

A modeling study of ≥ 2 MeV electron fluxes in GEO at different prediction time scales based on LSTM and transformer networks

Xiaojing Sun^{1,2,3,4} , Dedong Wang², Alexander Drozdov⁵, Ruilin Lin^{1,*}, Artem Smirnov^{2,6}, Yuri Shprits^{2,5,6}, Siqing Liu^{1,3}, Bingxian Luo^{1,3}, and Xi Luo⁷

¹ State Key Laboratory of Space Weather, National Space Science Center, Chinese Academy of Sciences, 100190 Beijing, China

² GFZ German Research Centre for Geosciences, 14473 Potsdam, Germany

³ University of Chinese Academy of Sciences, 101499 Beijing, China

⁴ System Research Institute, Deep Space Exploration Lab, 100043 Beijing, China

⁵ Department of Earth, Planetary, and Space Sciences, University of California, 138307 Los Angeles, CA, USA

⁶ Institute of Physics and Astronomy, University of Potsdam, 14469 Potsdam, Germany

⁷ Shandong Institute of Advanced Technology, 250100 Jinan, China

Received 16 October 2023 / Accepted 10 June 2024

Abstract—In this study, we develop models to predict the $\log_{10}$ of ≥ 2 MeV electron fluxes with 5-minute resolution at the geostationary orbit using the Long Short-Term Memory (LSTM) and transformer neural networks for the next 1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions. The data of the GOES-10 satellite from 2002 to 2003 are the training set, the data in 2004 are the validation set, and the data in 2005 are the test set. For different prediction time scales, different input combinations with 4 days as best offset time are tested and it is found that the transformer models perform better than the LSTM models, especially for higher flux values. The best combinations for the transformer models for next 1-hour, 3-hour, 6-hour, 12-hour, 1-day predictions are ($\log_{10}$ Flux, MLT), ($\log_{10}$ Flux, *Bt*, *AE*, *SYM-H*), ($\log_{10}$ Flux, *N*), ($\log_{10}$ Flux, *N*, *Dst*, *Lm*), and ($\log_{10}$ Flux, *Pd*, *AE*) with PE values of 0.940, 0.886, 0.828, 0.747, and 0.660 in 2005, respectively. When the low flux outliers of the ≥ 2 MeV electron fluxes are excluded, the prediction efficiency (PE) values for the 1-hour and 3-hour predictions increase to 0.958 and 0.900. By evaluating the prediction of ≥ 2 MeV electron daily and hourly fluences, the PE values of our transformer models are 0.857 and 0.961, respectively, higher than those of previous models. In addition, our models can be used to fill the data gaps of ≥ 2 MeV electron fluxes.

Keywords: Prediction model of ≥ 2 MeV electron fluxes / Geostationary orbit / Machine learning / Transformer / LSTM

1 Introduction

The ≥ 2 MeV electron flux at geostationary (GEO) orbit is a critical parameter for GEO satellites and is used as a significant indicator of increased risk of internal charging (Wrenn & Sims, 1996; Gubby & Evans, 2002; Wrenn et al., 2002; Romanova et al., 2005; Pilipenko et al., 2006; Horne et al., 2013; Lai et al., 2018). Energetic electrons can cause significant anomalies, leading to temporary or permanent loss of satellite functions, such as interruption of communications and degradation of navigation precision (Reagan et al., 1983; Baker et al., 1987; Violet & Frederickson, 1993; Lanzerotti et al., 1998; Lucci et al., 2005; Ryden et al., 2008; Lohmeyer et al., 2015;

Singh et al., 2021). The prediction of ≥ 2 MeV electron daily fluences for the next 3 days is one of the indispensable contents of space environment predictions. The alerts of relativistic electron enhancement events are triggered when ≥ 2 MeV electron daily fluences at GEO orbit exceed $10^8 \text{ cm}^{-2} \cdot \text{sr}^{-1} \cdot \text{day}^{-1}$.

It is essential to understand the distribution of high-energy electrons and make reliable predictions of the radiation environment around spacecraft. With the increasing number of satellites and their growing importance to our lives, a lot of effort over the last two decades has been devoted to understanding the distribution of high-energy electrons and to making reliable predictions. There have been several studies conducted on the mechanisms for the acceleration of relativistic electrons in the radiation belt, such as the ONERA Salammbô model (Beutier & Boscher, 1995; Varotsou et al., 2005, 2008; Maget et al., 2007; Bourdarie

*Corresponding author: linrl@nsssc.ac.cn

& Maget, 2012), the British Antarctic Survey (BAS) Radiation Belt model (Glauret et al., 2014a, b; Kersten et al., 2014; Allison et al., 2019), the Versatile Electron Radiation Belt (VERB) model (Shprits et al., 2008a, b, 2009, 2013, 2015; Subbotin & Shprits, 2009; Kim et al., 2011, 2012; Subbotin et al., 2010, 2011; Pakhotin et al., 2014; Drozdov et al., 2017, 2021; Wang & Shprits, 2019; Wang et al., 2020) and the Dynamic Radiation Belt Environment Assimilation (DREAM-3D) Model (Reeves et al., 2012; Tu et al., 2013). Radial diffusion processes driven by ULF (Ultra-low Frequency) waves and localized electron acceleration due to resonant interactions with whistler mode chorus waves are important mechanisms for electron acceleration (Horne & Thorne, 1998; Summers et al., 1998; Brautigam & Albert, 2000; Friedel et al., 2002; Meredith et al., 2002; Li et al., 2001; Li, 2004; Li et al., 2006, 2007, 2011; Miyoshi et al., 2003; Horne et al., 2005; Shprits et al., 2006a, b, 2008a, b, 2009, 2018, 2022; Albert, 2007, 2008; Millan & Thorne, 2007; Anderson et al., 2015; Li et al., 2016; Ma et al., 2018).

In addition to the physics-based models, empirical methods have also been employed for the prediction of energetic electrons at GEO orbit, such as Paulikas & Blake (1979), Baker et al. (1990), O'Brien et al. (2001), Burin des Roziers & Li (2006), Li et al. (2001), Rigler et al. (2004), Turner & Li (2008), Reeves et al. (2011), He et al. (2013), Sakaguchi et al. (2013), Potapov et al. (2014), Li et al. (2017), Qian et al. (2020), and Landis et al. (2022). The empirical models often require solar wind parameters, geomagnetic indices, and low-energy or medium-energy electron fluxes in the previous days to predict ≥ 2 MeV electron daily fluences for the next 1–3 days.

With the rapid progress of artificial intelligence, machine learning models have been applied to predict high-energy electron fluxes at GEO orbit. Many machine-learning methods have also been applied, such as Fukata et al. (2002), Ukhorskiy et al. (2004), Xue & Ye (2004), Ling et al. (2010), Balikhin et al. (2011), Balikhin et al. (2016), Wei et al. (2011), Wang & Shi (2012), Boynton et al. (2013), Boynton et al. (2015), Guo et al. (2013), Pakhotin et al. (2014), Ganushkina et al. (2014), Ganushkina et al. (2015), Shin et al. (2016), Wei et al. (2018), Zhang et al. (2020), Katsavrias et al. (2022), Landis et al. (2022), Saikin et al. (2021), and Son et al. (2022). Wei et al. (2018) and Sun et al. (2023) used the Long Short-Term Memory (LSTM) network to predict the ≥ 2 MeV electron daily fluence at GEO orbit for the next day and the next 3 days, respectively.

The previous prediction models are mainly focused on the prediction of ≥ 2 MeV electron daily fluences. Limited studies have been focused on the prediction of ≥ 2 MeV electron fluxes with a 5-minute resolution. Li et al. (2017) developed the model by the Empirical Orthogonal Function (EOF) to give ≥ 2 MeV electron fluxes with 5-minute resolution on the following day. The EOF coefficients are fitted by the solar wind parameters and geomagnetic indices. The prediction efficiency (PE) from January 2003 to June 2006 is 0.67. Landis et al. (2022) created a NARX (nonlinear autoregressive with exogenous input) neural network to model ≥ 0.8 MeV electron fluxes with 5-minute resolution from GOES-15 satellite. The NARX model performs well from June 2013 to June 2016, with a linear correlation (LC) coefficient of 0.68 and a PE value of 0.39.

Observation data from satellites often contain data gaps for ≥ 2 MeV electron fluxes. However, continuous data sets are not only useful for analyzing the dynamic distribution of electrons

in the radiation belt but also for model validation. Unfortunately, there is little attention on filling the data gaps of ≥ 2 MeV electron fluxes, but instead, research has focused on filling the data gaps in solar wind parameters and geomagnetic indices. Qin et al. (2007) developed a decorrelation-time-based approach to interpolate the solar-wind characteristics across data gaps and to evaluate parameters needed for global empirical magnetic models (Tsyganenko & Sitnov, 2005). Kondrashov et al. (2005) and Kondrashov & Ghil (2006) developed a gap-filling method based on the Singular Spectrum Analysis (SSA) method, which is mainly on the presence of significant oscillatory modes in the time series. Kondrashov et al. (2010, 2011, 2014) and Shprits et al. (2012, 2013) also applied the SSA method to reconstruct the solar wind data set covering the long-term time interval of the Combined Release and Radiation Effects Satellite (CRRES) by using the combination of solar wind factors, IMF data, and geomagnetic indices as inputs, and the method was used in several data assimilation and statistical studies of the radiation belts.

In this study, we develop models to predict ≥ 2 MeV electron fluxes with 5-minute resolution using the LSTM and transformer networks. Various combinations of ≥ 2 MeV electron fluxes, solar wind parameters, and geomagnetic indices as inputs are discussed, and the model performances are evaluated with different prediction time scales in 2005. These models can fill the data gaps of ≥ 2 MeV electron fluxes at GEO orbit. LSTM and transformer networks are two commonly used deep learning methods in processing sequential data. LSTM, a type of recurrent neural network, uses gate units to control the flow of information, thereby effectively avoiding the problem of vanishing or exploding gradients. In contrast, the transformer is a neural network that can achieve encoding and decoding of sequences based on attention mechanisms. When processing long sequences, the performance of the transformer model is usually better than that of the LSTM model. Considering the excellent performances of LSTM and transformer networks in processing sequential data, we used the two methods to develop the models.

This paper is structured as follows: Section 2 introduces the data, the LSTM and transformer networks, and the indices for model evaluation. In Section 3.1, we evaluate the performance of models with different offset times as inputs and determine the best offset time for modeling. Sections 3.2 and 3.3 focus on the performances of the LSTM and the transformer models. We evaluate the performances of the models with different prediction time scales and with different parameters as inputs and compare our models with other models in Section 3.4. Section 4 discusses the performances of the prediction models in different situations. The summary and conclusions are given in Section 5.

2 Data and methods

2.1 Data

The data used in this study include ≥ 2 MeV electron fluxes from GOES satellites, solar wind parameters, geomagnetic disturbance indices, magnetopause subsolar distance (R_0), L-shell (L_m), and magnetic local time (MLT) values between 2002 and 2005. The data of GOES satellites with 5-minute resolution

were provided by the National Centers for Environmental Information (NCEI) (<https://www.ngdc.noaa.gov/stp/satellite/goes/>). Solar wind parameters and geomagnetic disturbance indices were provided by OMNI (<https://cdaweb.gsfc.nasa.gov/>). The $R0$ indices were calculated by Lin et al. (2010) model. The Lm and MLT values were calculated using the international radiation environment modeling software library (IRBEM) (<https://github.com/PRBEM/IRBEM>).

The GOES satellites are a series of geostationary environmental satellites, continuously monitoring the energetic electron fluxes from 1974 (Grubb, 1975), which are managed by the National Oceanic and Atmospheric Administration (NOAA). Based on the ≥ 2 MeV electron fluxes with 5-minute resolution from GOES satellites, the proportion of missing data in different GOES satellites is analyzed. The data missing ratios of ≥ 2 MeV electron fluxes from GOES-8, GOES-9, GOES-10, GOES-11, and GOES-12 satellites are 6.6%, 6.9%, 5.4%, 4.7%, and 13.9%, respectively. The data missing ratios for eastward (westward) detectors from GOES-13 to GOES-15 satellites are 0.53% (0.45%), 1.26% (0.72%), and 0.43% (0.43%), respectively. The data qualities of GOES-13 to GOES-15 satellites are significantly better than those of previous GOES satellites. Considering that the model will also be used to fill the data gaps, the data will be from GOES-8 to GOES-12 satellites. We also considered the satellite location and the variation in ≥ 2 MeV electron fluxes at different longitudes. It is shown in Figure 1 of Sun et al. (2021) that most GOES satellites adjusted their locations during operation. For instance, the GOES-10 satellite operated from July 1998 to December 2009, and it shifted from around 135°W to about 60°W between July 2006 and November 2006. Sun et al. (2021) also showed that the ratios of ≥ 2 MeV electron daily fluences ($\text{cm}^{-2} \cdot \text{sr}^{-1} \cdot \text{day}^{-1}$) from GOES-10 at about 135°W to those from GOES-12 at about 75°W are mainly in the range of 1.0–4.0, with an average of 1.92. Due to the long duration of the GOES-10 satellite operating at the same fixed longitude (135°W), we ultimately chose the GOES-10 satellite.

