Modeling radiation belt electrons with information theory informed neural network

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, Maryland, USA

[2]Andrews University, Berrien Spring, Michigan, 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, AL), and L*. The model outputs radiation belt electron phase space density (PSD). The model is operational from L* = 3 to 6.5. The model is constructed with neural network 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. This baseline model complements well a class of empirical models that input data from Low Earth Orbit (LEO). 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.

Plain Language Summary

An empirical model of radiation belt relativistic electrons with energy 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 = 6378 km). The model outputs radiation belt electron phase space density (PSD). The model is operational from L* = 3 to L* 6.5. The model is constructed with information theory informed neural network. 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.

ORCID ID: 0000-0001-9342-1813

Keywords: radiation belt, relativistic electrons, solar wind drivers, machine learning, information theory, empirical model, phase space density.

Index terms: 2774, 2784, 2720, 2730, 4499

Broad Implications: A radiation belt relativistic electron model based on neural network assisted by information theory is developed. The model performs well and complements a class of empirical models that input observations from LEO.

Key points: (1) An empirical model to predict state of radiation belt relativistic electrons is developed; (2) The model PE increases with increasing L* with a max of 0.76 at L* = 6.5; (3) The model complements a class of empirical models that input observations from LEO.

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 ~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 = 6378 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 variability of the solar wind (e.g., *Baker et al*., 1990, 2018; 2019; *Li et al*., 2001; 2005; *Reeves*, 2007; *Ukhorskiy et al*., 2004; *Reeves et al.*, 2013; *Xiang et al*., 2017; *Pinto et al*., 2018; *Zhao et al*., 2017, *Alves 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., *Wing et al*., 2016; *Wing and Johnson*, 2019; *Borovsky*, 2018; 2020; *Maggiolo et al*., 2017; *Wing et al*., 2021). For example, solar wind velocity (V_{sw}) positively correlates with radiation belt electron fluxes (J_e) (e.g., *Baker et al*., 1990; *Reeves et al*., 2011; *Balikhin et al*., 2011; *Paulikas and Blake*, 1979; *Li et al*., 2001; 2005; *Wing et al*., 2016; 2020). Solar wind density (n_{sw}) negatively correlates with radiation belt J_e (e.g., Li et al., 2005; *Lyatsky and Kazanov*, 2008a; *Kellerman and Shprits*, 2012; *Rigler et al*., 2007; *Balikhin et al*., 2011; *Wing et al*., 2016; 2020). However, V_{sw} negatively correlates with n_{sw} (e.g., *Wing et al*., 2016; 2021; *Borovsky*, 2020).

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., *Reeves et al*., 1998; *Baker et al.*, 2019; *Lyatsky and Khazanov*, 2008b; *Borovsky and Denton*, 2014; *Tang et al*., 2017b; *Borovsky*, 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 and Sugiura*, 1966). SYM-H is similar to Disturbance Storm Time (Dst) index (*Dessler and Parker*, 1959), except that SYM-H index has one minute time resolution whereas Dst index has one 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*., 2021).

Wing et al. (2016; 2021) 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*., 2021). 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) has 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 and Lundstedt*, 1997), source regions of particle precipitation (*Newell et al*., 1990; 1991), high-frequency (HF) backscattered signals (*Wing et al*., 2003). NN have also been used to construct empirical radiation belt models (e.g., *Koons and Gorney*, 1991; *Perry et al*., 2010; *Ling et al*., 2010; *Smirnov et al*., 2020; *Claudepierre and O'Brien*, 2020; *Pires de Lima et al*., 2020; *Chen et al*., 2019; *Simms and Engebretson*, 2020). These empirical models generally complement physics-based models, e.g., DREAM (*Reeves et al.*, 2012), SPACECAST (*Horne et al*., 2013), VERB (*Shprits et al*. 2009) and other empirical models that use different approaches, e.g., NARMAX (*Wei et al*., 2011; *Balikhin et al*., 2016), Kalman filter (*Coleman et al*., 2018), linear prediction filter (*Baker et al*., 1990; *Kellerman et al*., 2012; *Chen et al*., 2019). 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, e.g., 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 radiation belt electron PSD using an information-theory informed NN as the core of the model (*Johnson and 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*. (2021) information theoretic analysis of the solar wind and magnetospheric drivers of PSD.

2. Data set

The NASA's Van Allen Probe (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). MagEIS measured the energy range of 30 keV to 4 MeV for electrons and 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 of the three by their adiabatic invariants and PSD (μ, I, 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 and 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 and Lanzerotti*, 1974).

The radiation belt electron PSD from MagEIS is calculated at 1 min resolution using TS04 magnetic field model (*Tsyganenko and Sitnov*, 2005) and a method similar to that used in *Turner et al*., 2014a; 2014b). 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 dataset were used and provided by NASA (http://omniweb.gsfc.nasa.gov/). Both the PSD and OMNI data 2013-2018 are averaged with 30 min sliding window.

We merge each OMNI solar wind parameter (V_{sw}, n_{sw} etc.) with the RBSP electron PSD (data from both RBSP A and B are used). The merged dataset 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*. (2021) ranked the solar wind and magnetospheric parameters based on the information transfer to the PSD (see Table 1 in *Wing et al*. (2021)). We select the top 8 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 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 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 multi-layered perceptrons (MLP). The NN architecture has 5 layers: 1 input layer (531 nodes), 1 output layer (1 node), and 3 hidden layers (each has 800 nodes). The model is developed using python and Tensorflow 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 left as shown in Figure 3a. In order to get higher performance, log PSD is used rather than PSD. Log PSD (Figure 3b) 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: (1) training set and (2) test set. The training set consists of data in the time intervals (2013.5–2015.5), (2016–2017), (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.

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: (1) 2013 April 27 – May 13 and (2) 2017 Mar 13 – 29. These intervals are selected also because they exhibit a wide range of solar wind driving as well as geomagentic 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.

Figures 4 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 quasi-periodic substorms (minimum AL > –400 nT) throughout the interval, which is fairly typical (*Borovsky and 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 SYM-H ~ –60 nT) and 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 magnetopause shadowing due to the sharp rise in solar wind density and dynamic pressure (e.g., *Li et al.*, 2001; *Kellerman and Shprits*, 2012; *Turner et al.*, 2012; *Ukhorskiy et al.*, 2006). However, the PSD seems 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 quasi periodic 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*, 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 4f shows that most of the time the observed and model PSD are within the same order of magnitude of each other, $|\Delta \log \mathrm{PSD}| < 1$ $(c^3\ MeV^{-3}\ cm^{-3})^{-1}$. Large $|\Delta \log \mathrm{PSD}|$ generally corresponds to low PSD and low L* that is in the slot region. In order to show this, several dotted vertical red lines are drawn to connect some of the largest $|\Delta \log \mathrm{PSD}|$ in Figure 4f to their corresponding PSD in Figure 4g. This trend can be seen throughout Figure 4. When PSD is low, a little discrepancy from the observed value would lead to large $|\Delta \log \mathrm{PSD}|$. 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 4, most of the time, the error is small, $|\Delta \log \mathrm{PSD}| < 1$ $(c^3\ MeV^{-3}\ cm^{-3})^{-1}$, for high and low PSD,

Figure 5 presents the interval 2013 May 05–13, which is the second half of the first event, in the same format as Figure 4. As in Figure 4, panel d shows quasi periodic moderate and small susbstorms (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 $(c^3\ MeV^{-3}\ cm^{-3})^{-1}$) on panel f. Unlike the density/pressure pulse in Figure 4, 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 $(c^3\ MeV^{-3}\ cm^{-3})^{-1}$) 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 $(c^3\ MeV^{-3}\ cm^{-3})^{-1}$) during this interval, except near the end at low L* (L* < 4) where Δ log PSD > 1 $(c^3\ MeV^{-3}\ cm^{-3})^{-1}$. It is not clear what causes the model to underestimate PSD at this time. As in Figure 4, 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 large and small PSD.

