RESEARCH ARTICLE
SPACE PHYSICS: MAGNETOSPHERIC PHYSICS

*Earth and Planetary Physics*
*6*: 350–358, 2022
doi: 10.26464/epp2022036

# A machine-learning-based electron density (MLED) model in the inner magnetosphere

**QingHua Zhou[1,2], YunXiang Chen[1,2], FuLiang Xiao[1,2*], Sai Zhang[1,2], Si Liu[1,2], Chang Yang[1,2], YiHua He[1,2], and ZhongLei Gao[1,2]**

[1]School of Physics and Electronic Sciences, Changsha University of Science and Technology, Changsha 410114, China;

[2]Hunan Provincial Key Laboratory of Flexible Electronic Materials Genome Engineering, Changsha University of Science and Technology, Changsha 410114, China

**Key Points:**

- A machine-learning-based electron density model in the inner magnetosphere using Van Allen Probes density data is developed.
- This model has explicit expressions with few parameters and presents results that are consistent with empirical observations.
- This model can be incorporated straightforwardly into previously developed radiation belt models. It promises to be helpful for space weather forecasting.

**Citation:** Zhou, Q. H., Chen, Y. X., Xiao, F. L., Zhang, S., Liu, S., Yang, C., He, Y. H., and Gao, Z. L. (2022). A machine-learning-based electron density (MLED) model in the inner magnetosphere. *Earth Planet. Phys.*, *6*(4), 350–358. http://doi.org/10.26464/epp2022036

**Abstract:** Plasma density is an important factor in determining wave-particle interactions in the magnetosphere. We develop a machine-learning-based electron density (MLED) model in the inner magnetosphere using electron density data from Van Allen Probes between September 25, 2012 and August 30, 2019. This MLED model is a physics-based nonlinear network that employs fundamental physical principles to describe variations of electron density. It predicts the plasmapause location under different geomagnetic conditions, and models separately the electron densities of the plasmasphere and of the trough. We train the model using gradient descent and backpropagation algorithms, which are widely used to deal effectively with nonlinear relationships among physical quantities in space plasma environments. The model gives explicit expressions with few parameters and describes the associations of electron density with geomagnetic activity, solar cycle, and seasonal effects. Under various geomagnetic conditions, the electron densities calculated by this model agree well with empirical observations and provide a good description of plasmapause movement. This MLED model, which can be easily incorporated into previously developed radiation belt models, promises to be very helpful in modeling and improving forecasting of radiation belt electron dynamics.

**Keywords:** background electron density; inner magnetosphere; machine learning; Van Allen Probes observation

## 1. Introduction

The background electron density can affect the propagation and instability of electromagnetic waves in the magnetosphere (e.g. Chen LJ et al., 2009; Xiao FL et al., 2013; Guo MY et al., 2020; Guan CY et al., 2020; Sauer K et al., 2020). The plasmasphere is composed of low-energy particles, forming a sphere-like reservoir of very cold (~1 eV), fairly dense plasma (~50–$10^4$ $cm^{-3}$) that co-rotates with the Earth. Recently, the method for Extreme Ultraviolet (EUV) image reconstruction of the plasmasphere was improved by Huang Y et al. (2021). The low-density region outside the plasmasphere is called the trough. The boundary between the plasmasphere and trough regions is the "plasmapause". The plasmapause varies with magnetic local time (MLT) and geomagnetic activity (Chappell, 1972; Carpenter and Anderson, 1992; Larsen et al., 2007; Fu HS et al., 2010a). Observations (e.g. Goldstein et al., 2004; Darrouzet et al., 2008; Goldstein et al., 2014) have shown that the enhanced convection electric field during geomagnetic activity leads to erosion of the plasmasphere and formation of a high-density plume in the afternoon sector. In the early stage of its formation, the plasmaspheric plume is generally broad in MLT. It co-rotates with the Earth and becomes narrower as time goes on. In addition, the existence of density troughs inside the main body of the plasmasphere has been confirmed by observation (Fu HS et al., 2010b). Whistler-mode chorus mainly appears in the plasmatrough (Meredith et al., 2012); whistler-mode hiss appears mainly in the plasmasphere (Thorne et al., 1973; Wang JZ et al., 2020). Electrostatic electron cyclotron harmonic waves tend to be generated in the lower (upper) half of harmonic bands in the lower (higher) density region (Zhou QH et al., 2017). In addition, the diffusion coefficients controlling wave-particle interactions are closely related to the background density. Acceleration (scattering loss) of particles by waves generally occurs in the lower (higher) density region (e.g. Xiao FL et al., 2009, 2010, 2015; He JB

Correspondence to: F. L. Xiao, flxiao@126.com
Received 15 MAR 2022; Accepted 27 APR 2022.
Accepted article online 20 JUN 2022.
©2022 by Earth and Planetary Physics.

et al., 2021; Yang C et al., 2021).

Direct measurement of electron density is a challenging task. It can be derived from spacecraft potentials (Escoubet et al., 1997) or, alternativelty, from the upper hybrid resonance frequency $f_{uh}$ (Kurth et al., 2015). Van Allen Probes running in the inner magnetosphere have provided measurements of electron density (inferred from $f_{uh}$) that are more reliable than previous approaches. However, these in-situ measurements detect only the local electron density where the spacecraft is located. A number of empirical models have been developed to obtain a global distribution of the electron density in the inner magnetosphere under various magnetospheric conditions. The model by Carpenter and Anderson (1992) is derived from ISEE-1 data and valid for $L$ shells between 2.25 and 8. The model presents separate empirical functions for the equatorial densities of the plasmaspheric and the trough regions. In this model the trough density depends on $L$ shell; the plasmaspheric density depends on $L$ shell, the long-term solar cycle, and seasonal effects. The plasmapause location is determined by the maximum geomagnetic activity index ($Kp$) in the preceding 24 hours. The Global Core Plasma Model (GCPM) (Gallagher et al., 2000) is an improvement of several previous models (Carpenter and Anderson, 1992; Gallagher et al., 1995). It incorporates densities of the plasmasphere, the trough, and the polar cap, and considers the influences of the $Kp$ index and MLT on the plasmapause location. The model by Sheeley et al. (2001) is developed from density data from the Combined Release and Radiation Effects Satellite (CRRES). The plasmasphere density is a function of $L$ shell; the trough density is a function of $L$ shell and MLT, both of which are independent of geomagnetic activity. The model is valid for $L$ shells between 3 and 7. Bortnik et al. (2016) developed neural network models for predicting the inner magnetospheric state. They then have published 2D and 3D neural network models for electron density based on data from Time History of Events and Macroscale Interactions during Substorms (THEMIS) Probes (Chu XN et al., 2017a, b; Bortnik et al., 2018). Zhelavskaya et al. (2017) have presented a neural network model for electron density based on EMFISIS data from the Van Allen Probes mission. However, these neural network models have not provided explicit expressions and parameters.