Solar wind parameters include solar wind speed (V_{sw}), density (N), dynamic pressure (Pd), the total magnitude of interplanetary magnetic field (Bt), B_x , B_y , and B_z components of interplanetary magnetic field (IMF) in the GSM coordinates, electric field (E), temperature (T), and plasma beta (β). Geomagnetic disturbance indices consist of kp , AE , $SYM-H$, and Dst . Solar wind parameters and AE index are with 5-minute resolution. The Dst index is with 1-hour resolution, and the kp index is with 3-hour resolution. Dst and kp indices are converted to 5-minute resolution by keeping the same values within the 1-hour or 3-hour interval. The solar wind data for calculating $R0$ are described above. The Lm and MLT values are calculated by International Reference Geomagnetic Field (IGRF) + Tsyganenko (T89) models using IRBEM. The IGRF (Macmillan & Finlay, 2010; Thébaud et al., 2015) is the internal geomagnetic field model and the T89 model (Tsyganenko, 1989) is the external geomagnetic field model. Missing values are filled by the linear interpolation method.

The distribution of the missing data of the GOES-10 satellite is shown in Figure 1. There are 2432 data gaps during the operation period of the GOES-10 satellite, and 61.7% of these are 5-minute data gaps, meaning that only one individual value is absent. 94.0% of the data gaps are under an hour, 96.5% are

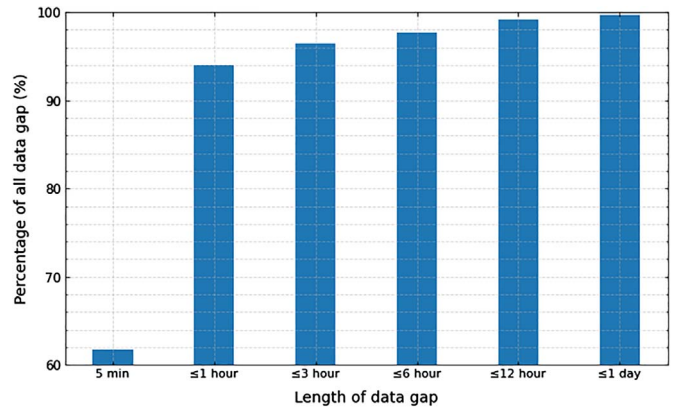


Figure 1. The distribution of the missing data of ≥ 2 MeV electron fluxes from the GOES-10 satellite.

under 3 h, and 99.6% are under 24 h. Therefore, a model for a 1-day prediction is sufficient to fill most of the gaps in the GOES data, and it is more important to improve the performance of the model for 5-minute or 1-hour predictions.

Figure 2a displays the ≥ 2 MeV electron fluxes from the GOES-10 satellite between 1999 and 2009, which is the main operation period of the GOES-10 satellite. The red dashed line indicates that ≥ 2 MeV electron fluxes are equal to $1157 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$, and the corresponding daily fluence is $10^8 \text{ cm}^{-2} \cdot \text{sr}^{-1} \cdot \text{day}^{-1}$, which is the threshold value of a relativistic electron enhancement event. Figures 2b–2d show the longitude of the GOES-10 satellite, the number of days with relativistic electron enhancement events, the number of relativistic electron enhancement events, and the percentage of missing data for each year.

It is obvious that the longitudes of GOES-10 remained stable from 1999 to 2005 as shown in Figure 2b, and the quality of the ≥ 2 MeV electron fluxes from the GOES-10 satellite between 2001 and 2005 is significantly better than that of other years as shown in Figure 2d. To correctly train the machine learning model, the training set should consist of different kinds of space weather phenomena. There are a lot of relativistic electron enhancement events between 2003 and 2005. In addition, the ≥ 2 MeV electron fluxes were relatively low and with relatively few relativistic electron enhancement events in 2002 as shown in Figures 2a and 2c. Finally, we chose the data from 2002 to 2005 as the data set. We use the data from 2002 to 2003 as the training set, the data in 2004 as the validation set, and the data in 2005 as the test set.

2.2 Long short-term memory (LSTM)

The LSTM network is a Recurrent Neural Network (RNN) based architecture that can preserve more input information and significantly reduces the vanishing gradient problem of conventional neural networks (Hochreiter & Schmidhuber, 1997; Graves & Schmidhuber, 2005; Graves, 2012). There are a series of “gates” in the LSTM network, including the Input Gate, Output Gate, and Forget Gate, which can manage to keep, forget, or ignore historical information based on a probabilistic model. The LSTM network is suited for propagating information through long sequences due to its unique structure. Therefore,

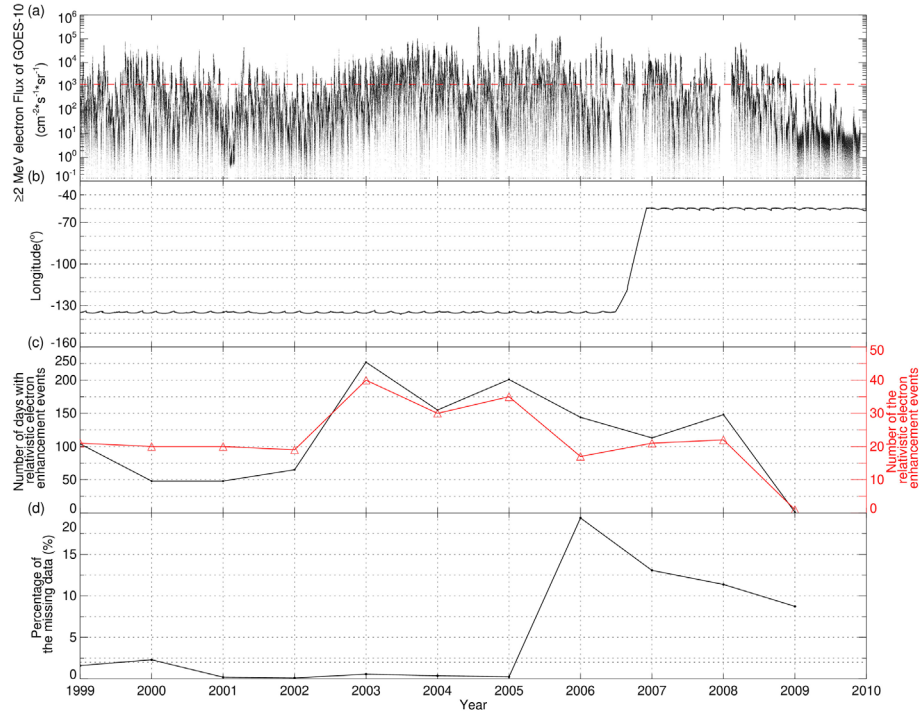


Figure 2. (a) The ≥ 2 MeV electron fluxes from the GOES-10 satellite, (b) the longitudes of GOES-10 satellite, (c) the number of the relativistic electron enhancement events (red line), and the number of days with relativistic electron enhancement events (black line), and (d) the percentage of missing data of the GOES-10 satellite from 1999 to 2010.

it is widely used in natural language processing (NLP) and time series predictions (Gers et al., 2000; Kai et al., 2013; Huang et al., 2015; Cai & Liu, 2016; Greff et al., 2016).

In this study, we develop models by LSTM for filling the data gaps of ≥ 2 MeV electron fluxes from the GOES-10 satellite. The loss function is mean-square error (MSE), the optimizer is AdamOptimizer (Kingma & Ba, 2014), the batch size of the training set is 64, and the learning rate is 1×10^{-4} . Each sliding window is used as an input sequence to predict the next time step's value based on the preceding time steps. Subsequently, the loss is computed by comparing these predicted values against the actual ones, and the model's weights are updated via backpropagation. This iterative training continues until every time window in the training set has been processed. We do not set a fixed epoch value to prevent overfitting. When the model performance stops improving for 10 consecutive epochs, the training stops.

The LSTM models target different prediction time scales, containing 1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions. The 1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions contain the next 12, 36, 72, 144, and 288 data points with 5-minute resolution.

2.3 Transformer

The transformer network is one of the sequence modeling architectures. It has undergone significant progress in recent years, displaying unmatched performance in a variety of applications, including NLP, speech recognition, and computer vision (Vaswani et al., 2017). The transformer network can digest vast sequences of data due to its multi-head self-attention

mechanism, which excels at finding semantic correlations between items in a lengthy sequence. Consequently, it is capable of learning time series data with complicated dynamics that pose a challenge for sequence models (Vaswani et al., 2017; Devlin et al., 2018; Dong et al., 2018; Liu et al., 2021; Zeng et al., 2022).

The transformer network was introduced in 2017 by a Google Brain team and is becoming increasingly popular for NLP issues including machine translation and time series prediction. Transformer network requires a lot of data for training. Moreover, more data is often used during training, which usually leads to better results. Considering their versatility and wide range of applications, we also develop models using the transformer network for filling the data gaps of ≥ 2 MeV electron fluxes from the GOES-10 satellite.

We build a four-layer transformer network model that contains two linear layers, a transformer layer, and an output layer. The loss function is MSE, the optimizer is Adam Optimizer (Kingma & Ba, 2014), and the learning rate is 1×10^{-4} . The transformer models also target different prediction time scales, which are the same as those in Section 2.2.

2.4 Model evaluation

The model performance is evaluated by the PE and the root mean square error (RMSE). They are defined as

$$PE = 1 - \frac{\sum_{i=1}^n (m_i - p_i)^2}{\sum_{i=1}^n (m_i - \bar{m})^2}, \quad (1)$$

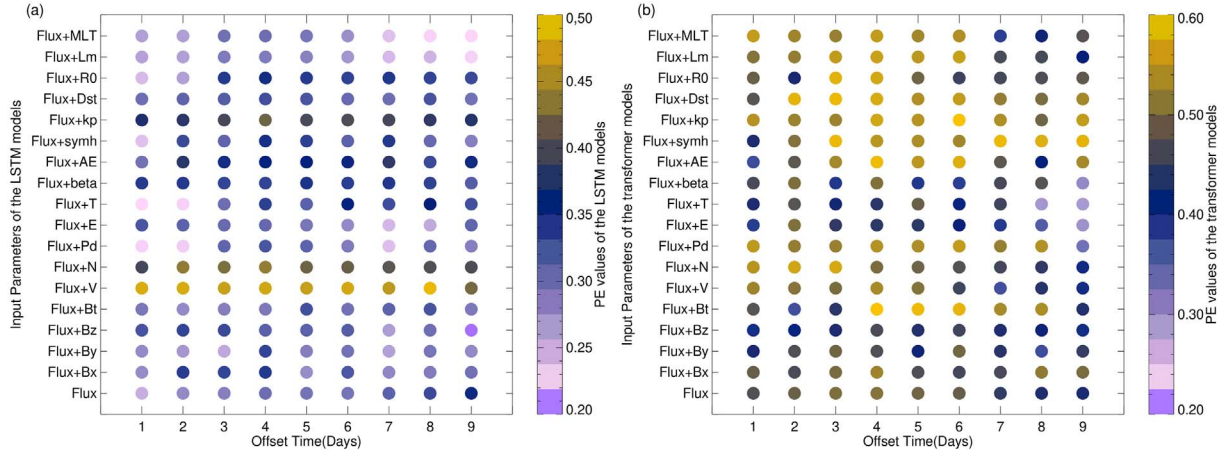


Figure 3. The PE values (color-coded) of the LSTM models (a) and the transformer models (b) for the 1-day prediction with different offset times and different input parameters.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_i - m_i)^2}, \quad (2)$$

where n is the total number of samples, $\bar{m}$ is the mean value of all observation samples, and m_i and p_i are the i th observation and prediction, respectively. The better the model, the lower the RMSE and the higher the PE values. In this study, the m_i is the $\log_{10} (\geq 2 \text{ MeV electron fluxes})$ from observations, and the p_i is the $\log_{10} (\geq 2 \text{ MeV electron fluxes})$ predicted by models.

In addition, the LC coefficient and bias are also used to assess model performances. The bias represents the divergence between the observations and predictions. We calculate the differences between m_i and p_i , add up all the differences, and divide the total value by the total number of predictions to get the bias.