Figure 6 presents the interval 2017 Mar 13–21, which is the first half of the second event in the same format as Figures 4 and 5. This interval shows the worst model performance out of the four intervals presented herein and one of the worst intervals seen in the entire test set. As in the previous intervals, there are quasi periodic 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 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 (e.g., *Li et al.*, 2001; *Kellerman and Shprits*, 2012; *Turner et al*., 2012; *Ukhorskiy et al*., 2006; *Wing et al*., 2016; 2021) and then increase due to the expected storm-time acceleration and substorm injections (e.g,, *Baker et al*., 1996; *Tang et al.*, 2017a; *Boyd et al*., 2016; *Wing et al*., 2016; 2021; *Meredith et al*., 2001; *Li et al*., 2009). It is not clear why this expected behavior is not observed in the RBSP PSD. 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 7 presents the interval 2017 Mar 21–29, which is the second half of the second event in the same format as Figure 6. 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. 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 (c^3 MeV^{-3} $cm^{-3})^{-1}$. 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 due to the magnetopause shadowing. 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 (c^3 MeV^{-3} $cm^{-3})^{-1}$) 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 ~ – 80nT) and three large substorms (two with minimum AL ~ –1000 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 of the day on Mar 28. The model PSD tracks the observed PSD fairly well during this highly disturbed period (Δ log PSD < 1 (c^3 MeV^{-3} $cm^{-3})^{-1}$) as shown in panels f and g.

Figures 4–7 show that the model performs well and the error is small for large and small PSD. There are instances when the error is large, |Δ log PSD| > 1 (c^3 MeV^{-3} $cm^{-3})^{-1}$, 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. Based on the evaluation of model PSD for the entire test set: root mean square (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 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 7 bins with each bin having 0.5. Figure 8a plots the PE as a function of L*, which ranges from from –0.043 for L* = 3 to 0.76 for L* = 6.5. Figure 8b 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 because this L* range has the most data points as shown in Figure 8b.

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 this location that at L* > 4. Indeed, *Wing et al*. (2021) 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 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*., 2021) 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. As with DREAM, our model performs better than AE8min (*Vette*, 1991) and CRRESELE (*Brautigam and Bell*, 1995) models. For many years, AE8 series model was considered standard for engineering applications. (AE8min model is superseded by a newer model, AE9, (*Ginet et al*., 2013), but like AE8, AE9 is a statistical model that is not relevant to individual event-based prediction).

We have also compared our model PE with that of PreMevE 2.0, which inputs solar wind velocity, POES and Los Alamos National Laboratory (LANL) geosynchronous satellite observations of MeV electrons at LEO and GEO, respectively, (*Pires de Lima et al*., 2020). The comparison is inexact because PreMevE 2.0 uses 5-hour time resolution, and forecasts 100 keV – 2MeV electron fluxes one day ahead. 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 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.

5. Discussion and conclusion

The radiation belt electron PSD has dependences 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 persistence 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., *Pires de Lima*, 2020; *Ling et al*., 2010). 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, e.g., sudden arrival of CMEs or density/pressure pulses. This persistence 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; *Pires de Lima*, 2020).

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 (μ = 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 persistence behavior where the output PSD would simply mimic the observed PSD with a time lag. Also, this new model can operate independent 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 of the radiation belt electron variability. 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 hours, used in some input parameters (see Table 1) are justified because the results from information theory analysis reveals long range linear and nonlinear causal relationship between these parameters and PSD (*Wing et al*., 2021). 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 increasing amount of efforts 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 architecture, namely feed forward-backpropagation or MLP architecture. Although the neural network 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 AE8min (*Vette*, 1991) and CRRESELE (*Brautigam and Bell*, 1995) models. Moreover, in our model, the error is generally small, $|\Delta \log \mathrm{PSD}| < 1\ (c^3\ \mathrm{MeV}^{-3}\ \mathrm{cm}^{-3})^{-1}$. 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 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, 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 PrevMevE 2.0 (*Pires de Lima et al*., 2020) generally decrease with increasing L because the models input POES data. PreMevE inputs only observations from POES at LEO and LANL at GEO (*Chen et al*., 2019). The lower performance with increasing L is also seen in another model, SHELLS, which inputs POES data (and Kp) (*Claudepierre and O'Brien*, 2020). They suggested that this behavior can be explained by (1) pitch angle scattering rate, which is proportional to |B|, decreases with increasing L; (2) rate of radial diffusion increases with L; and (3) low to high altitude mapping accuracy decreases with increasing L due to deviation from dipolar field. Thus, it can be seen that based on the performance as a function of L or L*, our model can 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., *Luo et al*., 2013; Li et al., 2007; *Weigel et al*., 1999; *Amariutei et al*., 2012). 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., *Temerin and Li*, 2006; *Cai et al*., 2009; *Bhaskar and Vichare*, 2019; *Chandorkar et al*., 2017; Siciliano et al., 2021). 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.

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

Acknowledgments. The solar wind, SYM-H, and AL dataset 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. Simon Wing acknowledges support of NASA Van Allen Probe Contract NNN16AA09T and NASA Grants NNX16AQ87G, 80NSSC20K0704, 80NSSC19K0843, 80NSSC19K0822, 80NSSC20K0188, 80NSSC20K1279, and 80NSSC20K1271. This work has benefitted from discussions within the International Space Science Institute (ISSI) Team # 455 "Complex Systems Perspectives Pertaining to the Research of the Near-Earth Electromagnetic Environment."

References

Abadi, M., A. Agarwal, P. Barham 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. *et al.* The Role of Solar Wind Structures in the Generation of ULF Waves in the Inner Magnetosphere. *Sol Phys* **292,** 92 (2017). https://doi.org/10.1007/s11207-017-1113-4

Amariutei, O. A. and Ganushkina, N. Yu.: On the prediction of the auroral westward electrojet index, Ann. Geophys., 30, 841–847, https://doi.org/10.5194/angeo-30-841-2012, 2012

Baker, D. N., R. L. McPherron, T. E. Cayton, and R. W. Klebesadel (1990), Linear prediction filter analysis of relativistic electron properties at 6.6 R_E , J. Geophys. Res., 95(A9),15133–15140, doi:10.1029/JA095iA09p15133.

