8-29-2022

Modeling Radiation Belt Electrons With Information Theory Informed Neural Networks

Simon Wing
Johns Hopkins University, simon.wing@jhuapl.edu

Drew L. Turner
Johns Hopkins University, drew.turner@jhuapl.edu

Aleksandr Y. Ukhorskiy
Johns Hopkins University, aleksandr.ukhorskiy@jhuapl.edu

Jay R. Johnson
Andrews University, jrj@andrews.edu

Thomas Sotirelis
Johns Hopkins University, tom.sotirelis@jhuapl.edu

See next page for additional authors

Authors

Simon Wing, Drew L. Turner, Aleksandr Y. Ukhorskiy, Jay R. Johnson, Thomas Sotirelis, Romina Nikoukar, and Giuseppe Romeo

Key Points:
- An empirical model to predict state of radiation belt relativistic electrons is developed
- The model prediction efficiency increases with increasing $L*$ with a PE > 0.6 at $L* > 5$
- The model can potentially complement a class of empirical models that input observations from low earth orbit

Correspondence to:
S. Wing,
simon.wing@jhuapl.edu

Author Contributions:
Conceptualization: Simon Wing, Aleksandr Y. Ukhorskiy
Data curation: Simon Wing, Thomas Sotirelis, Giuseppe Romeo
Formal analysis: Simon Wing
Funding acquisition: Simon Wing, Aleksandr Y. Ukhorskiy
Investigation: Simon Wing
Methodology: Simon Wing, Jay R. Johnson
Project Administration: Simon Wing
Resources: Simon Wing
Software: Simon Wing
Supervision: Simon Wing
Validation: Simon Wing
Visualization: Simon Wing, Romina Nikoukar
Writing – original draft: Simon Wing

Modeling Radiation Belt Electrons With Information Theory Informed Neural Networks

Simon Wing[1] , Drew L. Turner[1] , Aleksandr Y. Ukhorskiy[1] , Jay R. Johnson[2] , Thomas Sotirelis[1], Romina Nikoukar[1] , and Giuseppe Romeo[1]

[1]The Johns Hopkins University, Applied Physics Laboratory, Laurel, MD, USA, [2]Andrews University, Berrien Spring, MI, USA

Abstract An empirical model of radiation belt relativistic electrons (μ = 560–875 MeV G^{-1} and I = 0.088–0.14 R_E G$^{0.5}$) with average energy ~1.3 MeV is developed. The model inputs solar wind parameters (velocity, density, interplanetary magnetic field (IMF) |B|, Bz, and By), magnetospheric state parameters (SYM-H and AL), and $L*$. The model outputs the radiation belt electron phase space density (PSD). The model is operational from $L*$ = 3 to 6.5. The model is constructed with neural networks assisted by information theory. Information theory is used to select the most effective and relevant solar wind and magnetospheric input parameters plus their lag times based on their information transfer to the PSD. Based on the test set, the model prediction efficiency (PE) increases with increasing $L*$, ranging from −0.043 at $L*$ = 3 to 0.76 at $L*$ = 6.5. The model PE is near 0 at $L*$ = 3–4 because at this $L*$ range, the solar wind and magnetospheric parameters transfer little information to the PSD. Using solar wind observations at L1 and magnetospheric index (AL and SYM-H) models solely driven by solar wind, the radiation belt model can be used to forecast PSD 30–60 min ahead. This baseline model can potentially complement a class of empirical models that input data from low earth orbit (LEO).

Plain Language Summary An empirical model of radiation belt relativistic electrons with an energy of 1–2 MeV is developed. The model inputs solar wind parameters, magnetospheric state parameters, and $L*$. $L*$ gives a measure of radial distance from the center of the Earth with a unit of R_E (radius of the Earth = 6,378 km). The model outputs the radiation belt electron phase space density (PSD). The model is operational from $L*$ = 3 to $L*$ 6.5. The model is constructed with an information theory informed neural networks. Information theory is used to select the relevant solar wind and magnetospheric parameters and their lag times based on the amount of information they provide to the radiation belt electrons. The model performance increases with increasing radial distance ($L*$) because at distances close to Earth ($L*$ = 3–4), the solar wind and magnetospheric parameters provide little information about the radiation belt electron PSD. The model can be used to forecast radiation belt PSD 30–60 min ahead.

1. Introduction

The Earth's radiation belts are populated by electrons having energies of hundreds of keVs to several MeVs or even higher. These electrons are hazardous to satellites that encounter them in the inner-magnetosphere $r \sim$ 1.2–8 R_E, including at the geosynchronous orbit (GEO), and at their foot points at low earth orbit (LEO) in the ionosphere, where 1 R_E = radius of the Earth = 6,378 km. The MeV electrons can penetrate deep into spacecraft systems, leading to anomalous system, subsystem, or payload malfunctions, while those with energies <1 MeV can accumulate on or near the surface of the spacecraft structure, leading to potentially hazardous electrical discharges.

It has long been recognized that the variabilities of the radiation belt electrons, to a large extent, are driven ultimately by the variability of the solar wind (e.g., Alves et al., 2017; Baker et al., 1990, 2018, 2019; Li et al., 2001, 2005; Pinto et al., 2018; Reeves, 2007; Reeves et al., 2013; Ukhorskiy et al., 2004; Xiang et al., 2017; Zhao et al., 2017). However, many solar wind parameters positively and negatively correlate with one another, which can complicate the interpretation as to which solar wind parameters are the real drivers and which parameters are only coincidentally correlated with the radiation belt electrons (e.g., Borovsky, 2018, 2020; Maggiolo et al., 2017; Wing & Johnson, 2019; Wing et al., 2016; Wing et al., 2022). For example, solar wind velocity (V_{sw}) positively correlates with radiation belt electron fluxes (J_e) (e.g., Baker et al., 1990; Balikhin et al., 2011; Li et al., 2001, 2005; Paulikas & Blake, 1979; Reeves et al., 2011; Wing et al., 2016, 2022). Solar wind density

Writing – review & editing: Simon Wing, Drew L. Turner, Jay R. Johnson, Romina Nikoukar

(n_{sw}) negatively correlates with radiation belt J_e (e.g., Kellerman & Shprits, 2012; Li et al., 2005; Lyatsky & Khazanov, 2008b; Rigler et al., 2007; Wing et al., 2016, 2022). However, V_{sw} negatively correlates with n_{sw} (e.g., Borovsky, 2020; Wing et al., 2016, 2022).

Radiation belt electrons also have strong dependences on the magnetospheric state, which can be proxied by geomagnetic activity indices, such as SYM-H and AL (e.g., Baker et al., 2019; Borovsky, 2017; Borovsky & Denton, 2014; Lyatsky & Khazanov, 2008a; Tang, Wang, Ni, Su, et al., 2017; Zhao et al., 2017). SYM-H index gives a measure of the strength of the ring current and geomagnetic storms (Iyemori, 1990), while AL gives a measure of the strength of the westward auroral electrojets and substorms (Davis & Sugiura, 1966). SYM-H is similar to the disturbance storm time (Dst) index (Dessler & Parker, 1959), except that the SYM-H index has 1 minute time resolution whereas the Dst index has 1 hour resolution. Unfortunately, SYM-H and AL both also correlate with solar wind parameters, which raises the question how much additional unique information these two magnetic indices provide to the radiation belt electrons and what their response lag times may be, given the solar wind parameters (Wing et al., 2022).

Wing et al. (2016, 2022) showed that information theoretic tool such as conditional mutual information can be quite useful to untangle the intertwined solar wind and magnetospheric drivers of the radiation belt electrons. They were able to isolate the effect of individual drivers and their response lag times. Moreover, they ranked the solar wind and magnetospheric parameters based on the information transfer of these parameters to the radiation belt J_e (Wing et al., 2016) and, more recently, electron phase space density (PSD) (Wing et al., 2022). Thus, those studies provided relevant and useful information for radiation belt modeling.

Machine learning algorithms such as Neural Networks (NN) and deep learning (Rumelhart and McClelland, 1987; Schmidhuber, 2015) have found wide applications in space weather, particularly in empirical modeling. For example, NN have been used to develop models for Kp (e.g., Boberg et al., 2000; Wing et al., 2005; Wintoft et al., 2017), geomagnetic storm (Wu & Lundstedt, 1997), source regions of particle precipitation (Newell et al., 1990, 1991), and high-frequency (HF) backscattered signals (Wing et al., 2003). NN have also been used to construct empirical radiation belt models (e.g., Chen et al., 2019; Chu et al., 2021; Claudepierre & O'Brien, 2020; Koons & Gorney, 1991; Ling et al., 2010; Perry et al., 2010; Pires de Lima et al., 2020; Simms & Engebretson, 2020; Smirnov et al., 2020; Tang et al., 2022). These empirical models generally complement physics-based models, for example DREAM (Reeves et al., 2012), SPACECAST (Horne et al., 2013), VERB (Shprits et al., 2009), and other empirical models that use different approaches, for example, NARMAX (Balikhin et al., 2016; Wei et al., 2011), Kalman filter (Coleman et al., 2018), and linear prediction filter (Baker et al., 1990; Chen et al., 2019; Kellerman et al., 2013). For operational purpose, one may need to consider trade-offs among accuracy, computational speed, computing resource requirements, availability of input parameters, ease of use, etc.

The Van Allen Probes or Radiation Belt Storm Probe (RBSP) mission ended in 2019 and there is no dedicated follow-on mission to the equatorial radiation belts planned in the near future. The Polar Operational Environmental Satellite (POES) program, which provides observations of the precipitating radiation belt electrons, may end in the next several years, and there is no current plan to replace those assets. Moreover, as discussed later, NN models that input the past values of the output parameters tend not to be able to respond accurately and timely to sudden changes in the solar wind drivers, for example, sudden arrivals of density/pressure pulses or coronal mass ejections (CMEs) (e.g., Wing et al., 2005).

The present study develops an empirical model of the radiation belt electron PSD using an information-theory informed NN as the core of the model (Johnson & Wing, 2018). From the consideration of the versatility of running the model in real time and the aforementioned challenges, our model inputs only solar wind and magnetospheric state parameters (proxied by geomagnetic indices) and outputs outer radiation belt electron PSD. The input parameters and their lag times are determined from Wing et al. (2022) information theoretic analysis of the solar wind and magnetospheric drivers of PSD.

