Key Points:
- Neural network models are derived for geosynchronous electron flux and an upper boundary scaling factor for radiation belt simulations
- Included are the magnetic local time, solar wind speed, Dst, solar wind dynamic pressure, AE, and Kp inputs improved model results
- The neural network derived upper boundary scaling factor may be used to reconstruct radiation belt dynamics without in-situ observations

NARX Neural Network Derivations of the Outer Boundary Radiation Belt Electron Flux

D. A. Landis[1,2], A. A. Saikin[1] [image], I. Zhelavskaya[2] [image], A. Y. Drozdov[1] [image], N. Aseev[2] [image], Y. Y. Shprits[1,2,3] [image], M. F. Pfitzer[2], and A. G. Smirnov[2,3] [image]

[1]Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, CA, USA, [2]Helmoltz Centre Potsdam- GFZ German Research Centre for Geosciences, Potsdam, Germany, [3]Institute of Physics and Astronomy, University of Potsdam, Potsdam, Germany

Abstract We present two new empirical models of radiation belt electron flux at geostationary orbit. GOES-15 measurements of 0.8 MeV electrons were used to train a Nonlinear Autoregressive with Exogenous input (NARX) neural network for both modeling GOES-15 flux values and an upper boundary condition scaling factor (BF). The GOES-15 flux model utilizes an input and feedback delay of 2 and 2 time steps (i.e., 5 min time steps) with the most efficient number of hidden layers set to 10. Magnetic local time, Dst, Kp, solar wind dynamic pressure, AE, and solar wind velocity were found to perform as predicative indicators of GOES-15 flux and therefore were used as the exogenous inputs. The NARX-derived upper boundary condition scaling factor was used in conjunction with the Versatile Electron Radiation Belt (VERB) code to produce reconstructions of the radiation belts during the period of July–November 1990, independent of *in-situ* observations. Here, Kp was chosen as the sole exogenous input to be more compatible with the VERB code. This Combined Release and Radiation Effects Satellite-era reconstruction showcases the potential to use these neural network-derived boundary conditions as a method of hindcasting the historical radiation belts. This study serves as a companion paper to another recently published study on reconstructing the radiation belts during Solar Cycles 17–24 (Saikin et al., 2021, https://doi.org/10.1029/2020sw002524), for which the results featured in this paper were used.

Plain Language Summary Earth's radiation belts are comprised of two highly dynamic regions consisting of very energetic charged particles (protons and electrons). This paper presents two models that predict electron fluxes at geosynchronous orbit (i.e., the outer radiation belt) and create a scaling factor that can be used in simulations of the radiation belt. Both models are derived using satellite measurements of energetic electrons and a neural network-based machine-learning algorithm, the Nonlinear Autoregressive with Exogenous input (NARX). Common geomagnetic activity indices are used as driving inputs for the model. We compare our geosynchronous electron flux model to satellite observations to showcase their performance. Using our NARX-derived scaling factor, we reconstruct the radiation belts between July and November 1990, and compare it with contemporaneous satellite measurements to show how our model can reproduce observations. Our model allows us to simulate the historical radiation belts from as early as the 1930s, when satellite measurements were unavailable.

1. Introduction

Earth's radiation belts consist of two torus-shaped rings of highly energetic electrons and protons (Van Allen & Frank, 1959). The inner belt, located between $L = 1$ and 3 (Roederer, 1970), is composed of ~100–700 keV electrons and ~10–100 MeV protons (Fennell et al., 2015), while the outer radiation belt, $L = 2.8$–8, is composed of primarily 0.5–10 MeV electrons. Though the inner radiation belt remains generally stable, the outer radiation belt may experience dynamic variations, especially during periods of enhanced geomagnetic activity (Baker et al., 1986, 1997; Craven, 1966; Reeves et al., 1998; Rothwell & Mcilwain, 1960).

