

Application of Recurrent Neural Network to Modeling Earth's Global Electron Density

Sheng Huang¹, Wen Li¹, Xiao-Chen Shen¹, Qianli Ma^{1,2}, Xiangning Chu³, Donglai Ma², Jacob Bortnik², Luisa Capannolo¹, Yukitoshi Nishimura¹, and Jerry Goldstein^{4,5}

¹Center for Space Physics, Boston University, Boston, MA, USA.

²Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA, USA.

³Laboratory for Atmospheric and Space Physics, University of Colorado Boulder, Boulder, CO, USA.

⁴Space Science and Engineering Division, Southwest Research Institute, San Antonio, TX, USA

⁵Department of Physics and Astronomy, University of Texas, San Antonio, TX, USA

Corresponding author: Sheng Huang (hs2015@bu.edu); Wen Li (wenli77@bu.edu)

Key Points:

- An encoder-decoder model (recurrent neural network) is developed to reconstruct Earth's global electron density.
- By separating geomagnetic indices from satellite location as inputs, we directly model total electron density with geomagnetic indices.
- The model captures the spatial variation of density by predicting stable and evident plumes, roughly consistent with global observations.

Abstract

The total electron density is a fundamental quantity in the Earth's magnetosphere and plays an important role in a number of physical processes, but its dynamic global evolution is not fully quantified yet. We present an implementation of a specific type of recurrent neural network (encoder-decoder), which is distinct from previous models, to construct global electron density based on the multi-year data from Van Allen Probes. The history of geomagnetic indices is first encoded into a hidden state $\mathbf{H}$, then together with auxiliary information (satellite location), they are decoded into the quantity of interest (total electron density in this study). In this process the input of historical geomagnetic indices is detangled from the satellite location and is processed chronologically by the encoder. As a result, time evolution of geomagnetic indices is explicitly embedded in the structure and the encoded hidden state $\mathbf{H}$ can be viewed as the representation of the inner magnetospheric state. The magnetospheric state is then decoded to predict global electron density evolution. Our results show that the model can capture the dynamical evolution of total electron density with the formation and evolution of stable and evident plume configurations that roughly agree with global observations. Our findings demonstrate the importance of applying recurrent neural networks to specify the inner magnetospheric state in a novel way, which will potentially improve our fundamental understanding of wave and particle dynamics in the Earth's magnetosphere.

Plain Language Summary

The global evolution of total electron density in the Earth's magnetosphere is important for understanding energetic particle dynamics and their associated wave-particle interactions. Although physics-based models show reasonable evolution, they do not fully capture all the dynamical evolution on a global scale. Therefore, data-driven methods by utilizing machine learning techniques have been developed recently. The recurrent neural network is a variant of artificial neural networks, and it has been applied to many time-series data. It naturally handles the time dependence by processing the data chronologically. We develop an encoder-decoder model based on recurrent neural network and show that by properly designing the architecture, the model can separate the spacecraft orbital variation and the temporal evolution. We apply it to global total electron density modeling trained on the Van Allen Probes observation, and the result shows fairly good performance with stable and evident spatial structures of the plasmasphere and plumes.

1 Introduction

The plasmasphere is a region in the near-Earth space, extending up to ~ 7 Earth radii (R_E), filled with cold (0.1-10 eV) and dense (up to 10^4 cm^{-3}) plasma. It is populated mostly by the outflow of ions and electrons from the ionosphere and typically has an irregular toroidal shape around the Earth (Carpenter & Anderson, 1992) with an exponentially decreasing density profile as the radial distance increases. Previous studies have suggested that the shape of the plasmasphere is correlated with geomagnetic activity (Carpenter & Seely, 1976; Goldstein, 2007), and the global distribution of cold plasma can become complex during geomagnetically disturbed periods (Carpenter & Lemaire, 1997; Darrouzet et al., 2009) and highly structured during quiet periods (Goldstein & Sandel, 2005; Moldwin et al., 1994, 1995, 2003). During

geomagnetic storms, the outer plasmasphere is often eroded due to an enhanced convection electric field, forming a region of sharp gradient called “plasmopause”. The erosion starts from the nightside and propagates to the dayside (Gallagher & Adrian, 2007; Goldstein & Sandel, 2005), where the plasma is convected further away from the plasmasphere and transported sunward toward the magnetopause via a high-density structure, called “plasmaspheric plume”. Plasmaspheric plumes at the dayside magnetopause may modify the reconnection rate there (McFadden et al., 2008; Walsh et al., 2014). As the convection subsides, the corotation dominates and the plasmasphere gradually recovers through plasma refilling from the ionosphere during the recovery phase (Dent et al., 2006; Reinisch et al., 2004).

Quantification of the total electron density in the Earth’s magnetosphere is critical, as it affects many processes, including populating the plasmasphere and the plasmasheet (Elphic et al., 1997), affecting magnetopause stability (Borovsky & Denton, 2006; Zou et al., 2021), plasma wave generation and propagation (Chen et al., 2012; Li et al., 2013; Shi et al., 2019), and the loss and acceleration of energetic particles (Bortnik et al., 2007; Li et al., 2019; Liu et al., 2016; Malaspina et al., 2020; Thorne et al., 2016). Therefore, it is important to construct an accurate model of global electron density. Many empirical (Reinisch et al., 2009; Sheeley et al., 2001) and physics-based models (De Pascuale et al., 2018; Pierrard & Lemaire, 2004; Pierrard et al., 2009) have been developed. However, due to the lack of direct observation on a global scale, these models are either simplified with a limited number of parameters, thus lack complex spatio-temporal variation, or miss some key processes (e.g., the complex electric field evolution). By taking full advantage of long-term satellite observations, we aim to model the global electron density using a machine learning technique.

Machine learning is a field with a long and rich history (Bishop, 2006; Cortes & Vapnik, 1995; Van Otterlo & Wiering, 2012). It can be loosely defined as to enable the computers to learn from data without being explicitly programmed (Samuel, 1959). Deep learning (LeCun et al., 2015) as a subfield of machine learning benefits from the rapidly growing computing power and has gained its popularity after AlexNet (Krizhevsky et al., 2017) won several image classification contests in 2012. Over the recent years, there has been an explosive growth of deep learning in theory (E et al., 2017; Han et al., 2018; He & Tao, 2021), model structures (He et al., 2018; Liu et al., 2021; Vaswani et al., 2017; Zhou et al., 2021) and applications (Liu & Tegmark, 2021; Sirignano & Spiliopoulos, 2018). In the field of space physics, the application of machine learning has started in 1990s, when the early form of neural network is used to provide relativistic electron fluxes at geosynchronous orbit (Koons and Gorney, 1991), geomagnetic storm forecasts (Wu and Lundstedt, 1997), and Kp forecasts (Wing et al., 2005). More recently, deep learning has been applied to solar flare prediction (Nishizuka et al., 2021; Wang et al., 2020), plasma region classification (Breuillard et al., 2020), electron flux modeling (Chu et al., 2021; Ma et al., 2021; McGranaghan et al., 2021) and so on. The idea of modeling total electron density with deep learning is not new and previous studies have demonstrated the ability of deep learning to capture the plasmaspheric response to geomagnetic activity (Bortnik et al., 2016; Chu et al., 2017a, 2017b; Zhelavskaya et al., 2017). However, all of the previous methods use fully connected neural networks to model the relation, which is not necessarily the most natural way to handle time order information. Because the geomagnetic indices used to model the relation are time-series and all satellite observations are partly time-series (also contain the location dependence), we view the problem as time-series nowcasting/forecasting. Compared with traditional modeling methods including autoregressive moving average models (ARMA; Chen et al., 1995) and further variant with exogenous variables included (ARMAX; Huang et al., 2005),

deep learning models show certain advantages in modeling complex non-linear relation without prior knowledge on the time-series itself and outperform traditional methods in many applications (Lara-Benítez et al., 2021). In the present study, we apply an encoder-decoder architecture based on recurrent neural network to model the total electron density with explicit time evolution in the model structure.

Description of the data and the modeling method is provided in Sections 2 and 3. In Section 4, we examine the overall model performance and apply the model to two observations of plumes. Finally, we discuss our current understanding of similar modeling tasks and summarize the findings in Sections 5 and 6.

2 Data

In this study, we use the in-situ total electron density data from Van Allen Probes (Mauk et al., 2013), also known as Radiation Belt Storm Probes (RBSP), covering from January 2013 to October 2019. Van Allen Probes are twin satellites (RBSP-A and B) in a highly elliptical orbit in the Earth's inner magnetosphere ($< \sim 5.8 R_E$). We use the satellite location of L shell, magnetic local time (MLT) and magnetic latitude (MLAT), and the total electron density (N_e) inferred from upper hybrid resonance frequency (Kurth et al., 2015) measured by the High Frequency Receiver (HFR) onboard the Electric and Magnetic Field Instrument Suite and Integrated Science (EMFISIS) instrument (Kletzing et al., 2013). The MLT is converted into $\sin(\text{MLT}/12\pi)$ and $\cos(\text{MLT}/12\pi)$ as model input parameters. The time resolution of satellite data is processed to be of one minute by simple linear interpolation.

In this study, geomagnetic indices are used to represent the magnetospheric state. The solar wind parameters are not used since their time coverage is not continuous with intermittent gaps. Here we use SML, Hp30, and SYM-H indices as they capture the disturbance of the geomagnetic field in auroral zone, mid-latitude, and low-latitude, respectively. Thus, they are good indicators of the overall inner magnetospheric activity. The SML index (Gjerloev, 2012; Newell & Gjerloev, 2011), obtained from SuperMAG Web Service, is similar to the widely used AL index but provides full-time coverage during the Van Allen Probe era. The Hp30 index (Matzka et al., 2021) is from GFZ German Research Centre for Geosciences and is used to replace the 3-hour Kp index, providing higher temporal resolution of 30 min. These indices are interpolated at 1-min time resolution to match with the satellite data.