2. Data Set

The NASA's Van Allen Probes (RBSP) mission, which was launched in 2012, consisted of two identically instrumented spacecraft in near-equatorial orbit (about 10° inclination) with perigee at 600 km altitude and apogee at 5.8 R_E geocentric (Mauk et al., 2013). The MAGnetic Electron Ion Spectrometer (MagEIS) is part of the Energetic particle, Composition, and Thermal plasma Suite (ECT) instrument on board of RBSP (Spence et al., 2013).
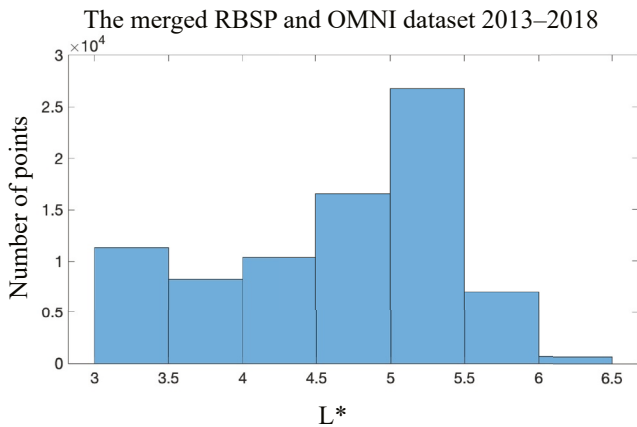
The merged RBSP and OMNI dataset 2013–2018

Figure 1. The distribution of the merged RBSP and OMNI data set 2013–2018.

MagEIS measured the energy range from 30 keV to 4 MeV for electrons and from 20 keV to 1 MeV for ions (Blake et al., 2013).

Radiation belt electron dynamics can often be well-organized by electron PSD as a function its three adiabatic invariants: PSD (μ, I, and L^*) where μ = the first adiabatic invariant related to the gyromotion perpendicular to the magnetic field line, I = the second adiabatic invariant related to the bounce motion along the field line (some studies use K instead of I, but they are related) (Green & Kivelson, 2004), and L^* = the third adiabatic invariant related to the curvature and gradient drift motion around the Earth (actually L^* is inversely proportional to the traditional third invariant Φ) (Roederer, 1970; Schulz & Lanzerotti, 1974).

The radiation belt electron PSD from MagEIS is calculated at 1 min resolution using the TS04 magnetic field model (Tsyganenko & Sitnov, 2005) and a method similar to that used in Turner, Angelopoulos, Li, et al., 2014, Turner, Angelopoulos, Morley, et al., 2014. We select the electrons with μ = 560–875 MeV G^{-1} and I = 0.088–0.14 R_E G$^{0.5}$. These electrons have an average energy of about 1.3 MeV over L^* = 2.9–6.5 and are concentrated near the magnetic equator (i.e., mirroring at low magnetic latitudes); thus, they are representative of the core population of relativistic electrons in Earth's outer radiation belt.

The solar wind, AL, and SYM-H data 2013–2018 at 1 min resolution from the OMNI data set were used and provided by NASA (http://omniweb.gsfc.nasa.gov/). Both the PSD and OMNI data 2013–2018 are averaged with a 30 min sliding window.

We merge each OMNI solar wind parameters (V_{sw}, n_{sw}, etc.) with the RBSP electron PSD (data from both RBSP A and B are used). The merged data set has ~64,500 points distributed from L^* = 2.9 to 6.5. However, the distribution is not uniform across L^*, as shown in Figure 1.

3. Methodology

It has been increasingly popular to use NN, including deep learning, to develop empirical space weather models, including radiation belt models. However, a novelty with our approach is that we use information theory to assist with the modeling. Figure 2 shows the schematic of the model.

The model inputs solar wind, magnetospheric parameters, and L^*; and outputs radiation belt electron PSD. Wing et al. (2022) ranked the solar wind and magnetospheric parameters based on the information transfer to the PSD
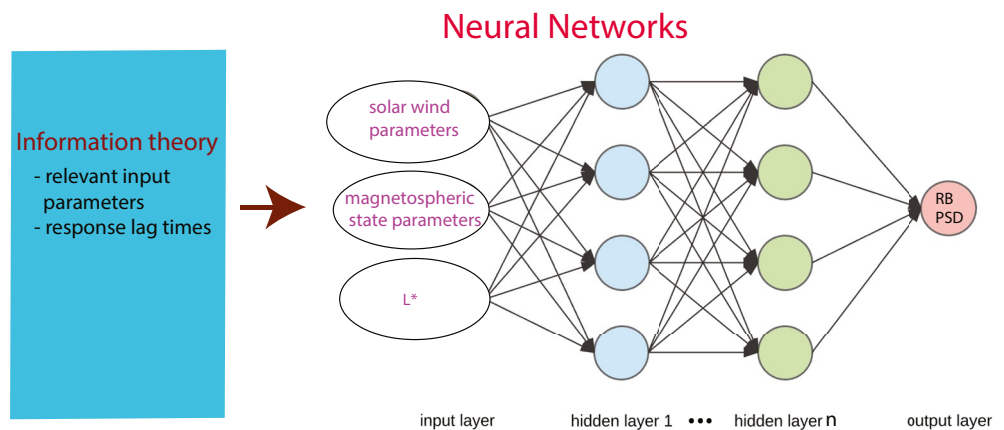
Figure 2. Schematic of the model that combines information theory and neural networks. The neural networks inputs the solar wind and magnetospheric parameters and L^*; and outputs phase space density (PSD) (see Table 1). Information theory is used to select and rank solar wind and magnetospheric parameters and their lag times based on information transfer to the radiation belt electron PSD. The model operates at L^* range from 3 to 6.5.

Table 1
Input and Output Parameters of the Model

	Input parameters	Output parameter
1	$V_{sw}(t)$ to $V_{sw}(t–72$ hr$)$	PSD (t)
2	$n_{sw}(t)$ to $n_{sw}(t–12$ hr$)$	
3	IMF $|B(t)|$ to $|B(t–10$ hr$)|$	
4	IMF $Bz(t)$ to $Bz(t–10$ hr$)$	
5	IMF $By(t)$ to $By(t–10$ hr$)$	
6	SYM-H(t) to SYM-H(t–72 hr)	
7	AL(t) to AL(t–72 hr)	

Note. V_{sw} = solar wind velocity. n_{sw} = solar wind density. IMF (By, Bz) = GSM y and z component of the interplanetary magnetic field, respectively.

(see Table 4 in Wing et al. (2022)). We select the top eight parameters as the model input parameters, namely solar wind velocity, SYM-H, AL, solar wind dynamic pressure, IMF |B|, IMF Bz, solar wind density, and IMF By (in decreasing order by the amount of information transferred from the parameter to the radiation belt electron PSD). The solar wind dynamic pressure usually tracks the solar wind density fairly well and the information content in the dynamic pressure is entirely captured by the solar wind speed and density, so we omit solar wind dynamic pressure. The input parameters and their lag times are listed in Table 1. The solar wind, magnetospheric, and PSD data have gaps. Those input and output parameters that have gaps are discarded. The model outputs PSD with no time lag with respect to the arrival time of the solar wind at the magnetosphere (nowcast).

The NN architecture used is the standard feedforward–backpropagation network, which is sometimes referred to as multilayered perceptrons (MLP). The NN architecture has five layers: one input layer (531 nodes), one output layer (1 node), and three hidden layers (each has 800 nodes). The model is developed using Python and Tensor-Flow machine learning package, which is an open source package (Abadi et al., 2016).

All the input and output parameters are normalized. The PSD distribution is skewed to the low values. In order to get higher performance, log PSD is used rather than PSD. Log PSD reduces the skewness in the original PSD distribution, which would help training the NN. Both RBSP A and B data are split into two sets: (a) training set and (b) test set. The training set consists of data in the time intervals (2013.5–2015.5), (2016–2017), and (2017.5–2018.5), while the test set consists of (2013–2013.5), (2015.5–2016), (2017–2017.5), and (2018.5–2019.0). Staggering the training and test sets ensures no systematic temporal bias (e.g., solar cycle dependencies) are present in the resulting model. Also, 20% of the training set is set aside as a validation set to avoid overfitting.

4. Results

In order to show the model performance, we select two long events from the test set where there are at least two weeks of continuous solar wind observations, AL and SYM-H records, and RBSP electron PSD observations: (a) 2013 April 27–May 13 and (b) 2017 Mar 13–29. These intervals are selected also because they exhibit a wide range of solar wind driving as well as geomagnetic storm and substorm dynamics. Thus, they are intended to show how well the model can perform under average and unusual solar wind and magnetospheric conditions. They are certainly not intended to show the best examples of the model performance.

