://doi.org/10.1029/2002gl016321>
- Kennel, C. F., & Petschek, H. E. (1966). Limit on stably trapped particle fluxes. *Journal of Geophysical Research*, 71(1), 1–28. <https://doi.org/10.1029/JZ071i001p00001>
- Kennel, C. F., & Thorne, R. M. (1967). Unstable growth of unducted whistlers propagating at an angle to geomagnetic field. *Journal of Geophysical Research*, 72(3), 871–878. <https://doi.org/10.1029/JZ072i003p00871>
- Kim, H. J., Lyons, L. R., Zou, S., Boudouridis, A., Lee, D. Y., Heinselman, C., & McCreedy, M. (2009). Evidence that solar wind fluctuations substantially affect the strength of dayside ionospheric convection. *Journal of Geophysical Research*, 114(A11), A11305. <https://doi.org/10.1029/2009ja014280>
- King, J. H., & Papitashvili, N. E. (2005). Solar wind spatial scales in and comparisons of hourly wind and ace plasma and magnetic field data. *Journal of Geophysical Research*, 110(A2), A02104. <https://doi.org/10.1029/2004ja010649>
- Kingma, D. P., & Ba, J. L. (2017). Adam: A method for stochastic optimization. *arXiv e-prints*, arXiv:1412.6980. Retrieved from <https://arxiv.org/abs/1412.6980>
- Kistler, L. M., Mouikis, C. G., Asamura, K., Yokota, S., Kasahara, S., Miyoshi, Y., et al. (2019). Cusp and nightside auroral sources of O<sup>+</sup> in the plasma sheet. *Journal of Geophysical Research: Space Physics*, 124(12), 10036–10047. <https://doi.org/10.1029/2019ja027061>
- Koons, H. C., Mazur, J. E., Selesnick, R. S., Blake, J. B., Fennell, J. F., Roeder, J. L., & Anderson, P. C. (1999). *The impact of the space environment on space systems* (Report No. TR-99(1670)-1). Aerospace Corporation.
- Kronberg, E. A., Haaland, S. E., Daly, P. W., Grigorenko, E. E., Kistler, L. M., Franz, M., & Dandouras, I. (2012). Oxygen and hydrogen ion abundance in the near-Earth magnetosphere: Statistical results on the response to the geomagnetic and solar wind activity conditions. *Journal of Geophysical Research*, 117(A12), A12208. <https://doi.org/10.1029/2012ja018071>
- Le Chat, G., Issautier, K., & Meyer-Vernet, N. (2012). The solar wind energy flux. *Solar Physics*, 279(1), 197–205. <https://doi.org/10.1007/s11207-012-9967-y>
- Li, W., Thorne, R., Bortnik, J., McPherron, R., Nishimura, Y., Angelopoulos, V., & Richardson, I. G. (2012). Evolution of chorus waves and their source electrons during storms driven by corotating interaction regions. *Journal of Geophysical Research*, 117(A8), A08209. <https://doi.org/10.1029/2012ja017797>
- Li, W., Thorne, R. M., Meredith, N. P., Horne, R. B., Bortnik, J., Shprits, Y. Y., & Ni, B. (2008). Evaluation of whistler mode chorus amplification during an injection event observed on CRRES. *Journal of Geophysical Research*, 113(A9), A09210. <https://doi.org/10.1029/2008ja013129>
- Liemohn, M. W., Moore, T. E., Craven, P. D., Maddox, W., Nagy, A. F., & Kozyra, J. U. (2005). Occurrence statistics of cold, streaming ions in the near-Earth magnetotail: Survey of polar-tide observations. *Journal of Geophysical Research*, 110(A7), A07211. <https://doi.org/10.1029/2004ja010801>
- Liemohn, M. W., Shane, A. D., Azari, A. R., Petersen, A. K., Swiger, B. M., & Mukhopadhyay, A. (2021). RMSE is not enough: Guidelines to robust data-model comparisons for magnetospheric physics. *Journal of Atmospheric and Solar-Terrestrial Physics*, 218, 105624. <https://doi.org/10.1016/j.jastp.2021.105624>
- Lockwood, M. (2019). Does adding solar wind Poynting flux improve the optimum solar wind-magnetosphere coupling function? *Journal of Geophysical Research: Space Physics*, 124(7), 5498–5515. <https://doi.org/10.1029/2019ja026639>