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

Baker, D.N., Erickson, P.J., Fennell, J.F. *et al.* Space Weather Effects in the Earth's Radiation Belts. *Space Sci Rev* **214,** 17 (2018). 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, 2588– 2602. https://doi.org/10.1029/2018JA026259

Balikhin, M. A., R. J. Boynton, S. N. Walker, J. E. Borovsky, S. A. Billings, and H. L. Wei(2011), Using the NARMAX approach to model the evolution of energetic electrons fluxes at geostationary orbit, Geophys. Res. Lett., 38, L18105, doi:10.1029/2011GL048980.

Balikhin, M. A., Rodriguez, J. V., Boynton, R. J., Walker, S. N., Aryan, H., Sibeck, D. G., and Billings, S. A. (2016), Comparative analysis of NOAA REFM and SNB^{3}GEO tools for the forecast of the fluxes of high-energy electrons at GEO, *Space Weather*, 14, 22– 31, doi:10.1002/2015SW001303

Bhaskar, Ankush, and Geeta Vichare (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, J. Space Weather Space Clim., 9, A12, DOI: 10.1051/swsc/2019007

Blake, J.B., Carranza, P.A., Claudepierre, S.G. *et al.* (2013), The *Magnetic Electron Ion Spectrometer* (MagEIS) Instruments Aboard the Radiation Belt Storm Probes (RBSP) Spacecraft. *Space Sci Rev* **179,** 383–421, https://doi.org/10.1007/s11214-013-9991-8

Boberg, F., P. Wintoft, and H. Lundstedt (2000), Real time Kp prediction from solar wind data using neural networks, Phys. Chem. Earth, 25, 275-280.

Borovsky, J. E., and 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, *J. Geophys. Res. Space Physics*, 119, 4307– 4334, doi:10.1002/2014JA019876.

Borovsky, J. E., and Yakymenko, K. (2017), Substorm occurrence rates, substorm recurrence times, and solar wind structure, *J. Geophys. Res. Space Physics*, 122, 2973– 2998, doi:10.1002/2016JA023625.

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, 11,961– 11,990. 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, 20– 29. https://doi.org/10.1002/2017JA024650

Borovsky, J. E. (2020), What magnetospheric and ionospheric researchers should know about the solar wind, *J. Atmos. Sol. Phys*., 204, 105271, https://doi.org/10.1016/j.jastp.2020.105271.

Boyd, A. J., Spence, H. E., Huang, C.-L., Reeves, G. D., Baker, D. N., Turner, D. L., Claudepierre, S. G., Fennell, J. F., Blake, J. B., and Shprits, Y. Y. (2016), Statistical properties of the radiation belt seed population, *J. Geophys. Res. Space Physics*, 121, 7636– 7646, doi:10.1002/2016JA022652.

Brautigam, D. H., and J. T. Bell (1995), CRRESELE documentation, Rep. PL-TR-95-2128, Phillips Lab., Hanscom AFB, Mass.

Cai, L., Ma, S., Cai, H. *et al.* (2009), Prediction of SYM-H index by NARX neural network from IMF and solar wind data. *Sci. China Ser. E-Technol. Sci.* **52,** 2877–2885, https://doi.org/10.1007/s11431-009-0296-9

Chandorkar, M., Camporeale, E., and Wing, S. (2017), Probabilistic forecasting of the disturbance storm time index: An autoregressive Gaussian process approach, *Space Weather*, 15, 1004– 1019, doi: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, 438– 454. https://doi.org/10.1029/2018SW002095

Claudepierre S. G., & O'Brien, T. P. (2020). Specifying high-altitude electrons using low-altitude LEO systems: The SHELLS model. *Space Weather*, 18, 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., and Sugiura, M. (1966), Auroral electrojet activity index *AE* and its universal time variations, *J. Geophys. Res.*, 71(3), 785– 801, doi:10.1029/JZ071i003p00785.

Dessler, A., and E. Parker (1959), Hydromagnetic theory of geomagnetic storms, *J. Geophys. Res.*, **64**(12), 2239–2252, doi:10.1029/JZ064i012p02239.

Ginet, G.P., O'Brien, T.P., Huston, S.L. *et al.* (2013), AE9, AP9 and SPM: New Models for Specifying the Trapped Energetic Particle and Space Plasma Environment. *Space Sci Rev* **179,** 579–615, https://doi.org/10.1007/s11214-013-9964-y

Green, J. C., and M. G. Kivelson (2004), Relativistic electrons in the outer radiation belt: Differentiating between acceleration mechanisms, J. Geophys. Res., 109, A03213, doi:10.1029/2003JA010153.

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

Iyemori, T. (1990), Storm-time magnetospheric currents inferred from mid-latitude geomagnetic field variations, J. Geomag. Geoelectr., 42, 1249–1265.

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., and Y. Y. Shprits (2012), On the influence of solar wind conditions on the outer-electron radiation belt, J. Geophys. Res., 117, A05217, doi:10.1029/2011JA017253.

Kellerman, A. C., Y. Y. Shprits, and D. L. Turner (2013), A Geosynchronous Radiation-belt Electron Empirical Prediction (GREEP) model, SpaceWeather, 11, doi:10.1002/swe.20074.

Koons, H. C., and Gorney, D. J. (1991), A neural network model of the relativistic electron flux at geosynchronous orbit, *J. Geophys. Res.*, 96(A4), 5549– 5556, doi:10.1029/90JA02380.

Li, W., Thorne, R. M., Angelopoulos, V., Bonnell, J. W., McFadden, J. P., Carlson, C. W., LeContel, O., Roux, A., Glassmeier, K. H., and Auster, H. U. (2009), Evaluation of whistler-mode chorus intensification on the nightside during an injection event observed on the THEMIS spacecraft, *J. Geophys. Res.*, 114, A00C14, doi:10.1029/2008JA013554.

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

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

Li, X., K. S. Oh, and M. Temerin (2007), Prediction of the AL index using solar wind parameters, J. Geophys. Res., 112, A06224, doi:10.1029/2006JA011918.

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

Luo, B., Li, X., Temerin, M., and Liu, S. (2013), Prediction of the *AU*, *AL*, and *AE* indices using solar wind parameters, *J. Geophys. Res. Space Physics*, 118, 7683– 7694, doi:10.1002/2013JA019188.

Lyatsky, W., and G. V. Khazanov (2008a), Effect of solar wind density on relativistic electrons at geosynchronous orbit, Geophys. Res. Lett., 35, L03109, doi:10.1029/2007GL032524.

Lyatsky, W., and G. V. Khazanov (2008b), Effect of geomagnetic disturbances and solar wind density on relativistic electrons at geostationary orbit, J. Geophys. Res., 113, A08224, doi:10.1029/2008JA013048.