Figure 3A plots solar wind velocity (a), density (b), SYM-H (c), AL (d), L^* and model PSD (e), Δ log PSD = log(observed PSD)–log(model PSD) (f), and observed and model PSD (g) for the first half of the first event, 2013 April 27–May 5. Panel d shows quasiperiodic substorms (minimum AL > −400 nT) throughout the interval, which is fairly typical (Borovsky & Yakymenko, 2017). However, an unusual feature of this interval is that there is a sharp density pulse (maximum ~15 cm^{-3}) that is followed by a moderate storm (minimum
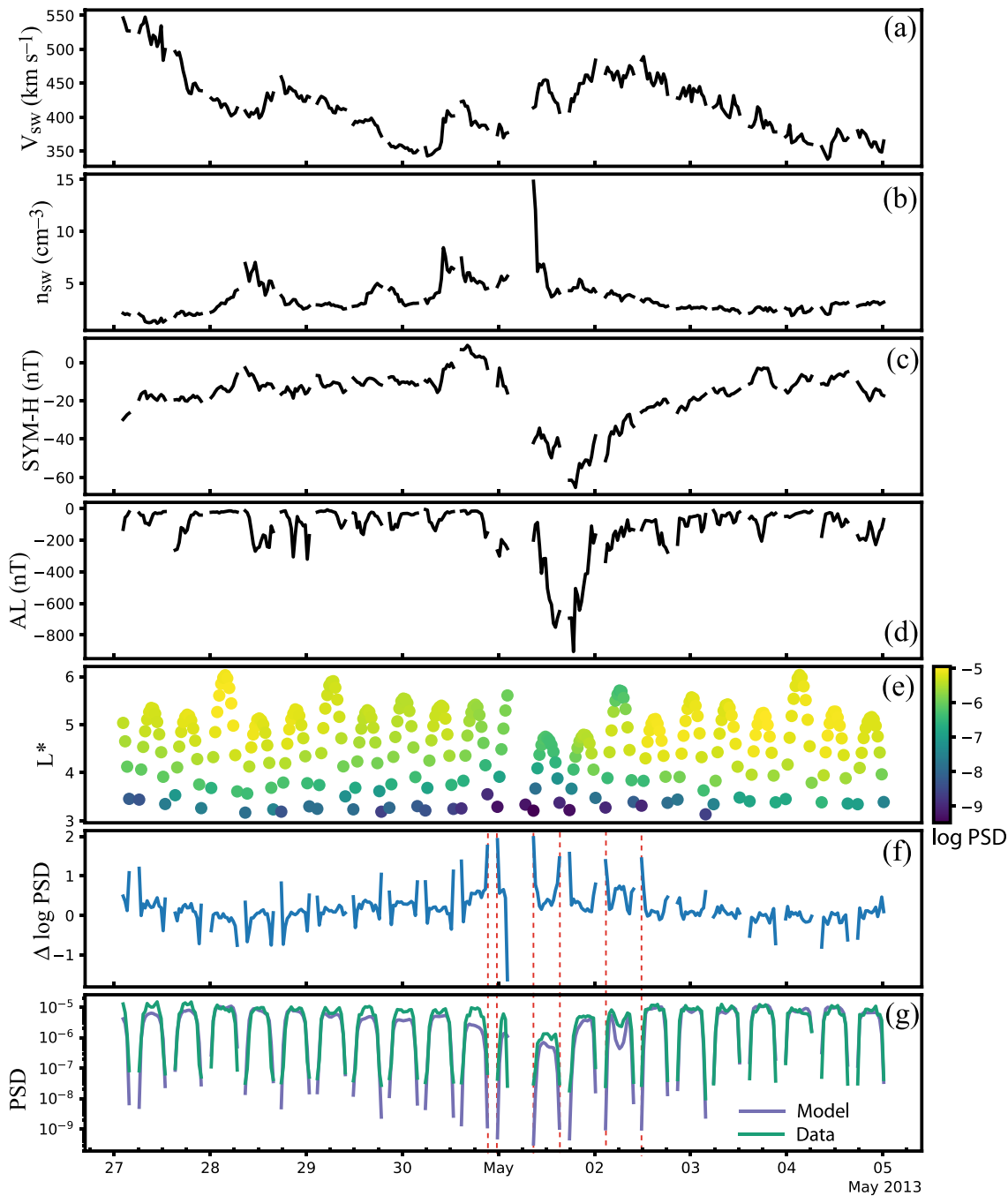
Figure 3A. Solar wind velocity (a), solar wind density (b), SYM-H (c), AL (d), L^* and log model phase space density (PSD) (e), Δ log PSD = log(observed PSD) – log(model PSD) (f), and observed (green curve) and model PSD (blue curve) (g) for 27 April–5 May 2013, which is the first half of the first event. The unit for PSD is (c^3 MeV^{-3} cm^{-3}). In panels (f) and (g), dotted vertical red lines are drawn to show that generally large |Δ log PSD| can be associated with low PSD.

SYM-H ~ -60 nT) and a large substorm (minimum AL ~ -900 nT) on May 1. Panel g shows that there is a drop in PSD on May 1, which may be attributed to the sharp rise in solar wind density and dynamic pressure through processes such as magnetopause shadowing, radial diffusion, or precipitation to the atmosphere (e.g., Borovsky & Denton, 2009; Kellerman & Shprits, 2012; Li et al., 2001; Thorne, 2010; Turner et al., 2012; Ukhorskiy et al., 2006). However, the PSD seemed to have recovered by the end of May 2. Panels f and g show that the model generally performs reasonably well throughout this interval even in the presence of quasiperiodic substorms, but it does not do as well around the density/pressure pulse and the storm and substorm on May 1–2. At high L^*,
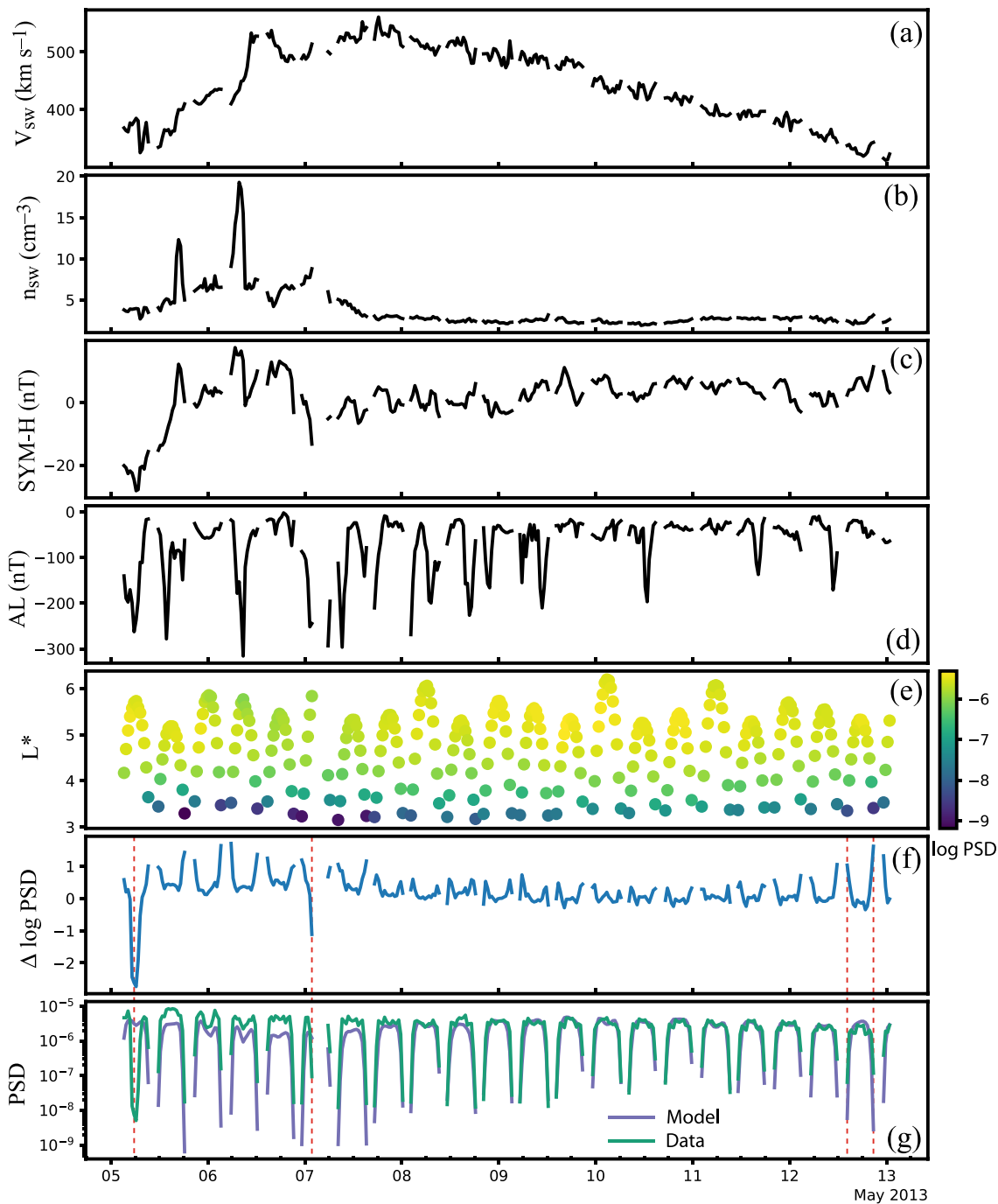
Figure 3B. Solar wind velocity (a), solar wind density (b), SYM-H (c), AL (d), $L*$ and log model phase space density (PSD) (e), Δ log PSD = log(observed PSD) – log(model PSD) (f), and observed (green curve) and model PSD (blue curve) (g) for 5–13 May 2013, which is the second half of the first event. The unit for PSD is (c^3 MeV^{-3} cm^{-3}). In panels (f) and (g), dotted vertical red lines are drawn to show that generally large $|\Delta$ log PSD$|$ can be associated with low PSD.

$L* > 4$, the model PSD appears to track the decrease and then the increase of the observed PSD reasonably well. However, at low $L*$, $L* < 4$, the model PSD decreases significantly, by more than an order of magnitude, whereas the observed PSD does not appear to be affected much by the density pulse.

Figure 3A panel f shows that most of the time the observed and model PSD are within the same order of magnitude of each other, $|\Delta$ log PSD$| < 1$. Large $|\Delta$ log PSD$|$ generally corresponds to low PSD and low $L*$ that is near the slot region. In order to show this, several dotted vertical red lines are drawn to connect some of the largest $|\Delta$

log PSD| in panel f to their corresponding PSD in panel g. This trend can be seen throughout Figure 3A. When PSD is low, a little discrepancy from the observed value would lead to a large |Δ log PSD|. Low PSD may be less relevant for space weather than high PSD within the outer radiation belt. It should be noted that as shown in Figure 3A, most of the time, the error is small, |Δ log PSD| < 1, for high and low PSD.

Figure 3B presents the interval 05–13 May 2013, which is the second half of the first event, in the same format as Figure 3A. As in Figure 3A, panel d shows quasiperiodic moderate and small substorms (minimum AL > ∼ −300 nT) throughout the interval. This interval starts out with a small storm (minimum SYM-H ∼ −28 nT) on May 5, and a narrow density pulse (maximum density ∼ 19 cm⁻³) on May 6. There is a brief PSD decrease that occurs at or just before the storm onset on May 5, but the model misses this brief drop in PSD (panel g), resulting in a brief large discrepancy (Δ log PSD < −2) on panel f. Unlike the density/pressure pulse in Figure 3A, the density/pressure pulse on May 6 does not seem to affect the observed PSD that much, but the model responds by decreasing its PSD, particularly at $L^* < 4$, resulting in a brief large discrepancy (Δ log PSD > 1) on May 6 (panel g). The rest of the interval has no storm, but there are small and moderate substorms (minimum AL > −300 nT). The model performs well (|Δ log PSD| < 1) during this interval, except near the end at low L^* ($L^* < 4$) where Δ log PSD > 1. It is not clear what causes the model to underestimate PSD at this time. As in Figure 3A, several dotted vertical red lines from some of the largest |Δ log PSD| are drawn in panels f and g to show that generally large |Δ log PSD| corresponds to low PSD, but most of the time the error is small for high and low PSD.