3 The models for predicting $\geq 2 \text{ MeV electron fluxes at GEO orbit}$

The models for predicting $\geq 2 \text{ MeV electron fluxes}$ by the LSTM or transformer network with different prediction time scales are developed by using the $\geq 2 \text{ MeV electron fluxes}$ between 2002 and 2004 from the GOES-10 satellite as inputs and tested in the year 2005. The prediction time scales are 1-hour (12 data points), 3-hour (36 data points), 6-hour (72 data points), 12-hour (144 data points), and 1-day (288 data points), respectively. The offset time is always 4 days as selected in Section 3.1. The data from 2002 to 2003 are the training set, the data in 2004 are the validation set, and the data in 2005 are the test set.

We only use $\log_{10} (\geq 2 \text{ MeV electron fluxes})$ or the combinations of it with other external parameters as inputs to develop models for different prediction time scales. There are 17 other external parameters, namely Bx , By , Bz , Bt , V_{sw} , N , Pd , E , T , β , AE , $SYM-H$, kp , Dst , $R0$, Lm , and MLT , as listed in Section 2.1. The best combinations of input parameters for modeling are determined by the model's PE values.

In Section 3.1, we use $\log_{10} (\geq 2 \text{ MeV electron fluxes})$ only or the combination of $\log_{10} (\geq 2 \text{ MeV electron fluxes})$ and other single external parameters as the inputs. In other sections, we use the combinations of $\log_{10} (\geq 2 \text{ MeV electron fluxes})$ with other external parameters as inputs, and the number of input parameters for models is controlled within five.

3.1 The selection of the best offset time

The length of the time series of model inputs is the offset time. For example, if the offset time is 4 days, the model will use the consecutive data of the last 4 days as inputs. In this study, the models are aimed at filling the 1-hour (12 data points), 3-hour (36 data points), 6-hour (72 data points), 12-hour (144 data points), and 1-day (288 data points) intervals.

The most suitable offset time will be determined by the PE values. We use the prediction of the $\log_{10} (\geq 2 \text{ MeV electron fluxes})$ the next day (the 1-day prediction with 288 data points) for an example. The $\log_{10} (\geq 2 \text{ MeV electron fluxes})$ is abbreviated as Flux. Three training processes are carried out for each input combination, and the average PE values of the three runs are used.

Figure 3 shows the PE values of the LSTM model (Fig. 3a) and transformer model (Fig. 3b) for the 1-day prediction with different input parameters and different offset times. The offset time ranges from 1 to 9 days and the input parameters are listed on the left of both panels. The colors in the panels represent PE values.

It is clear by the colors in Figure 3 that when only using one external parameter as input, the most important external factors for the LSTM models are V_{sw} , N , and kp , while more external factors also have an impact on transformer models, such as Bz , V_{sw} , N , Pd , AE , $SYM-H$, kp , Dst , $R0$, Lm and MLT . In general, the offset time of the models with solar wind parameters as input parameters is shorter than those with geomagnetic indices. The PE values of the most of models with various input parameters reach their peaks when the offset time is between 3 and 5 days, regardless of whether they are transformer or LSTM models. In addition, we tried some combinations of Flux with two or three external parameters as inputs, the peak of PE values also varied from 3 to 5 days. Moreover, the PE values of the

model are not significantly different with offset time between 3 and 5 days. Finally, the offset time for the predictions in the later study is set to 4 days.

3.2 The LSTM models for predicting ≥ 2 MeV electron fluxes with different prediction time scales

We develop models for different prediction time scales, including the 1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions. The PE values of the LSTM models only using $\log_{10}(\geq 2$ MeV electron fluxes) as input for the 1-hour, 3-hour, 6-hour, 12-hour, and 1-day are 0.913, 0.801, 0.619, 0.421, and 0.349, respectively, and those of the Persistence models are 0.913, 0.798, 0.572, 0.207, and 0.201, separately. The LSTM models only using $\log_{10}(\geq 2$ MeV electron fluxes) as input usually behave better than the Persistence models.

Some combinations as inputs can improve model performances by comparison with the models only using $\log_{10}(\geq 2$ MeV electron fluxes) as input. For the 1-hour and 3-hour predictions, *Bt*, *Vsw*, *N*, and *Dst* are the most important external parameters, because they have the highest frequency of occurrence among the input combinations ranked in the top 100 of PE values. For the 6-hour and 12-hour predictions, *Bt*, *Vsw*, *N*, *Pd*, *AE*, *kp*, *Dst*, *Lm*, and *MLT* can help improve PE values. The addition of *Vsw*, *N*, and *kp* improves the performances of the models for the 1-day prediction when only using $\log_{10}(\geq 2$ MeV electron fluxes) as input. Note that *Bt*, *Vsw*, *N*, *kp*, and *Dst* have the greatest influence on the ≥ 2 MeV electron fluxes at the GEO orbit. These parameters are also often used by previous researchers.

The best input combinations for the 1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions are (Flux, *Dst*), (Flux, *Bt*, *Vsw*), (Flux, *N*, *Dst*, *Lm*), (Flux, *Vsw*, *N*, *Lm*), and (Flux, *Vsw*, *N*), with PE values of 0.919, 0.811, 0.773, 0.554, and 0.490, respectively. Figure 4 shows the comparisons of ≥ 2 MeV electron fluxes between the observations from the GOES-10 satellite and the predictions of the LSTM models with the best combinations for different prediction time scales. The black dots in Figures 4a–Figures 4e represent the ≥ 2 MeV electron fluxes from the GOES-10 satellite in 2005. The red dots in Figures 4a–Figures 4e are the predictions of the LSTM models with (Flux, *Dst*), (Flux, *Bt*, *Vsw*), (Flux, *N*, *Dst*, *Lm*), (Flux, *Vsw*, *N*, *Lm*), and (Flux, *Vsw*, *N*) as inputs for 1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions from top to bottom, respectively. The blue dashed lines in each panel indicate that ≥ 2 MeV electron fluxes are equal to $1157 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$. The data in Figures 4a–Figures 4e are plotted in the flux-flux coordinates in Figures 4f–Figures 4j with black dots on their respective right sides to show the linear relationship of observations and model results. The blue lines, $y = x$, indicate the situation when the observations are completely consistent with the predictions.

As shown in Figure 4, the LC values of ≥ 2 MeV electron fluxes between the observations from the GOES-10 satellite and the predictions of the LSTM models with the best combinations as inputs are 0.953, 0.902, 0.879, 0.750, and 0.702 for different prediction time scales, respectively. The LSTM models perform worse as the prediction time scales increase. Due to their propensity to provide average values to assure better overall performance, the LSTM models' capability to portray in detail the peaks and valleys decreases as prediction time scales increase.

3.3 The transformer models for predicting ≥ 2 MeV electron fluxes with different prediction time scales

Due to the versatility of the transformer network, it has been widely used in machine learning. In order to compare the performance of the transformer network with the LSTM network on processing time series, we also develop models to predict the ≥ 2 MeV electron fluxes by using the transformer method. The prediction time scales are the same as those in Section 3.2.

The PE values of the transformer models only using $\log_{10}(\geq 2$ MeV electron fluxes) as input for the 1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions are 0.931, 0.855, 0.791, 0.677, and 0.554, respectively, which are all higher than those of the LSTM models. In addition, the transformer models can better capture the effect of external parameters on ≥ 2 MeV electron fluxes. *Bt*, *Vsw*, *N*, *Pd*, *AE*, *SYM-H*, *kp*, *Dst*, *Lm*, and *MLT* have significant impacts on the PE values of the transformer models with different prediction time scales. These external parameters are all common parameters in previous prediction models.

The best combinations for 1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions are (Flux, *MLT*), (Flux, *Bt*, *AE*, *SYM-H*), (Flux, *N*), (Flux, *N*, *Dst*, *Lm*), and (Flux, *Pd*, *AE*) with PE values of 0.940, 0.886, 0.828, 0.747, and 0.660, respectively. The transformer models perform better than the LSTM models at the same prediction time scales.

The comparisons of ≥ 2 MeV electron fluxes between the observations from the GOES-10 satellite and the predictions of the transformer models with the best combinations for different prediction time scales are shown in Figure 5. The format is the same as in Figure 4.

It is shown that the transformer models always perform better than the LSTM models in terms of the LC values and biases between observations from the GOES-10 satellite and forecast results, as well as the precision of peak and valley predictions. Therefore, the transformer models are used in the following study.

The PE values inevitably decrease as prediction time scales increase, but the prediction for a longer period has always been a challenge. Additionally, the transformer models perform rather poorly in the predictions of low fluxes, especially when the ≥ 2 MeV electron fluxes are less than $1 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$. The lowest limit of the GOES satellites' detectors prevents them from picking up ≥ 2 MeV electron fluxes below $0.133 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$ and the fluxes below $0.133 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$ are recorded as $0.133 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$ consistently, which will affect the forecast accuracy. Meanwhile, the transformer models require a large amount of data for training, however, the training set only contains 2.65% of the whole data set when the fluxes are under $1 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$. These factors work together to cause poor performances in low fluxes. All data are used for the computations of PE values in this section, and PE values of the ≥ 2 MeV electron fluxes higher than $0.133 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$ will be discussed in Section 4.

3.4 The comparisons with different models

Furthermore, we compare our models with several other models. We calculate the ≥ 2 MeV electron hourly fluences based on the sum of the 12 values of our transformer model

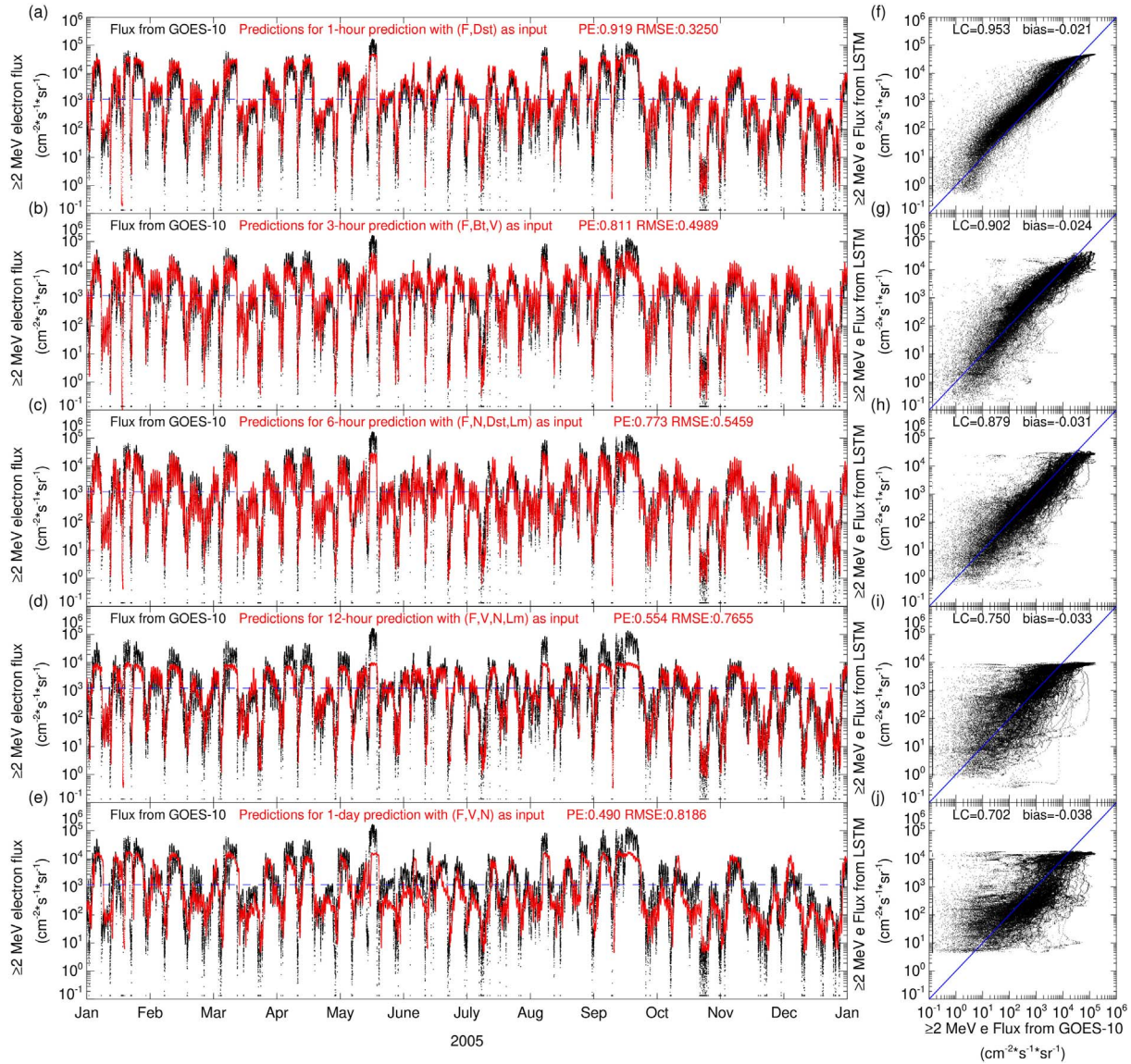


Figure 4. The comparisons of ≥ 2 MeV electron fluxes between the observations from GOES-10 satellite (black dots) and the predictions of the LSTM models (red dots) (a) with (Flux, Dst) as inputs for 1-hour prediction, (b) with (Flux, Bt , Vsw) as inputs for 3-hour prediction, (c) with (Flux, N , Dst , Lm) as inputs for 6-hour prediction, (d) with (Flux, Vsw , N , Lm) as inputs for 12-hour prediction, (e) with (Flux, Vsw , N) as inputs for 1-day prediction, respectively.

for the 1-hour prediction and compute ≥ 2 MeV electron daily fluences based on the sum of the 288 values of our transformer model for the 1-day prediction. The PE and RMSE values for the ≥ 2 MeV electron hourly fluences of our transformer model for the 1-hour prediction, the Persistence model, the LSTM model by Wei et al. (2018), the MLP (multilayer perceptron) model by Son et al. (2022), and the MLP model by Shin et al. (2016) are listed in Table 1. The prediction models above all can provide the prediction of ≥ 2 MeV electron hourly fluence in the next hour. The PE and RMSE values for ≥ 2 MeV electron daily fluences of our transformer model, the Persistence model, the Geomagnetic pulsation model by He et al. (2013), the EOF model by Li et al. (2017), and the EMD (empirical mode decomposition) model by Qian et al. (2020) are listed in Table 2.