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, 11,109– 11,127. https://doi.org/10.1002/2016JA023793

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

Meredith, N. P., Horne, R. B., and Anderson, R. R. (2001), Substorm dependence of chorus amplitudes: Implications for the acceleration of electrons to relativistic energies, *J. Geophys. Res.*, 106(A7), 13165– 13178, doi:10.1029/2000JA900156.

Newell, P. T., S. Wing, C. I. Meng, and V. Sigillito (1990), A neural network based system for monitoring the aurora, Johns Hopkins APL Tech. Dig., 11 (3 and 4), 291-299.

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

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

Paulikas, G. A., and J. B. Blake (1979), Effects of the solar wind on magnetospheric dynamics: Energetic electrons at the synchronous orbit, in *Quantitative Modeling of Magnetospheric Processes*, *Geophys. Monogr. Ser*., Vol 21, pp. 180-202, AGU, Washington D.C.

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, 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, e2019SW002399. https://doi.org/10.1029/2019SW002399

Reeves, G. D. (1998), Relativistic electrons and magnetic storms: 1992–1995, Geophys. Res. Lett., 25, 1817–1820, doi:10.1029/98GL01398.

Reeves, G. D. (2007), Radiation Belt Storm Probes: A New Mission for Space Weather Forecasting. Space Weather, 5: n/a. doi:10.1029/2007SW000341

Reeves, G. D., S. K. Morley, R. H. W. Friedel, M. G. Henderson, T. E. Cayton, G. Cunningham, J.

B. Blake, R. A. Christensen, and D. Thomsen (2011), On the relationship between relativistic electron flux and solar wind velocity: Paulikas and Blake revisited, J. Geophys. Res., 116, A02213, doi:10.1029/2010JA015735.

Reeves, G. D., Chen, Y., Cunningham, G. S., Friedel, R. W. H., Henderson, M. G., Jordanova, V. K., Koller, J., Morley, S. K., Thomsen, M. F., and Zaharia, S. (2012), Dynamic Radiation Environment Assimilation Model: DREAM, *Space Weather*, 10, S03006, doi:10.1029/2011SW000729

Reeves, G., S. Morley, and G. Cunningham (2013), Long-term variations in solar wind velocity and radiation belt electrons, J. Geophys. Res. Space Physics, 118, 1040–1048, doi:10.1002/jgra.50126.

Rigler, E. J., M. Wiltberger, and D. N. Baker (2007), Radiation belt electrons respond to multiple solar wind inputs, J. Geophys. Res., 112, A06208, doi:10.1029/2006JA012181

Roederer, J. G. (1970), Dynamics of Geomagnetically Trapped Radiation, Springer, New York, https://doi.org/ 10.1007/978-3-642-49300-3

Rumelhart, D. E., and J. L. McClelland (Eds.) (1987), Parallel Distributed Processing, Vol. I, MIT Press, Cambridge, MA.

Schulz, M., and L.J. Lanzerotti (1974) , Particle Diffusion in the Radiation Belts, Springer, Berlin, 1974, https:// doi.org/10.1007/978-3-642-65675-0

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, 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, e2019JA027357. https://doi.org/10.1029/2019JA027357

Shprits, Y. Y., D. Subbotin, and B. Ni (2009), Evolution of electron fluxes in the outer radiation belt computed with the VERB code, J. Geophys. Res., 114, A11209, doi:10.1029/2008JA013784.

Schmidhuber, J. (2015), Deep Learning in Neural Networks: An Overview, *Neural Networks*, **61**, 85–117, arXiv:1404.7828. doi:10.1016/j.neunet.2014.09.003

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, e2020SW002532. https://doi.org/10.1029/2020SW002532

Spence, H.E., Reeves, G.D., Baker, D.N. 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 Sci Rev 179, 311–336, https://doi.org/10.1007/s11214-013-0007-5

Tang, C. L., Wang, Y. X., Ni, B., Zhang, J.-C., Reeves, G. D., Su, Z. P., Baker, D. N., Spence, H. E., Funsten, H. O., and Blake, J. B. (2017a), Radiation belt seed population and its association with the relativistic electron dynamics: A statistical study, *J. Geophys. Res. Space Physics*, 122, 5261– 5276, doi:10.1002/2017JA023905.

Tang, C. L., Wang, Y. X., Ni, B., Su, Z. P., Reeves, G. D., Zhang, J.-C., Baker, D. N., Spence, H. E., Funsten, H. O., and Blake, J. B. (2017b), The effects of magnetospheric processes on relativistic electron dynamics in the Earth's outer radiation belt, *J. Geophys. Res. Space Physics*, 122, 9952– 9968, doi:10.1002/2017JA024407.

Temerin, M., and Li, X. (2006), *Dst* model for 1995–2002, *J. Geophys. Res.*, 111, A04221, doi:10.1029/2005JA011257

Tsyganenko, N. A., and Sitnov, M. I. (2005), Modeling the dynamics of the inner magnetosphere during strong geomagnetic storms, *J. Geophys. Res.*, 110, A03208, doi:10.1029/2004JA010798.

Turner, D. L., Y. Shprits, M. Hartinger, and V. Angelopoulos (2012), Explaining sudden losses of outer radiation belt electrons during geomagnetic storms, Nat. Phys., 8, 208–212, doi:10.1038/nphys2185.

Turner, D. L., et al. (2014a), On the cause and extent of outer radiation belt losses during the 30 September 2012 dropout event, *J. Geophys. Res. Space Physics*, 119, 1530– 1540, doi:10.1002/2013JA019446.

Turner, D. L., et al. (2014b), 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, *J. Geophys. Res. Space Physics*, 119, 1960– 1979, doi:10.1002/2014JA019770.

Ukhorskiy, A. Y., M. I. Sitnov, A. S. Sharma, B. J. Anderson, S. Ohtani, and A. T. Y. Lui (2004), Data-derived forecasting model for relativistic electron intensity at geosynchronous orbit,Geophys. Res. Lett., 31, L09806, doi:10.1029/2004GL019616.

Ukhorskiy, A. Y., B. J. Anderson, P. C. Brandt, and N. A. Tsyganenko (2006), Storm time evolution of the outer radiation belt: Transport and losses, J. Geophys. Res., 111, A11S03, doi:10.1029/2006JA011690.

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

Wei, H.-L., Billings, S. A., Surjalal Sharma, A., Wing, S., Boynton, R. J., and Walker, S. N. (2011), Forecasting relativistic electron flux using dynamic multiple regression models, Ann. Geophys., 29, 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.

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

Wing, S., Johnson, J. R., Jen, J., Meng, C.-I., Sibeck, D. G., Bechtold, K., Freeman, J., Costello, K., Balikhin, M., and Takahashi, K. (2005), Kp forecast models, *J. Geophys. Res.*, 110, A04203, doi:10.1029/2004JA010500.