Figure 4A presents the interval 13–21 March 2017, which is the first half of the second event in the same format as Figure 3. This interval shows the worst model performance out of the four intervals presented herein and is one of the worst intervals seen in the entire test set. As in the previous intervals, there are quasiperiodic small and moderate substorms (minimum AL > −350 nT) in panel d. The solar wind velocity fluctuates but is lower than average, <400 km s⁻¹, throughout the interval. There is a broad density pulse (maximum ∼ 23 cm⁻³) on Mar 15, which is followed by a small storm (minimum SYM-H ∼ −20 nT) and a moderate substorm (minimum AL ∼ −350 nT) near the beginning of Mar 16. There is no significant change in the observed PSD that can be attributed to these solar wind parameters and magnetospheric activity indices (storm and substorm). However, the increase of solar wind density/pressure followed by substorm injections cause the model PSD to first decrease due to the expected magnetopause shadowing/radial diffusion/precipitation (e.g., Borovsky & Denton, 2009; Kellerman & Shprits, 2012; Li et al., 2001; Turner et al., 2012; Ukhorskiy et al., 2006; Wing et al., 2016, 2022) and then increase due to the expected storm time acceleration and substorm injections (e.g., Baker et al., 1996; Boyd et al., 2016; Li et al., 2009; Meredith et al., 2001; Tang, Wang, Ni, Zhang, et al., 2017; Wing et al., 2016, 2022). Because the model significantly decreases its PSD while the observed PSD does not significantly change, the model PSD severely underestimates the observed PSD at all L^* as seen in panels f and g. As before, several dotted vertical red lines from some of the largest |Δ log PSD| are drawn in panels f and g to show that large |Δ log PSD| fairly consistently corresponds to low PSD.

Figure 4B presents the interval 21–29 March 2017, which is the second half of the second event in the same format as Figure 3. The solar wind velocity is higher than average, >500 km s⁻¹, throughout most of the interval. This interval has two interesting features, one at the beginning and one at the end of the interval. At the beginning of the interval, there is a density pulse (maximum ∼ 32 cm⁻³), which is followed by a large substorm (minimum AL ∼ −750), but there is no indication of a corresponding geomagnetic storm (minimum SYM-H ∼ −25 nT). In response to the density/pressure increase, both the observed and model PSDs first decrease and then increase on Mar 21–22. However, the model PSD decreases more than the observed PSD, resulting in a large discrepancy with Δ log PSD > 2. However, the model PSD increases quickly, such that by the end of Mar 21, it has more or less caught up with the observed PSD. Thereafter, the model PSD tracks the observed PSD fairly well as they are both recovering from the electron loss. The PSD completely recovers by the middle of the day on Mar 22 and thereafter, the model PSD generally performs well (Δ log PSD < 1) as shown in panels f and g. As before, several dotted vertical red lines from some of the largest |Δ log PSD| are drawn to show that large |Δ log PSD| fairly consistently corresponds to low PSD.

At the end of the interval, there is another density pulse (maximum ∼22 cm⁻³) that is followed by a large or moderate storm (minimum SYM-H ∼ − 80 nT) and three large substorms (two with minimum AL ∼ −1,000 nT one with minimum AL ∼ −750 nT) on Mar 27. In response, the observed PSD decreases soon after the density/pressure pulse in the first half of Mar 27 and then increases. The observed PSD completely recovers by the middle
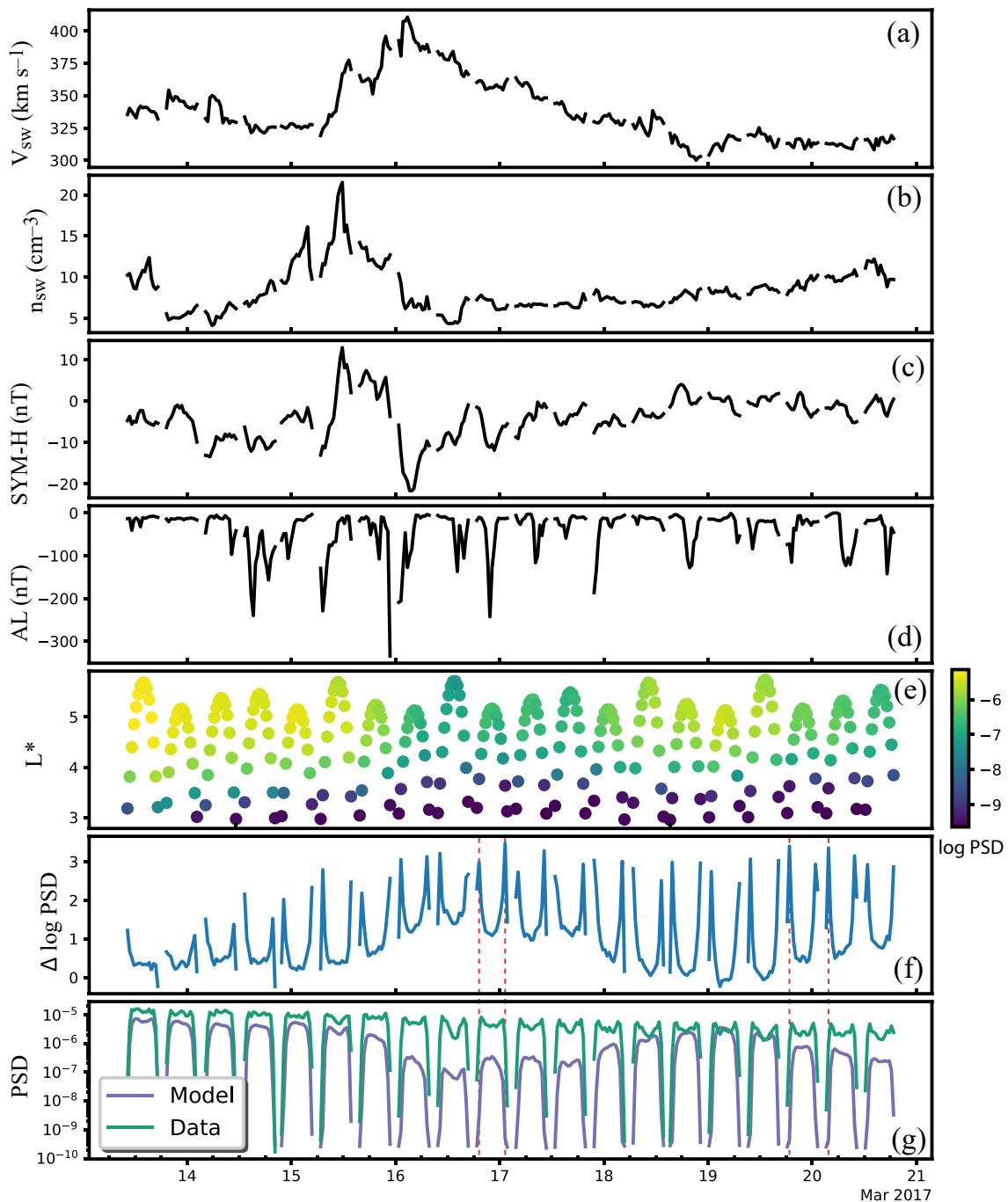
Figure 4A. Solar wind velocity (a), solar wind density (b), SYM-H (c), AL (d), $L*$ and log model phase space density (PSD) (e), Δ log PSD = log(observed PSD) – log(model PSD) (f), and observed (green curve) and model PSD (blue curve) (g) for 13–21 March 2017, which is the first half of the second event. The unit for PSD is $(c^3 \, MeV^{-3} \, cm^{-3})$. In panels (f) and (g), dotted vertical red lines are drawn to show that generally large $|\Delta$ log PSD$|$ can be associated with low PSD.

of the day on Mar 28. The model PSD tracks the observed PSD fairly well during this highly disturbed period (Δ log PSD < 1), as shown in panels f and g.

Figures 3 and 4 show that the model performs well and the error is small for high and low PSD. There are instances when the error is large, $|\Delta$ log PSD$| > 1$, but these points are usually associated with low PSD.

The model performance has also been evaluated statistically. There are 23,853 number of points in the test set. Figure 5 plots the model PSD versus the observed PSD in the entire test set. Consistent with the events shown
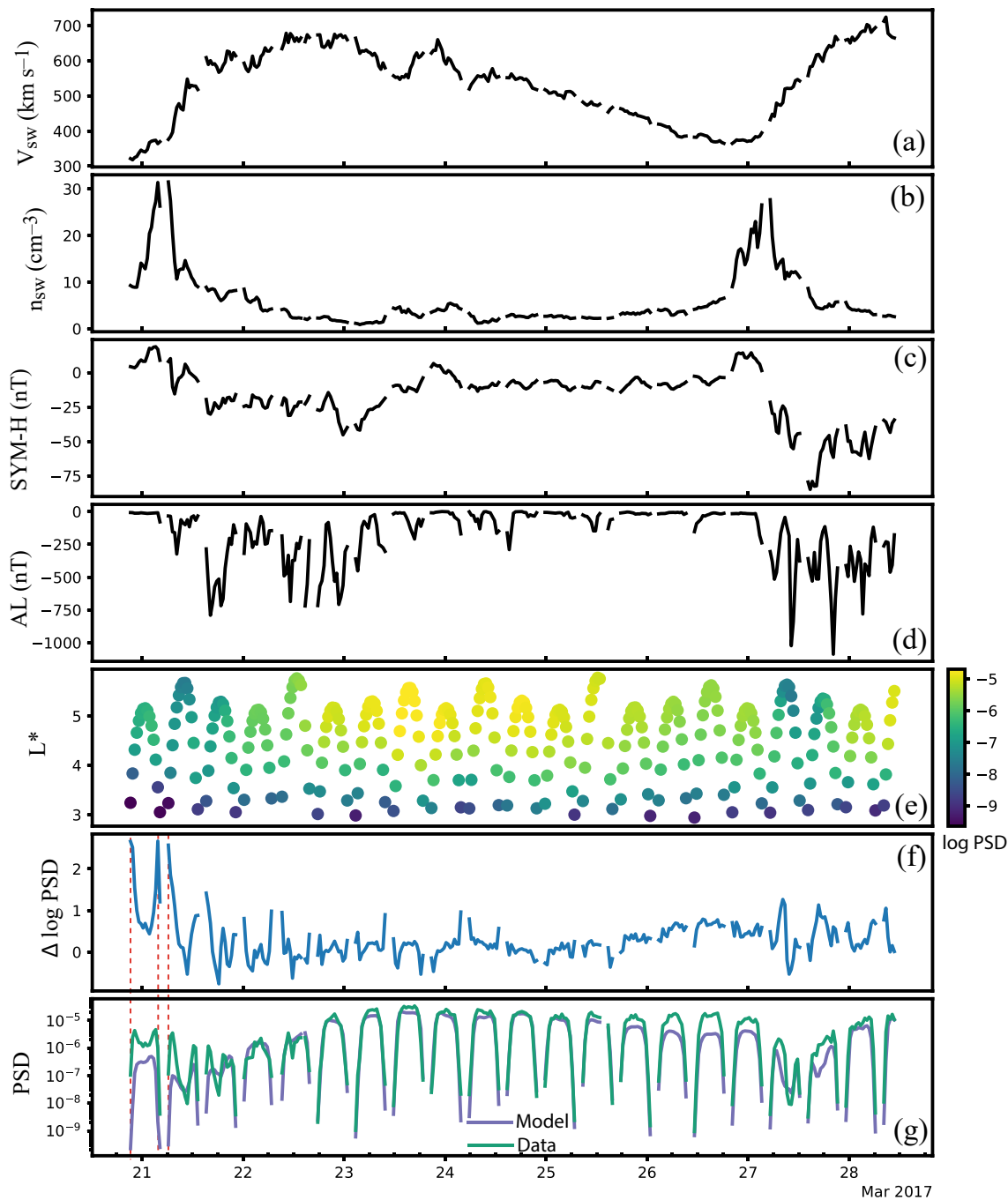
Figure 4B. Solar wind velocity (a), solar wind density (b), SYM-H (c), AL (d), $L*$ and log model PSD (e), Δ log PSD = log(observed PSD)–log(model PSD) (f), and observed (green curve) and model PSD (blue curve) (g) for 2017 Mar 21–29, which is the second half of the second event. The unit for PSD is (c^3 MeV^{-3} cm^{-3}). In panels (f) and (g), dotted vertical red lines are drawn to show that generally large |Δ log PSD| can be associated with low PSD.