There are currently few prediction models for ≥ 2 MeV electron hourly fluences, and most models provide ≥ 2 MeV electron hourly fluences for the next 24 h. There are a few models that calculate PE values for the next 1 h, and the test data for most models are not in 2005. But Sun et al. (2023) noted that PE values show the solar cycle dependence, so we cannot directly compare our model with other models. It can be seen in Figure 4 of Sun et al. (2023) that the PE values of most models in 2005 at $135^\circ W$ are lower than those in 2008, so our model performs better than the LSTM model by Wei et al. (2018). Wei et al. (2018) also pointed out that the PE value of their LSTM model is improved significantly compared to some earlier models. Moreover, most models show relatively high PE values in the next 1 hour, and the PE value of our model is higher than that of the Persistence model. Considering

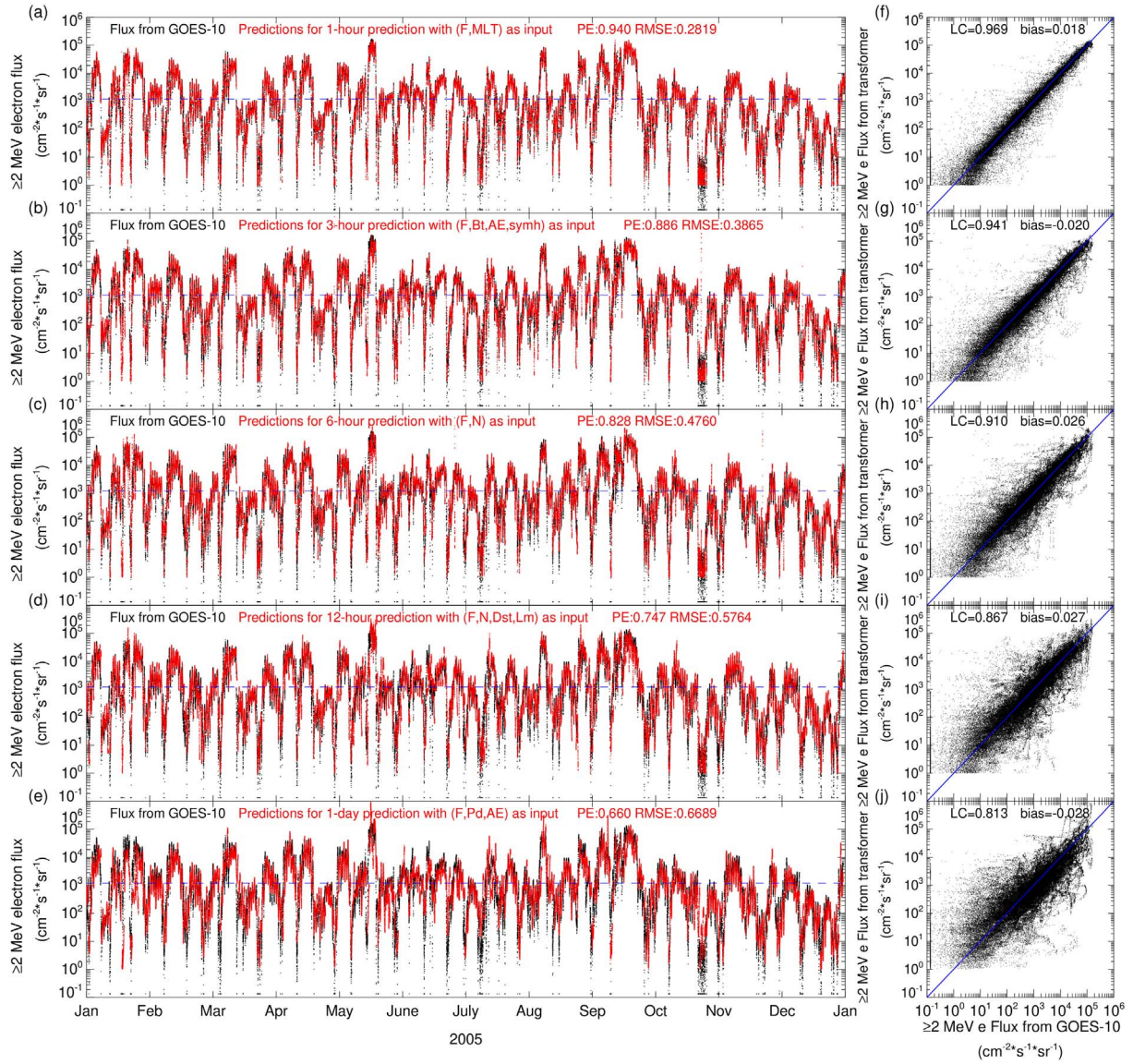


Figure 5. The comparisons of ≥ 2 MeV electron fluxes between the observations from GOES-10 satellite (black dots) and the predictions of the transformer models (red dots) (a) with (Flux, MLT) as inputs for 1-hour prediction, (b) with (Flux, *Bt*, *AE*, *SYM-H*) as inputs for 3-hour prediction, (c) with (Flux, *N*) as inputs for 6-hour prediction, (d) with (Flux, *N*, *Dst*, *Lm*) as inputs for 12-hour prediction, (e) with (Flux, *Pd*, *AE*) as inputs for 1-day prediction, respectively.

Table 1. The comparisons of prediction efficiencies of ≥ 2 MeV electron hourly fluences of different models.

Model	Our transformer model	Persistence model	LSTM model by Wei et al. (2018)	MLP model by Son et al. (2022)	MLP model by Shin et al. (2016)
PE	0.961	0.955	0.900	0.970	0.960
RMSE	0.2136	0.2155	0.2855	–	–
Model input	Flux + MLT	≥ 2 MeV electron hourly fluence 1 h ago	Flux + <i>Kp</i> + <i>R0</i>	Flux + <i>V_{sw}</i> + <i>N</i> + <i>T</i> + <i>B</i> + <i>B_z</i> + <i>Dst</i> + <i>Kp</i>	A total of 17 input parameters
Model output	≥ 2 MeV electron hourly fluence	≥ 2 MeV electron hourly fluence	≥ 2 MeV electron hourly fluence	≥ 2 MeV electron hourly fluence	≥ 2 MeV electron hourly fluence
Satellites	GOES-10	GOES-10	GOES-8, GOES-10, GOES-11, GOES-13	GOES-15, GOES-16	GOES-15
Test time	2005	2005	2008	October to December from 2011 to 2021	2014

Table 2. The comparisons of prediction efficiencies of ≥ 2 MeV electron daily fluences of different models.

Model	Our transformer model	Persistence model	Geomagnetic pulsation model by He et al. (2013)	EOF model by Li et al. (2017)	EMD model by Qian et al. (2020)
PE	0.857	0.507	0.720	0.664	0.790
RMSE	0.3301	0.6122	–	–	0.3800
Model input	Flux + MLT	≥ 2 MeV electron daily fluences 1 day ago	Flux + $PiI2$ + $Pc5$	Flux + V_{sw} + N + Pd + Bz + Dst + ap + AE	Flux + solar wind parameters + geomagnetic indices
Model output	≥ 2 MeV electron daily fluences	≥ 2 MeV electron daily fluences	≥ 2 MeV electron daily fluences	≥ 2 MeV electron daily fluences	≥ 2 MeV electron daily fluences
Satellites	GOES-10	GOES-10	GOES-12	GOES-10	GOES-10
Test time	2005	2005	2005	2005	2005

that our model provides 12 data points for the prediction within the next 1 h at a time instead of only giving one point of ≥ 2 MeV electron hourly fluence as in previous models, our model has an advantage in terms of time resolution.

The PE values for predicting ≥ 2 MeV electron daily fluences on the next day have solar cycle dependence, and they are different in different years. We choose the models whose testing years include 2005, as shown in Table 2. For the comparison of ≥ 2 MeV electron daily fluences, the PE value of our transformer model in 2005 is higher than those of other previous models. Moreover, our model can provide 288 data points for the next day at a time, which will describe the distribution of ≥ 2 MeV electron fluxes in more detail. As shown in Figures 5e and 5j, our transformer models perform relatively well during the high-flux periods. Considering that ≥ 2 MeV electron daily fluences are calculated by the 1-day prediction model, the PE values will increase with shorter prediction time scales. The PE value of ≥ 2 MeV electron daily fluences can reach 0.975 based on the 1-hour prediction model.

Our transformer model performs better than previous prediction models in predicting ≥ 2 MeV electron daily fluences or ≥ 2 MeV electron hourly fluences. In the future, the described methodology can be applied to train new models to predict ≥ 2 MeV electron fluxes at different prediction time scales or to fill in data gaps.

4 Discussion of the model performance

In this section, we discuss the distributions of the differences between predictions and observations, and the overall performance of the models.

Figure 6a shows the PE values of the transformer models for each month in 2005. It is clear that the transformer models with different time scales have the same trend of PE values changing with time in 2005 and the maximum amplitude of variations become increasingly obvious as the prediction time scales increase. The numbers of the relativistic electron enhancement events in each month are displayed in Figure 6b. It is found that the monthly PE values with a high number of relativistic electron enhancement events are usually lower than the average monthly PE value.

The PE values of the transformer model for the 1-hour prediction and the data number in different flux ranges are shown in Figures 6c and 6d, respectively. The relationships between PE values from other models with different prediction scales

and ≥ 2 MeV electron fluxes are similar to those from the 1-hour prediction model. It can be seen in Figure 6c that PE values are lower at low fluxes, especially when ≥ 2 MeV electron fluxes are below $10 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$. It can be concluded that models do not perform well during periods of low fluxes or relativistic electron enhancement events.

4.1 Discussion of the model performances during the relativistic electron enhancement events

In this section, we discuss the model performances during relativistic enhancement events. The relativistic electron enhancement event during 8–16 February 2005, which lasted for 9 days in total is taken as an example.

Figures 7a–7f show the observations from the GOES-10 satellite (black dots) and the predictions of the 1-hour transformer models (red dots) with different external parameters as inputs from 6 February to 17 February 2005. The combinations as inputs are Flux, (Flux, Bt), (Flux, N), (Flux, Pd), (Flux, Dst), and (Flux, MLT) from top to bottom, respectively. The external parameters used in modeling are shown in Figures 7g and 7h.

Compared to the model only using $\log_{10}(\geq 2 \text{ MeV electron fluxes})$ as input during the relativistic electron enhancement event, the addition of appropriate external parameters can help to improve the prediction efficiencies. The PE values of the models with Flux, (Flux, Bt), (Flux, N), (Flux, Pd), (Flux, Dst), (Flux, AE), (Flux, $R0$), (Flux, Lm), and (Flux, MLT) as inputs are 0.950, 0.965, 0.966, 0.965, 0.965, 0.961, 0.961, 0.961, and 0.967. These external parameters can reflect the changes in the solar wind and geomagnetic index and should be used as the key indicators of relativistic electron enhancement events, especially Bt , N , Pd , and Dst .