Geomagnetic storms play a crucial role for electron enhancement events within the radiation belts (Bingham et al., 2018; Drozdov et al., 2019; Reeves et al., 2003; Turner et al., 2014, 2019), as they can facilitate wave generation (e.g., EMIC, chorus, etc.). The following resonant wave-particle interactions, resulting from this wave generation, impact the dynamic processes within the radiation belts (e.g., Jordanova et al., 2008; Lyons et al., 1972; Summers et al., 2007a; Summers & Throne, 2003, etc.). Radiation belt electrons may be accelerated

or scattered at certain energies, depending on the type of wave activity [for example, hiss (Shprits et al., 2008), EMIC (e.g., Drozdov et al., 2017; Summers & Throne, 2003; Zhang et al., 2016), and chorus (e.g., Summers et al., 2007a, 2007b), etc.]. Understanding in detail how these electron acceleration and loss processes affect the outer radiation belt remains a dominant topic of research within the space science community. This introduction will provide a summary overview of magnetospheric drivers, previous models (and their inputs [e.g., geomagnetic indices, solar wind conditions, and electron flux observations]), and the techniques used to develop these models. The goal of this study is to develop two new models independent of in situ observations.

The dynamics of the Earth's magnetosphere are heavily susceptible to and driven by external solar wind conditions (e.g., solar wind dynamic pressure (Pdyn), solar wind speed (Vsw), etc.), and the level of geomagnetic activity (e.g., *Kp*, Dst, Sym-H, and AE, etc.; Reeves et al., 2013; Saikin et al., 2016). Hence, these indices are often considered as proxies to explain electron enhancements and/or depletions (at their respective energies) in the radiation belts. Long-term recorded observations of geomagnetic activity exist for some indices as early as the 1860s (i.e., the aa index; Bernoux & Maget, 2020).

However, direct observations of the radiation belts are limited by the uneven spatial and time distribution of previous missions and may bias the interpretation of the data (Abel et al., 1994; Saikin et al., 2015). Therefore, simulations and modeling have become the prominent methods of exploring the physics that govern the radiation belts in a controlled environment. The use of *Kp*-driven models has become more prevalent over the last few decades (e.g., Beutier & Boscher, 1995; Drozdov et al., 2015, 2017; Glauert et al., 2014; Ma et al., 2016; Reeves et al., 2012; Shprits & Ni, 2009; Shprits et al., 2008; Subbotin & Shprits, 2009), while other models, like DREAM (Reeves et al., 2012) or the Salammbô electron code (Maget et al., 2006) may incorporate both geomagnetic data (i.e., Dst or *Kp*) along with data observations from satellite missions. Although these models and methods have proven successful in replicating the respective mission's observations, the necessity of *in-situ* data observations to produce those simulations constricts the time periods we may examine to those when near-continuous observations exist (i.e., post-1986 (Glauert et al., 2018)).

In recent years, machine learning has become a more prominent tool in its role to help model the radiation belts (Balikhin et al., 2016; Ganushkina et al., 2015; Saikin et al., 2021) or to forecast geomagnetic indices (Ayala Solares et al., 2016; Camporeale, 2019). Specifically, the use of the nonlinear autoregressive with exogenous input (NARX) models has been successfully employed to predict and model geomagnetic indices (such as Dst (Balikhin et al., 2011; Boynton et al., 2011; Wei et al., 2004) and *Kp* (Ayala Solares et al., 2016)). NARX is a machine learning algorithm that uses both current and past time step inputs to determine current and future outputs. Unlike other machine learning algorithms, with NARX, the time history is automatically incorporated and therefore does not require the additional construction of time series for historical values as additional inputs. Another machine learning algorithm, the Nonlinear AutoRegressive Moving Average with eXogenous inputs (or NARMAX), is like the NARX model, except that it incorporates a moving average of the error. NARX, for simplicity, reduces this moving average error term to a singular noise term.

Though most studies have focused on the forecasting of geomagnetic indices, it has been suggested that a NARX neural network could be employed to determine boundary conditions used in simulation studies (Camporeale, 2019). Given NARX's predictive abilities, we can train an empirical model based on the relationship between radiation belt measurements and geomagnetic activity. These neural networks will allow us to simulate and explore the radiation belts independently of *in-situ* observations and expand the available radiation belt data sets used for analysis (provided the geomagnetic or solar wind drivers are provided). Recently, the results and techniques presented in this study were used to perform a series of radiation belt reconstructions. In the companion paper, Saikin et al. (2021) employed the use of the NARX neural network-derived upper boundary condition presented in this study, along with the VERB-3D code, to hindcast the radiation belt electrons from 1933 to 2017 (i.e., Solar Cycles 17–24) at 0.5, 1.0, and 2.0 MeV. This combination of empirical and physics models into gray-box models (i.e., models which incorporate a data-driven set of parameterization inputs with a physics-based construction) is particularly promising for space weather applications and modeling radiation belt electrons and should be continuously explored and refined (Camporeale, 2019).