in Figures 3 and 4 panels f and g, Figure 5 shows that high PSD, which has higher space weather impacts, can be modeled more accurately than low PSD. Most of the low PSD points come from low $L*$ as can be seen in Figures 3 and 4 panel e (further model evaluation as a function of $L*$ is given below). Figure 5 also shows that the model tends to underestimate the observed PSD, especially for low PSD.

Based on the evaluation of model PSD for the entire test set: root mean square error (rmse) = 3.1×10^{-6} c^3 MeV^{-3} cm^{-3}; the mean absolute percent error (mape) = 115%; the median absolute percent error = 57%; and
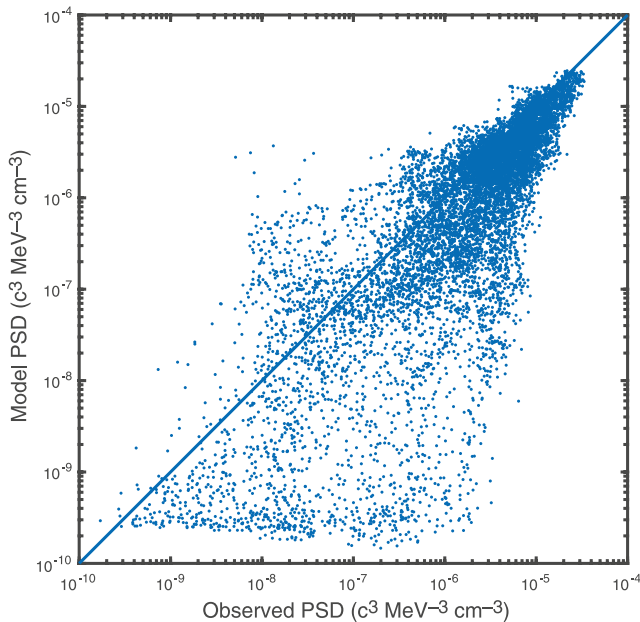
Figure 5. Model phase space density (PSD) versus observed PSD for the entire test set. High PSD, which has higher space weather impacts, can be modeled more accurately than low PSD.

the prediction efficiency (PE) = 0.62. PE is defined as $PE = 1 - \frac{\sum_1^n (o_i - m_i)^2}{\sum_1^n (o_i - <o>)^2}$ where o = observed PSD, m = model PSD, $<o>$ = mean observed PSD. PE = 1 indicates the model PSD exactly matches the observed PSD, while PE = 0 indicates the model simply outputs the mean value. PE < 0 indicates the model output is worse than simply outputting the mean for each point in the test set.

The model performance has a dependence on $L*$. The data are binned from $L* = 3$ to 6.5 into seven bins, with each bin having width = 0.5. Figure 6a plots the PE as a function of $L*$, which ranges from −0.043 for $L* = 3$ to 0.76 for $L* = 6.5$. Figure 6b shows the histogram of the number of points in each bin. The $L* = 6$–6.5 bin has the fewest points, $n = 227$ and hence the PE for this bin may be less accurate than those for other $L*$ bins. The PE for the entire test set (0.62) is close to that obtained for $L* = 4.5$–5.5 probably because this $L*$ range has the most data points as shown in Figure 6b.

The model PSD accuracy generally increases with increasing distance from the Earth (increasing $L*$). The model PE for $L* = 3$–4 is nearly 0 because the solar wind and magnetospheric drivers have less influence on the PSD at $L* < 4$. Indeed, Wing et al. (2022) showed that the solar wind density transfers information to PSD only at $L* > 4.5$. Solar wind velocity and AL transfer information to PSD at $L* > 4$ and only a small amount of information at $L* = 3.5$–4. Out of all the parameters that are inputted to the model, only SYM-H transfers information to PSD all the way to $L* = 3$, but the amount of information transfer at $L* = 3$–3.5 is small. Conversely, the input parameters (solar wind parameters, SYM-H, and AL) provide significant information about PSD at $L* > 4$ (Wing et al., 2022), and consequently, the model performance improves at this $L*$ range.

The model PE is similar to that obtained by DREAM (Reeves et al., 2012) at $L* > 4.5$ and slightly better than that obtained by DREAM at $L* < 4.5$. The model comparison is inexact because the PE for DREAM was calculated for 1 MeV electron fluxes whereas our model PE is for PSD with an average energy of 1.3 MeV. As with DREAM, our model performs better than AE 8 min (Vette, 1991) and CRRESELE (Brautigam & Bell, 1995) models. For many years, the AE8 series model was considered standard for engineering applications. (AE8 min model is superseded by a newer model, AE9 (Ginet et al., 2013), but like AE8, AE9 is a statistical model i.e., not relevant to individual event-based prediction).
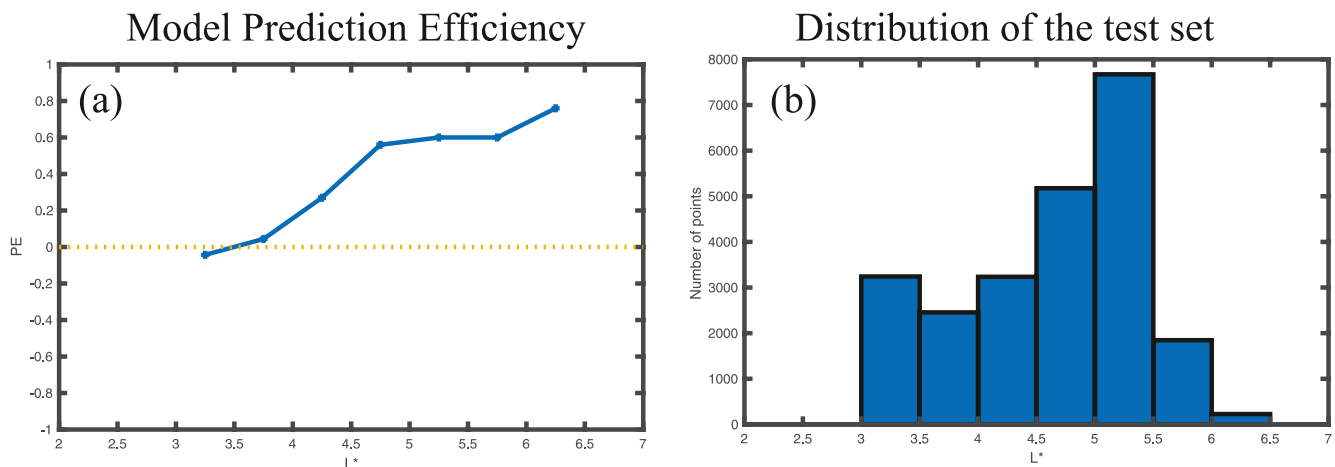

Figure 6. (a) The model prediction efficiency (PE) of the test set as a function of $L*$. The PE is lower at $L* < 4$ or 4.5 because solar wind and magnetospheric parameters transfer little information to phase space density at these $L*$. (b) The distribution of the test set.

We have also compared our model PE with that of PreMevE (Chen et al., 2019) and PreMevE 2.0 (Pires de Lima et al., 2020). PreMevE inputs POES and Los Alamos National Laboratory (LANL) geosynchronous satellite observations of MeV electrons at LEO and GEO, respectively, whereas PreMevE 2.0 inputs solar wind velocity, POES, and LANL data. The comparison is inexact because of the differences in the time cadence, forecasted time range, and predicted quantities. PreMevE 2.0 uses 5 hr time resolution and forecasts 100 keV–2 MeV electron fluxes 1 day ahead (PreMevE 2.0 does not have a nowcast model). Moreover, the PE for PreMevE 2.0 was calculated using log (electron flux), whereas our model PE is calculated using PSD. If these differences can be ignored, PreMevE 2.0 performs better than our model at $L = 2.8$–4.5 (PE = 0.6–0.8), but not as well at $L = 4.5$–6 (PE = 0.4–0.6). Their high PE at $L < 4.5$ can be attributed to the model inputting of POES data. As noted by the authors, PreMevE 2.0 forecasted values often lag behind the observations when the fluxes suddenly jump in response to the sudden change in the solar wind drivers (Pires de Lima et al., 2020), presumably because the NN assigns more weight to the POES electron fluxes than to the solar wind velocity as discussed in the next section.

Chen et al. (2019) developed two nowcast PreMevE models: SubModel 1 and 2. Our model (PE = 0–0.3) performs worse than SubModel 1 (PE = 0.6–0.65) at $L \leq 4$, but our model (PE = 0.6–0.76) performs better than SubModel 1 (PE = −0.2 to 0.2) at $L \geq 5$. Our model performs worse than SubModel 2 (PE = 0.9) at $L \leq 4$ and performs comparably with SubModel 2 (PE = 0.6–0.8) at $L \geq 5$.