We calculate the ≥ 2 MeV electron daily fluences based on our transformer model for the 1-day prediction with the best combination as input. There were 35 relativistic electron enhancement events in 2005. The PE values of the first day, the second day, and the end day of the relativistic electron enhancement events are -0.946 , 0.754 , and 0.576 , respectively. It is obvious that the first days of the relativistic electron enhancement events are difficult to predict.

4.2 Discussion of the model performances during low-flux periods

In this section, we discuss the model performance during low-flux periods. There is a protracted low-flux period between

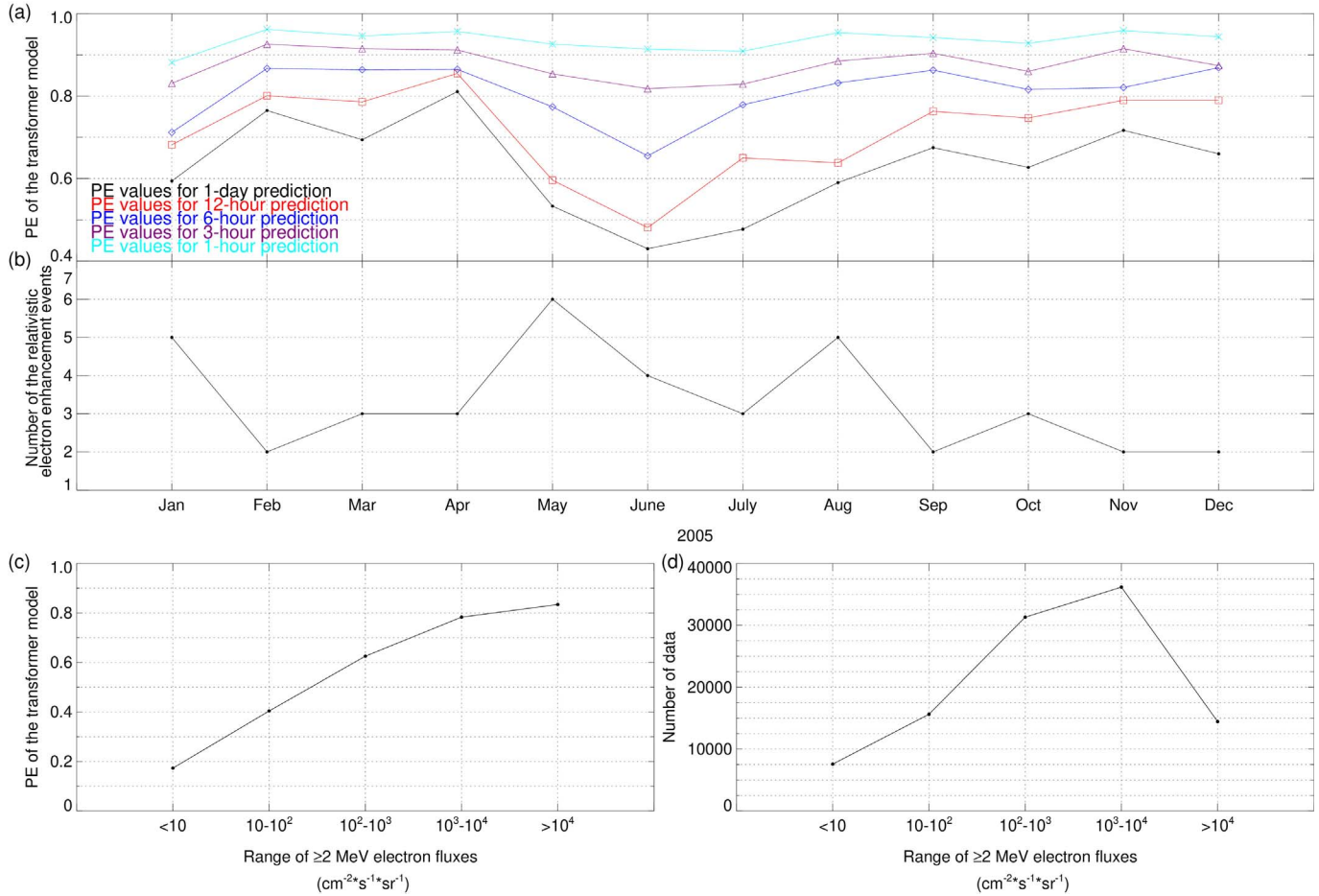


Figure 6. (a) The PE values of the transformer models with different prediction time scales in each month in 2005, (b) the numbers of the relativistic electron enhancement events, (c) the PE values of the transformer model for the 1-hour prediction in different flux ranges, and (d) the data number in different flux ranges.

22 October and 27 October 2005, and we used this period as an example to illustrate the improvements of the model.

Adding suitable external parameters at the prediction moment is an effective way of improving the prediction efficiencies during low-flux periods, as shown in Figure A.1 in [Supplementary material](#). Using Pd , $R0$, AE , $SYM-H$, Dst , Lm , and MLT as inputs aids in proving more accurate predictions compared to the model only using $\log_{10}(\geq 2 \text{ MeV electron fluxes})$ as input during low-flux periods, and the addition of parameters at prediction moment results in a higher PE value by comparing Figure A.1 with Figure A.2 in [Supplementary material](#). The models in Figure A.1 ([Supplementary material](#)) did not use the parameters at the prediction moment when modeling, while the models in Figure A.2 ([Supplementary material](#)) did.

The low $\geq 2 \text{ MeV}$ electron fluxes are due to the movement of the magnetopause. $R0$, the magnetopause subsolar distance is useful for low-flux predictions. The solar wind dynamic pressure (Pd) is one of the main factors causing the compression of the magnetopause. The compression of the magnetopause can cause obvious geomagnetic field disturbances, which can be reflected by AE , $SYM-H$, and Dst . The parameters related to the magnetopause and geomagnetic field can improve the models' performances during low-flux periods.

The $\geq 2 \text{ MeV}$ electron fluxes are not physically accurate when they are equal to $0.133 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$. When these data are removed, the PE values rise. The PE values for 5-minute (discussed later), 1-hour, 3-hour, 6-hour, 12-hour, or 1-day predictions without the $\geq 2 \text{ MeV}$ electron fluxes with $0.133 \text{ cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$ increase to 0.987, 0.958, 0.900, 0.838, 0.759, and 0.667 from 0.974, 0.940, 0.886, 0.828, 0.747, and 0.660, respectively.

4.3 Discussion of the model performances for the 5-minute prediction

Except for the 1-hour, 3-hour, 6-hour, 12-hour, or 1-day prediction models, we also developed the 5-minute prediction model using LSTM and transformer networks, which only give one forecast value for the next five minutes in turn.

The best offset time for the 5-minute prediction is selected at first. The performances of the Persistence model and the linear model are adequate for the 5-minute predictions, with the PE value reaching 0.966 and 0.968, respectively. The Persistence model uses the value at the current moment as the prediction at the following time step, and the linear model employs the linear equation created by the previous two data to produce the subsequent data. The 10-minute and 5-minute intervals are

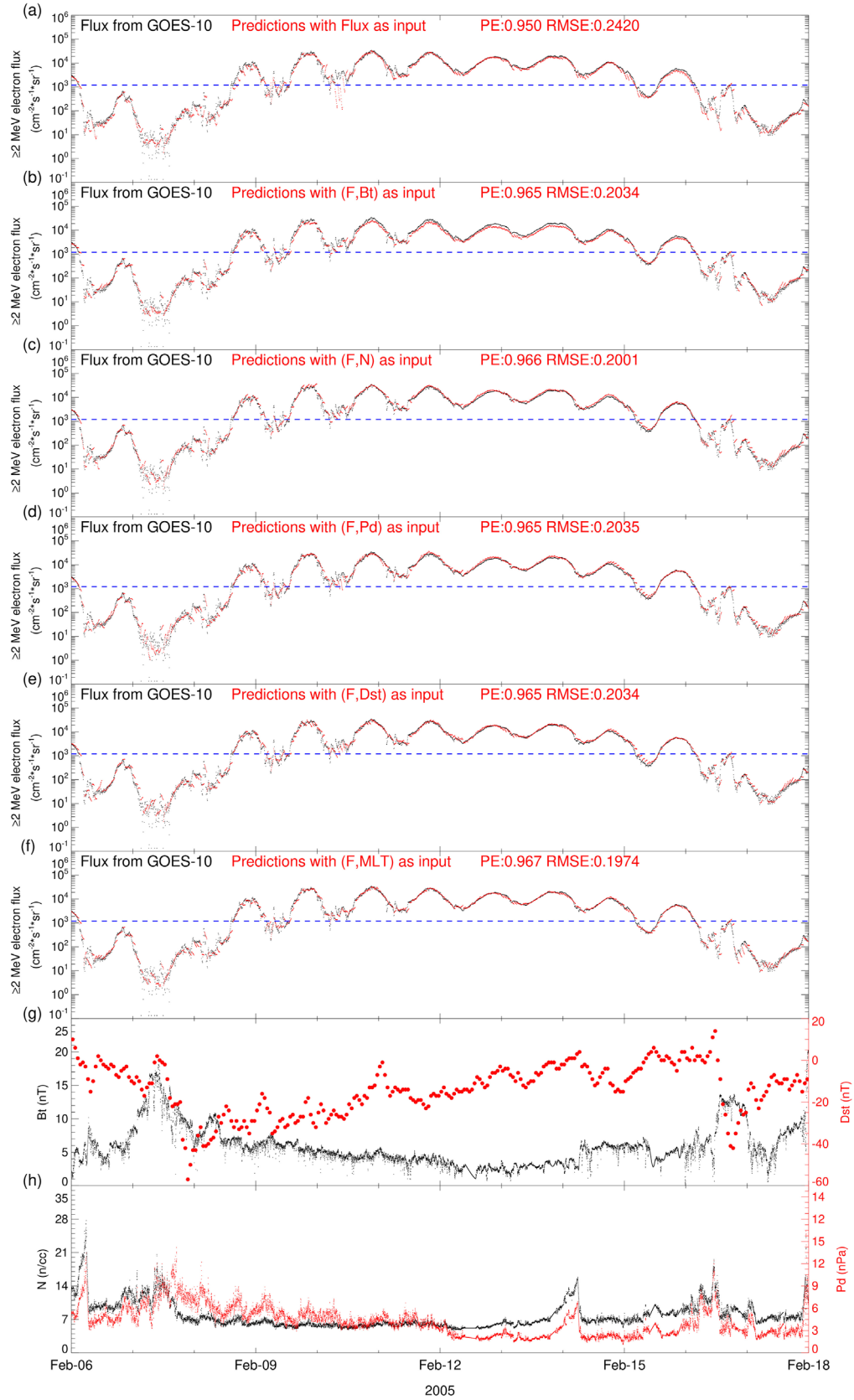


Figure 7. (a)–(f) The comparisons of the prediction results of the 1-hour transformer models with different inputs (red dots) and observations from the GOES-10 satellite (black dots) during the relativistic electron enhancement event, and (g)–(h) B_t , N , Pd , and Dst between 6 and 17 February 2005.

tested as offset time to evaluate the model performance, and the model with 10 minutes as offset time performs better than the model with 5 minutes as offset time. The offset time for the 5-minute predictions in this study is 10 minutes.

For predicting ≥ 2 MeV electron fluxes in the next five minutes in 2005, the PE and RMSE values of the LSTM model only using $\log_{10}$ (≥ 2 MeV electron fluxes) as input are 0.968 and 0.2053, respectively. We tested the addition of V_{sw} , kp , Lm , or MLT and found that those parameters improve model performance. The model with (Flux, Lm) as inputs performs best with the PE and RMSE values 0.970 and 0.2001, respectively. The model with (Flux, MLT) ranks second with the PE and RMSE values 0.969 and 0.2025, separately. Lm and MLT , which are associated with the geomagnetic field structure and not symmetrical about local time, are more significant for the 5-minute prediction model than solar wind parameters and geomagnetic indices. If only one external parameter is added in inputs for the 5-minute prediction model based on the LSTM method, Lm or MLT is recommended.

For the 5-minute prediction of ≥ 2 MeV electron fluxes in 2005, the PE and RMSE values of the transformer model only using $\log_{10}$ (≥ 2 MeV electron fluxes) as input are 0.972 and 0.1912, respectively. Model performances are improved when external parameters are added. The model with (Flux, Lm) as inputs perform best with the PE and RMSE values 0.974 and 0.1863, respectively, with the rank the same as the LSTM model. It demonstrated again that Lm and MLT , have more influence on the 5-minute prediction of ≥ 2 MeV electron fluxes than solar wind parameters and geomagnetic indices.