- Lockwood, M., Bentley, S. N., Owens, M. J., Barnard, L. A., Scott, C. J., Watt, C. E., & Allanson, O. (2019). 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. <https://doi.org/10.1029/2018sw001856>

- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, et al. (Eds.), *Advances in neural information processing systems* (Vol. 30). Curran Associates, Inc. Retrieved from <https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf>
- Luo, B., Tu, W., Li, X., Gong, J., Liu, S., Burin des Roziers, E., & Baker, D. N. (2011). On energetic electrons (>38 keV) in the central plasma sheet: Data analysis and modeling. *Journal of Geophysical Research*, *116*(A9), A09220. <https://doi.org/10.1029/2011ja016562>
- Lyons, L. R., Kim, H. J., Xing, X., Zou, S., Lee, D. Y., Heinselman, C., et al. (2009). Evidence that solar wind fluctuations substantially affect global convection and substorm occurrence. *Journal of Geophysical Research*, *114*(A11), A11306. <https://doi.org/10.1029/2009ja014281>
- Malykhin, A. Y., Grigorenko, E. E., Shklyar, D. R., Panov, E. V., Le Contel, O., Avakov, L., & Giles, B. (2021). Characteristics of resonant electrons interacting with whistler waves in the nearest dipolarizing magnetotail. *Journal of Geophysical Research: Space Physics*, *126*(7), e2021JA029440. <https://doi.org/10.1029/2021ja029440>
- Matéo-Vélez, J. C., Sicard, A., Payan, D., Ganushkina, N., Meredith, N. P., & Sillanpää, I. (2018). Spacecraft surface charging induced by severe environments at geosynchronous orbit. *Space Weather*, *16*(1), 89–106. <https://doi.org/10.1002/2017sw001689>
- Mayaud, P. N. (1980). Planetary indices derived from *k* indices (*Kp*, *am*, and *aa*). In *Derivation, meaning, and use of geomagnetic indices* (pp. 40–85). American Geophysical Union (AGU). <https://doi.org/10.1002/9781118663837.ch5>
- McFadden, J. P., Carlson, C. W., Larson, D., Bonnell, J., Mozer, F., Angelopoulos, V., et al. (2008). THEMIS ESA first science results and performance issues. *Space Science Reviews*, *141*(1–4), 477–508. <https://doi.org/10.1007/s11214-008-9433-1>
- McFadden, J. P., Carlson, C. W., Larson, D., Ludlam, M., Abiad, R., Elliott, B., et al. (2008). The THEMIS ESA plasma instrument and in-flight calibration. *Space Science Reviews*, *141*(1–4), 277–302. <https://doi.org/10.1007/s11214-008-9440-2>
- Morley, S. K., Brito, T. V., & Welling, D. T. (2018). Measures of model performance based on the log accuracy ratio. *Space Weather*, *16*(1), 69–88. <https://doi.org/10.1002/2017sw001669>
- Motoba, T., Ohtani, S., Gkioulidou, M., Ukhorskiy, A. Y., Lanzerotti, L. J., & Claudepierre, S. G. (2021). Superposed epoch analysis of dispersionless particle injections inside geosynchronous orbit. *Journal of Geophysical Research: Space Physics*, *126*(8), e2021JA029546. <https://doi.org/10.1029/2021JA029546>
- Mukhopadhyay, A., Jia, X., Welling, D. T., & Liemohn, M. W. (2021). Global magnetohydrodynamic simulations: Performance quantification of magnetopause distances and convection potential predictions. *Frontiers in Astronomy and Space Sciences*, *8*, 637197. <https://doi.org/10.3389/fspas.2021.637197>
- Nagata, D., Machida, S., Ohtani, S., Saito, Y., & Mukai, T. (2008). Solar wind control of plasma number density in the near-Earth plasma sheet: Three-dimensional structure. *Annales Geophysicae*, *26*(12), 4031–4049. <https://doi.org/10.5194/angeo-26-4031-2008>
- Newell, P. T., Sotirelis, T., Liou, K., Meng, C. I., & Rich, F. J. (2007). A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables. *Journal of Geophysical Research*, *112*(A1), A01206. <https://doi.org/10.1029/2006ja012015>