3 Methodology

3.1 Model Structure

We use a well-established architecture called encoder-decoder to model the global electron density. The architecture differs from the conventional models composed of several fully connected layers applied in previous studies (Bortnik et al., 2016; Chu et al., 2017a; Zhelavskaya et al., 2017). Encoder-decoder was first proposed to deal with natural language processing tasks (Cho et al., 2014) and was later extended to general sequence-to-sequence modeling (Sutskever et al., 2014). Since geomagnetic indices are time-series and satellite observations also vary with time, we adapt the architecture to our project, as illustrated in Figure 1a. The history of geomagnetic indices is passed to an encoder to obtain the magnetospheric state. Then, this encoded state and auxiliary satellite information (satellite location) are passed to

the decoder to obtain the quantity of interest Y (total electron density in our case). The key component of the encoder-decoder structure is an architecture called Long Short-Term Memory (LSTM; Hochreiter & Schmidhuber, 1997), which is a modified version of Recurrent Neural Networks (RNN; Elman, 1990). RNN is first introduced as a type of Artificial Neural Networks (ANN) to deal with time-series data, where the output of the hidden layer is used as the input of next time step. As a result, RNN is able to model the temporal dependence of the data and was later improved with many variants to deal with longer time pattern dependence (Cho et al., 2014; Gers et al., 1999; Huang et al., 2015). Among all these variants, LSTM is the most popular one and appears to perform well in many time-series modeling tasks (Lara-Benítez et al., 2021). By introducing different gates, LSTM can mitigate some limitations of vanilla RNN (e.g., the gradient vanishing problem), thus is able to sustain deeper architectures and capture longer time dependence. The detailed description and discussion of LSTM can be found in Staudemeyer & Morris (2019) and references therein.

To explain the workflow in detail, in the encoder (green block in Figure 1a), at any given time t (time is denoted in the superscript hereafter), the model output of each LSTM cell (Figure 1b) can be described as $\mathbf{H}^t = f(\mathbf{H}^{t-1}, \mathbf{X}^t)$, where f denotes the processes by the LSTM cell, $\mathbf{H}^t$ is the LSTM output (also called hidden state) at the current time t . $\mathbf{H}^{t-1}$ is the LSTM output from the previous time step (1 min in this study to match the resolution of the satellite measured density), which is also used as the input at the current time step. $\mathbf{X}^t$ represents the geomagnetic indices at the current time. By iteratively updating hidden state $\mathbf{H}^t$ with its last state $\mathbf{H}^{t-1}$ and current geomagnetic indices $\mathbf{X}^t$ over time, the history of geomagnetic indices is encoded into $\mathbf{H}$, which is one of the most intuitive ways to handle the time evolution. In the decoder (orange block in Figure 1a), apart from updating the hidden state with the current geomagnetic indices, we concatenate the hidden state $\mathbf{H}$ with location $\mathbf{Loc}$ ($L, \sin(\text{MLT}/12\pi), \cos(\text{MLT}/12\pi), \text{MLAT}$) and pass them to a two-layer fully connected neural network, with the final output to be the total electron density (N_e). This process is consistent with the previous models (e.g., Bortnik et al., 2016; Chu et al., 2017a, 2017b; Zhelavskaya et al., 2017).

The main difference between our model and previous models is in treating the time sequence of geomagnetic indices and satellite location. Specifically, in the fully connected neural network, the information from geomagnetic indices and satellite location are entangled (they are concatenated together before fed to the model). Moreover, in fully connected neural networks, the “corresponding” hidden state of magnetospheric state $\mathbf{H}$ at different times is derived separately from a sequence of historical geomagnetic indices without an explicit time-dependence. As a result, the neuron outputs of different times do not have a clear dependence on each other. In our encoder-decoder model (specifically in the encoder section), the hidden state $\mathbf{H}^t$ is updated from the last state $\mathbf{H}^{t-1}$ together with geomagnetic indices at time t (Figure 1b), thus it explicitly depends on the previous hidden state (at time $t-1$) and can ultimately be viewed as a representation of the magnetospheric state. In addition, the dependence of $\mathbf{H}^t$ on $\mathbf{H}^{t-1}$ leads to a more continuous $\mathbf{H}$ over time (possibly producing a smoothly varying total electron density evolution) and ensures to explicitly take into account the time evolution of the system.

3.2 Data Processing

To process data, we first divide the RBSP observations into inbound and outbound orbits. Since each bound consists of ~ 270 minutes of observation, we set the length of the decoder to be

270, fill it with NaN if the bound is shorter, and truncate it if the bound is longer. Geomagnetic indices, satellite location, and N_e in the decoder are thus 270 minutes long, with a 1-min time resolution. The dynamics of the plasmasphere occur on timescales that are much longer than a minute, so 1-min resolution provides sufficient resolution to capture any variations in the large-scale density distribution. Note that the same resolution was used by Zhelavskaya et al. (2017). The 270-min length is chosen because it has complete coverage from low/high L shell to high/low L shell to provide sufficient variability in total electron density for the model to train on. In the encoder, we set the geomagnetic indices to be 200-hour long with a 10-min time resolution after the hyperparameter tuning described in section 3.3. In total, we end up with $\sim 26,000$ inbounds and outbounds. To divide them into different datasets (for training, validation, and testing), we first assign calendar dates into 7-day blocks and randomly pick 70% of the blocks as the training set, 20% as the validation set, and 10% as the test set. We then match the RBSP-A and RBSP-B sequence to the split time blocks. Note that 7 days is longer than the plasmaspheric system "memory time", so by doing this we take into account long-term seasonal and solar cycle variation while avoiding data leakage.

To summarize, the model takes as input 200 hours of geomagnetic indices in the encoder and 270 minutes of geomagnetic indices and satellite location (L , $\sin(\text{MLT}/12\pi)$, $\cos(\text{MLT}/12\pi)$, MLAT) in the decoder. The model outputs with 270 minutes of the predicted total electron density. Note that the 200-hour encoder length and 270-min decoder length are only for the training process and are totally extendable.

3.3 Optimization

The model performance is evaluated on the validation set, based on the difference between $\log_{10} Ne$ and $\log_{10} Ne^P$, where P denotes the predicted electron density from our model. We choose the mean square error (MSE) as the loss function and r-square (r^2 , where r denotes the Pearson correlation coefficient) as the evaluation metrics.

We optimize the hyperparameters in our model using the steps described below.

(a) We choose one hyperparameter from the list: LSTM hidden size, fully connected layer size, encoder length and encoder time resolution;

(b) We set a range for the hyperparameter. Using Newton's iteration, we find the hyperparameter that results in the best model performance, and choose it as the optimized parameter;

(c) We loop through all hyperparameters, and the above steps (a-b) are performed for each hyperparameter iteratively;

(d) We repeat the above steps (a-c) until the loss does not improve by 0.002.

Using this method, our final hyperparameters are determined to be the following; (1) LSTM hidden size is 512; (2) geomagnetic indices in the encoder are 200-hour long with 10-minute time resolution; (3) fully connected network in the decoder has two layers of 16 neurons each, with $\tanh$ as the activation function.

4 Results

In this section, we evaluate the performance of the final model that we optimized. Figure 2 shows the comparison between satellite observation and model prediction on (a) training, (b) validation, and (c) test dataset. For each point on the plot, the x - and y -axis represent $\log_{10} Ne$ and $\log_{10} Ne^P$ respectively. A perfect model would be a diagonal line ($y=x$) marked with a white dashed line. For the test dataset, our model accuracy metrics are $MSE=0.114$ and $r^2=0.834$, which is close to those in the training ($MSE=0.089$ and $r^2=0.872$) and validation sets ($MSE=0.115$ and $r^2=0.841$), and is comparable to the previous studies (Bortnik et al., 2016; Chu et al., 2017a, 2017b; Zhelavskaya et al., 2017). However, we should point out that a direct one-to-one comparison between the performance of different models can be very challenging because the datasets for training were processed differently and were obtained from different satellites. Below we examine the global performance of our model during two events; both are out-of-sample datasets that were not included in the training process.

The first case occurred on 26 June 2000 during a modest geomagnetic storm (SYM-H down to ~ -80 nT with K_p up to 6) in association with a series of substorm activities (Figure 3a). The plume evolution was observed by global imaging (Goldstein et al., 2004) from the Imager for the Magnetopause-to-Aurora Global Exploration (IMAGE) mission, which was the first mission dedicated to imaging the Earth's magnetosphere (Burch, 2000). The Extreme Ultraviolet Imager (EUV) measured 30.4 nm light resonantly scattered from He^+ ions (Sandel et al., 2000) to estimate total plasma density. Figure 3 shows the comparison between the model prediction (panel b) and the EUV observations (panel c) at different times (1-4). Since the total density distribution from the EUV observations requires assumptions, we focus on the plasmaspheric shape which is identified with the largest He^+ density gradient (if any) of the EUV measurement. It is worth noting that the plasmopause identified in the EUV images (corresponding to the edge intersecting total electron density of $40 \pm 10 \text{ cm}^{-3}$) is a good proxy (within $0.25 R_E$) for the plasmopause in proton or electron density profiles (Goldstein et al., 2003; Harris et al., 1970; Horwitz et al., 1990). Statistically, the density profiles of protons and He^+ ions are morphologically similar (Goldstein et al., 2019b). The corresponding plasmopause of the modeled electron density is defined as the contour of 50 cm^{-3} (approximately consistent with the value of the edge in EUV images, noted above) as marked in red in Figure 3b, with a grey shadow covering a slightly broader area ($20\text{-}80 \text{ cm}^{-3}$) to denote the plume region. At the beginning of the day (time (1)), under quiet geomagnetic conditions, the plasmasphere was saturated with the plasmopause approaching the geosynchronous orbit ($L = 6.6$). After the main phase of the geomagnetic storm, at time (2), the plasmasphere showed a sunward surge where the cold plasma was dragged to the dusk sector due to enhanced convection electric field. During the storm recovery phase from time (3) to (4), the plume structure wrapped up as it corotated with the Earth, forming the final tail-like structure. During this process, the model captured the quick formation of the enhanced density region on the dusk side and the plume feature was stable over the next ten hours when the geomagnetic substorm activity was strong. Near the end of the day, the model predicted a gradually rotating plume as the convection subsided, as marked by the more quiet geomagnetic indices, which is consistent with the previous study (Goldstein et al., 2014b). We estimated the approximate rotation rate of the predicted plume from 1:30 to 5:00 UT to be 0.9 times of the Earth's rotation rate. During the entire period, the model showed complete plume evolution over 15 hours and the animation is included in Movie S1 and S2 in the Supporting Information.