This study focuses on the creation and utilization of a NARX neural network to model geosynchronous flux (with a 5-min cadence) and determine an upper boundary condition that may be used to model the radiation belts, independent of direct observations. The geosynchronous model is the first to predict with such high resolution
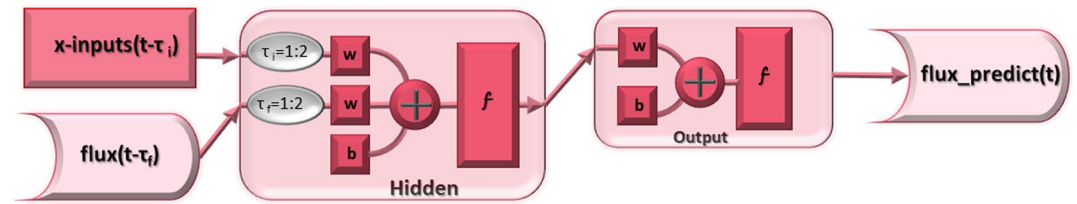
Figure 1. A diagram of the structure of the Nonlinear Autoregressive with Exogenous input model. The x-inputs (t - τ_i) are the exogenous data time series where τ_i indicates the delay from time step t. The gray circles indicate the input and feedback delays for the exogenous time series and the flux (t - τ_f), with their respective number of time steps (2 and 2 respectively, see Section 3.1 for details). Both flux and the x-inputs are weighted, w, while b denotes biases. The f represents a sigmoid function.

and to not recalibrate after the initial starting point (being driven solely by geomagnetic indices). This manuscript is organized as follows: in Section 2, we describe the NARX neural network in detail (its inputs and exports) followed by the data observations used in the construction of the neural network (i.e., geomagnetic, solar wind conditions, and GOES observations) and comparison (CRRES). Section 3 describes how the NARX neural network was optimized and how the exogenous inputs were selected (by providing correlation coefficients (CCs) based on the model performance against GOES-15 measurements). In Section 4, we apply our derived neural network GOES-15 measurements and our upper boundary condition to another period independent of the GOES observations (i.e., CRRES). Finally, in Sections 5 and 6, we discuss our results and summarize our conclusions, respectively.

2. Methodology

2.1. Description of NARX

The NARX model is a type of recurrent neural network (RNN) designed to utilize time series data to model their predictions. RNNs are unique from other neural network algorithms in that, when they examine the connection between two nodes, RNN expects to find causality in only one direction, preserving the temporal order of the nodes (i.e., a future node or event cannot be the cause of a previous one). Following the description in Ayala Solares et al. (2016), the NARX model can be formalized as follows:

$$y(k) = f\left(y(k-1), \ldots, y\left(k - n_y\right), u(k-1), \ldots, u\left(k - n_u\right)\right) + e(k) \tag{1}$$

where $f(y, u)$ is a function that will be determined from the data. The functions $u(k)$ and $y(k)$ are the system input and output signals, respectively. e(k) represents the error (i.e., system noise, with k = 1, 2, ….N), and n_u and n_y are the maximum lags for the input and output signals, respectively. By assuming that $f(y, u)$ can be approximated by a linear combination of a predefined set of functions $\Phi_i(\varphi(k))$, Equation 1 can be expressed in a linear-in-the-parameters form:

$$y(k) = \sum_{i=1}^{M} \theta_i \Phi_i(\varphi(k)) + e(k) \tag{2}$$

where θ_i are the coefficients to be determined (i.e., estimated), $\Phi_i(\varphi(k))$ are the predefined functions that depend on the regression vector $\varphi(k) = \left[y(k-1), \ldots, y\left(k - n_y\right), u(k-1), \ldots, u\left(k - n_u\right)\right]^T$ of past inputs and outputs, and where M is the number of functions used in the set. Figure 1 shows a basic configuration of the NARX model used in this paper. The exogenous inputs, "x-inputs(t-τ_i)", represent the data time series used to guide and estimate the output parameter, "flux(t-τ_f)", the past values of which are also inputted into the network. The initial training stage incorporates both input and feedback delays (in 2 and 2 time steps, respectively; see Section 3 for more detail) for both the inputs and past flux to determine the appropriate causal relationship. The input delay represents the number of past time steps (including the current time t) of the exogenous inputs (e.g., AE, Dst, MLT, and Kp, etc.) that are used to predict the current time step's target value (i.e., flux). Similarly, the feedback delay is, exclusively, the number of past time steps used for the target quantity (i.e., the flux) to predict the current value. The initial flux and input parameters are then weighted, w, and modified with the respective biases, b. These results are then filtered through a sigmoid function, f, which serves to create a newly trained flux model.