We construct a physics-based nonlinear network to describe the dynamic and nonlinear relationships between electron density and geomagnetic activity levels, solar cycle effects, and seasonal effects. Using data from Van Allen Probes mission during the period from September 25, 2012 to August 30, 2019, we perform machine learning with gradient descent and backpropagation algorithms to find the optimal parameters of the density model. We shall present explicit expressions with few parameters, which can be used straightforwardly by the radiation belt community.

## 2. Data and Methods

### 2.1 Data Set

The high-frequency receiver (HFR) of the Electric and Magnetic Field Instrument Suite and Integrated Science (EMFISIS) instrumentation suite on board the Van Allen Probes mission measures electric components of plasma waves in the frequency range of 10 to 500 kHz, and provides electron density ($N$) data inferred from $f_{uh}$. The density data are available from September 25, 2012 to August 30, 2019. We combined them with MagEphem data from Van Allen Probes and from OMNI (Operating Missions as Nodes on the Internet). Each data sample contains $N$, $Kp$ index, SYM-H index, the 13-month-average sunspot number ($\bar{R}$), the day of the year (DoY), and the position of the probe ($L$ Shell, MLT and magnetic latitude). The minimum time interval between each sample is 1 minute. We focus on the electron density near the magnetic equatorial plane and have chosen those samples within ±3° of magnetic

latitude to form a data set (1,017,586 observations in total). Samples in the data set are binned as a function of $L$ in steps of 0.5 $L$ and MLT in an interval of 1 h. We plot the corresponding global distribution of the samples in Figure 1. The number of samples increases as $L$ increases. In the region of $1.1 < L < 2.5$, there are fewer than 5000 samples in each bin. In the region of $5.0 < L < 6.0$, more than 8000 samples are obtained in each bin. Samples occur most frequently from the dusk-to-midnight side. The bin with the most samples (20,000 samples) is located in the region of $L$ = 5.5–6.0 and 15 < MLT < 17. O'Brien and Moldwin (2003) investigated the role of various geomagnetic indices in modeling the plasmapause location. They found that the recent minimum $Dst$ index provides a better model than the maximum $Kp$ index, which is used in GCPM. Therefore, we use the minimum SYM-H in the preceding 24 hours (SYM-$H_{m24}$) to represent the geomagnetic activity level. Figures 1b–1d show that the global distribution of samples under various geomagnetic activities is basically the same. The number of samples during SYM-$H_{m24} > -30$ nT, $-50$ nT < SYM-$H_{m24} \leq -30$ nT and SYM-$H_{m24} \leq -50$ nT are 791,340, 147,517 and 78,729, respectively. Under these geomagnetic activity conditions, the orbits of Van Allen Probes have good statistical coverage for $1.1 < L < 6.0$ and all magnetic local times near the equator.

### 2.2 Model Description and Training Algorithm

For the plasmaspheric density, we adopt equations similar to the GCPM (Gallagher et al., 2000):

$$N_{ps} = 10^{g \cdot h} - 1, \tag{1}$$

$$g = (k_1 L + k_2) + k_3 \left[ \left( \cos \frac{2\pi (\mathrm{DoY} + 9)}{365} + k_4 \cos \frac{4\pi (\mathrm{DoY} + 9)}{365} \right) + k_5 \bar{R} + k_6 \right] \mathrm{e}^{\frac{-(L-2)}{1.5}}, \tag{2}$$

$$h = \left[ 1 + \left( \frac{L}{L_{pp}} \right)^{2(a_9 - 1)} \right]^{\left( \frac{a_9}{1 - a_9} \right)}, \tag{3}$$

where $L_{pp}$ is the location of the midpoint of the plasmapause and $a_9$ controls the plasmapause gradient. Both terms are associated with $Q = \log_{10}|\text{SYM-}H_{m24}|$:

$$L_{pp} = (b_1 Q + b_2) \left[ 1 + \mathrm{e}^{\left( -1.5x^2 + 0.08x - 0.7 \right)} \right], \tag{4}$$

$$a_9 = (b_3 Q + b_4), \tag{5}$$

where

$$b_1 = k_7 \cos \left( \frac{\mathrm{MLT}}{12} \pi \right) + k_8, \tag{6}$$

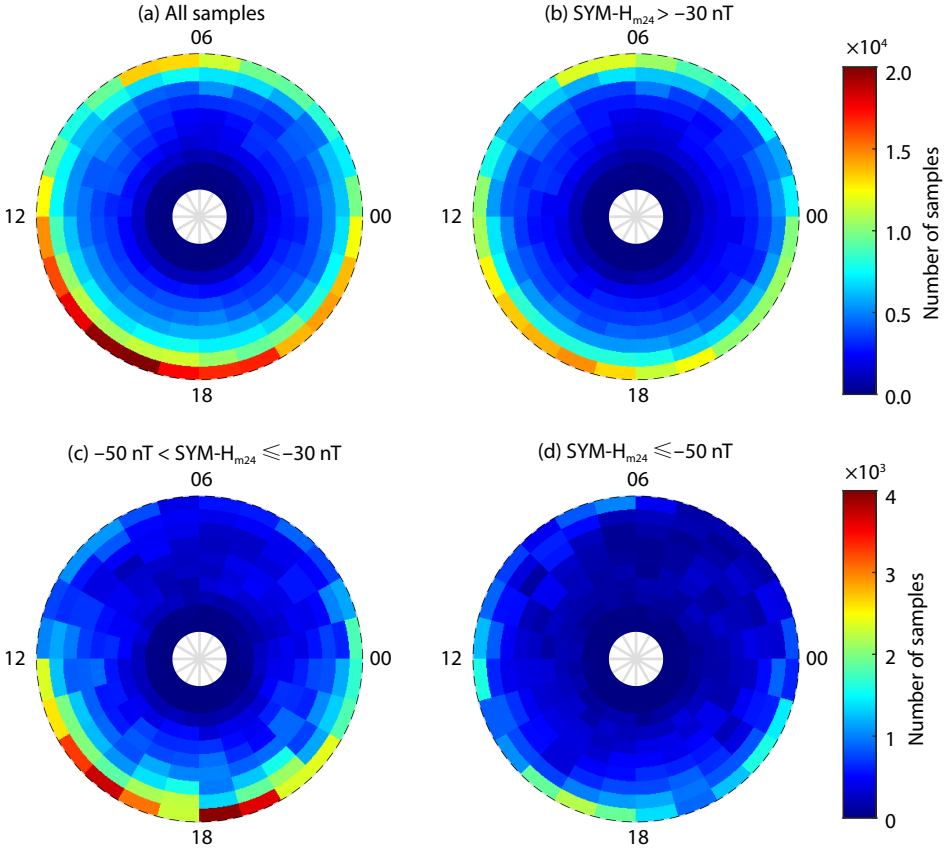


**Figure 1**. (a) The global distribution of electron density samples within ±3° of the equator during the period from September 25, 2012 to August 30, 2019 (b, c, d) for different levels of geomagnetic activity.