The comparisons of ≥ 2 MeV electron fluxes between the observations from the GOES-10 satellite and the 5-minute predictions of the LSTM and transformer models with different combinations as inputs are shown in Figure A.2 in [Supplementary Material](#). For the 5-minute prediction, the predictions of the LSTM models are basically in line with the observations from GOES satellites.

It can be seen that for the 5-minute predictions, the LSTM and the transformer models with (Flux, Lm) as inputs both performed well, with the PE values of 0.970 and 0.974 in 2005, respectively. The distributions of ≥ 2 MeV electron fluxes are directly impacted by the alteration in the form of a geomagnetic field, which also impacts the shape of the outer radiation belt. Lm and MLT , which are associated with the geomagnetic field structure, are more important for the 5-minute prediction than solar wind parameters and geomagnetic indices.

5 Conclusions

The variations in ≥ 2 MeV electron flux at GEO orbit are the result of the combined contribution of space environment parameters. In this study, we applied machine learning to predict ≥ 2 MeV electron fluxes at GEO orbit because machine learning can effectively deal with massive data samples and solve nonlinear problems. We developed models to predict ≥ 2 MeV electron fluxes with various prediction time scales using the transformer and LSTM networks. The main conclusions are as follows:

Based on the performances of models with different combinations as inputs, the best offset time is determined as four days for different prediction time scales (1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions). When only using one external

parameter as input, the most important external factors for the LSTM models are found to be V_{sw} , N , and kp , while the addition of more external factors also improves transformer model predictions, such as Bz , V_{sw} , N , Pd , AE , $SYM-H$, kp , Dst , RO , Lm , and MLT .

For the 1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions, the transformer models performed better than the LSTM models with different prediction time scales. Meanwhile, the predictions of the transformer models also showed more detailed temporal variations in fluxes than the LSTM models.

The best combinations of the LSTM models for 1-hour, 3-hour, 6-hour, 12-hour, and 1-day predictions are found to be (Flux, Dst), (Flux, Bt , V_{sw}), (Flux, N , Dst , Lm), (Flux, V_{sw} , N , Lm), and (Flux, V_{sw} , N) with PE values of 0.919, 0.811, 0.773, 0.554, and 0.490, respectively, and those of the transformer models are (Flux, MLT), (Flux, Bt , AE , $SYM-H$), (Flux, N), (Flux, N , Dst , Lm), and (Flux, Pd , AE) with PE values of 0.940, 0.886, 0.828, 0.747, and 0.660, respectively.

For the comparison of ≥ 2 MeV electron daily fluences, the PE value of our transformer model in 2005 is higher than those of other previous models. Moreover, our model can provide 288 data points for the next day at a time, which will describe the distribution of ≥ 2 MeV electron fluxes in more detail. In addition, our transformer model performs relatively well during the high-flux periods. The PE value of ≥ 2 MeV electron daily fluences can reach 0.965 based on the 1-hour prediction model.

We discussed the transformer model performances during the relativistic electron enhancement events and low-flux periods. The model performances during relativistic electron enhancement events can be improved by adding appropriate external parameters, such as Bt , N , Pd , RO , AE , $SYM-H$, Lm , and MLT . For the low-flux periods, adding one or more of the parameters of Pd , RO , AE , $SYM-H$, Dst , Lm , or MLT in the inputs improves the accuracy of predictions in low-flux periods.

Based on the evaluation results of our models above, it can be concluded that our models are suitable for filling the data gaps of ≥ 2 MeV electron fluxes.

Acknowledgements

This work was supported by grants from Project U2106201 of the National Natural Science Foundation of China (NSFC). The data used throughout this study are courtesy of NOAA/SWPC science teams. Thanks to the NOAA National Environmental Information Center (NCEI) for providing processed GOES series satellite data and OMNI for providing the solar wind parameters and geomagnetic disturbance indices. The authors also thank the IRBEM (international radiation environment modeling software library) for providing the internal geomagnetic field model and the external geomagnetic field model. Thanks to the China Scholarship Council (CSC) for providing the corresponding author the chance in GFZ section 2.7. We would like to thank Melanie Burns Allison for the paper polishing. Thanks to all the people in GFZ section 2.7 for the useful discussion, especially Stefano Bianco and Maximilian Pfitzer. The editor thanks Spiridon Kasapis, Richard Boynton and an anonymous reviewer for their assistance in evaluating this paper.

Supplementary material

Supporting Information for “A Modeling Study of ≥ 2 MeV Electron Fluxes at Different Prediction Time Scales Based on Both Networks”. The supplementary material is available at <https://www.swsc-journal.org/10.1051/swsc/2024021/olm>.

References

- Albert J. 2007. Simple approximations of quasi-linear diffusion coefficients. *J Geophys Res Space Phys* **112**(A12): A12202. <https://doi.org/10.1029/2007JA012551>.
- Albert J. 2008. Efficient approximations of quasi-linear diffusion coefficients in the radiation belts. *J Geophys Res Space Phys* **113**(A6): A06208. <https://doi.org/10.1029/2007JA012936>.
- Allison HJ, Horne RB, Glauert SA, Del Zanna G. 2019. On the importance of gradients in the low-energy electron phase space density for relativistic electron acceleration. *J Geophys Res Space Phys* **124**(4): 2628–2642. <https://doi.org/10.1029/2019JA026516>.
- Anderson B, Millan R, Reeves G, Friedel R. 2015. Acceleration and loss of relativistic electrons during small geomagnetic storms. *Geophys Res Lett* **42**(23): 10–113. <https://doi.org/10.1002/2015GL066376>.
- Baker D, Blake J, Gorney D, Higbie P, Klebesadel R, King J. 1987. Highly relativistic magnetospheric electrons: A role in coupling to the middle atmosphere? *Geophys Res Lett* **14**(10): 1027–1030. <https://doi.org/10.1029/GL014i010p01027>.
- Baker D, McPherron R, Cayton T, Klebesadel R. 1990. Linear prediction filter analysis of relativistic electron properties at 6.6 RE. *J Geophys Res Space Phys* **95**(A9): 15133–15140. <https://doi.org/10.1029/JA095iA09p15133>.
- Balikhin M, Rodriguez J, Boynton R, Walker S, Aryan H, Sibeck D, Billings S. 2016. Comparative analysis of NOAA REFM and SNB3GEO tools for the forecast of the fluxes of high-energy electrons at GEO. *Space Weather* **14**(1): 22–31. <https://doi.org/10.1002/2015SW001303>.
- Balikhin MA, Boynton RJ, Walker SN, Borovsky JE, Billings SA, Wei H-L. 2011. Using the NARMAX approach to model the evolution of energetic electrons fluxes at geostationary orbit. *Geophys Res Lett* **38**(18): L18105. <https://doi.org/10.1029/2011gl048980>.
- Beutier T, Boscher D. 1995. A three-dimensional analysis of the electron radiation belt by the Salammbô code. *J Geophys Res Space Phys* **100**(A8): 14853–14861. <https://doi.org/10.1029/94JA03066>.
- Bourdarie S, Maget V. 2012. Electron radiation belt data assimilation with an ensemble Kalman filter relying on the Salammbô code. *Ann Geophys* **30**: 929–943. <https://doi.org/10.5194/angeo-30-929-2012>.
- Boynton R, Balikhin M, Billings S. 2015. Online NARMAX model for electron fluxes at GEO. *Ann Geophys* **33**: 405–411. <https://doi.org/10.5194/ANGE0-33-405-2015>.
- Boynton R, Balikhin M, Billings S, Reeves G, Ganushkina N, Gedalin M, Amariutei O, Borovsky J, Walker S. 2013. The analysis of electron fluxes at geosynchronous orbit employing a NARMAX approach. *J Geophys Res Space Phys* **118**(4): 1500–1513. <https://doi.org/10.1002/jgra.50192>.
- Brautigam D, Albert J. 2000. Radial diffusion analysis of outer radiation belt electrons during the October 9, 1990, magnetic storm. *J Geophys Res Space Phys* **105**(A1): 291–309. <https://doi.org/10.1029/1999JA900344>.
- Burin des Roziers E, Li X. 2006. Specification of >2 MeV geosynchronous electrons based on solar wind measurements. *Space Weather* **4**(6): S6007. <https://doi.org/10.1029/2005SW000177>.
- Cai M, Liu J. 2016. Maxout neurons for deep convolutional and LSTM neural networks in speech recognition. *Speech Commun* **77**: 53–64. <https://doi.org/10.1016/j.specom.2015.12.003>.
- Devlin J, Chang M-W, Lee K, Toutanova K. 2018. BERT: pre-training of deep bidirectional transformers for language understanding. *arXiv preprint*. <https://doi.org/10.48550/arXiv.1810.04805>.
- Dong L, Xu S, Xu B. 2018. Speech-transformer: a no-recurrence sequence-to-sequence model for speech recognition. In: 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP), Calgary, AB, Canada, IEEE, pp. 5884–5888. <https://doi.org/10.1109/ICASSP.2018.8462506>.
- Drozov A, Shprits YY, Usanova M, Aseev N, Kellerman A, Zhu H. 2017. EMIC wave parameterization in the long-term VERB code simulation. *J Geophys Res Space Phys* **122**(8): 8488–8501. <https://doi.org/10.1002/2017ja024389>.
- Drozov AY, Allison HJ, Shprits YY, Elkington SR, Aseev NA. 2021. A comparison of radial diffusion coefficients in 1-D and 3-D long-term radiation belt simulations. *J Geophys Res Space Phys* **126**(8): e2020JA028707. <https://doi.org/10.1029/2020JA028707>.
- Friedel R, Reeves G, Obara T. 2002. Relativistic electron dynamics in the inner magnetosphere – a review. *J Atmos Sol-Terr Phys* **64**(2): 265–282. [https://doi.org/10.1016/S1364-6826\(01\)00088-8](https://doi.org/10.1016/S1364-6826(01)00088-8).
- Fukata M, Taguchi S, Okuzawa T, Obara T. 2002. Neural network prediction of relativistic electrons at geosynchronous orbit during the storm recovery phase: effects of recurring substorms. *Ann Geophys* **20**: 947–951. <https://doi.org/10.5194/angeo-20-947-2002>.
- Ganushkina NY, Amariutei O, Welling D, Heynderickx D. 2015. Nowcast model for low-energy electrons in the inner magnetosphere. *Space Weather* **13**(1): 16–34. <https://doi.org/10.1002/2014sw001098>.
- Ganushkina NY, Liemohn M, Amariutei O, Pitchford D. 2014. Low-energy electrons (5–50 keV) in the inner magnetosphere. *J Geophys Res Space Phys* **119**(1): 246–259. <https://doi.org/10.1002/2013ja019304>.
- Gers FA, Schmidhuber J, Cummins F. 2000. Learning to forget: Continual prediction with LSTM. *Neural Comput* **12**(10): 2451–2471. <https://doi.org/10.1049/cp:19991218>.
- Glauert SA, Horne RB, Meredith NP. 2014a. Simulating the Earth's radiation belts: Internal acceleration and continuous losses to the magnetopause. *J Geophys Res Space Phys* **119**(9): 7444–7463. <https://doi.org/10.1002/2014JA020092>.
- Glauert SA, Horne RB, Meredith NP. 2014b. Three-dimensional electron radiation belt simulations using the BAS Radiation Belt Model with new diffusion models for chorus, plasmaspheric hiss, and lightning-generated whistlers. *J Geophys Res Space Phys* **119**(1): 268–289. <https://doi.org/10.1002/2013JA019281>.
- Graves A. 2012. Long short-term memory. In: *Supervised sequence labelling with recurrent neural networks. Studies in computational intelligence*, vol. 385. Springer, Berlin, Heidelberg, pp. 37–45. https://doi.org/10.1007/978-3-642-24797-2_4.
- Graves A, Schmidhuber J. 2005. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw* **18**(5–6): 602–610. <https://doi.org/10.1016/j.neunet.2005.06.042>.
- Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J. 2016. LSTM: a search space odyssey. *IEEE Trans Neural Netw Learn Syst* **28**(10): 2222–2232. <https://doi.org/10.1109/TNNLS.2016.2582924>.
- Grubb R. 1975. The SMS/GOES space environment monitor subsystem. *NASA STI/Recon Technical Report No.* 76, 28260. Available at https://www.ngdc.noaa.gov/stp/satellite/goes/doc/ERL-SEL-42_SEM.pdf.
- Gubby R, Evans J. 2002. Space environment effects and satellite design. *J Atmos Sol-Terr Phys* **64**(16): 1723–1733. [https://doi.org/10.1016/S1364-6826\(02\)00122-0](https://doi.org/10.1016/S1364-6826(02)00122-0).
- Guo C, Xue B, Lin Z. 2013. Approach for predicting the energetic electron flux in geosynchronous earth orbit. *Chin J Space Sci* **33**(4): 418–426. <https://doi.org/10.11728/cjss2013.04.418>.