Through each iteration, the weights and bias are optimized. This process continues until the best performing neural network is produced (preferably before the model becomes overtrained).

For this study, we have created two NARX models derived from the GOES-15 measurements. The first model was designed to model GOES-15 flux measurements of 0.8 MeV electrons. The second model is to describe an upper boundary scaling factor (BF) at $L* = 6.6$, used with the VERB-3D code to hindcast simulations of the radiation belts. We use and train NARX to obtain this BF scaling factor that can be multiplied by an average spectrum at GEO J (E, L = 7; Shprits et al., 2006; Subbotin & Shprits, 2009; Subbotin, Shprits, & Ni, 2011). Following the approach of Brautigam and Albert (2000), it is assumed that the flux at $L* = 6.6$ is determined by J_{model} (GEO) = $J_{averaged}$(at GEO) * BF. For comparison, we will examine our hindcasted results with the CRRES mission observations during the July–November 1990 period to show that crafting a NARX model during the GOES-15 era can be used to aid simulations when contemporaneous *in-situ* measurements are unavailable, as was done in Saikin et al. (2021).

2.2. Observations

2.2.1. GOES

Observations from the Geostationary Operational Environmental Satellite-15 (GOES-15) were used in the development of the neural network. The Energetic Proton Electron and Alpha Detector provides 1-min measurements of electrons in two integral energy channels (>0.8 MeV and >2.0 MeV) at geostationary orbit (6.6 R_e). The east and west telescopes on board GOES-15 are averaged together. The full energy spectrum range was represented by assuming an exponential spectrum nature with the 0.8 and 2.0 MeV energy GOES-15 energy channels. The integral flux was converted to differential flux at the two energies (0.8 and 2.0 MeV) and with a linear interpolation of the logarithm of the flux between them. At each data point, $L*$ was calculated using the T89 magnetic field model (Tsyganenko, 1989). We determined the respective energy for each GOES $L*$ that corresponded to 1 MeV at $L* = 6.6$, using the conservation of the first adiabatic invariant (μ). This derived differential flux was then converted to phase space density (PSD), for the 1 MeV electrons at $L* = 6.6$ (assuming a flat PSD vs. $L*$). A sine function distribution for the pitch angle measurements (taken from the MAGnetopsheric Electron Detector) was used at the outer boundary, and we use the long-term averaged spectrum to represent the energy distribution of the outer boundary PSD. This method of determining PSD and differential flux from the GOES-15 measurements has proven to be an effective choice when comparing simulations to observations (Cervantes et al., 2020; Saikin et al., 2021).

2.2.2. CRRES

Launched in July 1990, the Combined Release and Radiation Effects Satellite (CRRES) executed an orbit with a perigee and apogee of 1.05 R_e and 6.26 R_e, respectively, with an orbital period of 9.4 hr at an inclination of 18°. On board the CRRES mission, electron measurements were taken by the Medium Electron A (MEA) instrument (Vampola et al., 1992) in a logarithmic distribution of energies from 0.15 to 1.58 MeV over a total of 17 channels. The CRRES mission provided *in-situ* measurements of the radiation belts for 14 months during Solar Cycle 22.

2.3. The VERB 3D Code

To simulate and model the relativistic electrons within the radiation belts, we employ the VERB-3D code to solve the Fokker-Planck equation numerically (Shprits et al., 2008; Subbotin & Shprits, 2009; Subbotin, Shprits, & Ni, 2011; Subbotin, Shprits, Gkioulidou, et al., 2011). The Fokker-Planck equation encompasses several processes to describe the dynamics and evolution of relativistic electron populations. Through wave-particle interactions (i.e., radial transport, local acceleration, or the loss of electrons to the atmosphere), chorus, plasmaspheric hiss, electromagnetic ion-cyclotron (EMIC) waves, and lightning whistler waves often impact the electron population through a combination of pitch-angle, energy, mixed diffusion, and radial diffusion (caused by ultralow frequency waves).