The second event was on 7 September 2015 when a major substorm occurred with the peak SML < -2,000 nT, associated with a geomagnetic storm with SYM-H down to ~ -100 nT and Kp up to ~6 (Figure 4a). RBSP-A observed a long-lived plume during the period, while the Global Positioning System (GPS) also measured strong enhancements of total electron content (TEC) near the dusk sector. By measuring the delay of GPS signals traveling through the ionosphere, the TEC along the signal path can be estimated and serves as a reference for the total electron density. Figure 4b shows the one-to-one comparison between the RBSP-A observation (darker color) and the model prediction (lighter color), and the plume periods (light blue shadow) are identified with a density change by a factor of 5 within $\Delta L \leq 0.5$ at $L > 4$. Figure 4c shows the comparison between the modeled density and the RBSP-B observation. It is worth noting that RBSP-B did not observe a similar trend as RBSP-A did during this period, indicative of a very confined spatial extent of the plume. The global view on the equatorial plane (Figure 4d) presents the modeled plasmopause (red) and plume (grey shadow), and measured equatorial TEC (color-coded). The GPS TEC measurements were gridded to 2 x 6 deg cells and mapped to the equatorial plane using the T01 magnetic field model (Tsyganenko, 2002a, 2002b). Green lines represent simulated plasmopause from test particle simulations (PTP; Goldstein et al. (2005, 2014a, 2014b, 2019a)), which started at 00 UT on 7 September 7 2015 using an initial plasmopause specified by O'Brien and Moldwin (2003) and solved with 10s time steps driven by Kp and the OMNI solar wind electric field with the scaling factor $f = 0.2$. The PTP simulation is useful for showing the overall plasmaspheric response to solar activity (Goldstein et al., 2019a), including sunward surge, plume formation and rotation with the Earth, and is a good reference for the model to compare with on a large scale, as the reliable observation of the plasmopause location on a global scale is not always available. The RBSP-A trajectory (where the plume was identified shown in Figure 4b) is plotted with the cyan line with the satellite location marked with star in Figure 4d. Starting from 16:00 UT, the substorm reached its peak activity as shown in Figure 4a, and RBSP-A encountered the enhanced density region on the dusk side. In the meantime, the equatorial TEC significantly increased in the noon-dusk sector. For the next 8 hours (time(3) to time(7)), the convection pattern was suggested to be stable from the PTP simulation, and RBSP-A entered the plume region again with fewer disturbances inside the plume. During the second plume period (00–04 UT), the RBSP-A trajectory overlapped with the modelled plume region, which suggests that the model captured the plume location accurately. The plume shape from the RNN model became narrower and rotated duskward, as the PTP simulation suggested.

It is interesting to note that although the model does not provide an accurate one-to-one agreement with the RBSP observation during some periods (Figure 4b), it predicts a stable plume at roughly the same location from a global view (Figure 4d; also see animation in Movie S3 and S4 in the Supporting Information), which means that the model learned a global response of total electron density evolution to geomagnetic activity fairly well. However, an accurate global density model with the perfect agreement with the observed plume location requires further investigations.

5 Discussions

Total electron density models using artificial neural networks have been developed in recent studies; however, some questions including how to process geomagnetic indices

considering its time-series nature remain unsolved. In the present study, we model the total electron density using encoder-decoder to evaluate how well we can construct the global density model by naturally taking into account the time order information. Since many global models use geomagnetic indices (e.g., Sym-H, HP30 and SML) as inputs and the modeled quantity also has the time-series format, the encoder-decoder structure could be extended to modeling other quantities observed by satellites.

The history of geomagnetic indices is typically used to drive the model, because the physical information we need is embedded in these indices. Therefore, the informativeness of the indices directly limits the model capability, if some fundamental processes are missed from the relation with these geomagnetic indices (e.g., Obana et al., 2019). While fully connected neural networks are flexible and proved to be able to approximate virtually any functions with any accuracy (e.g., Cybenko, 1989; Hornik et al., 1989, 1990), not all model structures can be optimized for a certain dataset. Therefore, a special network design might be more efficient in capturing specific structures, like RNN for time order dependence and Convolutional Neural Network (CNN) for locality variation. Since many space physics data are inherently time-series, we expect a potential growth of using RNN models in the field. However, predicting an accurate global evolution from the satellite orbital observation can be challenging in general, especially when the target variable has a strong spatial variation. In our study, the total electron density exhibits a strong spatial variation (sharp gradient over L /MLT), which can strongly impact the model performance. Finally, evaluating the performance of the model itself poses some challenges. Although commonly used loss functions are able to provide general evaluation of how the model performs statistically, they do not necessarily describe its performance in a certain aspect of interest. In our case, we are particularly interested in the plume evolution, yet we notice that models of similar loss performance can have totally different plume features. This suggests that more sophisticated metrics are required to evaluate the model performance in a more comprehensive way.

6 Summary

We applied an encoder-decoder structure based on recurrent neural network to in-situ Van Allen Probes data to construct the global electron density distribution in the Earth's magnetosphere on the equatorial plane. We introduced the encoder-decoder structure: first, an LSTM model converts the history of geomagnetic indices into a hidden magnetospheric state and updates it over time with evolving geomagnetic indices as inputs. Then, a fully connected neural network takes the magnetospheric state and satellite position as inputs and calculates the total electron density at any location. By doing this, we have time-order information explicitly embedded in the model structure.

Our model reaches comparable performance as the previous studies; however, a quantified comparison is difficult since the trained datasets are different and data were processed differently. We examine the model performance on the evolution of plumes during two individual events, and the model shows fairly good agreement with the qualitative global observations by IMAGE EUV and GPS TEC. In both cases, a plume structure formed during disturbed geomagnetic activities remained stable for an extended period, and finally rotated with roughly the speed of Earth's rotation rate as expected during the recovery stage. These findings suggest that the model learned a general response of global electron density to storm/substorm activity, with different stages of enhanced convection, stable convection and subsiding

convection. It is worth noting that the model did not present a sharp and narrow plume structure as the satellite observed (plume with sharp boundary in Figure 4b); however, improving the model performance on such a finer spatial resolution is beyond the scope of the present study, and thus is left as a future investigation.

In summary, our study reports a novel method: encoder-decoder, which is different from previous studies, and demonstrates its capability of reconstructing the dynamic total electron density evolution on a global scale from a time-series dataset. By separating geomagnetic indices from satellite location as inputs, our RNN model (encoder-decoder model) directly models total electron density in response to geomagnetic indices. The fairly good performance of our density model using RNN also demonstrates the potential importance of applying it to other physical quantities (e.g., magnetospheric waves, particle fluxes, etc.) that are partly time-series as well. The dynamic global models of these physical quantities using RNN will potentially improve our fundamental understanding of source, loss, and transport of cold plasmas, generation and propagation of various plasma waves, as well as acceleration, loss, and transport of energetic particles in the Earth's magnetosphere in a comprehensive way.

Data Availability Statement

The Van Allen Probes total electron density data from the EMFISIS instrument were obtained from <http://emfisis.physics.uiowa.edu/Flight/>. The geomagnetic indices used in the model training in the present study are publicly available at https://omniweb.gsfc.nasa.gov/form/omni_min.html (SYM-H); <https://supermag.jhuapl.edu> (SML); <https://www.gfz-potsdam.de/en/hpo-index/> (Hp30). GPS TEC was obtained at www.openmadrigal.org. All data used to produced figures are publicly available at <https://doi.org/10.6084/m9.figshare.19097078.v2>.

Acknowledgments

This research at Boston University is supported by the NASA grant 80NSSC20K0196, 80NSSC18K0657, 80NSSC20K0604, 80NSSC21K1321, NSF grant AGS-1847818, AGS-1907698, AFOSR grant FA9559-16-1-0364, as well as the Alfred P. Sloan Research Fellowship FG-2018-10936. SH gratefully acknowledges the NASA FINESST grant 80NSSC21K1385. We gratefully acknowledge the Van Allen Probes Mission, OMNI dataset, SuperMAG collaborators, GFZ collaborators and PyTorch team. We also acknowledge Longzhi Gan for his helpful insights.

References

- Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.
- Borovsky, J. E., & Denton, M. H. (2006). Effect of plasmaspheric drainage plumes on solar-wind/magnetosphere coupling. *Geophysical Research Letters*, 33(20), L20101. <https://doi.org/10.1029/2006GL026519>
- Bortnik, J., Thorne, R. M., & Meredith, N. P. (2007). Modeling the propagation characteristics of chorus using CRRES suprathermal electron fluxes: PROPAGATION CHARACTERISTICS OF CHORUS. *Journal of Geophysical Research: Space Physics*, 112(A8), n/a-n/a. <https://doi.org/10.1029/2006JA012237>