Wing, S., J. R. Johnson, E. Camporeale, and G. D. Reeves (2016), Information theoretical approach to discovering solar wind drivers of the outer radiation belt, J. Geophys. Res. Space Physics, 121, 9378–9399, doi:10.1002/2016JA022711

Wing, S, J. R. Johnson (2019), Applications of Information Theory in Solar and Space Physics, *Entropy*, 21(2):140, https://doi.org/10.3390/e21020140

Wing, S., J. R. Johnson, D. L. Turner, A. Y. Ukhorskiy, and A. J. Boyd (2021), Untangling the solar wind and magnetospheric drivers of the radiation belt electrons, under review, J. Geophys. Res.

Wintoft, P., Wik, M., Matzka, J., and Shprits, Y. (2017). Forecasting Kp from Solar Wind Data: Input Parameter Study Using 3-hour Averages and 3-hour Range Values. *J. Space Weather Space Clim.* 7, A29. doi:10.1051/swsc/2017027

Wu, J., and H. Lundstedt (1997), Geomagnetic storm predictions from solar wind data with the use of dynamic neural networks, J. Geophys. Res., 102, 14,255-14,268.

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, 9858– 9879. https://doi.org/10.1002/2017JA024487

Zhao, H., Baker, D. N., Jaynes, A. N., Li, X., Elkington, S. R., Kanekal, S. G., Spence, H. E., Boyd, A. J., Huang, C.-L., and Forsyth, C. (2017), On the relation between radiation belt electrons and solar wind parameters/geomagnetic indices: Dependence on the first adiabatic invariant and L^*, *J. Geophys. Res. Space Physics*, 122, 1624– 1642, doi:10.1002/2016JA023658.

Input and output parameters of the model

	Input parameters	Output parameter
1	Vsw(t) to Vsw(t–72 hr)	PSD (t)
2	nsw(t) to nsw(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)	

Table 1. Input and output parameters of the model. Vsw = solar wind velocity. nsw = solar wind density. IMF (By, Bz) = GSM y and z component of the interplanetary magnetic field, respectively.

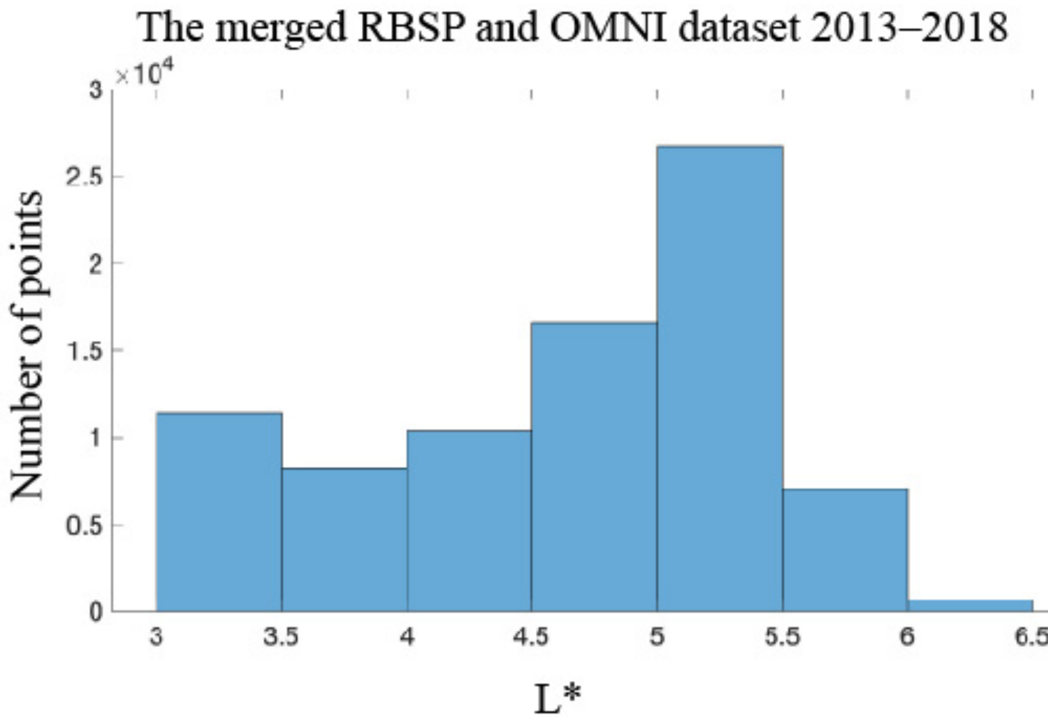

Figure 1. The distribution of the merged RBSP and OMNI dataset 2013–2018.

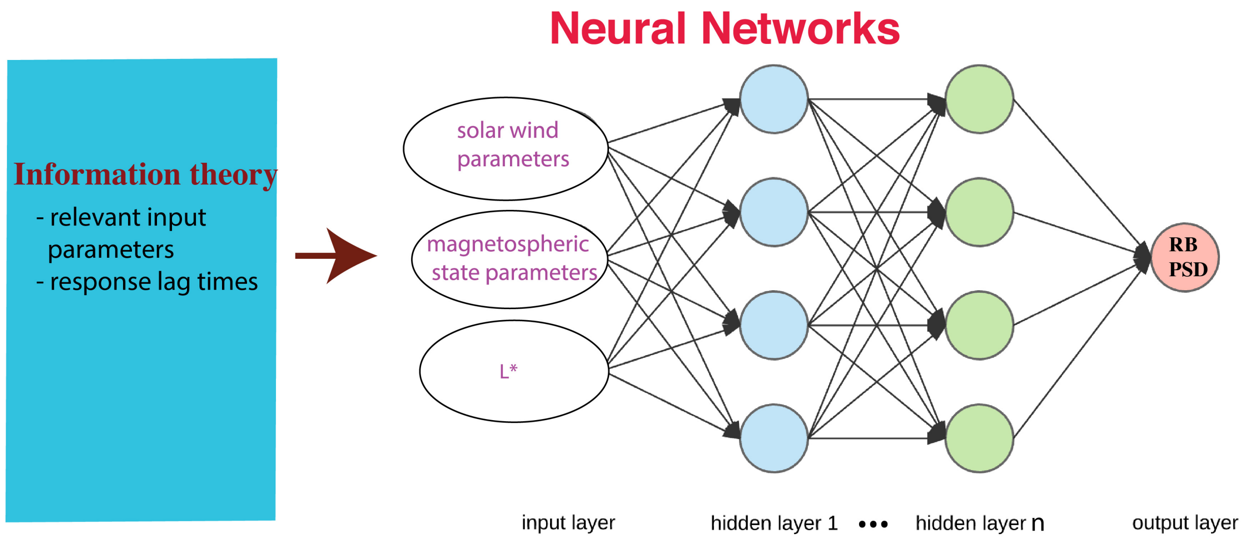

Figure 2. Schematic of the model that combines information theory and neural network. The neural network inputs the solar wind and magnetospheric parameters and L*; and outputs PSD (see Table 1). Information theory is used to select and rank solar wind and magnetospheric parameters and their lag times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