Following the works performed by Saikin et al. (2021) and Subbotin and Shprits (2012), we use a single grid configuration of modified adiabatic invariants K ($K = J/\sqrt{8m_0\mu}$, where μ and m_0 represent the first adiabatic invariant and the electron mass, respectively) and V ($V = \mu * (K + 0.5)^2$) to solve the 3D Fokker-Planck equation. The use of these parameterizations is convenient for numerical calculations and defining boundary

conditions because K is independent of the particle's energy, and V is only loosely dependent on the particle's pitch angle. This method allows us to sidestep the interpolation between numerical grids used in earlier VERB-3D simulations, which reduces unstable code behavior and numerical errors (Subbotin & Shprits, 2009). Using these modified invariants, the Fokker-Planck equation takes the form:

$$
\begin{aligned}
\frac{df}{dt} = {} & \frac{1}{G}\frac{\partial}{\partial L}G\left\langle D_{LL}\right\rangle\frac{\partial f}{\partial L} + \frac{1}{G}\frac{\partial}{\partial V}G\left(\left\langle D_{VV}\right\rangle\frac{\partial f}{\partial V} + \left\langle D_{VK}\right\rangle\frac{\partial f}{\partial K}\right) \\
& + \frac{1}{G}\frac{\partial}{\partial K}G\left(\left\langle D_{KV}\right\rangle\frac{\partial f}{\partial V} + \left\langle D_{KK}\right\rangle\frac{\partial f}{\partial K}\right) - \frac{f}{\tau}
\end{aligned}
\tag{3}
$$

where f represents the three-dimensional PSD and τ is the electron's lifetime. Here, $\langle D_{LL}\rangle$, denotes the radial diffusion coefficients while $\langle D_{VV}\rangle$, $\langle D_{KK}\rangle$, $and\ \langle D_{VK}\rangle$ represent the drift and bounce-averaged diffusion coefficients. The required Jacobian transformation from an adiabatic invariant system (μ, J, Φ) is denoted by G ($G = -2\pi B_0 R_E^2 L^{-2}\sqrt{8\mu m_0}/(K + 0.5)^2$). R_E represents the Earth radius and $B_0 = 0.3$ G (the magnetic field on the equator of the Earth's surface). Finally, f/τ is a loss term accounting for losses to the atmosphere and those caused by magnetopause shadowing.

All simulations were performed on an orthogonal grid of size $29 \times 62 \times 61$ points for L, V, and K, respectively. The boundary conditions for these simulations are set at $L^* = 6.6$ (with the total L^* range set from 1.0 to 6.6) for energies from 10 keV to 10 MeV (V) and pitch angles from 0.7° to 89.3° (K), respectively. At the equator, K is defined as 0 G^0.5*R_e. The V grid points are distributed on a logarithmic scale, while the L^* and pitch angle grid points are distributed linearly. Losses within the loss cone are denoted by the lower K boundary condition and do not represent the loss cone; effective loss within the loss cone is simulated through the f/τ term. PSD is set to zero at the lower K boundary, while the upper K boundary condition (also set to 0) is set to a zero-gradient PSD (which represents a flat distribution at 90°). PSD at the upper V boundary is set to zero, while PSD at the lower V boundary is set to an initial value and remains constant, representing the steady state balance between sources and losses of the low energy population. The VERB code does not include convection with previous studies that have shown that a lower energy boundary condition set at ~tens of keV has little effect on radiation belt electrons (Castillo et al., 2019; Subbotin, Shprits, Gkioulidou, et al., 2011). To represent losses to the atmosphere, the lower boundary condition in L^* is set to zero.

In keeping with previous VERB-3D code simulation studies (Kim et al., 2011; Subbotin, Shprits, & Ni, 2011), Kp-dependent electromagnetic (DLL_m) radial diffusion coefficients derived from Brautigam and Albert (2000) were used. Though initially designed for $Kp \leq 6$, we have used the functional dependence derived by Brautigam and Albert (2000) for the radial coefficient model to include the $Kp > 6$ values.