- 413 Bortnik, Jacob, Li, W., Thorne, Richard. M., & Angelopoulos, V. (2016). A unified approach to
414 inner magnetospheric state prediction. *Journal of Geophysical Research: Space Physics*,
415 *121*(3), 2423–2430. <https://doi.org/10.1002/2015JA021733>
- 416 Breuillard, H., Dupuis, R., Retino, A., Contel, O. L., Amaya, J., & Lapenta, G. (2020).
417 Automatic classification of plasma regions in near-Earth space with supervised machine
418 learning: application to Magnetospheric Multi Scale 2016-2019 observations. *Frontiers in*
419 *Astronomy and Space Sciences*, *7*, 55. <https://doi.org/10.3389/fspas.2020.00055>
- 420 Burch, J. L. (2000), IMAGE mission overview, *Space Sci. Rev.*, *91*, 1.
- 421 Carpenter, D. L., & Anderson, R. R. (1992). An ISEE/whistler model of equatorial electron
422 density in the magnetosphere. *Journal of Geophysical Research*, *97*(A2), 1097.
423 <https://doi.org/10.1029/91JA01548>
- 424 Carpenter, D. L., & Lemaire, J. (1997). Erosion and Recovery of the Plasmasphere in the
425 Plasmopause Region. In B. Hultqvist & M. Øieroset (Eds.), *Transport Across the*
426 *Boundaries of the Magnetosphere* (pp. 153–179). Dordrecht: Springer Netherlands.
427 https://doi.org/10.1007/978-94-009-0045-5_7
- 428 Carpenter, D. L., & Seely, N. T. (1976). Cross-L plasma drifts in the outer plasmasphere: Quiet
429 time patterns and some substorm effects. *Journal of Geophysical Research*, *81*(16), 2728–
430 2736. <https://doi.org/10.1029/JA081i016p02728>
- 431 Chen, J.-F., Wang, W.-M., & Huang, C.-M. (1995). Analysis of an adaptive time-series
432 autoregressive moving-average (ARMA) model for short-term load forecasting, *34*(3), 10.
433 [https://doi.org/10.1016/0378-7796\(95\)00977-1](https://doi.org/10.1016/0378-7796(95)00977-1)
- 434 Chen, L., Thorne, R. M., Li, W., Bortnik, J., Turner, D., & Angelopoulos, V. (2012). Modulation
435 of plasmaspheric hiss intensity by thermal plasma density structure. *Geophysical Research*
436 *Letters*, *39*(14), n/a-n/a. <https://doi.org/10.1029/2012GL052308>
- 437 Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., &
438 Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder-Decoder for
439 Statistical Machine Translation. *ArXiv:1406.1078 [Cs, Stat]*. Retrieved from
440 <http://arxiv.org/abs/1406.1078>
- 441 Chu, X., Ma, D., Bortnik, J., Tobiska, W. K., Cruz, A., Bouwer, S. D., et al. (2021). Relativistic
442 Electron Model in the Outer Radiation Belt Using a Neural Network Approach. *Space*
443 *Weather*, *19*(12). <https://doi.org/10.1029/2021SW002808>
- 444 Chu, X., Bortnik, J., Li, W., Ma, Q., Denton, R., Yue, C., et al. (2017a). A neural network model
445 of three-dimensional dynamic electron density in the inner magnetosphere. *Journal of*
446 *Geophysical Research: Space Physics*, *122*(9), 9183–9197.
447 <https://doi.org/10.1002/2017JA024464>
- 448 Chu, X., Bortnik, Jacob., Li, W., Ma, Q., Angelopoulos, V., & Thorne, Richard. M. (2017b).
449 Erosion and refilling of the plasmasphere during a geomagnetic storm modeled by a neural
450 network. *Journal of Geophysical Research: Space Physics*, *122*(7), 7118–7129.
451 <https://doi.org/10.1002/2017JA023948>
- 452 Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.
453 <https://doi.org/10.1007/BF00994018>
- 454 Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function, *12*.
- 455 Darrouzet, F., Gallagher, D. L., André, N., Carpenter, D. L., Dandouras, I., Décréau, P. M. E., et
456 al. (2009). Plasmaspheric Density Structures and Dynamics: Properties Observed by the
457 CLUSTER and IMAGE Missions. *Space Science Reviews*, *145*(1–2), 55–106.
458 <https://doi.org/10.1007/s11214-008-9438-9>

- De Pascuale, S., Jordanova, V. K., Goldstein, J., Kletzing, C. A., Kurth, W. S., Thaller, S. A., & Wygant, J. R. (2018). Simulations of Van Allen Probes Plasmaspheric Electron Density Observations. *Journal of Geophysical Research: Space Physics*, 123(11), 9453–9475. <https://doi.org/10.1029/2018JA025776>
- Dent, Z. C., Mann, I. R., Goldstein, J., Menk, F. W., & Ozeke, L. G. (2006). Plasmaspheric depletion, refilling, and plasmopause dynamics: A coordinated ground-based and IMAGE satellite study. *Journal of Geophysical Research*, 111(A3), A03205. <https://doi.org/10.1029/2005JA011046>
- E, W., Han, J., & Jentzen, A. (2017). Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics*, 5(4), 349–380. <https://doi.org/10.1007/s40304-017-0117-6>
- Elman, J. L. (1990). Finding Structure in Time. *Cognitive Science*, 14(2), 179–211. https://doi.org/10.1207/s15516709cog1402_1
- Elphic, R. C., Thomsen, M. F., & Borovsky, J. E. (1997). The fate of the outer plasmasphere. *Geophysical Research Letters*, 24(4), 365–368. <https://doi.org/10.1029/97GL00141>
- Gallagher, D. L., & Adrian, M. L. (2007). Two-dimensional drift velocities from the IMAGE EUV plasmaspheric imager. *Journal of Atmospheric and Solar-Terrestrial Physics*, 69(3), 341–350. <https://doi.org/10.1016/j.jastp.2006.05.028>
- Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to Forget: Continual Prediction with LSTM. *1999 Ninth International Conference on Artificial Neural Networks*, 2, 850–855 vol.2. <https://doi.org/10.1049/cp:19991218>
- Gjerloev, J. W. (2012). The SuperMAG data processing technique: TECHNIQUE. *Journal of Geophysical Research: Space Physics*, 117(A9), n/a-n/a. <https://doi.org/10.1029/2012JA017683>
- Goldstein, J. (2007). Plasmasphere Response: Tutorial and Review of Recent Imaging Results. *Space Science Reviews*, 124(1–4), 203–216. <https://doi.org/10.1007/s11214-006-9105-y>
- Goldstein, J., & Sandel, B. R. (2005). The global pattern of evolution of plasmaspheric drainage plumes. In J. Burch, M. Schulz, & H. Spence (Eds.), *Geophysical Monograph Series* (Vol. 159, pp. 1–22). Washington, D. C.: American Geophysical Union. <https://doi.org/10.1029/159GM02>
- Goldstein, J., Spasojevic, M., Reiff, P. H., Sandel, B. R., Forrester, W. T., Gallagher, D. L., & Reinisch, B. W. (2003). Identifying the plasmopause in IMAGE EUV data using IMAGE RPI in situ steep density gradients. *Journal of Geophysical Research*, 108(A4), 1147. <https://doi.org/10.1029/2002JA009475>
- Goldstein, J., Sandel, B. R., Thomsen, M. F., Spasojevic, M., & Reiff, P. H. (2004). Simultaneous remote sensing and in situ observations of plasmaspheric drainage plumes. *Journal of Geophysical Research*, 109(A3), A03202. <https://doi.org/10.1029/2003JA010281>
- Goldstein, J., Burch, J. L., & Sandel, B. R. (2005). Magnetospheric model of subauroral polarization stream: MAGNETOSPHERIC SAPS MODEL. *Journal of Geophysical Research: Space Physics*, 110(A9). <https://doi.org/10.1029/2005JA011135>
- Goldstein, J., Thomsen, M. F., & DeJong, A. (2014a). In situ signatures of residual plasmaspheric plumes: Observations and simulation. *Journal of Geophysical Research: Space Physics*, 119(6), 4706–4722. <https://doi.org/10.1002/2014JA019953>

- Goldstein, J., Pascuale, S. D., Kletzing, C., Kurth, W., Genestreti, K. J., Skoug, R. M., et al. (2014b). Simulation of Van Allen Probes plasmopause encounters. *Journal of Geophysical Research: Space Physics*, 119(9), 7464–7484. <https://doi.org/10.1002/2014JA020252>
- Goldstein, J., Pascuale, S., & Kurth, W. S. (2019a). Epoch-Based Model for Stormtime Plasmopause Location. *Journal of Geophysical Research: Space Physics*, 124(6), 4462–4491. <https://doi.org/10.1029/2018JA025996>
- Goldstein, J., Gallagher, D., Craven, P. D., Comfort, R. H., Genestreti, K. J., Mouikis, C., et al. (2019b). Temperature Dependence of Plasmaspheric Ion Composition. *Journal of Geophysical Research: Space Physics*, 124(8), 6585–6595. <https://doi.org/10.1029/2019JA026822>
- Han, J., Jentzen, A., & E, W. (2018). Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34), 8505–8510. <https://doi.org/10.1073/pnas.1718942115>
- Harris, K. K., Sharp, G. W., & Chappell, C. R. (1970). Observations of the plasmopause from OGO 5. *Journal of Geophysical Research*, 75(1), 219–224. <https://doi.org/10.1029/JA075i001p00219>
- He, F., & Tao, D. (2021). Recent advances in deep learning theory. *ArXiv:2012.10931 [Cs, Stat]*. Retrieved from <http://arxiv.org/abs/2012.10931>
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2018). Mask R-CNN. *ArXiv:1703.06870 [Cs]*. Retrieved from <http://arxiv.org/abs/1703.06870>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
- Hornik, K., Stinchcombe, M., & White, H. (1990). Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 3(5), 551–560. [https://doi.org/10.1016/0893-6080\(90\)90005-6](https://doi.org/10.1016/0893-6080(90)90005-6)
- Horwitz, J. L., Comfort, R. H., & Chappell, C. R. (1990). A statistical characterization of plasmasphere density structure and boundary locations. *Journal of Geophysical Research*, 95(A6), 7937. <https://doi.org/10.1029/JA095iA06p07937>
- Huang, C.-M., Huang, C.-J., & Wang, M.-L. (2005). A Particle Swarm Optimization to Identifying the ARMAX Model for Short-Term Load Forecasting. *IEEE Transactions on Power Systems*, 20(2), 1126–1133. <https://doi.org/10.1109/TPWRS.2005.846106>
- Huang, Z., Xu, W., & Yu, K. (2015). Bidirectional LSTM-CRF Models for Sequence Tagging. *ArXiv:1508.01991 [Cs]*. Retrieved from <http://arxiv.org/abs/1508.01991>
- Kletzing, C. A., Kurth, W. S., Acuna, M., MacDowall, R. J., Torbert, R. B., Averkamp, T., et al. (2013). The Electric and Magnetic Field Instrument Suite and Integrated Science (EMFISIS) on RBSP. *Space Science Reviews*, 179(1–4), 127–181. <https://doi.org/10.1007/s11214-013-9993-6>
- 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.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>