- Ni, B., Shprits, Y., Hartinger, M., Angelopoulos, V., Gu, X., & Larson, D. (2011). Analysis of radiation belt energetic electron phase space density using THEMIS SST measurements: Cross-satellite calibration and a case study. *Journal of Geophysical Research*, *116*(A3), A03208. <https://doi.org/10.1029/2010ja016104>
- Nishida, A., & Lyon, E. F. (1972). Plasma sheet at lunar distance: Structure and solar-wind dependence. *Journal of Geophysical Research*, *77*(22), 4086–4099. <https://doi.org/10.1029/JA077i022p04086>
- Ohtani, S., Nose, M., Christon, S. P., & Lui, A. T. Y. (2011). Energetic O<sup>+</sup> and H<sup>+</sup> ions in the plasma sheet: Implications for the transport of ionospheric ions. *Journal of Geophysical Research*, *116*(A10), A10211. <https://doi.org/10.1029/2011ja016532>
- Ohtani, S., Wing, S., Merkin, V. G., & Higuchi, T. (2014). Solar cycle dependence of nightside field-aligned currents: Effects of dayside ionospheric conductivity on the solar wind-magnetosphere-ionosphere coupling. *Journal of Geophysical Research: Space Physics*, *119*(1), 322–334. <https://doi.org/10.1002/2013ja019410>
- Perreault, P., & Akasofu, S. I. (1978). A study of geomagnetic storms. *Geophysical Journal International*, *54*(3), 547–573. <https://doi.org/10.1111/j.1365-246X.1978.tb05494.x>
- Richardson, J. D., & Paularena, K. I. (2001). Plasma and magnetic field correlations in the solar wind. *Journal of Geophysical Research*, *106*(A1), 239–251. <https://doi.org/10.1029/2000ja000071>
- Rostoker, G., & Fälthammar, C. G. (1967). Relationship between changes in the interplanetary magnetic field and variations in the magnetic field at the Earth's surface. *Journal of Geophysical Research*, *72*(23), 5853–5863. <https://doi.org/10.1029/JZ072i023p05853>
- Ruan, P., Fu, S. Y., Zong, Q. G., Pu, Z. Y., Cao, X., Liu, W. L., et al. (2005). Ion composition variations in the plasma sheet observed by cluster/rapid. *Geophysical Research Letters*, *32*(1), L01105. <https://doi.org/10.1029/2004gl021266>
- Runov, A., Angelopoulos, V., Sergeev, V. A., Glassmeier, K.-H., Auster, U., McFadden, J., et al. (2009). Global properties of magnetotail current sheet flapping: THEMIS perspectives. *Annales Geophysicae*, *27*(1), 319–328. <https://doi.org/10.5194/angeo-27-319-2009>
- Schild, M. A., Freeman, J. W., & Dessler, A. J. (1969). A source for field-aligned currents at auroral latitudes. *Journal of Geophysical Research*, *74*(1), 247–256. <https://doi.org/10.1029/JA074i001p00247>
- Sergeev, V., Runov, A., Baumjohann, W., Nakamura, R., Zhang, T. L., Volwerk, M., et al. (2003). Current sheet flapping motion and structure observed by Cluster. *Geophysical Research Letters*, *30*(6), 1327. <https://doi.org/10.1029/2002GL016500>
- Sergeev, V. A., Sormakov, D. A., & Angelopoulos, V. (2014). A missing variable in solar wind-magnetosphere-ionosphere coupling studies. *Geophysical Research Letters*, *41*(23), 8215–8220. <https://doi.org/10.1002/2014gl062271>
- Shapley, L. S. (1953). A value for *n*-person games. In K. Harold William & T. Albert William (Eds.), *Contributions to the Theory of Games (AM-28), Volume II* (Vol. 2, pp. 307–218). Princeton University Press. <https://doi.org/10.1515/9781400881970>
- Shrikumar, A., Greenside, P., & Kundaje, A. (2017). Learning important features through propagating activation. In D. Precup & Y. W. Teh (Eds.), *34th International Conference on Machine Learning Research* (Vol. 70, pp. 3145–3153). PMLR.