Wave particle interactions through gyro-resonance are also incorporated using previously derived parameterizations of waves occurring both inside and outside the plasmasphere. Within the plasmasphere, the simulation includes wave-particle parameterizations from Very Low Frequency (VLF) transmitter-generated whistler waves (Subbotin, Shprits, & Ni, 2011; Subbotin, Shprits, Gkioulidou, et al., 2011), plasmaspheric hiss (Zhu et al., 2019), and lightning waves (Subbotin, Shprits, & Ni, 2011; Subbotin, Shprits, Gkioulidou, et al., 2011). Beyond the plasmasphere, chorus waves were parameterized using wave amplitude and frequency measurements from the Van Allen Probes, corresponding to the dayside and nightside magnetosphere, respectively (Orlova et al., 2016). Following Saikin et al. (2021), EMIC wave-related diffusion coefficients (Drozdov et al., 2017) were included when the Kp index exceeded 2 with an assumed wave power of 0.1 nT². Diffusion coefficients were derived using the Full Diffusion Code (Orlova & Shprits, 2011; Shprits & Ni, 2009), based on the parameters obtained from those previous studies. They were then converted into the modified adiabatic invariants (V and K) following the methodology described by Subbotin and Shprits (2012). Those coefficients were used to perform 90-year simulations in the companion paper (Saikin et al., 2021). For this simulation, the plasmapause boundary was modeled using the Carpenter and Anderson (1992) model.

3. Developing the NARX Model

3.1. Optimizing Structure for the NARX Flux Model

To begin developing a NARX neural network for both flux measurements and BF, we must first optimize the architecture. This is done by determining the appropriate time series input delay and feedback delay with the highest correlations compared to the test data set, while also finding the appropriate number of hidden layers.
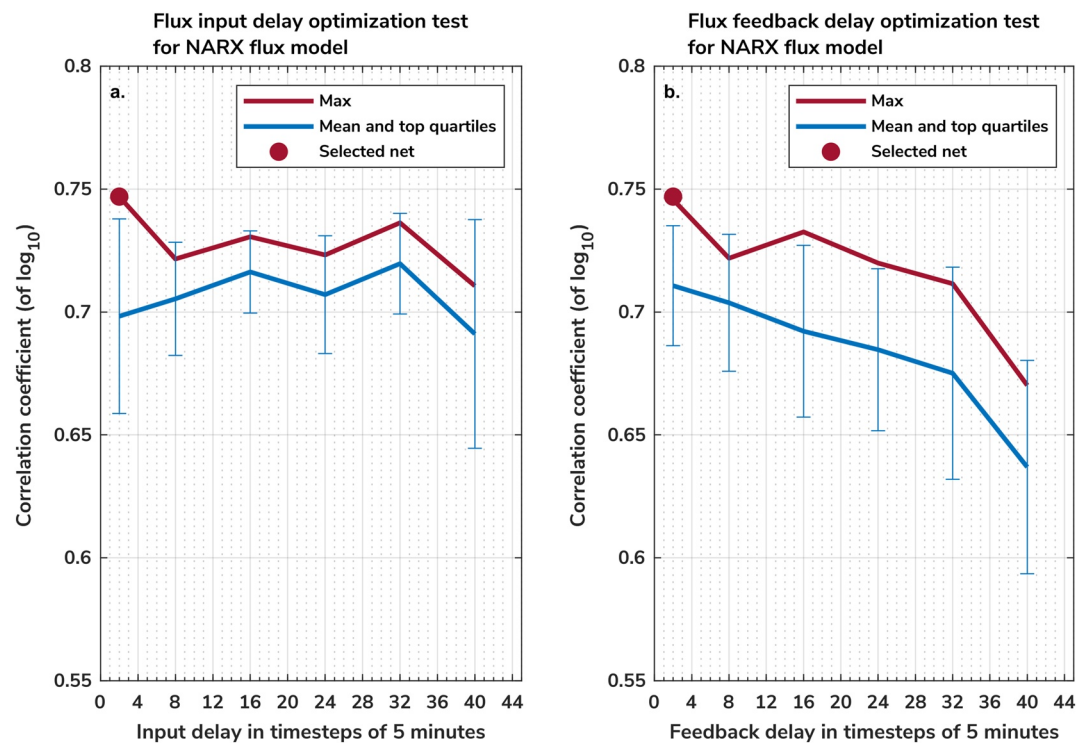
Figure 2. The correlation coefficients (CCs) of log_{10} results associated with testing the input (a) and feedback delay (b) time steps used in constructing the Nonlinear Autoregressive with Exogenous input flux model. Each CC represents the maximum derived value (maroon) and the average and top quartiles of the 20 independent neural network trials (blue), with each time step representing 5 min. The selected delays are denoted by the maroon dot.