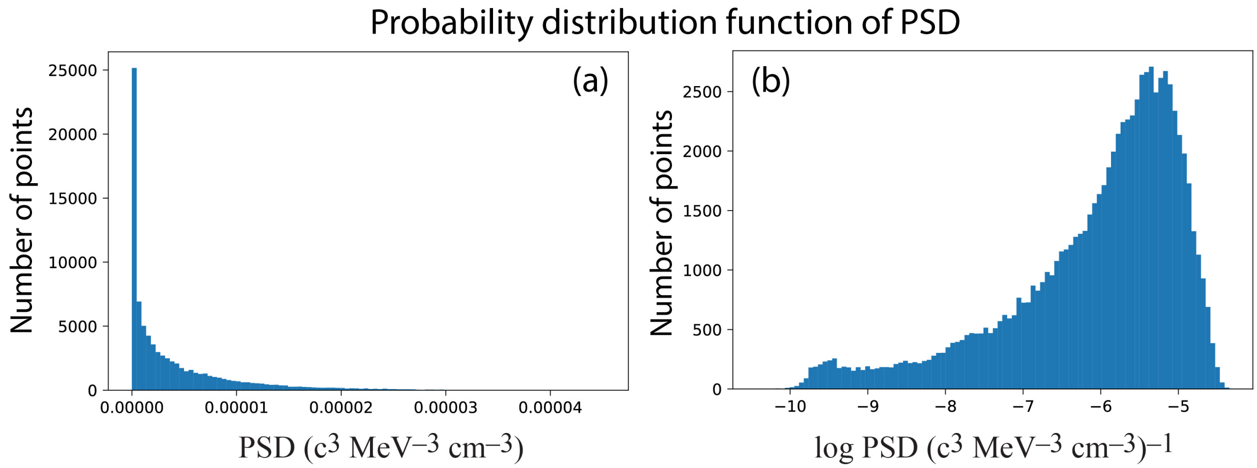

Figure 3. The distribution of PSD (a) and log PSD (b).

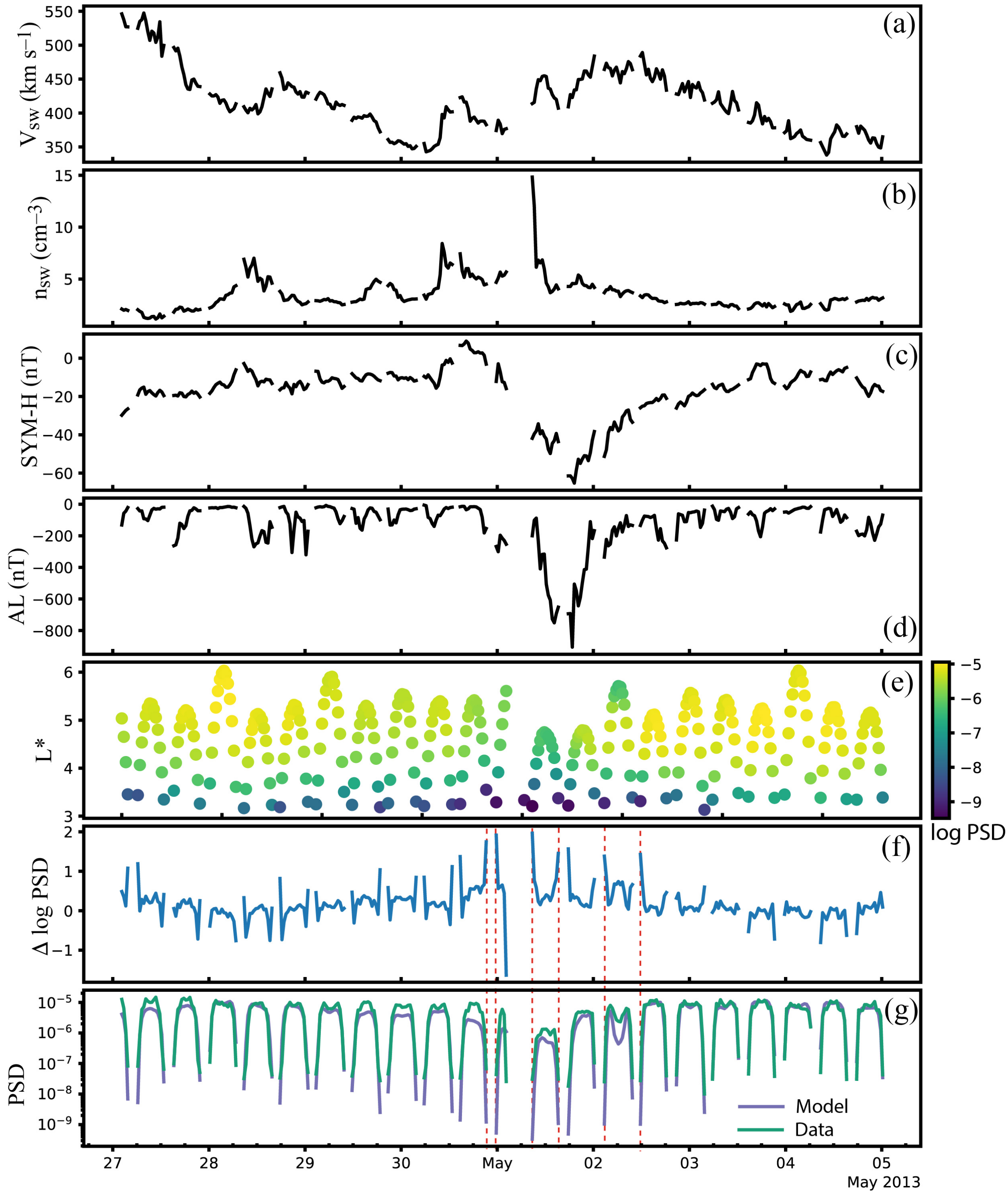

Figure 4. Solar wind velocity (a), solar wind density (b), SYM-H (c), AL (d), L* and log model PSD (e), Δ log PSD = log(observed PSD) – log(model PSD) (f), and observed (green curve) and model PSD (blue curve) (g) for 2013 April 27 – May 5, which is the first half of the first event. The unit for PSD and ΔPSD 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.