- Kurth, W. S., De Pascuale, S., Faden, J. B., Kletzing, C. A., Hospodarsky, G. B., Thaller, S., & Wygant, J. R. (2015). Electron densities inferred from plasma wave spectra obtained by the Waves instrument on Van Allen Probes. *Journal of Geophysical Research: Space Physics*, 120(2), 904–914. <https://doi.org/10.1002/2014JA020857>
- Lara-Benítez, P., Carranza-García, M., & Riquelme, J. C. (2021). An Experimental Review on Deep Learning Architectures for Time Series Forecasting. *International Journal of Neural Systems*, 31(03), 2130001. <https://doi.org/10.1142/S0129065721300011>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Li, W., Thorne, R. M., Bortnik, J., Reeves, G. D., Kletzing, C. A., Kurth, W. S., et al. (2013). An unusual enhancement of low-frequency plasmaspheric hiss in the outer plasmasphere associated with substorm-injected electrons. *Geophysical Research Letters*, 40(15), 3798–3803. <https://doi.org/10.1002/grl.50787>
- Li, W., Shen, X.C., Ma, Q., Capannolo, L., Shi, R., Redmon, R. J., et al. (2019). Quantification of Energetic Electron Precipitation Driven by Plume Whistler Mode Waves, Plasmaspheric Hiss, and Exohiss. *Geophysical Research Letters*, 46(7), 3615–3624. <https://doi.org/10.1029/2019GL082095>
- Liu, H., Dai, Z., So, D. R., & Le, Q. V. (2021). Pay Attention to MLPs. *ArXiv:2105.08050 [Cs]*. Retrieved from <http://arxiv.org/abs/2105.08050>
- Liu, W., Tu, W., Li, X., Sarris, T., Khotyaintsev, Y., Fu, H., et al. (2016). On the calculation of electric diffusion coefficient of radiation belt electrons with in situ electric field measurements by THEMIS. *Geophysical Research Letters*, 43(3), 1023–1030. <https://doi.org/10.1002/2015GL067398>
- Liu, Z., & Tegmark, M. (2021). AI Poincaré: Machine Learning Conservation Laws from Trajectories. *Physical Review Letters*, 126(18), 180604. <https://doi.org/10.1103/PhysRevLett.126.180604>
- Ma, D., Chu, X., Bortnik, J., Claudepierre, S. G., Tobiska, W. K., Cruz, A., et al. (2021). Modeling the dynamic variability of sub-relativistic outer radiation belt electron fluxes using machine learning. *ArXiv:2109.08647 [Physics]*. Retrieved from <http://arxiv.org/abs/2109.08647>
- Malaspina, D. M., Zhu, H., & Drozdov, A. Y. (2020). A Wave Model and Diffusion Coefficients for Plasmaspheric Hiss Parameterized by Plasmopause Location. *Journal of Geophysical Research: Space Physics*, 125(2). <https://doi.org/10.1029/2019JA027415>
- Matzka, Jürgen; Bronkalla, Oliver; Kervalishvili, Guram; Rauberg, Jan; Stolle, Claudia (2021): Geomagnetic Hpo index. V. 1.0. GFZ Data Services. <https://doi.org/10.5880/Hpo.0001>
- 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>
- McFadden, J. P., Carlson, C. W., Larson, D., Bonnell, J., Mozer, F. S., Angelopoulos, V., et al. (2008). Structure of plasmaspheric plumes and their participation in magnetopause reconnection: First results from THEMIS. *Geophysical Research Letters*, 35(17), L17S10. <https://doi.org/10.1029/2008GL033677>
- McGranaghan, R. M., Ziegler, J., Bloch, T., Hatch, S., Camporeale, E., Lynch, K., et al. (2021). Toward a Next Generation Particle Precipitation Model: Mesoscale Prediction Through Machine Learning (a Case Study and Framework for Progress). *Space Weather*, 19(6). <https://doi.org/10.1029/2020SW002684>

- Moldwin, M B, Sandel, B. R., Thomsen, M. F., & Elphic, R. C. (2003). QUANTIFYING GLOBAL PLASMASPHERIC IMAGES WITH IN SITU OBSERVATIONS, *109*, 47. <https://doi.org/10.1023/B:SPAC.0000007512.69979.8f>
- Moldwin, Mark B., Thomsen, M. F., Bame, S. J., McComas, D. J., & Moore, K. R. (1994). An examination of the structure and dynamics of the outer plasmasphere using multiple geosynchronous satellites. *Journal of Geophysical Research*, *99*(A6), 11475. <https://doi.org/10.1029/93JA03526>
- Moldwin, Mark B., Thomsen, M. F., Bame, S. J., McComas, D., & Reeves, G. D. (1995). The fine-scale structure of the outer plasmasphere. *Journal of Geophysical Research*, *100*(A5), 8021. <https://doi.org/10.1029/94JA03342>
- Newell, P. T., & Gjerloev, J. W. (2011). Evaluation of SuperMAG auroral electrojet indices as indicators of substorms and auroral power: SUPERMAG AURORAL ELECTROJET INDICES. *Journal of Geophysical Research: Space Physics*, *116*(A12), n/a-n/a. <https://doi.org/10.1029/2011JA016779>
- Nishizuka, N., Kubo, Y., Sugiura, K., Den, M., & Ishii, M. (2021). Operational solar flare prediction model using Deep Flare Net. *Earth, Planets and Space*, *73*(1), 64. <https://doi.org/10.1186/s40623-021-01381-9>
- Obana, Y., Maruyama, N., Shinbori, A., Hashimoto, K. K., Fedrizzi, M., Nosé, M., et al. (2019). Response of the Ionosphere-Plasmasphere Coupling to the September 2017 Storm: What Erodes the Plasmasphere so Severely? *Space Weather*, *17*(6), 861–876. <https://doi.org/10.1029/2019SW002168>
- O’Brien, T. P., & Moldwin, M. B. (2003). Empirical plasmopause models from magnetic indices. *Geophysical Research Letters*, *30*(4), 2002GL016007. <https://doi.org/10.1029/2002GL016007>
- Olah, C. (2015, August 27). Understanding LSTM. Retrieved from <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- van Otterlo, M., & Wiering, M. (2012). Reinforcement Learning and Markov Decision Processes. In M. Wiering & M. van Otterlo (Eds.), *Reinforcement Learning: State-of-the-Art* (pp. 3–42). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-27645-3_1
- Pierrard, V., & Lemaire, J. F. (2004). Development of shoulders and plumes in the frame of the interchange instability mechanism for plasmopause formation: PLASMAPAUSE DEFORMATIONS. *Geophysical Research Letters*, *31*(5), n/a-n/a. <https://doi.org/10.1029/2003GL018919>
- Pierrard, Viviane, Goldstein, J., André, N., Jordanova, V. K., Kotova, G. A., Lemaire, J. F., et al. (2009). Recent Progress in Physics-Based Models of the Plasmasphere. *Space Science Reviews*, *145*(1–2), 193–229. <https://doi.org/10.1007/s11214-008-9480-7>
- Reinisch, B., Moldwin, M. B., Denton, R. E., Gallagher, D. L., Matsui, H., Pierrard, V., & Tu, J. (2009). Augmented Empirical Models of Plasmaspheric Density and Electric Field Using IMAGE and CLUSTER Data. *Space Science Reviews*, *145*(1–2), 231–261. <https://doi.org/10.1007/s11214-008-9481-6>
- Reinisch, B. W., Huang, X., & Peng, S. (2004). Plasmaspheric mass loss and refilling as a result of a magnetic storm. *Journal of Geophysical Research*, *109*(A1), A01202. <https://doi.org/10.1029/2003JA009948>
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers, *3*, 210–229.

- Sandel, B. R., Broadfoot, A. L., Curtis, C. C., King, R. A., Stone, T. C., Hill, R. H., et al. (2000). The Extreme Ultraviolet Imager Investigation for the Image Mission. In J. L. Burch (Ed.), *The Image Mission* (pp. 197–242). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-011-4233-5_7
- Sheeley, B. W., Moldwin, M. B., Rassoul, H. K., & Anderson, R. R. (2001). An empirical plasmasphere and trough density model: CRRES observations. *Journal of Geophysical Research: Space Physics*, 106(A11), 25631–25641. <https://doi.org/10.1029/2000JA000286>
- Shi, R., Li, W., Ma, Q., Green, A., Kletzing, C. A., Kurth, W. S., et al. (2019). Properties of Whistler Mode Waves in Earth’s Plasmasphere and Plumes. *Journal of Geophysical Research: Space Physics*, 124(2), 1035–1051. <https://doi.org/10.1029/2018JA026041>
- Sirignano, J., & Spiliopoulos, K. (2018). DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics*, 375, 1339–1364. <https://doi.org/10.1016/j.jcp.2018.08.029>
- Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks. *ArXiv:1909.09586 [Cs]*. Retrieved from <http://arxiv.org/abs/1909.09586>
- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. *ArXiv:1409.3215 [Cs]*. Retrieved from <http://arxiv.org/abs/1409.3215>
- Thorne, R. M., Bortnik, J., Li, W., Chen, L., Ni, B., & Ma, Q. (2016). How Whistler-Mode Waves and Thermal Plasma Density Control the Global Distribution of the Diffuse Aurora and the Dynamical Evolution of Radiation Belt Electrons. In C. R. Chappell, R. W. Schunk, P. M. Banks, J. L. Burch, & R. M. Thorne (Eds.), *Geophysical Monograph Series* (pp. 115–125). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781119066880.ch9>
- Tsyganenko, N. A. (2002a). A model of the near magnetosphere with a dawn-dusk asymmetry 1. Mathematical structure: A NEW MAGNETOSPHERE MAGNETIC FIELD MODEL, 1. *Journal of Geophysical Research: Space Physics*, 107(A8), SMP 12-1-SMP 12-15. <https://doi.org/10.1029/2001JA000219>
- Tsyganenko, N. A. (2002b). A model of the near magnetosphere with a dawn-dusk asymmetry 2. Parameterization and fitting to observations: A NEW MAGNETOSPHERE MAGNETIC FIELD MODEL, 2. *Journal of Geophysical Research: Space Physics*, 107(A8), SMP 10-1-SMP 10-17. <https://doi.org/10.1029/2001JA000220>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention Is All You Need. *ArXiv:1706.03762 [Cs]*. Retrieved from <http://arxiv.org/abs/1706.03762>
- Walsh, B. M., Phan, T. D., Sibeck, D. G., & Souza, V. M. (2014). The plasmaspheric plume and magnetopause reconnection. *Geophysical Research Letters*, 41(2), 223–228. <https://doi.org/10.1002/2013GL058802>
- Wang, X., Chen, Y., Toth, G., Manchester, W. B., Gombosi, T. I., Hero, A. O., et al. (2020). Predicting Solar Flares with Machine Learning: Investigating Solar Cycle Dependence. *The Astrophysical Journal*, 895(1), 3. <https://doi.org/10.3847/1538-4357/ab89ac>
- 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.