$$b_2 = k_9\cos\left(\frac{\mathrm{MLT}}{12}\pi\right) + k_{10}, \tag{7}$$

$$b_3 = k_{11}\cos\left(\frac{\mathrm{MLT}}{12}\pi\right) + k_{12}, \tag{8}$$

$$b_4 = k_{13}\cos\left(\frac{\mathrm{MLT}}{12}\pi\right) + k_{14}, \tag{9}$$

$$x = \begin{cases} |\mathrm{MLT} - \Phi_B|\,\pi/12, |\mathrm{MLT} - \Phi_B| < 12, \\ (24 - |\mathrm{MLT} - \Phi_B|)\,\pi/12, |\mathrm{MLT} - \Phi_B| \geq 12, \end{cases} \tag{10}$$

$$\Phi_B = \frac{k_{15}}{Q + k_{16}} + k_{17}. \tag{11}$$

For the trough, the variation of density with $L$ shell and MLT is described following the form of the trough model proposed by (Sheeley et al., 2001):

$$N_{tr} = k_{18}\left(\frac{3}{L}\right)^{4.0} + k_{19}\left(\frac{3}{L}\right)^{3.5}\cos\left(\left\{\mathrm{MLT} - \left[7.7\left(\frac{3}{L}\right)^2 + 12\right]\right\}\frac{\pi}{12}\right). \tag{12}$$

By integrating plasmaspheric and trough densities, the electron density is represented by:

$$N = u_1 N_{tr} + u_2 N_{ps}, \tag{13}$$

where $u_1$ and $u_2$ are functions that connect one regional description to another:

$$u_1 = 0.5\tanh\left[3.4534\frac{(L - L_{pp})}{0.1}\right] + 0.5, \tag{14}$$

$$u_2 = 0.5\tanh\left[-3.4534\frac{(L - L_{pp})}{0.1}\right] + 0.5. \tag{15}$$

It should be mentioned that the electron density varies with season by a cycle of 365 days and with MLT by a cycle of 24 hours. Since both cycles are symmetrically distributed, previous studies (Gallagher et al., 2000; Sheeley et al., 2001) have used cosine functions to describe the density variation with DoY and MLT. Therefore, we choose similar expressions of density variation with DoY and MLT. Meanwhile, some parameter values in Equations (1)–(15) can reasonably describe the density variation with season and MLT. To reduce the computational time and the number of parameters that need to be fitted, the values of these parameters are adopted directly from those previous studies (Gallagher et al., 2000; Sheeley et al., 2001).

We obtain optimal parameters of the density model ($k_1$–$k_{19}$) using standard methods for optimizing classical neural networks, i.e., gradient descent and backpropagation algorithms (Rumelhart et al., 1986; Hecht-Nielsen, 1989). We use the mean square error (MSE) as the cost function:

$$J(k_1, k_2, \cdots, k_{19}) = \mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{N}^{(i)} - N^{(i)}\right)^2, \tag{16}$$

where $n$ is the sample size of the training data set,and $\hat{N}^{(i)}$ and $N^{(i)}$ are the calculated and measured values of $N$ respectively. MSE is a metric of the difference between calculated and measured values.

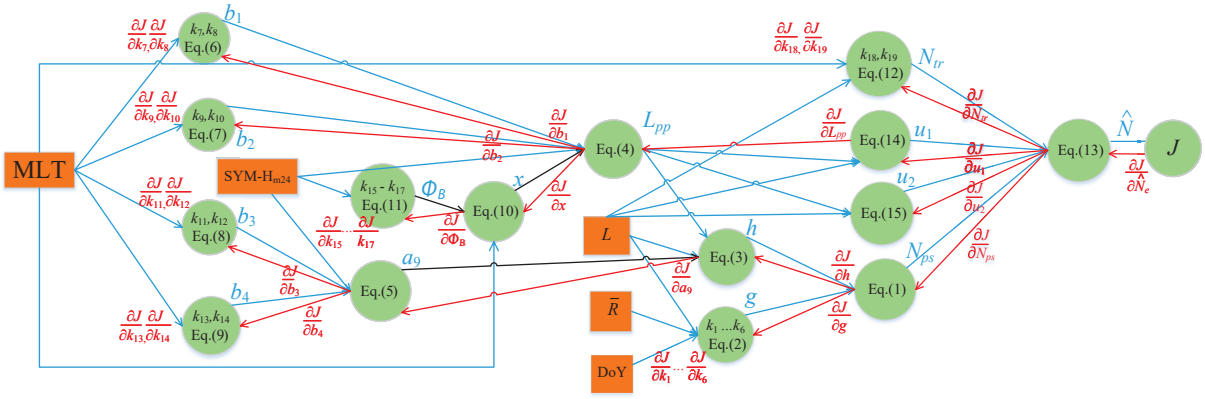


**Figure 2**. The structure of the network to build the density model. The orange nodes represent the input parameters. The green nodes represent intermediate variables and the output of the model. The numbers of the equations corresponding to each node and the parameters to be optimized are shown in the node.

Figure 2 shows the constructed network of neuron-like units according to Equations (1)–(15). This is a nonlinear network because these equations are nonlinear. The orange nodes represent the input parameters of the density model. The intermediate variables and the output of the model are represented by the green nodes. We follow the blue arrows to calculate the density and cost function of the model. The derivatives of the cost function with respect to the parameter $k_i$ ($i$ = 1–19) are calculated using the backpropagation algorithm (Rumelhart et al., 1986; Hecht-Nielsen, 1989) following the red arrows. The Levenberg-Marquardt algorithm is sensitive to the local optimum when training neural networks, i.e., the results obtained from different initial values vary widely. We adopt the mini-batch and Adam algorithms commonly used in non-convex optimization problems to find the global optimum efficiently. The metric used to evaluate the density model performance is the Adjusted R-squared coefficient ($R_a^2$):

$$R_a^2 = 1 - \frac{n-1}{n-p-1}\frac{\sum_i \left(\hat{N}_i - N_i\right)^2}{\sum_i \left(N_i - \overline{N}\right)^2}, \tag{17}$$

where $\overline{N}$ is the mean density of data set and $p$ is the number of fitting parameters in the model ($p$ = 19 for our model). The value range of $R_a^2$ is $(-\infty, 1]$; the larger the $R_a^2$ is, the better the model fits.

In order to avoid the contingency of results caused by the random division of the data set, 5-fold cross-validation (Ojala and Garriga, 2009) is used. The dataset is divided into five blocks randomly and evenly, without overlap between any of them. We perform five training sessions, taking one block in each turn as the test set and the remaining four blocks as the training set. One thousand iterations are performed in each training session; the mini-batch of each iteration is 200 samples randomly selected from the training data set. We obtain five different models from the five training sessions and calculate $R_a^2$ for each model based on the test set of each training session. The larger $R_a^2$ is, the more accurate is the model on the test set. We choose the model with the largest $R_a^2$ as the optimal one.