- Shukhtina, M. A., Dmitrieva, N. P., Popova, N. G., Sergeev, V. A., Yahnin, A. G., & Despirak, I. V. (2005). Observational evidence of the loading-unloading substorm scheme. *Geophysical Research Letters*, *32*(17), L17107. <https://doi.org/10.1029/2005gl023779>
- Southwood, D. J., & Wolf, R. A. (1978). An assessment of the role of precipitation in magnetospheric convection. *Journal of Geophysical Research*, *83*(A11), 5227. <https://doi.org/10.1029/JA083iA11p05227>
- 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*, 1929–1958.
- Stepanov, N. A., Sergeev, V. A., Sormakov, D. A., Andreeva, V. A., Dubyagin, S. V., Ganushkina, N., et al. (2021). Superthermal proton and electron fluxes in the plasma sheet transition region and their dependence on solar wind parameters. *Journal of Geophysical Research: Space Physics*, *126*(4), e2020JA028580. <https://doi.org/10.1029/2020ja028580>
- Swiger, B. (2022). briswi/ESWPSNN: v0.1-initial-public-release (v0.1-initial-public-release). Zenodo. <https://doi.org/10.5281/zenodo.7083352>
- Swiger, B. M., Liemohn, M. W., & Ganushkina, N. Y. (2020). Improvement of plasma sheet neural network accuracy with inclusion of physical information. *Frontiers in Astronomy and Space Sciences*, *7*, 42. <https://doi.org/10.3389/fspas.2020.00042>

- Swiger, B. M., Liemohn, L. W., Ganushkina, N. Y., & Dubyagin, S. V. (2022). Dataset for "Energetic Electron Flux Predictions in the near-Earth Plasma Sheet from Solar Wind Driving" [Dataset]. University of Michigan - Deep Blue Data. <https://doi.org/10.7302/araa-6f62>
- Terasawa, T., Fujimoto, M., Mukai, T., Shinohara, I., Saito, Y., Yamamoto, T., et al. (1997). Solar wind control of density and temperature in the near-Earth plasma sheet: WIND/GEOTAIL collaboration. *Geophysical Research Letters*, *24*(8), 935–938. <https://doi.org/10.1029/96gl04018>
- Thomsen, M. F., Henderson, M. G., & Jordanova, V. K. (2013). Statistical properties of the surface-charging environment at geosynchronous orbit. *Space Weather*, *11*(5), 237–244. <https://doi.org/10.1002/swe.20049>
- Troshichev, O. A., Podorozhkina, N. A., & Janzhura, A. S. (2011). Invariability of relationship between the polar cap magnetic activity and geoeffective interplanetary electric field. *Annales Geophysicae*, *29*(8), 1479–1489. <https://doi.org/10.5194/angeo-29-1479-2011>
- Tsutomu, T., & Teruki, M. (1976). Flapping motions of the tail plasma sheet induced by the interplanetary magnetic field variations. *Planetary and Space Science*, *24*(2), 147–159. [https://doi.org/10.1016/0032-0633\(76\)90102-1](https://doi.org/10.1016/0032-0633(76)90102-1)
- Turner, D. L., Claudepierre, S. G., Fennell, J. F., O'Brien, T. P., Blake, J. B., Lemon, C., et al. (2015). Energetic electron injections deep into the inner magnetosphere associated with substorm activity. *Geophysical Research Letters*, *42*(7), 2079–2087. <https://doi.org/10.1002/2015gl063225>