- Wu, J.-G., and Lundstedt, H. (1997), Geomagnetic storm predictions from solar wind data with the use of dynamic neural networks, *J. Geophys. Res.*, 102(A7), 14255– 14268, doi:10.1029/97JA00975.
- Zhelavskaya, Irina S., Shprits, Y. Y., & Spasojević, M. (2017). Empirical Modeling of the Plasmasphere Dynamics Using Neural Networks. *Journal of Geophysical Research: Space Physics*, 122(11), 11,227-11,244. <https://doi.org/10.1002/2017JA024406>
- Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. *ArXiv:2012.07436 [Cs]*. Retrieved from <http://arxiv.org/abs/2012.07436>
- Zou, Y., Walsh, B. M., Shi, X., Lyons, L., Liu, J., Angelopoulos, V., et al. (2021). Geospace Plume and Its Impact on Dayside Magnetopause Reconnection Rate. *Journal of Geophysical Research: Space Physics*, 126(6). <https://doi.org/10.1029/2021JA029117>

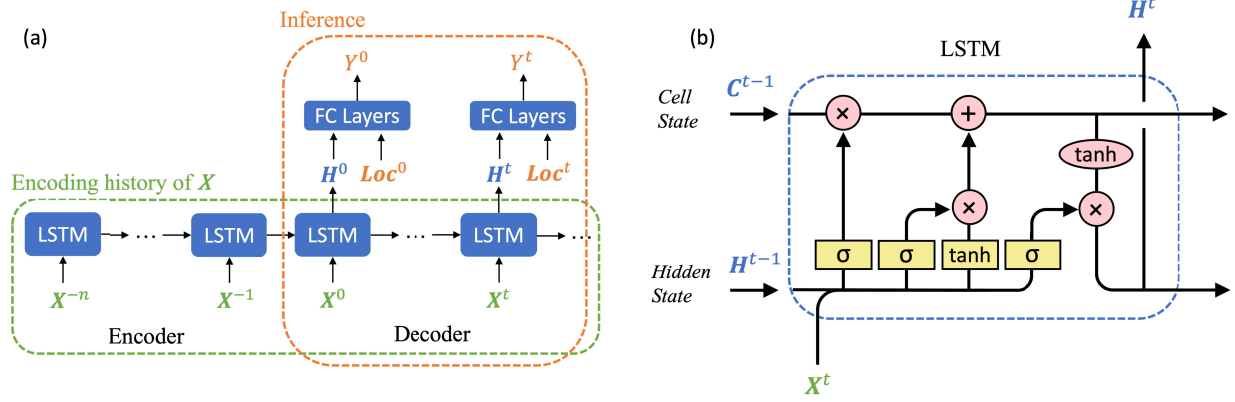


Figure 1. Schematics of the proposed encoder-decoder network. (a) The architecture of an encoder-decoder network with geomagnetic indices X (from X^{-n} to X^{-1}) encoded into the hidden state H , and later decoded into total electron density Y at location Loc . (b) The architecture of a typical LSTM cell at time t , which consists of cell input X^t , cell state C^t , hidden state/cell output H^t and four gates (yellow boxes) to control the information flow (modified from Olah (2015)).

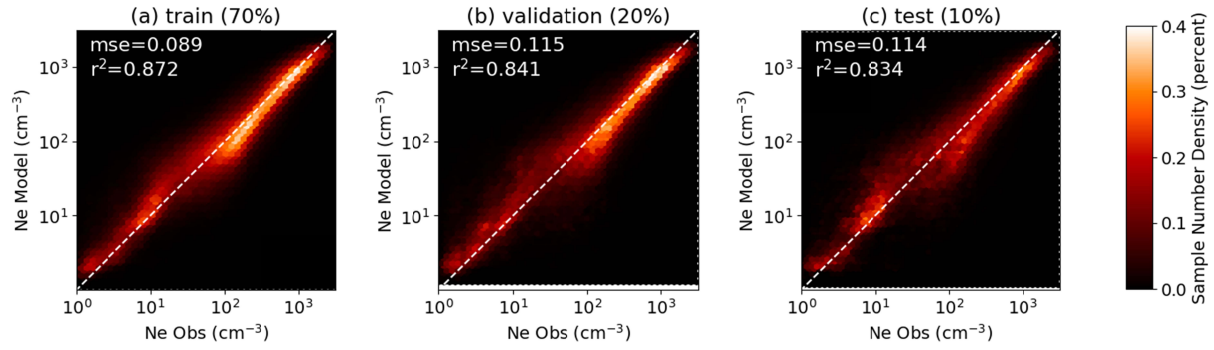


Figure 2. Model performance for different datasets of total electron density. (a) Comparison of observed (x-axis) and modeled electron density (y-axis) for the training set, color-coded by the percentage of data in the corresponding region; (b) same as (a) but for the validation set; (c) same as (a) but for the test set.

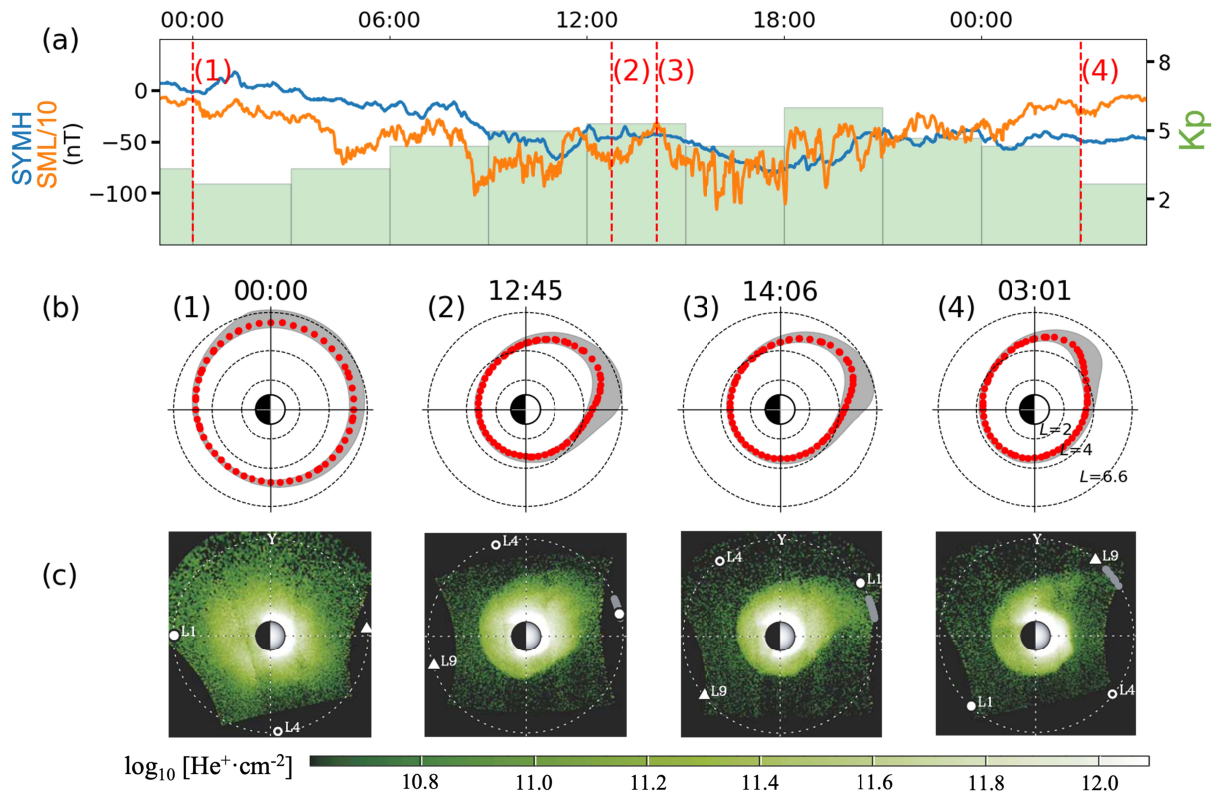


Figure 3. An event study of the modeled electron density on 26 June 2000. (a) SYM-H (blue), SML (orange), and Kp (green) of the event. The times of four snapshots are marked with red dotted vertical lines. (b) The Earth is at the center of each plot and the contour of 50 cm⁻³ for the modeled electron density indicates the nominal plasmopause (red) on the equatorial plane. Grey shadow shows a broader region of 20-80 cm⁻³ indicating plume. (c) IMAGE EUV observation modified from Goldstein et al. (2004). The color intensity is EUV equatorial He⁺ abundance, as shown by the colorbar at the bottom. The dotted lines indicate the geosynchronous orbit.