- He T, Liu S, Shen H, Gong J. 2013. Quantitative prediction of relativistic electron flux at geosynchronous orbit with geomagnetic pulsations parameters. *Chin J Space Sci* **33**(1): 20–27. <https://doi.org/10.11728/cjss2013.01.020>.
- Hochreiter S, Schmidhuber J. 1997. Long short-term memory. *Neural Comput* **9**(8): 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- Horne R, Glauert S, Meredith N, Boscher D, Maget V, Heynderickx D, Pitchford D. 2013. Space weather impacts on satellites and forecasting the Earth's electron radiation belts with SPACECAST. *Space weather* **11**(4): 169–186. <https://doi.org/10.1002/swe.20023>.
- Horne RB, Thorne RM. 1998. Potential waves for relativistic electron scattering and stochastic acceleration during magnetic storms. *Geophys Res Lett* **25**(15): 3011–3014. <https://doi.org/10.1029/98GL01002>.
- Horne RB, Thorne RM, Glauert SA, Albert JM, Meredith NP, Anderson RR. 2005. Timescale for radiation belt electron acceleration by whistler mode chorus waves. *J Geophys Res Space Phys* **110**(A3): A03225. <https://doi.org/10.1029/2004JA010811>.
- Huang Z, Xu W, Yu K. 2015. Bidirectional LSTM-CRF models for sequence tagging. *arXiv preprint*. <https://doi.org/10.48550/arXiv.1508.01991>.
- Kai Y, Lei J, Chen Y, Wei X. 2013. Deep learning: Yesterday, today, and tomorrow. *J Comput Res Dev* **50**(9): 1799–1804. <https://icad.ict.ac.cn/en/article/id/1340>.
- Katsavrias C, Aminalragia-Giamini S, Papadimitriou C, Daglis IA, Sandberg I, Jiggins P. 2022. Radiation belt model including semi-annual variation and solar driving (sentinel). *Space Weather* **20**(1): e2021SW002936. <https://doi.org/10.1029/2021SW002936>.
- Kersten T, Horne RB, Glauert SA, Meredith NP, Fraser BJ, Grew RS. 2014. Electron losses from the radiation belts caused by EMIC waves. *J Geophys Res Space Phys* **119**(11): 8820–8837. <https://doi.org/10.1002/2014JA020366>.
- Kim K-C, Shprits Y, Subbotin D, Ni B. 2011. Understanding the dynamic evolution of the relativistic electron slot region including radial and pitch angle diffusion. *J Geophys Res Space Phys* **116**(A10): A10214. <https://doi.org/10.1029/2011JA016684>.
- Kim K-C, Shprits Y, Subbotin D, Ni B. 2012. Relativistic radiation belt electron responses to GEM magnetic storms: Comparison of CRRES observations with 3-D VERB simulations. *J Geophys Res Space Phys* **117**(A8): A08221. <https://doi.org/10.1029/2011JA017460>.
- Kingma DP, Ba J. 2014. Adam: A method for stochastic optimization. *arXiv preprint*. <https://doi.org/10.48550/arXiv.1412.6980>.
- Kondrashov D, Denton R, Shprits Y, Singer H. 2014. Reconstruction of gaps in the past history of solar wind parameters. *Geophys Res Lett* **41**(8): 2702–2707. <https://doi.org/10.1002/2014GL059741>.
- Kondrashov D, Feliks Y, Ghil M. 2005. Oscillatory modes of extended Nile River records (AD 622–1922). *Geophys Res Lett* **32**(10): L10702. <https://doi.org/10.1029/2004GL022156>.
- Kondrashov D, Ghil M. 2006. Spatio-temporal filling of missing points in geophysical data sets. *Nonlinear Proc Geophys* **13**(2): 151–159. <https://doi.org/10.5194/npg-13-151-2006>.
- Kondrashov D, Ghil M, Shprits Y. 2011. Lognormal Kalman filter for assimilating phase space density data in the radiation belts. *Space Weather* **9**(11): S11006. <https://doi.org/10.1029/2011SW000726>.
- Kondrashov D, Shprits Y, Ghil M. 2010. Gap filling of solar wind data by singular spectrum analysis. *Geophys Res Lett* **37**(15): L15101. <https://doi.org/10.1029/2010GL044138>.
- Lai ST, Cahoy K, Lohmeyer W, Carlton A, Aniceto R, Minow J. 2018. Deep dielectric charging and spacecraft anomalies. In: *Extreme events in geospace*, Elsevier, 2018, pp. 419–432. <https://doi.org/10.1016/b978-0-12-812700-1.00016-9>.
- Landis D, Saikin A, Zhelavskaya I, Drozdov A, Aseev N, Shprits Y, Pfister M, Smirnov A. 2022. NARX neural network derivations of the outer boundary radiation belt electron flux. *Space Weather* **20**(5): e2021SW002774. <https://doi.org/10.1029/2020SW002524>.
- Lanzerotti L, Breglia C, Maurer D, Johnson III G, MacLennan C. 1998. Studies of spacecraft charging on a geosynchronous telecommunications satellite. *Adv Space Res* **22**(1): 79–82. [https://doi.org/10.1016/S0273-1177\(97\)01104-6](https://doi.org/10.1016/S0273-1177(97)01104-6).
- Li S, Huang W, Liu S, Zhong Q. 2017. Dynamic prediction model of relativistic electron differential fluxes at the geosynchronous orbit. *Chin J Space Sci* **37**(3): 298–311. <https://doi.org/10.11728/cjss2017.03.298>.
- Li W, Ma Q, Thorne R, Bortnik J, Zhang X-J, et al. 2016. Radiation belt electron acceleration during the 17 March 2015 geomagnetic storm: Observations and simulations. *J Geophys Res Space Phys* **121**(6): 5520–5536. <https://doi.org/10.1002/2016JA022400>.
- Li W, Shprits Y, Thorne R. 2007. Dynamic evolution of energetic outer zone electrons due to waveparticle interactions during storms. *J Geophys Res Space Phys* **112**(A10): A10220. <https://doi.org/10.1029/2007JA012368>.
- Li X. 2004. Variations of 0.7–6.0 MeV electrons at geosynchronous orbit as a function of solar wind. *Space Weather* **2**(3): S03006. <https://doi.org/10.1029/2003sw000017>.
- Li X, Baker D, O'Brien T, Xie L, Zong Q. 2006. Correlation between the inner edge of outer radiation belt electrons and the innermost plasmapause location. *Geophys Res Lett* **33**(14): L14107. <https://doi.org/10.1029/2006GL026294>.
- Li X, Temerin M, Baker D, Reeves G. 2011. Behavior of MeV electrons at geosynchronous orbit during last two solar cycles. *J Geophys Res Space Phys* **116**(A11): A11207. <https://doi.org/10.1029/2011JA016934>.
- Li X, Temerin M, Baker D, Reeves G, Larson D. 2001. Quantitative prediction of radiation belt electrons at geostationary orbit based on solar wind measurements. *Geophys Res Lett* **28**(9): 1887–1890. <https://doi.org/10.1029/2000GL012681>.
- Lin R, Zhang X, Liu S, Wang Y, Gong J. 2010. A three-dimensional asymmetric magnetopause model. *J Geophys Res Space Phys* **115**(A4): A04207. <https://doi.org/10.1029/2009JA014235>.
- Ling A, Ginot G, Hilmer R, Perry K. 2010. A neural network-based geosynchronous relativistic electron flux forecasting model. *Space Weather* **8**(9): S09003. <https://doi.org/10.1029/2010sw000576>.
- Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B. 2021. Swin transformer: Hierarchical vision transformer using shifted windows. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, Canada, IEEE, 10012–10022. <https://doi.org/10.48550/arXiv.2103.14030>.
- Lohmeyer W, Carlton A, Wong F, Bodeau M, Kennedy A, Cahoy K. 2015. Response of geostationary communications satellite solid-state power amplifiers to high-energy electron fluence. *Space Weather* **13**(5): 298–315. <https://doi.org/10.1002/2014SW001147>.
- Lucci N, Levitin A, Belov A, Eroshenko E, Puitsyna N, et al. 2005. Space weather conditions and spacecraft anomalies in different orbits. *Space Weather* **3**(1): 01001. <https://doi.org/10.1029/2003SW000056>.
- Ma Q, Li W, Bortnik J, Thorne R, Chu X, et al. 2018. Quantitative evaluation of radial diffusion and local acceleration processes during GEM challenge events. *J Geophys Res Space Phys* **123**(3): 1938–1952. <https://doi.org/10.1002/2017JA025114>.
- Macmillan S, Finlay C. 2010. The international geomagnetic reference field. In: *Geomagnetic observations and models*, Mandaia M, Korte M (Eds.), *Geomagnetic observations and models. IAGA Special Sopron Book Series*, vol. 5, Springer. pp. 265–276. https://doi.org/10.1007/978-90-481-9858-0_10.