**Table 1.** The optimal values of parameters in Equations (1)–(12).

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| $k_1$ | −0.43 | $k_{11}$ | 2.1726 |
| $k_2$ | 4.4 | $k_{12}$ | −1.5138 |
| $k_3$ | 0.09867 | $k_{13}$ | −7.9149 |
| $k_4$ | −0.2555 | $k_{14}$ | 44.2847 |
| $k_5$ | 0.0022 | $k_{15}$ | 14.4557 |
| $k_6$ | −1.956 | $k_{16}$ | 1.2235 |
| $k_7$ | 0.015 | $k_{17}$ | 11.5 |
| $k_8$ | −1.5226 | $k_{18}$ | 187.7 |
| $k_9$ | 0.0319 | $k_{19}$ | 88.2 |
| $k_{10}$ | 5.7689 | | |

## 3. Results and Discussion

Table 1 lists the optimal values of parameters in Equations (1)–(12). Figure 3 shows the regression of the modeled ($N_{\rm model}$) and observed ($N_{\rm obs}$) electron densities for various data sets. The classes of each dataset and the corresponding correlation coefficients ($R$) are shown in the panel headers. The data points are essentially distributed close to the dashed line representing the perfect fit, suggesting that the MLED model fits the observations well. The samples near the perfect fit line in Figure 3 are much larger than those away from the perfect fit line. The correlation of observations with MLED predictions is greater than 0.87 in all three datasets. Thus, the MLED model appears to offer promising prediction performance. The correlation coefficient on the test dataset is 0.90, implying that the MLED model has good generalizability. In machine learning tasks, overfitting may occur when the model is too complex and the metrics of the training set are much better than those of the test set. The concise network used here avoids overfitting. For such concise networks, it is possible that the metrics of the test set are slightly higher than those of the training set, as shown in Figure 3.

In Table 2 we present the $R_a^2$ coefficient and the root mean square error (RMSE) for the training and test sets under different

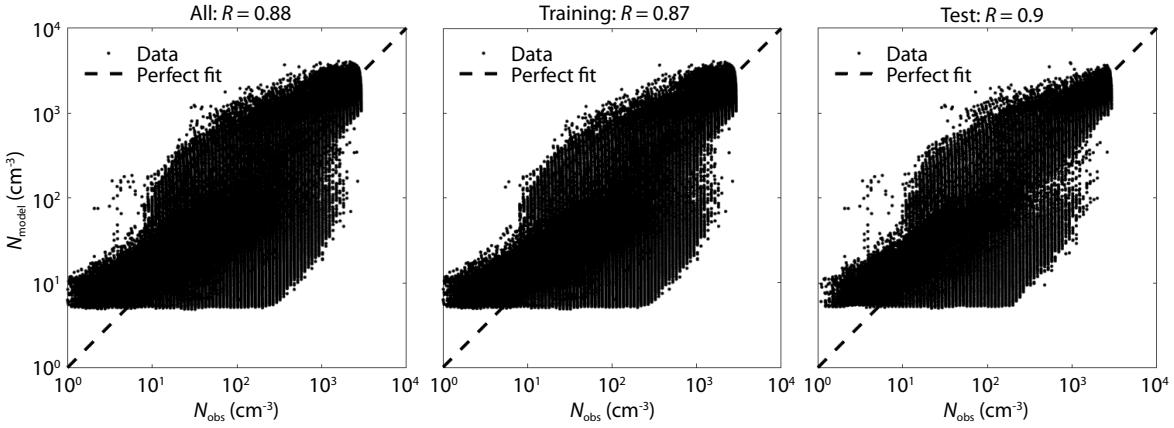


**Figure 3**. Regressions between the observed electron density and the simulation results of the MLED model for various data sets. The dashed line represents a perfect model fit.

**Table 2.** The $R_a^2$ coefficients and RMSE of both models under various geomagnetic activities.

| | | Training | | Test | |
|---|---|---|---|---|---|
| | | MLED | GCPM | MLED | GCPM |
| All samples | $R_a^2$ | 0.7383 | −0.2282 | 0.7702 | 0.1126 |
| | RMSE | 259.0 | 561.1 | 242.9 | 477.3 |
| SYM-$H_{m24}$ > −30 nT | $R_a^2$ | 0.7485 | −0.1423 | 0.8085 | 0.3448 |
| | RMSE | 254.7 | 542.9 | 227.9 | 450.8 |
| −50 nT < SYM-$H_{m24}$ ≤ −30 nT | $R_a^2$ | 0.7075 | −0.4212 | 0.7661 | 0.1904 |
| | RMSE | 266.2 | 586.7 | 242.3 | 421.6 |
| SYM-$H_{m24}$ ≤ −50 nT | $R_a^2$ | 0.5769 | −1.5331 | 0.5666 | −1.2918 |
| | RMSE | 301.8 | 738.4 | 291.7 | 670.8 |

geomagnetic activities. The $R_a^2$ coefficient and RMSE of the MLED model are better than those of the GCPM model in both the training and test sets. Figure 4 shows the equatorial electron densities derived from the MLED model. Since the magnitudes of $k_3$–$k_5$ are very small, the electron density is insensitive to variations in DoY and $\overline{R}$. In Figure 4 we set DoY = 30 and $\overline{R}$ = 90 in the calculation of *N*. The results show that *N* is very sensitive to the variation of *L*. When *L* varies from 1.2 to 7.0, *N* drops by 3 orders of magnitude. The electron density changes dramatically at the plasmapause, from greater than 1000 to less than 50 as *L* increases by 0.1. At fixed *L*, the electron density is less sensitive to the change in MLT. The electron density in the plume region is roughly one order of magnitude higher than that outside the plume region. The SYM-H index has a significant effect on the locations of the plasmapause and the plume region. During low geomagnetic activity (e.g. SYM-$H_{m24}$ = −5 nT), the plasmapause is located at $L \approx 4.8$ in the sector