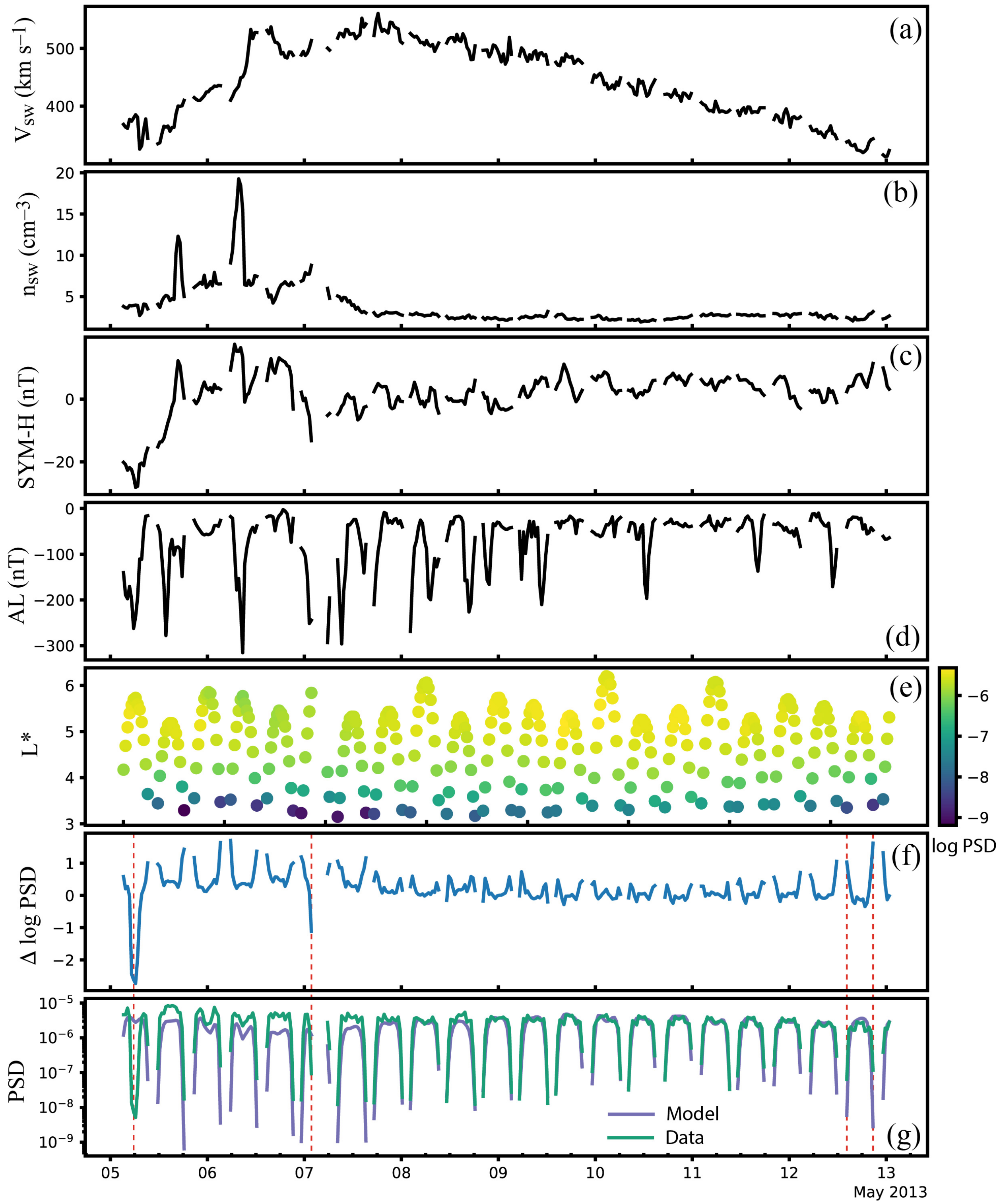

Figure 5. Solar wind velocity (a), solar wind density (b), SYM-H (c), AL (d), L* and log model PSD (e), Δ log PSD = log(observed PSD) – log(model PSD) (f), and observed (green curve) and model PSD (blue curve) (g) for 2013 May 5 – 13, which is the second half of the first event. The unit for PSD and ΔPSD 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.

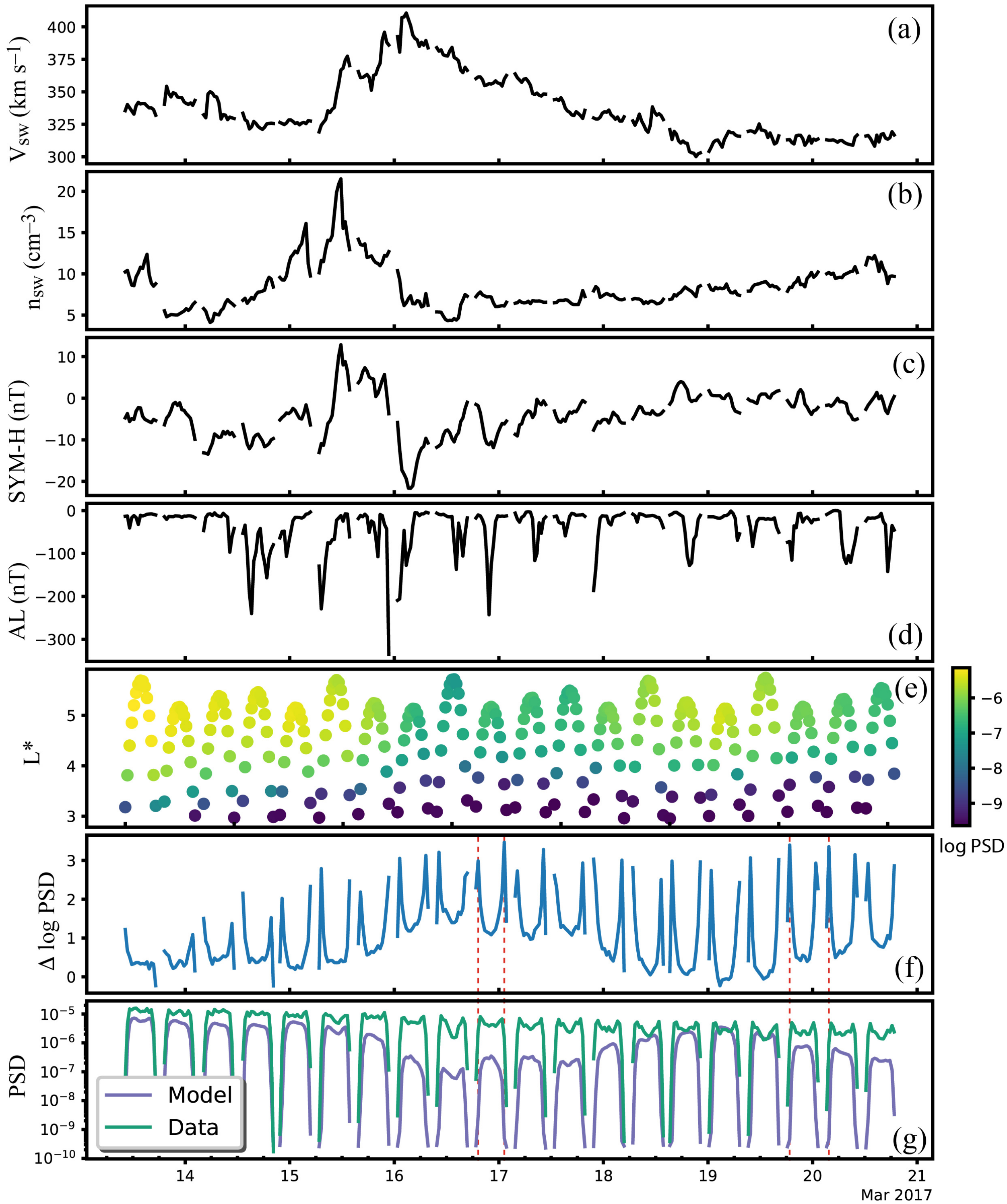

Figure 6. Solar wind velocity (a), solar wind density (b), SYM-H (c), AL (d), L* and log model PSD (e), Δ log PSD = log(observed PSD) – log(model PSD) (f), and observed (green curve) and model PSD (blue curve) (g) for 2017 Mar 13 – 21, which is the first half of the second event. The unit for PSD and ΔPSD 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.