We began with determining the input and feedback delays. Using the GOES-15 > 800 keV electron measurements from June 2013 to June 2016, we trained 20 independent neural networks. These independently trained networks were then run on a 2-month test period (July–August 2016). Figure 2 shows the average of the top five CCs between the predicted and observed fluxes associated with the corresponding input (Figure 2a) and feedback delay (Figure 2b) for those 20 independent runs on the July–August 2016 test period (where each delay corresponds to a 5-min time step). The input delay for these test runs reaches its highest CC (∼0.74) when the input delay is set to 2 time steps (i.e., 10 min, the minimum of time steps possible for a NARX neural network) and proceeds to decrease down to a correlation of ∼0.71 at 40 time steps. However, the feedback delay similarly begins with a CC of ∼0.75 and begins to increase to ∼0.66 at 40 time steps (i.e., 200 min). From this point, the feedback delay CCs decrease with increasing time step inputs. With these results, we decided to use our model assuming an input and feedback delay of 2 and 2 time steps, respectively.

The next step in optimizing the NARX structure is to determine the appropriate number of hidden layer neurons. Although the number of hidden layer neurons can be determined by the number of exogenous inputs included in the model and the number of outputs, experiments can yield better results. To determine the appropriate number of hidden neurons to include, using the optimal input and feedback delays found in Figure 2, we trained 20 different NARX neural networks with the same training period. Once trained, we compared and correlated each model against the test period, again July–August 2016. After completing these correlations, we increased the number of hidden layer neurons and repeated the process. Figure 3 shows the averaged top five and maximum CCs determined from each of the variable number of hidden neurons chosen. The highest correlation (∼0.74) occurred when the number of hidden neurons was set at 16. While this number of hidden neurons is less than the number of neurons suggested by heuristic methods, higher values might cause overfitting. For the purposes of modeling the GOES-15 flux measurements, the input delay, the feedback delay, and the number of hidden layer neurons chosen for the neural network were set at 2 time steps (i.e., 10 min), 2 time steps (i.e., 10 min), and 16 hidden neurons, respectively, to produce a GOES flux model with a 5-min cadence.
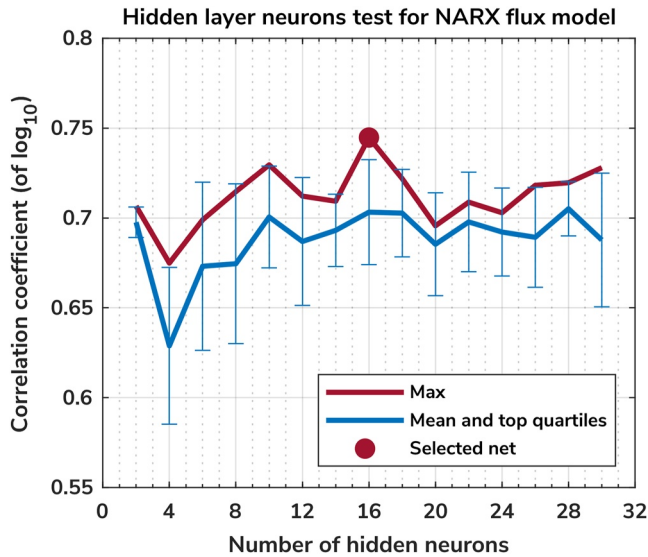
Figure 3. The maximum $\log_{10}$ correlation coefficient (maroon) over 20 trials, determined from the testing period, against the number of hidden layer neurons (*x*-axis) used in the network. The mean top quartile (blue) is also plotted. The number of hidden layers initially began at 2 and was increased, in 2 step intervals, to a maximum of 30 hidden layers. The dot represents the selected number of hidden neurons.

3.2. Selection of Exogenous Inputs for NARX Flux Model

An essential component involved in the development of NARX models is the choice of exogenous inputs. While several inputs exist (e.g., *Kp*, Dst, AE, and Vsw, etc.), which can be used to train the neural network, it is important to verify which indices provide the best outcome when compared to the 0.8 MeV electrons measured by GOES. The inclusion of too much information (i.e., too many inputs/indices) may damage the performance of the neural network due to overfitting. The second stage of optimizing our NARX neural network therefore pertains to determining the most efficient selection of indices.

For this testing, we decided to focus on seven parameters (*Kp*, MLT, Dst, AE geomagnetic indices, solar wind speed [Vsw], and solar wind dynamic pressure) and one locational metric (magnetic local time [MLT]). All trial runs used the same training period (June 2013–June 2016) and the same architecture were derived from Section 3.1 and were tested against the same testing period of July–August 2016.