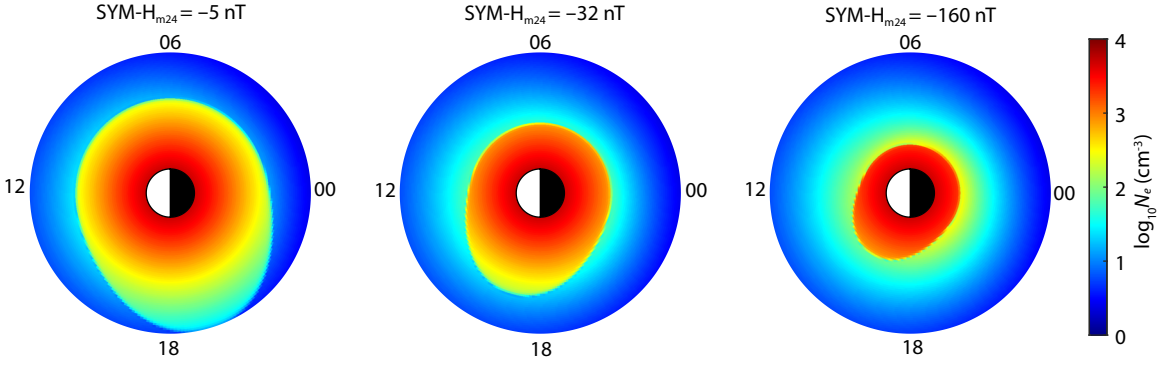


**Figure 4**. Simulated equatorial electron densities under different geomagnetic activities.

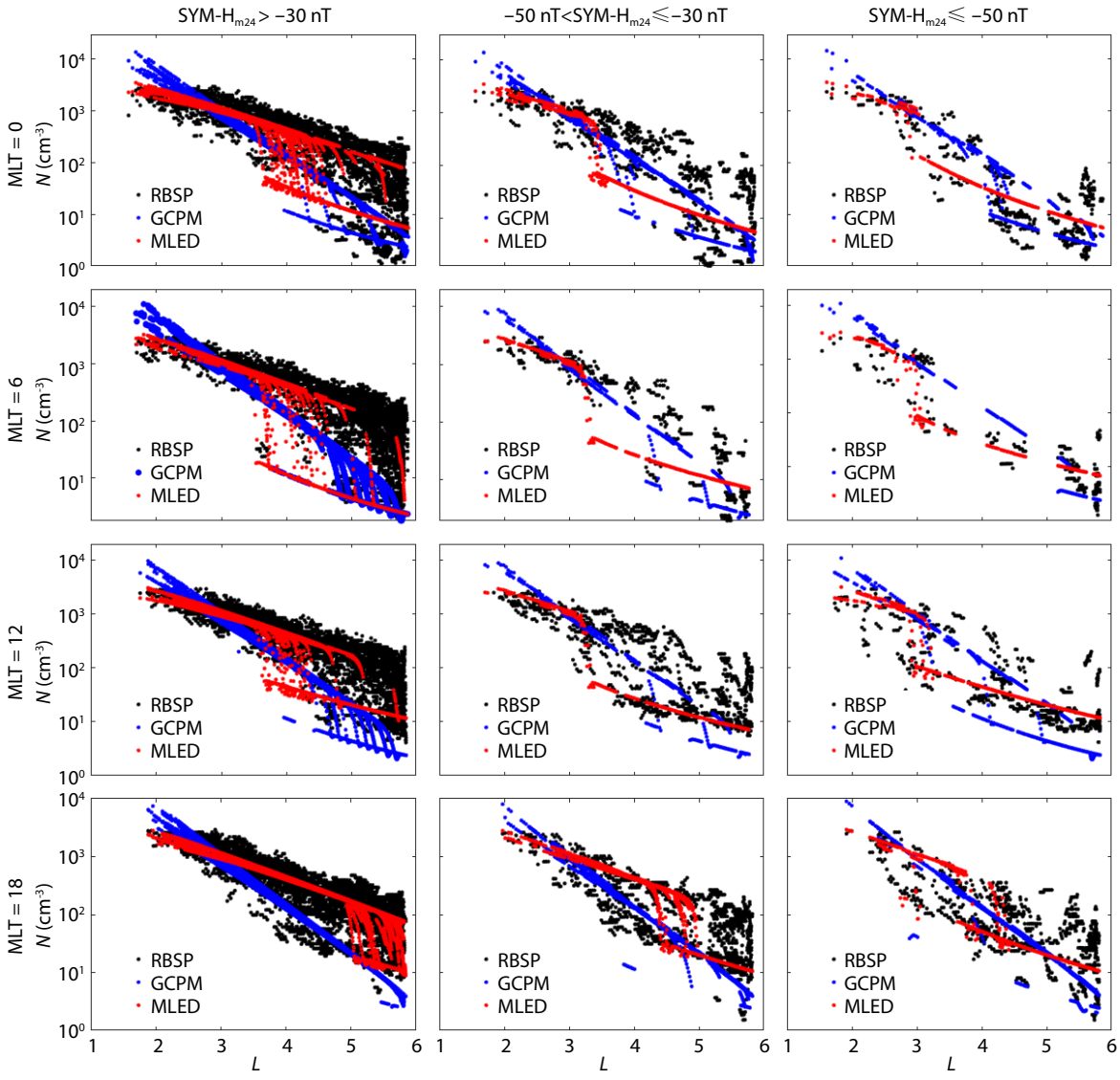


**Figure 5**. Comparison of the simulation results of MLED (red dots) and GCPM (blue dots) models with the measured densities (black dots) for four different magnetic local times under different geomagnetic activities.

**Table 3.** Comparison of the $R_a^2$ coefficient of MLED and GCPM models for different MLT.

| | | MLT = 0 | MLT = 6 | MLT = 12 | MLT = 18 |
|---|---|---|---|---|---|
| SYM-H$_{m24}$ > −30 nT | MLED | 0.7267 | 0.7444 | 0.7875 | 0.8102 |
| | GCPM | −0.6418 | −0.4475 | 0.2170 | 0.3847 |
| −50 nT < SYM-H$_{m24}$ ≤ −30 nT | MLED | 0.6873 | 0.7685 | 0.6784 | 0.6994 |
| | GCPM | −0.9493 | −1.1429 | 0.1843 | 0.3661 |
| SYM-H$_{m24}$ ≤ −50 nT | MLED | −0.7856 | 0.4437 | 0.4910 | 0.6311 |
| | GCPM | −2.2801 | −2.9358 | −2.0681 | −0.2622 |

from midnight to post-noon and there is a high-density plume on the duskside. During active geomagnetic storms, the plasmapause is compressed toward the Earth and the plume region rotates westward, with a pattern similar to that of the observational results.