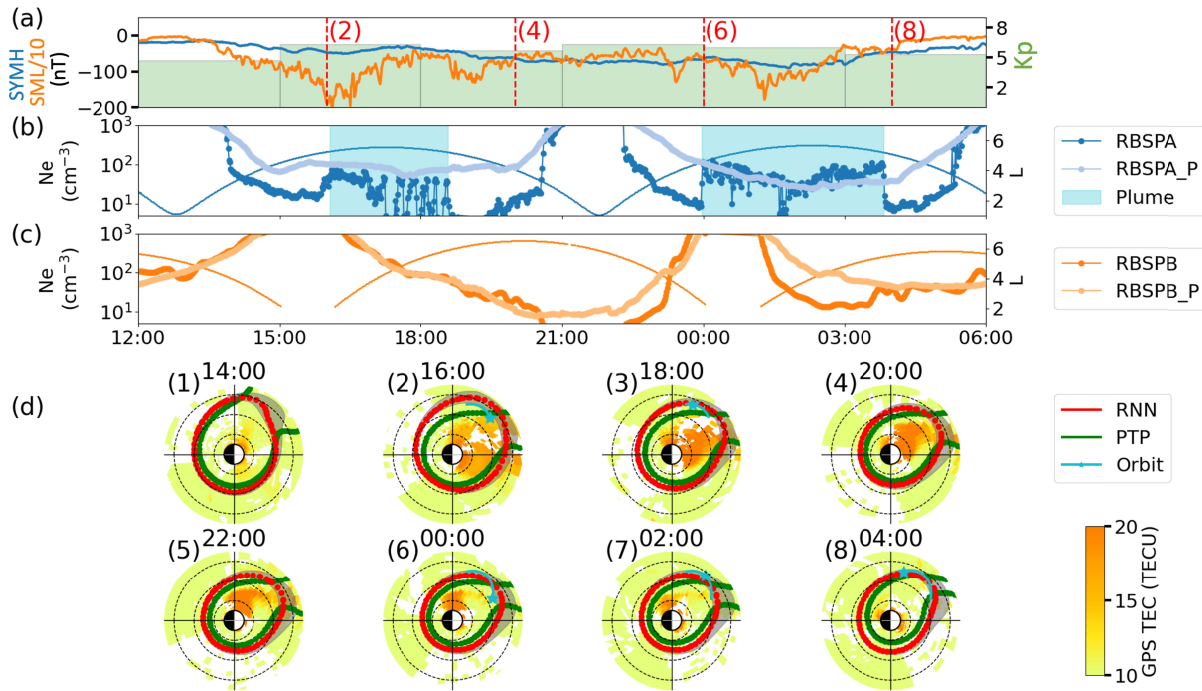
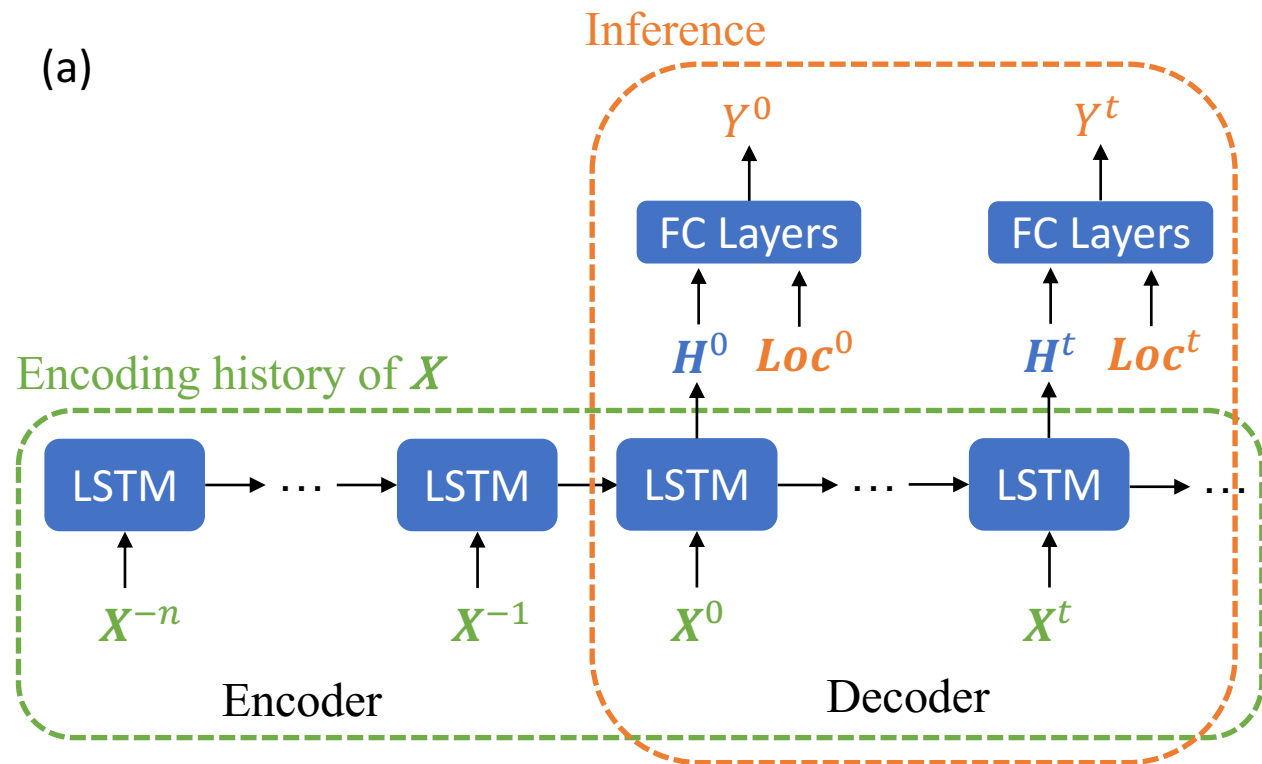


Figure 4. Case comparison on 7-8 September 2015. (a) SYM-H (blue), SML (orange), and Kp (green) of the event. Times of the four snapshots are marked with red dotted lines. (b) Line plot of the RBSP-A observation and the corresponding model prediction (lighter color with suffix "P"). L shell is plotted with smaller dots, and the plume period is marked with cyan shadow. (c) Same as (b) but for the RBSP-B observation. (d) Red dots represent the contour of 50 cm^{-3} for the modeled electron density as an indication of plasmopause. Grey shadow covers the regions of $20\text{--}80 \text{ cm}^{-3}$ indicating plume. Green dots (PTP) are simulated plasmopause. Cyan line represents the RBSP-A trajectory of two plume periods with the star symbol indicating the satellite location at that time. GPS TEC measurements (color-coded) are mapped to the equatorial plane.

Figure 1.

(a)



(b)

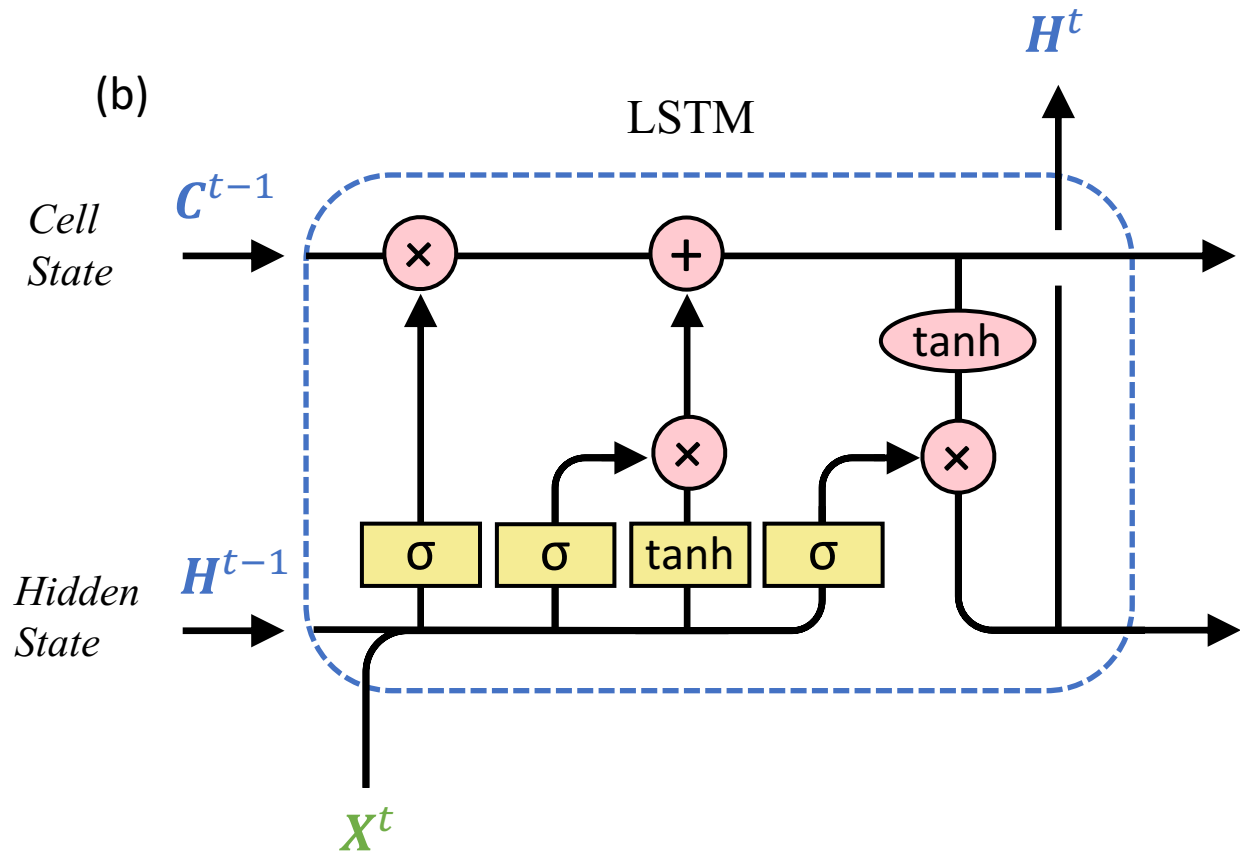


Figure 2.