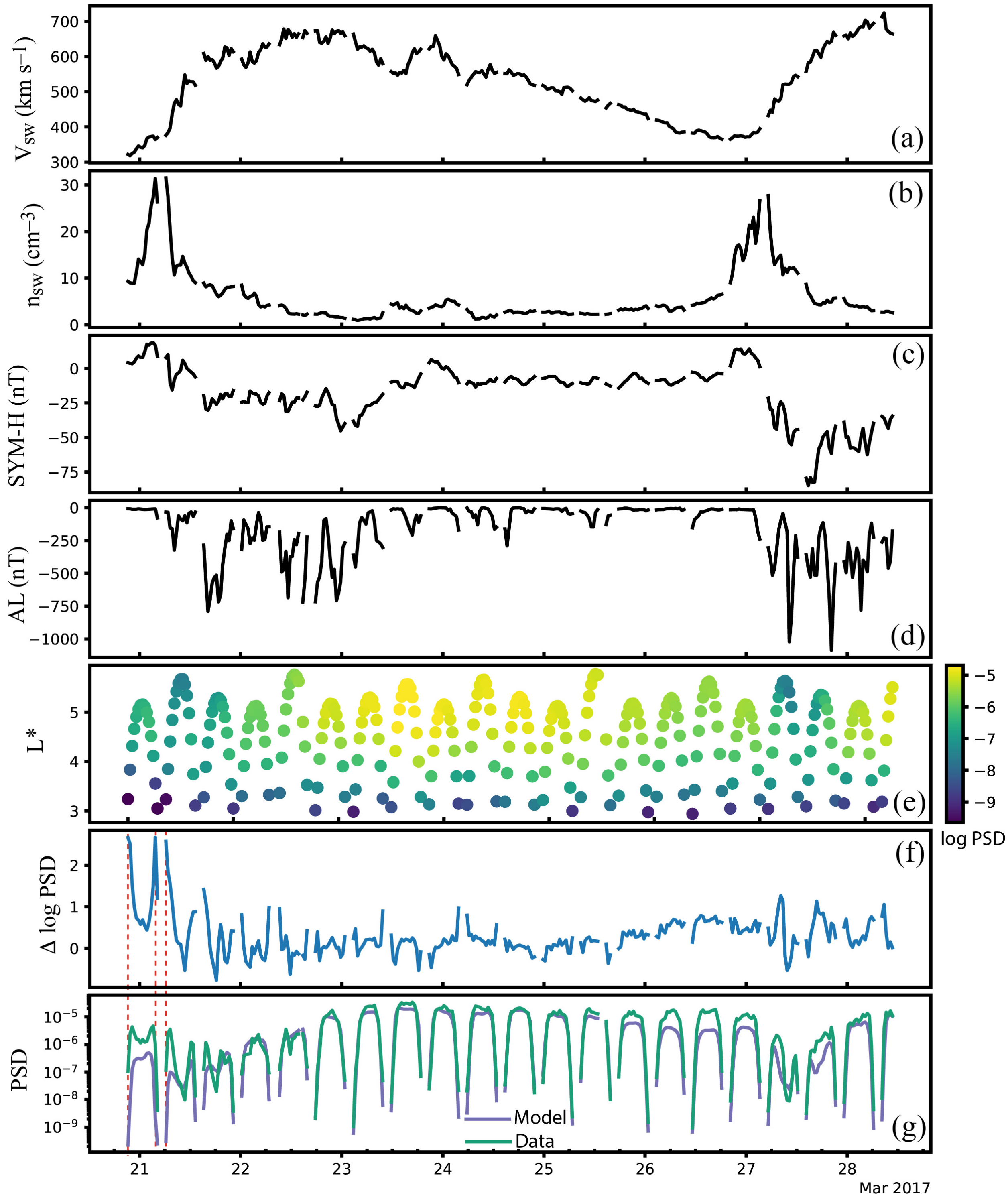

Figure 7. Solar wind velocity (a), solar wind density (b), SYM-H (c), AL (d), L* and log model PSD (e), Δ log PSD = log(observed PSD) – log(model PSD) (f), and observed (green curve) and model PSD (blue curve) (g) for 2017 Mar 21 – 29, which is the second half of the second event. The unit for PSD and ΔPSD 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.

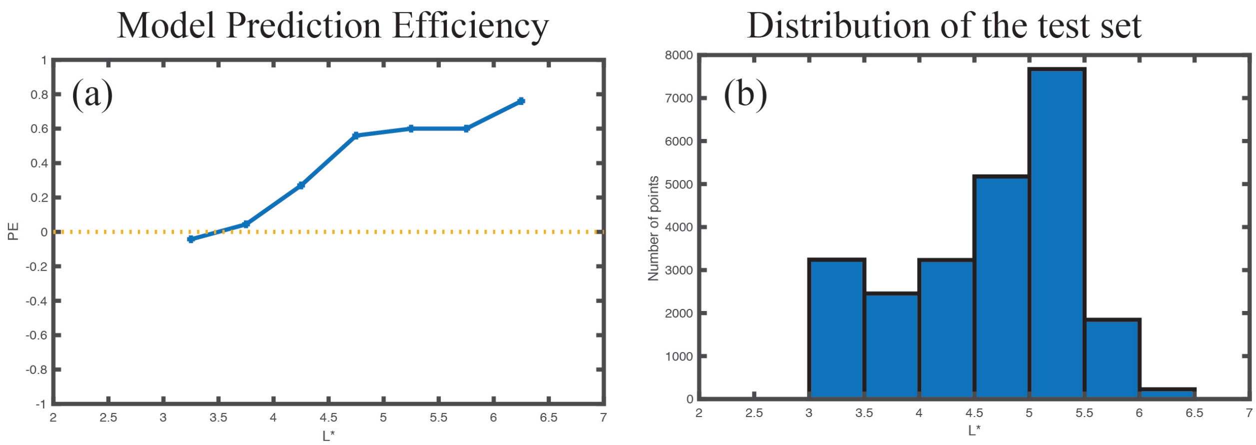

Figure 8. The model prediction efficiency (PE) of the test set as a function of L* (a). The PE is lower at L* < 4 or 4.5 because solar wind and magnetospheric parameters transfer little information to PSD at these L*. The distribution of the test set (b).