We extracted samples from the data set at midnight (MLT = 0), dawn (MLT = 6), noon (MLT = 12) and dusk (MLT = 18) during different levels of geomagnetic activity. We use the measured data ($L$, MLT, $\overline{R}$, DoY, $Kp$ index, and SYM-H index) as input parameters for the GCPM and MLED models, and compare the simulation results with the measured densities. As shown in Figure 5, in the region of $L$ < 2.5 the results of the MLED model are generally

**Table 4.** Comparison of the RMSE of MLED and GCPM models for different MLT.

| | | MLT = 0 | MLT = 6 | MLT = 12 | MLT = 18 |
|---|---|---|---|---|---|
| $SYM\text{-}H_{m24} > -30$ nT | MLED | 291.6 | 223.2 | 231.9 | 215.4 |
| | GCPM | 714.6 | 531.2 | 445.2 | 387.9 |
| $-50$ nT $< SYM\text{-}H_{m24} \leq -30$ nT | MLED | 304.5 | 305.6 | 257.8 | 211.8 |
| | GCPM | 760.3 | 929.8 | 410.6 | 307.6 |
| $SYM\text{-}H_{m24} \leq -50$ nT | MLED | 213.2 | 417.1 | 320.3 | 231.9 |
| | GCPM | 834.0 | 1109 | 786.3 | 428.9 |

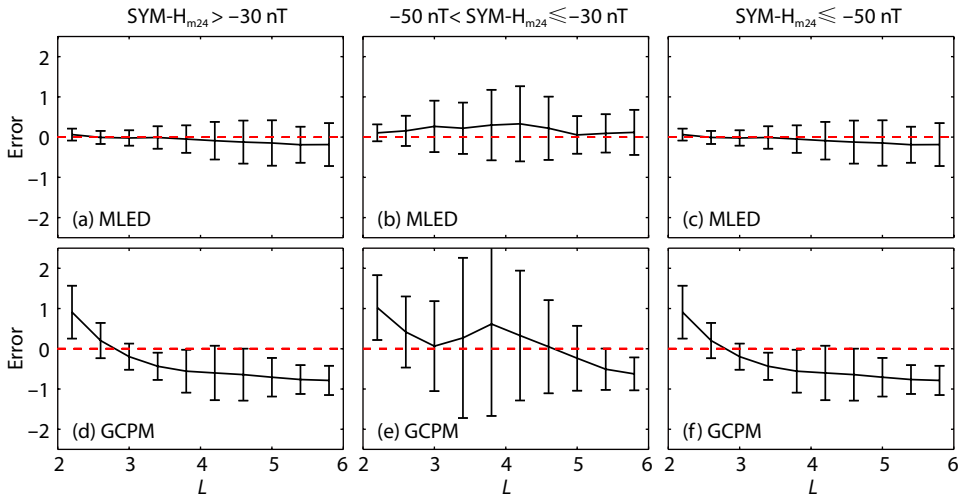


**Figure 6**. The distribution of the relative error as a function of *L* for (a–c) MLED and (d–f) GCPM models under different geomagnetic activities.

consistent with the measured densities while the GCPM model yields significantly higher densities. The MLED model better presents the pattern of decreasing density with increasing *L*. In contrast, the density calculated by the GCPM model decreases rapidly with increasing *L*, much lower than the measured density for $L > 5$ in most cases. Furthermore, the MLED model provides a better picture, than does the GCPM model, of how the plasmapause moves earthward with increasing levels of geomagnetic activity.

The $R_a^2$ coefficient and RMSE of GCPM and MLED models for MLT = 0, 6, 12 and 18 are listed in Tables 3 and 4, respectively. Under all conditions, the $R_a^2$ coefficient (RMSE) of the MLED model is higher (lower) than that of the GCPM model. The $R_a^2$ coefficient of the GCPM model is less than 0 in most cases and its maximum value is 0.3847. The minimum value of the $R_a^2$ coefficient of the MLED model is 0.4437 and the maximum value is 0.8102. The ratio of the RMSE of the MLED model to that of the GCPM model is in the range of 0.26 to 0.69.

Figure 6 shows the distribution of the relative error as a function of *L* for MLED and GCPM models under different geomagnetic activities. The mean relative errors and standard deviations are illustrated as bar plots. The relative error is defined as the ratio of the absolute error of the modeled value to the observed value: $(N_{model} - N_{obs})/N_{obs}$. The standard deviation is used to estimate the uncertainty of the density model predictions. The absolute values of mean relative errors of the MLED model are less than ~0.3 and significantly smaller than those of the GCPM model at different *L* shells. Figures 6d–6f further confirm the results of Figure 5 that the GCPM model produces overestimates at lower *L* shells and underestimates at higher *L* shells. The MLED model can be incorporated into the previously developed radiation belt models (e.g. Glauert and Horne, 2005; Xiao FL et al., 2009, 2010; Su ZP et al., 2010; Shprits et al., 2015) to better forecast the dynamic evolution of energetic electrons in the radiation belt.

## 4. Results and Discussion

During the period from September 25, 2012 to August 30, 2019, Van Allen Probes covered all magnetic local times in the low latitude region of $L = 1.1 - 6.0$ and provided a good opportunity to measure electron density on the global scale. We develop a machine-learning-based model of the equatorial electron density (MLED) in the inner magnetosphere based on density data from the Van Allen Probes mission. This MLED model is a physics-based nonlinear network. Gradient descent and backpropagation algorithms are used to find the optimal parameters of the nonlinear network. This approach can efficiently treat nonlinear relationships among physical quantities in space plasma environments. Compared to recently developed neural network density models (Chu XN et al., 2017a, b; Zhelavskaya et al., 2017; Bortnik et al., 2018), this MLED model provides explicit expressions with very few parameters and is therefore easily described and deployed.

The MLED model describes the variation of electron density in the

inner magnetosphere as a function of $L$, MLT, SYM-H index, DoY, and $\bar{R}$. Different from previous neural network models, it employs fundamental physical principles to describe the variation of electron density. It models the electron densities of the plasmasphere and trough separately, and predicts the plasmapause location under different geomagnetic activity levels. Moreover, this MLED model uses cosine functions to describe the density variation with DoY and MLT based on the periodic variation of electron density with season and MLT. Under different geomagnetic conditions, the electron densities calculated by this model agree remarkably well with the Van Allen Probes measurements and present a better picture of the inward movement of the plasmapause with increasing geomagnetic activity. This MLED model should improve the accuracy of forecasting the dynamic evolution of energetic electrons in the radiation belt.

It should be mentioned that the MLED model adopts a concise network structure with the input parameters $L$, MLT, $\bar{R}$, DoY, and SYM-$H_{m24}$ index. We use the minimum SYM-H in the preceding 24 hours (SYM-$H_{m24}$) to represent the prevailing geomagnetic activity level. This means that the MLED model can hardly reflect the plasmaspheric evolution on a short time scale, such as the highly dynamic variation of the plasmasphere on a timescale of several hours during the main phase of geomagnetic storms, and the plasmaspheric reformation on a timescale of several minutes during substorms (Su ZP et al., 2018). To reflect variations of the plasmasphere during geomagnetic storms and substorms, and thus make the results more accurate, it may be necessary to employ a more complex neural network (e.g., Recurrent Neural Network with short-term memory capabilities) with higher temporal resolution sequences of SYM-H and *AE* indices as input parameters. Increasing the number of input parameters and deepening the number of layers can be expected to improve the performance of the model, but doing so may not be cost-effective; for example, increasing $R$ from ~0.9 to ~0.95, especially when solving large-scale problems, may incur high computational costs while yielding limited performance improvement. In the future, we will investigate how additional input parameters and layers of the network would affect the performance of the model, carefully balancing the performance benefits of each modification against its computational cost.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China grants 42074198, 41774194, 41974212 and 42004141, Natural Science Foundation of Hunan Province 2021JJ20010, Science and Technology Innovation Program of Hunan Province 2021RC3098, and Foundation of Education Bureau of Hunan Province for Distinguished Young Scientists 20B004. All the Van Allen Probes data are publicly available at https://cdaweb.gsfc.nasa.gov/pub/data/rbsp/. The OMNI data are obtained online (https://spdf.gsfc.nasa.gov/pub/data/omni/).