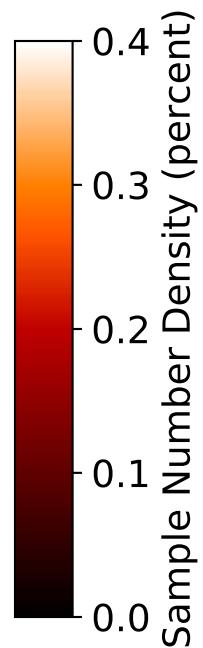
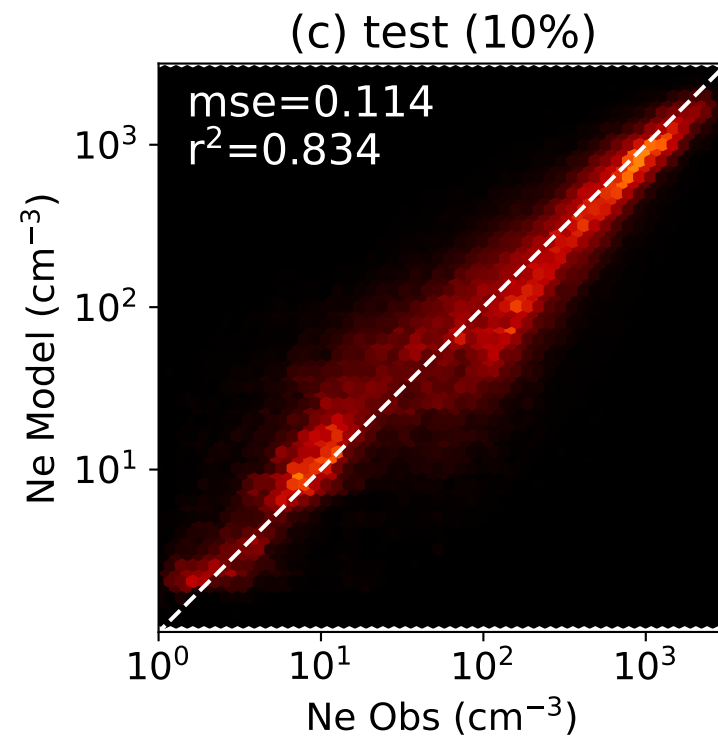
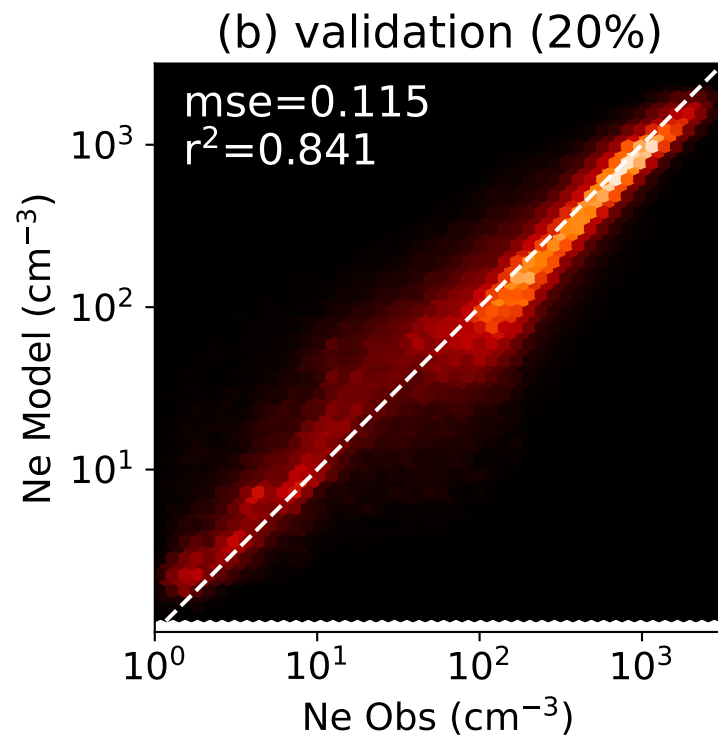
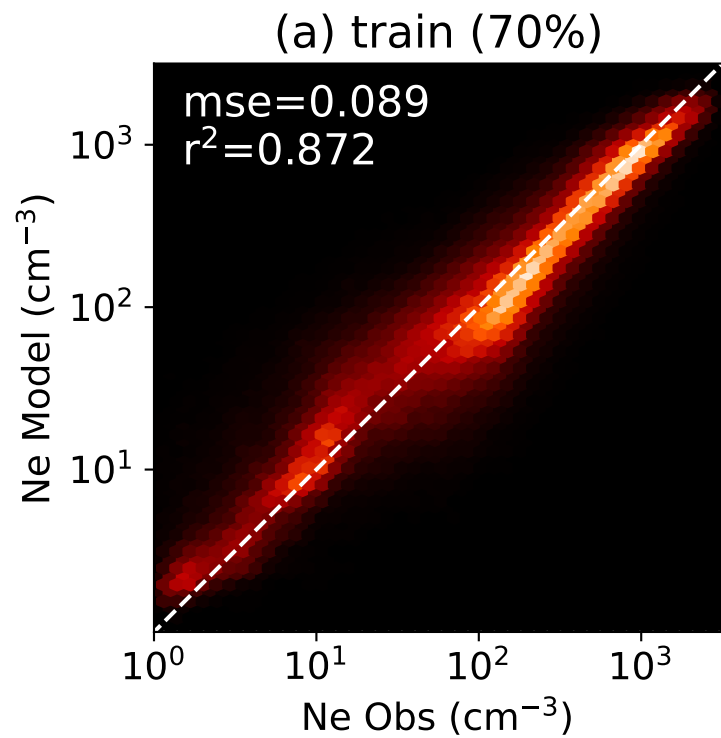


Figure 3.

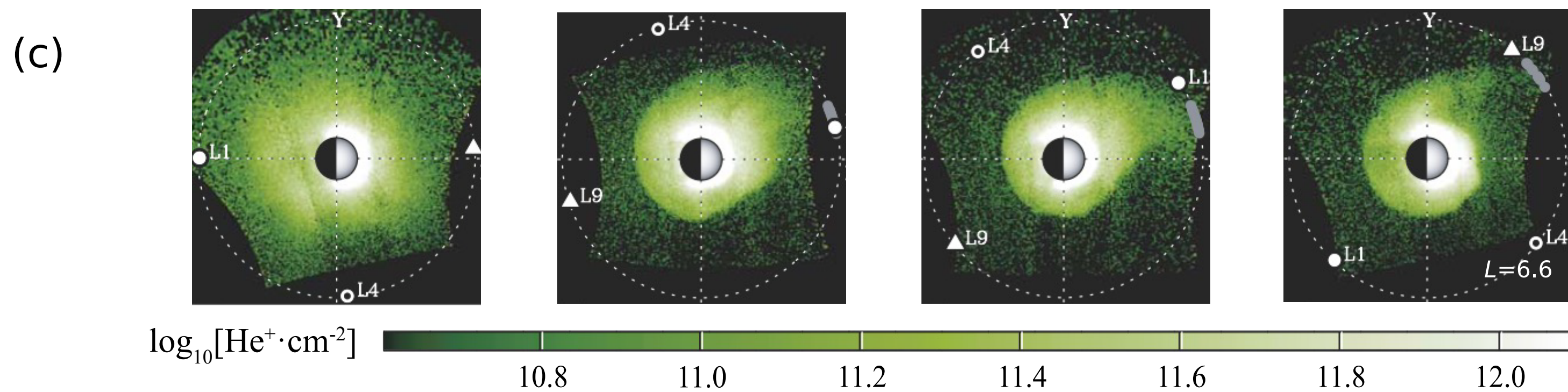
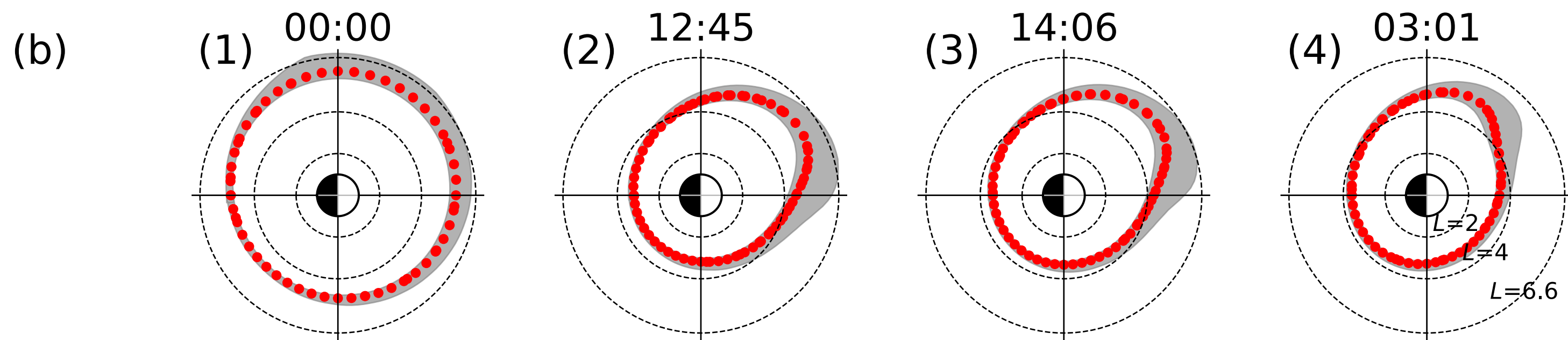
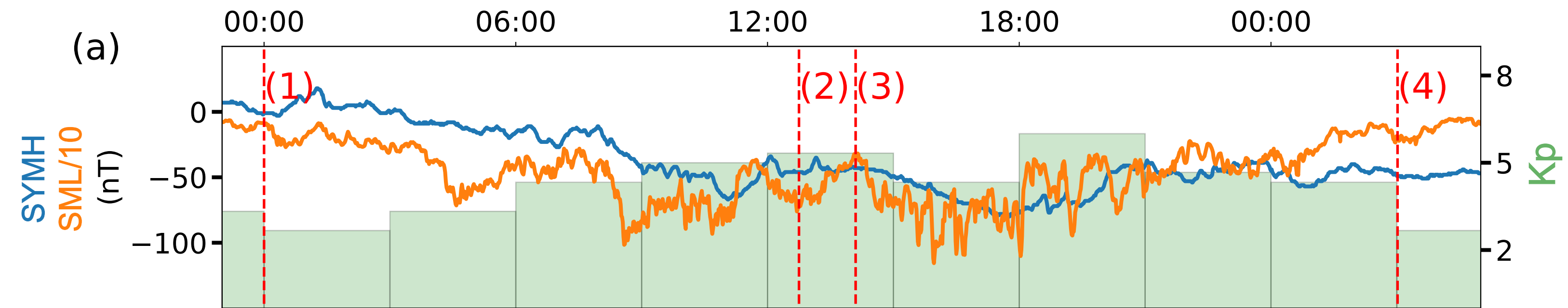


Figure 4.