- Maget V, Bourdarie S, Boscher D, Friedel R. 2007. Data assimilation of LANL satellite data into the Salammbô electron code over a complete solar cycle by direct insertion. *Space Weather* **5**(10): S10003. <https://doi.org/10.1029/2007SW000322>.
- Meredith NP, Horne RB, Iles RH, Thorne RM, Heynderickx D, Anderson RR. 2002. Outer zone relativistic electron acceleration associated with substorm-enhanced whistler mode chorus. *J Geophys Res Space Phys* **107**(A7): 1144. <https://doi.org/10.1029/2001JA900146>.
- Millan R, Thorne R. 2007. Review of radiation belt relativistic electron losses. *J Atmos Sol-Terr Phys* **69**(3): 362–377. <https://doi.org/10.1016/j.jastp.2006.06.019>.
- Miyoshi Y, Morioka A, Misawa H, Obara T, Nagai T, Kasahara Y. 2003. Rebuilding process of the outer radiation belt during the 3 November 1993 magnetic storm: NOAA and Exos-D observations. *J Geophys Res Space Phys* **108**(A1): 1004. <https://doi.org/10.1029/2001JA007542>.
- O'Brien T, Sornette D, McPherron R. 2001. Statistical asynchronous regression: Determining the relationship between two quantities that are not measured simultaneously. *J Geophys Res Space Phys* **106**(A7): 13247–13259. <https://doi.org/10.1029/2000JA900193>.
- Pakhotin I, Drozdov A, Shprits Y, Boynton R, Subbotin D, Balikhin M. 2014. Simulation of highenergy radiation belt electron fluxes using NARMAX-VERB coupled codes. *J Geophys Res Space Phys* **119**(10): 8073–8086. <https://doi.org/10.1002/2014JA020238>.
- Paulikas G, Blake J. 1979. Effects of the solar wind on magnetospheric dynamics: Energetic electrons at the synchronous orbit. *Quantitative modeling of magnetospheric processes* **21**: 180–202. <https://doi.org/10.1029/GM021p0180>.
- Pilipenko V, Yagova N, Romanova N, Allen J. 2006. Statistical relationships between satellite anomalies at geostationary orbit and high-energy particles. *Adv Space Res* **37**(6): 1192–1205. <https://doi.org/10.1016/j.asr.2005.03.152>.
- Potapov A, Tsegmed B, Ryzhakova L. 2014. Solar cycle variation of “killer” electrons at geosynchronous orbit and electron flux correlation with the solar wind parameters and ULF waves intensity. *Acta Astronaut* **93**: 55–63. <https://doi.org/10.1016/j.actaastro.2013.07.004>.
- Qian Y, Yang J, Zhang H, Shen C, Wu Y. 2020. An hourly prediction model of relativistic electrons based on empirical mode decomposition. *Space Weather* **18**(8): e2018SW0022078. <https://doi.org/10.1029/2018SW0022078>.
- Qin Z, Denton R, Tsyganenko N, Wolf S. 2007. Solar wind parameters for magnetospheric magnetic field modeling. *Space Weather* **5**(11). <https://doi.org/10.1029/2006SW000296>.
- Reagan J, Meyerott R, Gaines E, Nightingale R, Filbert P, Imhof W. 1983. Space charging currents and their effects on spacecraft systems. *IEEE Trans Electr Insul* **18**(3): 354–365. <https://doi.org/10.1109/TEI.1983.298625>.
- Reeves GD, Chen Y, Cunningham G, Friedel R, Henderson MG, Jordanova V, Koller J, Morley S, Thomsen M, Zaharia S. 2012. Dynamic radiation environment assimilation model: DREAM. *Space Weather* **10**(3): S03006. <https://doi.org/10.1029/2011SW000729>.
- Reeves GD, Morley SK, Friedel RH, Henderson MG, Cayton TE, Cunningham G, Blake JB, Christensen RA, Thomsen D. 2011. On the relationship between relativistic electron flux and solar wind velocity: Paulikas and Blake revisited. *J Geophys Res Space Phys* **116**(A2): A02213. <https://doi.org/10.1029/2010JA015735>.
- Rigler E, Baker D, Weigel R, Vassiliadis D, Klimas A. 2004. Adaptive linear prediction of radiation belt electrons using the Kalman filter. *Space Weather* **2**(3): S03003. <https://doi.org/10.1029/2003SW000036>.
- Romanova N, Pilipenko V, Yagova N, Belov A. 2005. Statistical correlation of the rate of failures on geosynchronous satellites with fluxes of energetic electrons and protons. *Cosm Res* **43**: 179–185. <https://doi.org/10.1007/s10604-005-0032-6>.
- Ryden KA, Morris PA, Ford KA, Hands AD, Dyer CS, et al. et al. 2008. Observations of internal charging currents in medium earth orbit. *IEEE Trans Plasma Sci* **36**(5): 2473–2481. <https://doi.org/10.1109/TPS.2008.2001945>.
- Saikin A, Shprits YY, Drozdov A, Landis DA, Zhelavskaya I, Cervantes S. 2021. Reconstruction of the radiation belts for solar cycles 17–24 (1933–2017). *Space Weather* **19**(3): e2020SW002524. <https://doi.org/10.1029/2020SW002524>.
- Sakaguchi K, Miyoshi Y, Saito S, Nagatsuma T, Seki K, Murata K. 2013. Relativistic electron flux forecast at geostationary orbit using Kalman filter based on multivariate autoregressive model. *Space Weather* **11**(2): 79–89. <https://doi.org/10.1002/swe.20020>.
- Shin D-K, Lee D-Y, Kim K-C, Hwang J, Kim J. 2016. Artificial neural network prediction model for geosynchronous electron fluxes: Dependence on satellite position and particle energy. *Space Weather* **14**(4): 313–321. <https://doi.org/10.1002/2015SW001359>.
- Shprits Y, Daae M, Ni B. 2012. Statistical analysis of phase space density buildups and dropouts. *J Geophys Res Space Phys* **117**(A1): A01219. <https://doi.org/10.1029/2011JA016939>.
- Shprits Y, Kellerman A, Kondrashov D, Subbotin D. 2013. Application of a new data operator-splitting data assimilation technique to the 3-D VERB diffusion code and CRRES measurements. *Geophys Res Lett* **40**(19): 4998–5002. <https://doi.org/10.1002/grl.50969>.
- Shprits Y, Thorne R, Horne R, Glauert S, Cartwright M, Russell C, Baker D, Kanekal S. 2006a. Acceleration mechanism responsible for the formation of the new radiation belt during the 2003 Halloween solar storm. *Geophys Res Lett* **33**(5): L05104. <https://doi.org/10.1029/2005GL024256>.
- Shprits YY, Allison HJ, Wang D, Drozdov A, Szabo-Roberts M, Zhelavskaya I, Vasile R. 2022. A new population of ultra-relativistic electrons in the outer radiation zone. *J Geophys Res Space Phys* **127**(5): e2021JA030214. <https://doi.org/10.1029/2021JA030214>.
- Shprits YY, Elkington SR, Meredith NP, Subbotin DA. 2008a. Review of modeling of losses and sources of relativistic electrons in the outer radiation belt I: Radial transport. *J Atmos Sol-Terr Phys* **70**(14): 1679–1693. <https://doi.org/10.1016/j.jastp.2008.06.008>.
- Shprits YY, Horne RB, Kellerman AC, Drozdov AY. 2018. The dynamics of Van Allen belts revisited. *Nat Phys* **14**(2): 102–103. <https://doi.org/10.1038/nphys4350>.
- Shprits YY, Kellerman AC, Drozdov AY, Spence HE, Reeves GD, Baker DN. 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>.
- Shprits YY, Subbotin D, Ni B. 2009. Evolution of electron fluxes in the outer radiation belt computed with the VERB code. *J Geophys Res Space Phys* **114**(A11): A11209. <https://doi.org/10.1029/2008ja013784>.
- Shprits YY, Subbotin DA, Meredith NP, Elkington SR. 2008b. Review of modeling of losses and sources of relativistic electrons in the outer radiation belt II: Local acceleration and loss. *J Atmos Sol-Terr Phys* **70**(14): 1694–1713. <https://doi.org/10.1016/j.jastp.2008.06.014>.
- Shprits YY, Thorne RM, Horne RB, Summers D. 2006b. Bounce-averaged diffusion coefficients for field-aligned chorus waves. *J Geophys Res Space Phys* **111**(A10): A10225. <https://doi.org/10.1029/2006ja011725>.

- Singh J, Omale SO, Inumoh LO, Ale F. 2021. Impact of radiation pressure and circumstellar dust on motion of a test particle in Manev's field. *Astrodynamics* **5**: 77–89. <https://doi.org/10.1007/s42064-020-0071-z>.
- Son J, Moon Y-J, Shin S. 2022. 72-hour time series forecasting of hourly relativistic electron fluxes at geostationary orbit by deep learning. *Space Weather* **20**(10): e2022SW003153. <https://doi.org/10.1029/2022SW003153>.
- Subbotin D, Shprits Y. 2009. Three-dimensional modeling of the radiation belts using the Versatile Electron Radiation Belt (VERB) code. *Space Weather* **7**(10): S10001. <https://doi.org/10.1029/2008SW000452>.
- Subbotin D, Shprits Y, Ni B. 2010. Three-dimensional VERB radiation belt simulations including mixed diffusion. *J Geophys Res Space Phys* **115**(A3): A03205. <https://doi.org/10.1029/2009JA015070>.
- Subbotin D, Shprits Y, Ni B. 2011. Long-term radiation belt simulation with the VERB 3-D code: comparison with CRRES observations. *J Geophys Res Space Phys* **116**(A12): A12210. <https://doi.org/10.1029/2011JA017019>.
- Summers D, Thorne RM, Xiao F. 1998. Relativistic theory of wave-particle resonant diffusion with application to electron acceleration in the magnetosphere. *J Geophys Res Space Phys* **103**(A9): 20487–20500. <https://doi.org/10.1029/98JA01740>.
- Sun X, Lin R, Liu S, He X, Shi L, Luo B, Zhong Q, Gong J. 2021. Modeling the relationship of 2 MeV electron fluxes at different longitudes in geostationary orbit by the machine learning method. *Remote Sens* **13**(17): 3347–. <https://doi.org/10.3390/rs13173347>.
- Sun X, Lin R, Liu S, Luo B, Shi L, Zhong Q, Luo X, Gong J, Li M. 2023. Prediction models of 2 MeV electron daily fluences for 3 days at GEO orbit using a long short-term memory network. *Remote Sens* **15**(10): 2538. <https://doi.org/10.3390/rs15102538>.
- Thébault E, Finlay CC, Beggan CD, Alken P, Aubert J, et al. 2015. International geomagnetic reference field: the 12th generation. *Earth Planet Space* **67**: 1–19. <https://doi.org/10.1186/s40623-015-0228-9>.
- Tsyganenko N, Sitnov M. 2005. Modeling the dynamics of the inner magnetosphere during strong geomagnetic storms. *J Geophys Res Space Phys* **110**(A3): A03208. <https://doi.org/10.1029/2004JA010798>.
- Tsyganenko NA. 1989. A magnetospheric magnetic field model with a warped tail current sheet. *Planet Space Sci* **37**(1): 5–20. [https://doi.org/10.1016/0032-0633\(89\)90066-4](https://doi.org/10.1016/0032-0633(89)90066-4).
- Tu W, Cunningham G, Chen Y, Henderson M, Camporeale E, Reeves G. 2013. Modeling radiation belt electron dynamics during GEM challenge intervals with the DREAM3D diffusion model. *J Geophys Res Space Phys* **118**(10): 6197–6211. <https://doi.org/10.1002/jgra.50560>.
- Turner DL, Li X. 2008. Quantitative forecast of relativistic electron flux at geosynchronous orbit based on low-energy electron flux. *Space Weather* **6**(5): 05005. <https://doi.org/10.1029/2007SW000354>.
- Ukhorskiy A, Sitnov M, Sharma A, Anderson B, Ohtani S, Lui A. 2004. Data-derived forecasting model for relativistic electron intensity at geosynchronous orbit. *Geophys Res Lett* **31**(9): L09806. <https://doi.org/10.1029/2004GL019616>.
- Varotsou A, Boscher D, Bourdarie S, Horne RB, Glauert SA, Meredith NP. 2005. Simulation of the outer radiation belt electrons near geosynchronous orbit including both radial diffusion and resonant interaction with Whistler-mode chorus waves. *Geophys Res Lett* **32**(19): L19106. <https://doi.org/10.1029/2005GL023282>.
- Varotsou A, Boscher D, Bourdarie S, Horne RB, Meredith NP, Glauert SA, Friedel RH. 2008. Three-dimensional test simulations of the outer radiation belt electron dynamics including electron-chorus resonant interactions. *J Geophys Res Space Phys* **113**(A12): A12212. <https://doi.org/10.1029/2007JA012862>.
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. 2017. Attention is all you need. *Advances in neural information processing systems* **30**. <https://arxiv.org/abs/1706.03762>.
- Violet M, Frederickson A. 1993. Spacecraft anomalies on the CRRES satellite correlated with the environment and insulator samples. *IEEE Trans Nucl Sci* **40**(6): 1512–1520. <https://doi.org/10.1109/23.273511>.
- Wang D, Shprits YY. 2019. On how high-latitude chorus waves tip the balance between acceleration and loss of relativistic electrons. *Geophys Res Lett* **46**(14): 7945–7954. <https://doi.org/10.1029/2019GL082681>.
- Wang D, Shprits YY, Zhelavskaya IS, Effenberger F, Castillo AM, Drozdov AY, Aseev NA, Cervantes S. 2020. The effect of plasma boundaries on the dynamic evolution of relativistic radiation belt electrons. *J Geophys Res Space Phys* **125**(5): e2019JA027422. <https://doi.org/10.1029/2019JA027422>.
- Wang R, Shi L. 2012. Study on the forecasting method of relativistic electron flux at geostationary orbit based on support vector machine. *Chin J Space Sci* **32**(3): 354–361. <https://doi.org/10.1007/s11783-011-0280-z>.
- Wei H-L, Billings S, Surjalal Sharma A, Wing S, Boynton R, Walker S. 2011. Forecasting relativistic electron flux using dynamic multiple regression models. *Ann Geophys* **29**: 415–420. <https://doi.org/10.5194/angeo-29-415-2011>.
- Wei L, Zhong Q, Lin R, Wang J, Liu S, Cao Y. 2018. Quantitative prediction of high-energy electron integral flux at geostationary orbit based on deep learning. *Space Weather* **16**(7): 903–916. <https://doi.org/10.1029/2018SW001829>.
- Wrenn G, Rodgers D, Ryden K. 2002. A solar cycle of spacecraft anomalies due to internal charging. *Ann Geophys* **20**: 953–956. <https://doi.org/10.5194/angeo-20-953-2002>.
- Wrenn G, Sims A. 1996. Internal charging in the outer zone and operational anomalies. *Radiation belts: models and standards* **97**: 275–278. <https://doi.org/10.1029/GM097P0275>.
- Xue B, Ye Z. 2004. Forecast of the enhancement of relativistic electron at the geo-synchronous orbit. *Chin J Space Sci* **24**(4): 283–288. <https://doi.org/10.1007/BF02911033>.
- Zeng A, Ju X, Yang L, Gao R, Zhu X, Dai B, Xu Q. 2022. Deciwatc: A simple baseline for 10× efficient 2D and 3D pose 53 estimation. In: *European Conference on Computer Vision. ECCV 2022*, Springer Nature Switzerland, Cham, Switzerland, pp. 607–624.
- Zhang H, Fu S, Xie L, Zhao D, Yue C, et al. 2020. Relativistic electron flux prediction at geosynchronous orbit based on the neural network and the quantile regression method. *Space Weather* **18**(9): e2020SW002445. <https://doi.org/10.1029/2020SW002445>.

Cite this article as: Sun X, Wang D, Drozdov A, Lin R, Smirnov A, et al. 2024. A modeling study of ≥ 2 MeV electron fluxes in GEO at different prediction time scales based on LSTM and transformer networks. *J. Space Weather Space Clim.* **14**, 25. <https://doi.org/10.1051/swsc/2024021>.