## References

Bortnik, J., Li, W., Thorne, R. M., and Angelopoulos, V. (2016). A unified approach to inner magnetospheric state prediction. *J. Geophys. Res.*: *Space Phys.*, *121*(3), 2423–2430. https://doi.org/10.1002/2015JA021733

Bortnik, J., Chu, X. N., Ma, Q. L., Li, W., Zhang, X. J., Thorne, R. M., Angelopoulos, V., Denton, R. E., Kletzing, C. A., … Baker, D. N. (2018). Artificial neural networks for determining magnetospheric conditions. In E. Camporeale, et al. (Eds.), *Machine Learning Techniques for Space Weather* (pp. 279-300). Amsterdam, Netherlands: Elsevier. https://doi.org/10.1016/B978-0-12-811788-0.00011-1

Carpenter, D. L., and Anderson, R. R. (1992). An ISEE/whistler model of equatorial electron density in the magnetosphere. *J. Geophys. Res.*: *Space Phys.*, *97*(A2), 1097–1108. https://doi.org/10.1029/91JA01548

Chappell, C. R. (1972). Recent satellite measurements of the morphology and dynamics of the plasmasphere. *Rev. Geophys.*, *10*(4), 951–979. https://doi.org/10.1029/RG010i004p00951

Chen, L. J., Thorne, R. M., and Horne, R. B. (2009). Simulation of EMIC wave excitation in a model magnetosphere including structured high-density plumes. *J. Geophys. Res.*: *Space Phys.*, *114*(A7), A07221. https://doi.org/10.1029/2009JA014204

Chu, X. N., Bortnik, J., Li, W., Ma, Q., Angelopoulos, V., and Thorne, R. M. (2017a). Erosion and refilling of the plasmasphere during a geomagnetic storm modeled by a neural network. *J. Geophys. Res.*: *Space Phys.*, *122*(7), 7118–7129. https://doi.org/10.1002/2017JA023948

Chu, X. N., Bortnik, J., Li, W., Ma, Q. L., Denton, R., Yue, C., Angelopoulos, V., Thorne, R. M., Darrouzet, F., … Menietti, J. (2017b). A neural network model of three-dimensional dynamic electron density in the inner magnetosphere. *J. Geophys. Res.*: *Space Phys.*, *122*(9), 9183–9197. https://doi.org/10.1002/2017JA024464

Darrouzet, F., De Keyser, J., Décréau, P. M. E., El Lemdani-Mazouz, F., and Vallières, X. (2008). Statistical analysis of plasmaspheric plumes with Cluster /WHISPER observations. *Ann. Geophys.*, *26*(8), 2403–2417. https://doi.org/10.5194/angeo-26-2403-2008

Escoubet, C. P., Pedersen, A., Schmidt, R., and Lindqvist, P. A. (1997). Density in the magnetosphere inferred from ISEE 1 spacecraft potential. *J. Geophys. Res.*: *Space Phys.*, *102*(A8), 17595–17609. https://doi.org/10.1029/97JA00290

Fu, H. S., Tu, J., Song, P., Cao, J. B., Reinisch, B. W., and Yang, B. (2010a). The nightside-to-dayside evolution of the inner magnetosphere: Imager for Magnetopause-to-Aurora Global Exploration Radio Plasma Imager observations. *J. Geophys. Res.*: *Space Phys.*, *115*(A4), A04213. https://doi.org/10.1029/2009JA014668

Fu, H. S., Tu, J., Cao, J. B., Song, P., Reinisch, B. W., Gallagher, D. L., and Yang, B. (2010b). IMAGE and DMSP observations of a density trough inside the plasmasphere. *J. Geophys. Res.*: *Space Phys.*, *115*(A7), A07227. https://doi.org/10.1029/2009JA015104

Gallagher, D. L., Craven, P. D., Comfort, R. H., and Moore, T. E. (1995). On the azimuthal variation of core plasma in the equatorial magnetosphere. *J. Geophys. Res.*: *Space Phys.*, *100*(A12), 23597–23605. https://doi.org/10.1029/95JA02100

Gallagher, D. L., Craven, P. D., and Comfort, R. H. (2000). Global core plasma model. *J. Geophys. Res.*: *Space Phys.*, *105*(A8), 18819–18833. https://doi.org/10.1029/1999JA000241

Glauert, S. A., and Horne, R. B. (2005). Calculation of pitch angle and energy diffusion coefficients with the PADIE code. *J. Geophys. Res.*: *Space Phys.*, *110*(A4), A04206. https://doi.org/10.1029/2004JA010851

Goldstein, J., Sandel, B. R., Thomsen, M. F., Spasojević, M., and Reiff, P. H. (2004). Simultaneous remote sensing and in situ observations of plasmaspheric drainage plumes. *J. Geophys. Res.*: *Space Phys.*, *109*(A3), A03202. https://doi.org/10.1029/2003JA010281

Goldstein, J., Thomsen, M. F., and DeJong, A. (2014). In situ signatures of residual plasmaspheric plumes: Observations and simulation. *J. Geophys. Res.*: *Space Phys.*, *119*(6), 4706–4722. https://doi.org/10.1002/2014JA019953

Guan, C. Y., Shang, X. J., Xie, Y. Q., Yang, C., Zhang, S., Liu, S., and Xiao, F. L. (2020). Generation of simultaneous $H^+$ and $He^+$ band EMIC waves in the nightside radiation belt. *Sci. China Technol. Sci.*, *63*(11), 2369–2374. https://doi.org/10.1007/s11431-019-1545-6

Guo, M. Y., Zhou, Q. H., Xiao, F. L., Liu, S., He, Y. H., and Yang, C. (2020). Upward propagation of lightning-generated whistler waves into the radiation belts. *Sci. China Technol. Sci.*, *63*(2), 243–248. https://doi.org/10.1007/s11431-018-9486-9

He, J. B., Jin, Y. Y., Xiao, F. L., He, Z. G., Yang, C., Xie, Y. Q., He, Q., Wang, C. Z.,

Shang, X. J., … Zhang, S. (2021). The influence of various frequency chorus waves on electron dynamics in radiation belts. *Sci. China Technol. Sci.*, *64*(4), 890–897. https://doi.org/10.1007/s11431-020-1750-6