5. Discussion and Conclusion

The radiation belt electron PSD has dependencies on the solar wind drivers and the state of the magnetosphere. The PSD also has a strong dependence on its past values because the magnetospheric dynamics can often be characterized, to a large extent, as being persistent. Because of this magnetospheric persistent characteristic, knowledge of the previous values of PSD (or J_e) either directly from in situ satellites or inferred from the precipitating electrons, would immensely help NN learn more easily and reduce the error of the output PSD (or J_e) significantly (e.g., Ling et al., 2010; Pires de Lima et al., 2020). However, a common problem for supervised learning NN models is that during the learning phase, the models would learn quickly that they would do very well if they assigned a lot of weight on the previous values and far less weight on the solar wind input parameters. As a result, the model output would, to some extent, mimic the input value with some time lag and would not be able to respond correctly and timely to sudden changes in the solar wind drivers, for example, sudden arrival of CMEs or density/pressure pulses. This persistent behavior is widely seen not just in the radiation belt models but also in other magnetospheric models that input past values of the predicted parameters (e.g., Wing et al., 2005).

The present study develops an empirical radiation belt model that inputs solar wind parameters, the magnetospheric state parameters as proxied by AL and SYM-H, and L^* (i.e., location in the radiation belts). The model outputs radiation belt electron PSD at a particular set of adiabatic invariant coordinates ($\mu = 560$–875 MeV G^{-1} and I = 0.088–0.14 R_E G$^{0.5}$, and user-input L^*). It is, of course, more challenging to model PSD without having its past values as a reference. On the other hand, the model PSD does not exhibit the undesired persistent behavior where the output PSD would simply mimic the observed PSD with a time lag. Also, this new model can operate independently of input data from any radiation belt observatories, whether they be in the near-equatorial plane (e.g., Van Allen Probes) or at LEO (e.g., POES). This renders the model robust for operational space weather purposes.

The study demonstrates how information theory can be used to assist empirical modeling. Information theory is used to select the solar wind parameters and magnetospheric indices (proxy for the magnetospheric state) and their optimal lag times. The rather large number of past values, up to 72 hr, used in some input parameters (see Table 1) are justified because the results from information theory analysis reveal long-range linear and nonlinear causal relationships between these parameters and PSD (Wing et al., 2022). Information theory analysis also helps explain the model performance, such as increasing PE with increasing L^* as discussed in Section 4. Recently, there has been an increasing amount of effort put into developing "explainable" models, which stems from the desire to build more confidence on the usage of black box models such as neural networks. The fact that all the input parameters and their lag times have been shown to transfer information to PSD (instead of choosing input parameters in an ad hoc manner) and the model performance falls within the expected behavior of information theory analysis, should help build confidence in our model.

Moreover, we have used one of the simplest neural network architectures, namely feed forward-backpropagation or MLP architecture. Although the neural networks dimension is wide and deep, the simple architecture allows for relatively quick training and development time (the model was developed on a laptop computer). However, despite the simple architecture, the model appears to perform well. Using PE as a metric, the model performs as well as or slightly better than DREAM (Reeves et al., 2012) and performs better than AE8 min (Vette, 1991) and CRRESELE (Brautigam & Bell, 1995) models. Moreover, in our model, the error is generally small, |Δ log PSD| < 1. There are instances when the error is large, but these points are usually associated with low PSD slot region, which is expected considering the very high and sharp gradient in PSD at the boundary between the outer belt and the slot. Also, low PSD may have smaller space weather impacts. The good performance can be attributed, at least partly, to the usage of information theory, which guides the selection of the input parameters and their lag times.

Interestingly, just like our model, the DREAM model's PE increases with increasing L^* but for a different reason. DREAM performs better at higher L^* because the model was developed using data at $L^* > 4.2$ (Reeves et al., 2012), whereas our model performs better at higher L^* because solar wind and magnetospheric indices (SYM-H, and AL) transfer more information to higher L^* than lower L^*. This behavior can be contrasted to a class of empirical models that input precipitating radiation belt electrons observed at LEO. For example, the PEs for PreMevE (Chen et al., 2019) and PreMevE 2.0 (Pires de Lima et al., 2020) generally decrease with increasing L because the models input POES data. The lower performance with increasing L is also seen in another model, SHELLS, which inputs POES data (and Kp) (Claudepierre & O'Brien, 2020). They suggested that this behavior can be explained by (a) pitch angle scattering rate, which is proportional to |B|, decreases with increasing L; (b) rate of radial diffusion increases with L; and (c) low to high altitude mapping accuracy decreases with increasing L due to deviation from the dipolar field. Thus, it can be seen that based on the performance as a function of L or L^*, our model can potentially complement a class of empirical models that input POES data or, in general, LEO satellite data.

For operational consideration, the model can input solar wind observations that are routinely available from the solar wind monitor at L1 and forecast PSD 30–60 min ahead. The input AL can be obtained from an AL forecast/nowcast model that is driven entirely by solar wind (e.g., Amariutei & Ganushkina, 2012; Li et al., 2007; Luo et al., 2013; Weigel et al., 1999). Likewise, the input SYM-H can be obtained from a SYM-H or Dst forecast/nowcast model that is driven entirely by the solar wind (e.g., Bhaskar & Vichare, 2019; Cai et al., 2009; Chandorkar et al., 2017; Siciliano et al., 2021; Temerin & Li, 2002, 2006). The Luo et al. (2013) AL and Temerin and Li (2006) Dst forecasts are routinely made available at the University of Colorado website http://lasp.colorado.edu/space_weather/dsttemerin/dsttemerin.html. Although our model PSD can be useful for comparisons with physics-based models such as DREAM (Reeves et al., 2012) or VERB (Shprits et al., 2009), for operational consideration, we plan to write a post-processing routine to convert PSD to electron fluxes, which are more useful for space weather users.

The present model, which uses a simple neural network architectures, is intended to serve as a baseline model. To follow up on the present study, we plan to use more sophisticated neural network architectures, long short-term memory (LSTM), which was designed to work with time series data, and hence holds promises for better performance.

Data Availability Statement

The solar wind, SYM-H, and AL data set were obtained from NASA OMNIweb https://omniweb.gsfc.nasa.gov/. The RBSP MagEIS Level 4 data can be obtained from https://spdf.gsfc.nasa.gov/pub/data/rbsp/ and the PSD data can be obtained from https://rbspgway.jhuapl.edu/psd. The model and the derived data are publicly available at the Zenodo Archive (Wing, 2022).

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2016). *TensorFlow: Large-Scale machine learning on heterogeneous distributed systems*, Technical Report, arXiv:1603.04467 .
Alves, L. R., Souza, V. M., Jauer, P. R., da Silva, L. A., Medeiros, C., Braga, C. R., et al. (2017). The role of solar wind structures in the generation of ULF waves in the inner magnetosphere. *Solar Physics*, 292(7), 92. https://doi.org/10.1007/s11207-017-1113-4

Amariutei, O. A., & Ganushkina, N. Y. (2012). On the prediction of the Auroral westward electrojet index. *Annales Geophysicae*, *30*(5), 841–847. https://doi.org/10.5194/angeo-30-841-2012

Baker, D. N., Erickson, P. J., Fennell, J. F., Foster, J. C., Jaynes, A. N., & Verronen, P. T. (2018). Space weather effects in the Earth's radiation belts. *Space Science Reviews*, *214*(1), 17. https://doi.org/10.1007/s11214-017-0452-7

Baker, D. N., Hoxie, V., Zhao, H., Jaynes, A. N., Kanekal, S., Li, X., & Elkington, S. (2019). Multiyear measurements of radiation belt electrons: Acceleration, transport, and loss. *Journal of Geophysical Research: Space Physics*, *124*(4), 2588–2602. https://doi.org/10.1029/2018JA026259

Baker, D. N., McPherron, R. L., Cayton, T. E., & Klebesadel, R. W. (1990). Linear prediction filter analysis of relativistic electron properties at 6.6 R_E. *Journal of Geophysical Research*, *95*(A9), 15133–15140. https://doi.org/10.1029/JA095iA09p15133

Baker, D. N., Pulkkinen, T. I., Angelopoulos, V., Baumjohann, W., & McPherron, R. L. (1996). Neutral line model of substorms: Past results and present view. *Journal of Geophysical Research*, *101*(A6), 12975–13010. https://doi.org/10.1029/95JA03753

Balikhin, M. A., Boynton, R. J., Walker, S. N., Borovsky, J. E., Billings, S. A., & Wei, H. L. (2011). Using the NARMAX approach to model the evolution of energetic electrons fluxes at geostationary orbit. *Geophysical Research Letters*, *38*(18), L18105. https://doi.org/10.1029/2011GL048980

Balikhin, M. A., Rodriguez, J. V., Boynton, R. J., Walker, S. N., Aryan, H., Sibeck, D. G., & Billings, S. A. (2016). Comparative analysis of NOAA REFM and SNB³GEO tools for the forecast of the fluxes of high-energy electrons at GEO. *Space Weather*, *14*(1), 22–31. https://doi.org/10.1002/2015SW001303

Bhaskar, A., & Vichare, G. (2019). Forecasting of SYMH and ASYH indices for geomagnetic storms of solar cycle 24 including St. Patrick's day, 2015 storm using NARX neural network. *Journal of Space Weather and Space Climate*, *9*, A12. https://doi.org/10.1051/swsc/2019007

Blake, J. B., Carranza, P. A., Claudepierre, S. G., Clemmons, J. H., Crain, W. R., Dotan, Y., et al. (2013). The magnetic electron ion spectrometer (MagEIS) instruments aboard the radiation belt storm probes (RBSP) spacecraft. *Space Science Reviews*, *179*(1–4), 383–421. https://doi.org/10.1007/s11214-013-9991-8

Boberg, F., Wintoft, P., & Lundstedt, H. (2000). Real time *K*p prediction from solar wind data using neural networks. *Physics and Chemistry of the Earth*, *25*(4), 275–280. https://doi.org/10.1016/s1464-1917(00)00016-7

Borovsky, J. E. (2017). Time-integral correlations of multiple variables with the relativistic-electron flux at geosynchronous orbit: The strong roles of substorm-injected electrons and the ion plasma sheet. *Journal of Geophysical Research: Space Physics*, *122*(12), 11961–11990. https://doi.org/10.1002/2017JA024476