- Turner, D. L., & Li, X. (2008). Quantitative forecast of relativistic electron flux at geosynchronous orbit based on low-energy electron flux. *Space Weather*, *6*(5), S05005. <https://doi.org/10.1029/2007sw000354>
- Vassiliadis, D., Klimas, A. J., Baker, D. N., & Roberts, D. A. (1995). A description of the solar wind-magnetosphere coupling based on nonlinear filters. *Journal of Geophysical Research*, *100*(A3), 3495–3512. <https://doi.org/10.1029/94ja02725>
- Wang, C.-P., Kim, H.-J., Yue, C., Weygand, J. M., Hsu, T.-S., & Chu, X. (2017). Effects of solar wind ultralow-frequency fluctuations on plasma sheet electron temperature: Regression analysis with support vector machine. *Journal of Geophysical Research: Space Physics*, *122*(4), 4210–4227. <https://doi.org/10.1002/2016ja023746>
- Wang, C.-P., Yue, C., Zaharia, S., Xing, X., Lyons, L., Angelopoulos, V., et al. (2013). Empirical modeling of plasma sheet pressure and three-dimensional force-balanced magnetospheric magnetic field structure: 1. Observation. *Journal of Geophysical Research: Space Physics*, *118*(10), 6154–6165. <https://doi.org/10.1002/jgra.50585>
- Weimer, D. R., & King, J. H. (2008). Improved calculations of interplanetary magnetic field phase front angles and propagation time delays. *Journal of Geophysical Research*, *113*(A1), A01105. <https://doi.org/10.1029/2007ja012452>
- Wicks, R. T., Owens, M. J., & Horbury, T. S. (2010). The variation of solar wind correlation lengths over three solar cycles. *Solar Physics*, *262*(1), 191–198. <https://doi.org/10.1007/s11207-010-9509-4>
- 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), A08205. <https://doi.org/10.1029/2005ja011086>
- Wolf, R. A., Spiro, R. W., Sazykin, S., & Toffoletto, F. R. (2007). How the Earth's inner magnetosphere works: An evolving picture. *Journal of Atmospheric and Solar-Terrestrial Physics*, *69*(3), 288–302. <https://doi.org/10.1016/j.jastp.2006.07.026>
- Yu, Y., Rastätter, L., Jordanova, V. K., Zheng, Y., Engel, M., Fok, M., & Kuznetsova, M. M. (2019). Initial results from the gem challenge on the spacecraft surface charging environment. *Space Weather*, *17*(2), 299–312. <https://doi.org/10.1029/2018sw002031>
- Yue, C., Wang, C.-P., Lyons, L., Wang, Y., Hsu, T.-S., Henderson, M., et al. (2015). A 2-D empirical plasma sheet pressure model for substorm growth phase using the support vector regression machine. *Journal of Geophysical Research: Space Physics*, *120*(3), 1957–1973. <https://doi.org/10.1002/2014ja020787>
- Zheng, H., Fu, S. Y., Zong, Q. G., Pu, Z. Y., Wang, Y. F., & Parks, G. K. (2012). Observations of ionospheric electron beams in the plasma sheet. *Physical Review Letters*, *109*(20), 205001. <https://doi.org/10.1103/PhysRevLett.109.205001>
- Zou, Z., Shprits, Y. Y., Ni, B., Aseev, N. A., Zuo, P., & Wei, F. (2020). An artificial neural network model of electron fluxes in the Earth's central plasma sheet: A THEMIS survey. *Astrophysics and Space Science*, *365*(6), 100. <https://doi.org/10.1007/s10509-020-03819-0>