Hecht-Nielsen, R. (1989). Theory of the backpropagation neural network. In *International 1989 Joint Conference on Neural Networks* (pp. 593-605). Washington, DC, USA: IEEE. https://doi.org/10.1109/IJCNN.1989.118638

Huang, Y., Dai, L., Wang, C., Xu, R. L., and Li, L. (2021). A new inversion method for reconstruction of plasmaspheric $He^+$ density from EUV images. *Earth Planet. Phys.*, *5*(2), 218–222. https://doi.org/10.26464/epp2021020

Kurth, W. S., De Pascuale, S., Faden, J. B., Kletzing, C. A., Hospodarsky, G. B., Thaller, S., and Wygant, J. R. (2015). Electron densities inferred from plasma wave spectra obtained by the Waves instrument on Van Allen Probes. *J. Geophys. Res.*: *Space Phys.*, *120*(2), 904–914. https://doi.org/10.1002/2014JA020857

Larsen, B. A., Klumpar, D. M., and Gurgiolo, C. (2007). Correlation between plasmapause position and solar wind parameters. *J. Atmos. Sol. -Terr. Phys.*, *69*(3), 334–340. https://doi.org/10.1016/j.jastp.2006.06.017

Meredith, N. P., Horne, R. B., Sicard-Piet, A., Boscher, D., Yearby, K. H., Li, W., and Thorne, R. M. (2012). Global model of lower band and upper band chorus from multiple satellite observations. *J. Geophys. Res.*: *Space Phys.*, *117*(A10), A10225. https://doi.org/10.1029/2012JA017978

O'Brien, T. P., and Moldwin, M. B. (2003). Empirical plasmapause models from magnetic indices. *Geophys. Res. Lett.*, *30*(4), 1152. https://doi.org/10.1029/2002GL016007

Ojala, M., and Garriga, G. C. (2009). Permutation tests for studying classifier performance. In *2009 Ninth IEEE International Conference on Data Mining* (pp. 908-913). Miami Beach, FL, USA: IEEE. https://doi.org/10.1109/ICDM.2009.108

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Sauer, K., Baumgärtel, K., and Sydora, R. (2020). Gap formation around $\Omega_e/2$ and generation of low-band whistler waves by Landau-resonant electrons in the magnetosphere: Predictions from dispersion theory. *Earth Planet. Phys.*, *4*(2), 138–150. https://doi.org/10.26464/epp2020020

Sheeley, B. W., Moldwin, M. B., Rassoul, H. K., and Anderson, R. R. (2001). An empirical plasmasphere and trough density model: CRRES observations. *J. Geophys. Res.*: *Space Phys.*, *106*(A11), 25631–25641. https://doi.org/10.1029/2000JA000286

Shprits, Y. Y., Kellerman, A. C., Drozdov, A. Y., Spence, H. E., Reeves, G. D., and Baker, D. N. (2015). Combined convective and diffusive simulations: VERB-4D comparison with 17 March 2013 Van Allen Probes observations. *Geophys. Res. Lett.*, *42*(22), 9600–9608. https://doi.org/10.1002/2015GL065230

Su, Z. P., Xiao, F. L., Zheng, H. N., and Wang, S. (2010). STEERB: A three-dimensional code for storm-time evolution of electron radiation belt. *J. Geophys. Res.*: *Space Phys.*, *115*(A9), A09208. https://doi.org/10.1029/2009JA015210

Su, Z. P., Liu, N. G., Zheng, H. N., Wang, Y. M., and Wang, S. (2018). Multipoint observations of nightside plasmaspheric hiss generated by substorm-injected electrons. *Geophys. Res. Lett.*, *45*(20), 10921–10932. https://doi.org/10.1029/2018GL079927

Thorne, R. M., Smith, E. J., Burton, R. K., and Holzer, R. E. (1973). Plasmaspheric hiss. *J. Geophys. Res.*, *78*(10), 1581–1596. https://doi.org/10.1029/JA078i010p01581

Wang, J. Z., Zhu, Q., Gu, X. D., Fu, S., Guo, J. G., Zhang, X. X., Yi, J., Guo, Y. J., Ni, B. B., and Xiang, Z. (2020). An empirical model of the global distribution of plasmaspheric hiss based on Van Allen Probes EMFISIS measurements. *Earth Planet. Phys.*, *4*(3), 246–265. https://doi.org/10.26464/epp2020034

Xiao, F. L., Su, Z. P., Zheng, H. N., and Wang, S. (2009). Modeling of outer radiation belt electrons by multidimensional diffusion process. *J. Geophys. Res.*: *Space Phys.*, *114*(A3), A03201. https://doi.org/10.1029/2008JA013580

Xiao, F. L., Su, Z. P., Zheng, H. N., and Wang, S. (2010). Three-dimensional simulations of outer radiation belt electron dynamics including cross-diffusion terms. *J. Geophys. Res.*: *Space Phys.*, *115*(A5), A05216. https://doi.org/10.1029/2009JA014541

Xiao, F. L., Zhou, Q. H., He, Z. G., Yang, C., He, Y. H., and Tang, L. J. (2013). Magnetosonic wave instability by proton ring distributions: Simultaneous data and modeling. *J. Geophys. Res.*: *Space Phys.*, *118*(7), 4053–4058. https://doi.org/10.1002/jgra.50401

Xiao, F. L., Yang, C., Su, Z. P., Zhou, Q. H., He, Z. G., He, Y. H., Baker, D. N., Spence, H. E., Funsten, H. O., and Blake, J. B. (2015). Wave-driven butterfly distribution of Van Allen belt relativistic electrons. *Nat. Commun.*, *6*, 8590. https://doi.org/10.1038/ncomms9590

Yang, C., Wang, Z. Q., Xiao, F. L., He, Z. G., Xie, Y. Q., Zhang, S., He, Y. H., Liu, S., and Zhou, Q. H. (2021). Correlated observations linking loss of energetic protons to EMIC waves. *Sci. China Technol. Sci.*, *65*(1), 131–138. https://doi.org/10.1007/s11431-021-1882-x

Zhelavskaya, I. S., Shprits, Y. Y., and Spasojević, M. (2017). Empirical modeling of the plasmasphere dynamics using neural networks. *J. Geophys. Res.*: *Space Phys.*, *122*(11), 11227–11244. https://doi.org/10.1002/2017JA024406

Zhou, Q. H., Xiao, F. L., Yang, C., Liu, S., He, Y. H., Baker, D. N., Spence, H. E., Reeves, G. D., and Funsten, H. O. (2017). Generation of lower and upper bands of electrostatic electron cyclotron harmonic waves in the Van Allen radiation belts. *Geophys. Res. Lett.*, *44*(11), 5251–5258. https://doi.org/10.1002/2017GL073051