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. (2020). What magnetospheric and ionospheric researchers should know about the solar wind. *Journal of Atmospheric and Solar-Terrestrial Physics*, *204*, 105271. https://doi.org/10.1016/j.jastp.2020.105271

Borovsky, J. E., & Denton, M. H. (2009). Relativistic-electron dropouts and recovery: A superposed epoch study of the magnetosphere and the solar wind. *Journal of Geophysical Research*, *114*(A2), A02201. https://doi.org/10.1029/2008JA013128

Borovsky, J. E., & Denton, M. H. (2014). Exploring the cross correlations and autocorrelations of the ULF indices and incorporating the ULF indices into the systems science of the solar wind-driven magnetosphere. *Journal of Geophysical Research: Space Physics*, *119*(6), 4307–4334. https://doi.org/10.1002/2014JA019876

Borovsky, J. E., & Yakymenko, K. (2017). Substorm occurrence rates, substorm recurrence times, and solar wind structure. *Journal of Geophysical Research: Space Physics*, *122*(3), 2973–2998. https://doi.org/10.1002/2016JA023625

Boyd, A. J., Spence, H. E., Huang, C.-L., Reeves, G. D., Baker, D. N., Turner, D. L., et al. (2016). Statistical properties of the radiation belt seed population. *Journal of Geophysical Research: Space Physics*, *121*(8), 7636–7646. https://doi.org/10.1002/2016JA022652

Brautigam, D. H., & Bell, J. T. (1995). CRRESELE documentation. In *Report PL-TR-95-2128, phillips lab*. Hanscom AFB.

Cai, L., Ma, S., Cai, H., Zhou, Y., & Liu, R. (2009). Prediction of SYM-H index by NARX neural network from IMF and solar wind data. *Science in China Series E: Technological Sciences*, *52*(10), 2877–2885. https://doi.org/10.1007/s11431-009-0296-9

Chandorkar, M., Camporeale, E., & Wing, S. (2017). Probabilistic forecasting of the disturbance storm time index: An autoregressive Gaussian process approach. *Space Weather*, *15*(8), 1004–1019. https://doi.org/10.1002/2017SW001627

Chen, Y., Reeves, G. D., Fu, X., & Henderson, M. (2019). PreMevE: New predictive model for megaelectron-volt electrons inside Earth's outer radiation belt. *Space Weather*, *17*(3), 438–454. https://doi.org/10.1029/2018SW002095

Chu, X., Ma, D., Bortnik, J., Tobiska, W. K., Cruz, A., Bouwer, S. D., et al. (2021). Relativistic electron model in the outer radiation belt using a neural network approach. *Space Weather*, *19*(12), e2021SW002808. https://doi.org/10.1029/2021SW002808

Claudepierre, S. G., & O'Brien, T. P. (2020). Specifying high-altitude electrons using low-altitude LEO systems: The SHELLS model. *Space Weather*, *18*(3), e2019SW002402. https://doi.org/10.1029/2019SW002402

Coleman, T., McCollough, J. P., Young, S., & Rigler, E. J. (2018). Operational nowcasting of electron flux levels in the outer zone of Earth's radiation belt. *Space*.

Davis, T. N., & Sugiura, M. (1966). Auroral electrojet activity index *AE* and its universal time variations. *Journal of Geophysical Research*, *71*(3), 785–801. https://doi.org/10.1029/JZ071i003p00785

Dessler, A., & Parker, E. (1959). Hydromagnetic theory of geomagnetic storms. *Journal of Geophysical Research*, *64*(12), 2239–2252. https://doi.org/10.1029/JZ064i012p02239

Ginet, G. P., O'Brien, T. P., Huston, S. L., Johnston, W. R., Guild, T. B., Friedel, R., et al. (2013). AE9, AP9 and SPM: New models for specifying the trapped energetic particle and space plasma environment. *Space Science Reviews*, *179*(1–4), 579–615. https://doi.org/10.1007/s11214-013-9964-y

Green, J. C., & Kivelson, M. G. (2004). Relativistic electrons in the outer radiation belt: Differentiating between acceleration mechanisms. *Journal of Geophysical Research*, *109*(A3), A03213. https://doi.org/10.1029/2003JA010153

Horne, R. B., Glauert, S. A., Meredith, N. P., Boscher, D., Maget, V., Heynderickx, D., & Pitchford, D. (2013). Space weather impacts on satellites and forecasting the Earth's electron radiation belts with SPACECAST. *Space Weather*, *11*(4), 169–186. https://doi.org/10.1002/swe.20023

Iyemori, T. (1990). Storm-time magnetospheric currents inferred from mid-latitude geomagnetic field variations. *Journal of Geomagnetism and Geoelectricity*, *42*(11), 1249–1265. https://doi.org/10.5636/jgg.42.1249

Johnson, J. R., & Wing, S. (2018). An information-theoretical approach to space weather. In *Machine learning techniques for space weather* (pp. 45–69). Elsevier (ISBN: 978-0-12-811788-0).

Kellerman, A. C., & Shprits, Y. Y. (2012). On the influence of solar wind conditions on the outer-electron radiation belt. *Journal of Geophysical Research*, *117*(A5), A05217. https://doi.org/10.1029/2011JA017253

Kellerman, A. C., Shprits, Y. Y., & Turner, D. L. (2013). A geosynchronous radiation-belt electron empirical prediction (GREEP) model. *Space Weather*, *11*(8), 463–475. https://doi.org/10.1002/swe.20074

Koons, H. C., & Gorney, D. J. (1991). A neural network model of the relativistic electron flux at geosynchronous orbit. *Journal of Geophysical Research*, *96*(A4), 5549–5556. https://doi.org/10.1029/90JA02380

Li, W., Thorne, R. M., Angelopoulos, V., Bonnell, J. W., McFadden, J. P., Carlson, C. W., et al. (2009). Evaluation of whistler-mode chorus intensification on the nightside during an injection event observed on the THEMIS spacecraft. *Journal of Geophysical Research*, *114*(A1), A00C14. https://doi.org/10.1029/2008JA013554

Li, X., Baker, D. N., Temerin, M., Reeves, G., Friedel, R., & Shen, C. (2005). Energetic electrons, 50 keV to 6 MeV, at geosynchronous orbit: Their responses to solar wind variations. *Space Weather*, *3*(4), S04001. https://doi.org/10.1029/2004SW000105

Li, X., Oh, K. S., & Temerin, M. (2007). Prediction of the AL index using solar wind parameters. *Journal of Geophysical Research*, *112*(A6), A06224. https://doi.org/10.1029/2006JA011918

Li, X., Temerin, M., Baker, D. N., Reeves, G. D., & Larson, D. (2001). Quantitative prediction of radiation belt electrons at geostationary orbit based on solar wind measurements. *Geophysical Research Letters*, *28*(9), 1887–1890. https://doi.org/10.1029/2000GL012681

Ling, A. G., Ginet, G. P., Hilmer, R. V., & Perry, K. L. (2010). A neural network–based geosynchronous relativistic electron flux forecasting model. *Space Weather*, *8*(9), S09003. https://doi.org/10.1029/2010SW000576

Luo, B., Li, X., Temerin, M., & Liu, S. (2013). Prediction of the *AU, AL,* and *AE* indices using solar wind parameters. *Journal of Geophysical Research: Space Physics*, *118*(12), 7683–7694. https://doi.org/10.1002/2013JA019188

Lyatsky, W., & Khazanov, G. V. (2008a). Effect of geomagnetic disturbances and solar wind density on relativistic electrons at geostationary orbit. *Journal of Geophysical Research*, *113*(A8), A08224. https://doi.org/10.1029/2008JA013048

Lyatsky, W., & Khazanov, G. V. (2008b). Effect of solar wind density on relativistic electrons at geosynchronous orbit. *Geophysical Research Letters*, *35*(3), L03109. https://doi.org/10.1029/2007GL032524

Maggiolo, R., Hamrin, M., De Keyser, J., Pitkänen, T., Cessateur, G., Gunell, H., & Maes, L. (2017). The delayed time response of geomagnetic activity to the solar wind. *Journal of Geophysical Research: Space Physics*, *122*(11), 11109–11127. https://doi.org/10.1002/2016JA023793

Mauk, B. H., Fox, N. J., Kanekal, S. G., Kessel, R. L., Sibeck, D. G., & Ukhorskiy, A. (2013). Science objectives and rationale for the radiation belt storm probes mission. *Space Science Reviews*, *179*(1–4), 3–27. https://doi.org/10.1007/s11214-012-9908-y

Meredith, N. P., Horne, R. B., & Anderson, R. R. (2001). Substorm dependence of chorus amplitudes: Implications for the acceleration of electrons to relativistic energies. *Journal of Geophysical Research*, *106*(A7), 13165–13178. https://doi.org/10.1029/2000JA900156

Newell, P. T., Wing, S., Meng, C. I., & Sigillito, V. (1990). A neural network based system for monitoring the aurora. *Johns Hopkins APL Technical Digest*, *11*(3–4), 291–299.

Newell, P. T., Wing, S., Meng, C.-I., & Sigillito, V. (1991). The Auroral oval position, structure, and intensity of precipitation from 1984 onward: An automated on-line data base. *Journal of Geophysical Research*, *96*(A4), 5877–5882. https://doi.org/10.1029/90JA02450

Paulikas, G. A., & Blake, J. B. (1979). Effects of the solar wind on magnetospheric dynamics: Energetic electrons at the synchronous orbit. In *Quantitative modeling of magnetospheric processes* (Vol. 21, pp. 180–202). Geophys. Monogr. Ser., AGU.

Perry, K. L., Ginet, G. P., Ling, A. G., & Hilmer, R. V. (2010). Comparing geosynchronous relativistic electron prediction models. *Space Weather*, *8*(12), S12002. https://doi.org/10.1029/2010SW000581

Pinto, V. A., Kim, H.-J., Lyons, L. R., & Bortnik, J. (2018). Interplanetary parameters leading to relativistic electron enhancement and persistent depletion events at geosynchronous orbit and potential for prediction. *Journal of Geophysical Research: Space Physics*, *123*(2), 1134–1145. https://doi.org/10.1002/2017JA024902

Pires de Lima, R., Chen, Y., & Lin, Y. (2020). Forecasting megaelectron-volt electrons inside earth's outer radiation belt: PreMevE 2.0 based on supervised machine learning algorithms. *Space Weather*, *18*(2), e2019SW002399. https://doi.org/10.1029/2019SW002399

Reeves, G., Morley, S., & Cunningham, G. (2013). Long-term variations in solar wind velocity and radiation belt electrons. *Journal of Geophysical Research: Space Physics*, *118*(3), 1040–1048. https://doi.org/10.1002/jgra.50126

Reeves, G. D. (2007). Radiation belt storm probes: A new mission for space weather forecasting. *Space Weather*, *5*(11). https://doi.org/10.1029/2007SW000341

Reeves, G. D., Chen, Y., Cunningham, G. S., Friedel, R. W. H., Henderson, M. G., Jordanova, V. K., et al. (2012). Dynamic radiation environment assimilation model: DREAM. *Space Weather*, *10*(3), S03006. https://doi.org/10.1029/2011SW000729

Reeves, G. D., Morley, S. K., Friedel, R. H. W., Henderson, M. G., Cayton, T. E., Cunningham, G., et al. (2011). On the relationship between relativistic electron flux and solar wind velocity: Paulikas and Blake revisited. *Journal of Geophysical Research*, *116*(A2), A02213. https://doi.org/10.1029/2010JA015735

Rigler, E. J., Wiltberger, M., & Baker, D. N. (2007). Radiation belt electrons respond to multiple solar wind inputs. *Journal of Geophysical Research*, *112*(A6), A06208. https://doi.org/10.1029/2006JA012181

Roederer, J. G. (1970). *Dynamics of geomagnetically trapped radiation*. Springer. https://doi.org/10.1007/978-3-642-49300-3

Rumelhart, D. E., & McClelland, J. L. (Eds.) (1987). *Parallel distributed processing*. (Vol. I). MIT Press.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117. arXiv:1404.7828. https://doi.org/10.1016/j.neunet.2014.09.003

Schulz, M., & Lanzerotti, L. J. (1974). *Particle diffusion in the radiation belts*. Springer. https://doi.org/10.1007/978-3-642-65675-0

Shprits, Y. Y., Subbotin, D., & Ni, B. (2009). Evolution of electron fluxes in the outer radiation belt computed with the VERB code. *Journal of Geophysical Research*, *114*(A11), A11209. https://doi.org/10.1029/2008JA013784

Siciliano, F., Consolini, G., Tozzi, R., Gentili, M., Giannattasio, F., & De Michelis, P. (2021). Forecasting *SYM-H* index: A comparison between long short-term memory and convolutional neural networks. *Space Weather*, *19*(2), e2020SW002589. https://doi.org/10.1029/2020SW002589

Simms, L. E., & Engebretson, M. J. (2020). Classifier neural network models predict relativistic electron events at geosynchronous orbit better than multiple regression or ARMAX models. *Journal of Geophysical Research: Space Physics*, *125*(5), e2019JA027357. https://doi.org/10.1029/2019JA027357

Smirnov, A. G., Berrendorf, M., Shprits, Y. Y., Kronberg, E. A., Allison, H. J., Aseev, N. A., et al. (2020). Medium energy electron flux in earth's outer radiation belt (MERLIN): A machine learning model. *Space Weather*, *18*(11), e2020SW002532. https://doi.org/10.1029/2020SW002532

Spence, H. E., Reeves, G. D., Baker, D. N., Blake, J. B., Bolton, M., Bourdarie, S., et al. (2013). Science goals and overview of the radiation belt storm probes (RBSP) energetic particle, composition, and thermal plasma (ECT) suite on NASA's Van allen probes mission. *Space Science Reviews*, *179*(1–4), 311–336. https://doi.org/10.1007/s11214-013-0007-5

Tang, C. L., Wang, Y. X., Ni, B., Su, Z. P., Reeves, G. D., Zhang, J.-C., et al. (2017). The effects of magnetospheric processes on relativistic electron dynamics in the Earth's outer radiation belt. *Journal of Geophysical Research: Space Physics*, *122*(10), 9952–9968. https://doi.org/10.1002/2017JA024407

Tang, C. L., Wang, Y. X., Ni, B., Zhang, J.-C., Reeves, G. D., Su, Z. P., et al. (2017). Radiation belt seed population and its association with the relativistic electron dynamics: A statistical study. *Journal of Geophysical Research: Space Physics*, *122*(5), 5261–5276. https://doi.org/10.1002/2017JA023905

Tang, R., Tao, Y., Li, J., Chen, Z., Deng, X., & Li, H. (2022). The short-time prediction of the energetic electron flux in the planetary radiation belt based on stacking ensemble-learning algorithm. *Space Weather*, *20*(2), e2021SW002969. https://doi.org/10.1029/2021SW002969

Temerin, M., & Li, X. (2002). A new model for the prediction of *Dst* on the basis of the solar wind. *Journal of Geophysical Research*, *107*(A12), 1472. https://doi.org/10.1029/2001JA007532

Temerin, M., & Li, X. (2006). *Dst* model for 1995–2002. *Journal of Geophysical Research*, *111*(A4), A04221. https://doi.org/10.1029/2005JA011257

Thorne, R. M. (2010). Radiation belt dynamics: The importance of wave-particle interactions. *Geophysical Research Letters*, *37*(22), L22107. https://doi.org/10.1029/2010GL044990

Tsyganenko, N. A., & Sitnov, M. I. (2005). Modeling the dynamics of the inner magnetosphere during strong geomagnetic storms. *Journal of Geophysical Research*, *110*(A3), A03208. https://doi.org/10.1029/2004JA010798

Turner, D. L., Angelopoulos, V., Li, W., Bortnik, J., Ni, B., Ma, Q., et al. (2014). Competing source and loss mechanisms due to wave-particle interactions in Earth's outer radiation belt during the 30 September to 3 October 2012 geomagnetic storm. *Journal of Geophysical Research: Space Physics*, *119*, 1960–1979. https://doi.org/10.1002/2014JA019770

Turner, D. L., Angelopoulos, V., Morley, S. K., Henderson, M. G., Reeves, G. D., Li, W., et al. (2014). On the cause and extent of outer radiation belt losses during the 30 September 2012 dropout event. *Journal of Geophysical Research: Space Physics*, *119*(3), 1530–1540. https://doi.org/10.1002/2013JA019446

Turner, D. L., Shprits, Y., Hartinger, M., & Angelopoulos, V. (2012). Explaining sudden losses of outer radiation belt electrons during geomagnetic storms. *Nature Physics*, *8*(3), 208–212. https://doi.org/10.1038/nphys2185

Ukhorskiy, A. Y., Anderson, B. J., Brandt, P. C., & Tsyganenko, N. A. (2006). Storm time evolution of the outer radiation belt: Transport and losses. *Journal of Geophysical Research*, *111*(A11), A11S03. https://doi.org/10.1029/2006JA011690

Ukhorskiy, A. Y., Sitnov, M. I., Sharma, A. S., Anderson, B. J., Ohtani, S., & Lui, A. T. Y. (2004). Data-derived forecasting model for relativistic electron intensity at geosynchronous orbit. *Geophysical Research Letters*, *31*(9), L09806. https://doi.org/10.1029/2004GL019616

Vette, J. I. (1991). The AE-8 trapped electron model environment. In *Rep. NSSDC/WDC-A-R&S 91–24*. NASA Goddard Space Flight Center.

Wei, H.-L., Billings, S. A., Surjalal Sharma, A., Wing, S., Boynton, R. J., & Walker, S. N. (2011). Forecasting relativistic electron flux using dynamic multiple regression models. *Annales Geophysicae*, *29*(2), 415–420. https://doi.org/10.5194/angeo-29-415-2011

Weigel, R., Horton, W., Tajima, T., & Detman, T. (1999). Forecasting auroral electrojet activity from solar wind input with neural networks. *Geophysical Research Letters*, *26*(10), 1353–1356. https://doi.org/10.1029/1999gl900280

Wing, S. (2022). Radiation belt model based on information theory informed neural network. In *Space weather (version 0)*. https://doi.org/10.5281/zenodo.6941624

Wing, S., Greenwald, R. A., Meng, C.-I., Sigillito, V. G., & Hutton, L. V. (2003). Neural networks for automated classification of ionospheric irregularities in HF radar backscattered signals. *Radio Science*, *38*(4), 1063. https://doi.org/10.1029/2003RS002869

Wing, S., & Johnson, J. R. (2019). Applications of information theory in solar and space physics. *Entropy*, *21*(2), 140. https://doi.org/10.3390/e21020140

Wing, S., Johnson, J. R., Camporeale, E., & Reeves, G. D. (2016). Information theoretical approach to discovering solar wind drivers of the outer radiation belt. *Journal of Geophysical Research: Space Physics*, *121*(10), 9378–9399. https://doi.org/10.1002/2016JA022711

Wing, S., Johnson, J. R., Jen, J., Meng, C.-I., Sibeck, D. G., Bechtold, K., et al. (2005). *K*p forecast models. *Journal of Geophysical Research*, *110*(A4), A04203. https://doi.org/10.1029/2004JA010500

Wing, S., Johnson, J. R., Turner, D. L., Ukhorskiy, A. Y., & Boyd, A. J. (2022). Untangling the solar wind and magnetospheric drivers of the radiation belt electrons. *Journal of Geophysical Research: Space Physics*, *127*(4), e2021JA030246. https://doi.org/10.1029/2021JA030246

Wintoft, P., Wik, M., Matzka, J., & Shprits, Y. (2017). Forecasting *K*p from solar wind data: Input parameter study using 3-hour averages and 3 hr range values. *Journal of Space Weather and Space Climate*, *7*, A29. https://doi.org/10.1051/swsc/2017027

Wu, J., & Lundstedt, H. (1997). Geomagnetic storm predictions from solar wind data with the use of dynamic neural networks. *Journal of Geophysical Research*, *102*(A7), 14255–14268. https://doi.org/10.1029/97ja00975

Xiang, Z., Tu, W., Li, X., Ni, B., Morley, S. K., & Baker, D. N. (2017). Understanding the mechanisms of radiation belt dropouts observed by Van Allen Probes. *Journal of Geophysical Research: Space Physics*, *122*(10), 9858–9879. https://doi.org/10.1002/2017JA024487

Zhao, H., Baker, D. N., Jaynes, A. N., Li, X., Elkington, S. R., Kanekal, S. G., et al. (2017). On the relation between radiation belt electrons and solar wind parameters/geomagnetic indices: Dependence on the first adiabatic invariant and *L**. *Journal of Geophysical Research: Space Physics*, *122*(2), 1624–1642. https://doi.org/10.1002/2016JA023658