Table 1 shows the CCs for the input tests when that respective metric is removed from the network training. Of the tested metrics, the best performing neural network with the highest CC is the one where none of the indices are omitted (~0.74). Conversely, the lowest CC (~0.25) appears when MLT is removed as an input, indicating the necessity of the MLT measurements. Since the omission of no index yielded the highest average correlation, for the purposes of this model, we have included all of the exogenous inputs: Vsw, MLT, *Kp*, Pdyn, AE, and Dst.

3.3. Optimizing Structure for NARX BF Model

The same procedure outlined above was also used in determining the appropriate architecture for the BF model (to be used in the VERB-3D simulations). BF was determined (as described in Section 2) from GOES-15 measurements with the period of 1 January 2011 to 25 July 2015, used as the training set. The testing period was set from 26 July 2015 to 30 April 2016. Figure 4a shows the matrix of CCs (correlated to the testing period) determined from varying combinations of input and feedback delays, each using intervals of 1-hr time steps. Figure 4b shows the corresponding standard deviation (over the 20 trial runs) associated with that respective input and feedback delay combination. The highest correlation (~0.53) was found when both the input and feedback delays were set at 60 time steps (60 hr). This combination of input and feedback delay also corresponded to the lowest (~0.022) standard deviation among all configurations examined.

We have repeated the same processes in determining the appropriate number of hidden layers to use for the BF model. Figure 5 shows the CCs as a function of the number of hidden layers used, in the same format as Figure 3. Unlike the hidden layers test for the flux model, which featured a singular peak number at 10 hidden layers followed by descending correlations, the number of hidden layers for the BF model peaks at its starting number of two hidden layers (the minimum possible number of hidden layers) and then proceeds to decline. As the number of hidden layers increases, so do the corresponding standard deviations. Since the VERB-3D code uses *Kp* as its driving geomagnetic index and the corresponding time data series, we have chosen only to use *Kp* as the exogenous input for this model, which may explain why the choice of two hidden layers, fewer than the other model,

Table 1
Correlation Coefficient by Removed Index

	MLT	*Kp*	Dst	AE	Vsw	Pressure	None
Max correlation coefficient	0.25	0.65	0.69	0.73	0.63	0.73	0.74
Upper quartile average	0.19	0.63	0.68	0.69	0.62	0.72	0.70

a. Correlation coefficient of delay groupings for NARX BF model

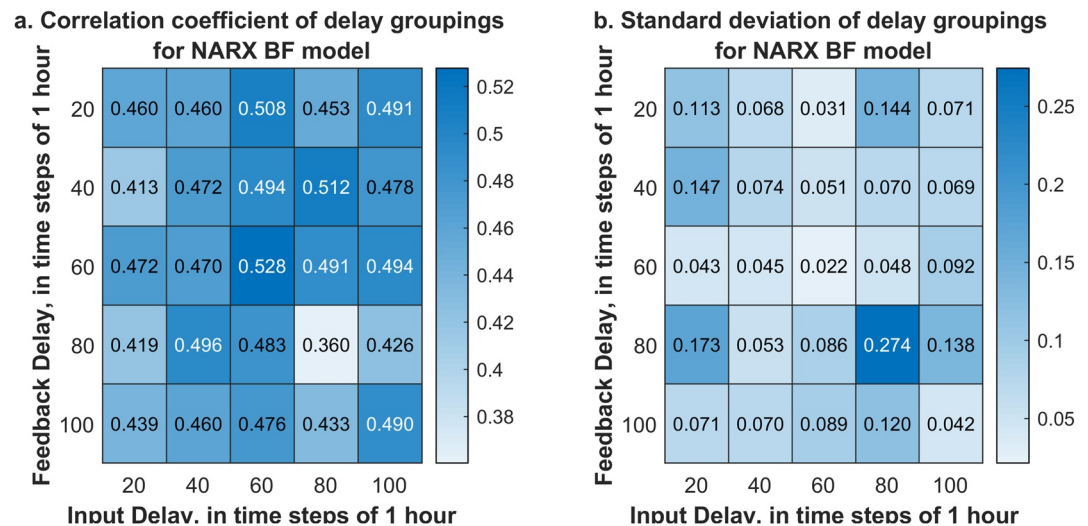**b. Standard deviation of delay groupings for NARX BF model**

Figure 4. The averaged (over 20 trials) correlation coefficients associated with variable input delays and feedback delays in 1-hr time steps for the (a) Nonlinear Autoregressive with Exogenous input-derived boundary scaling factor models, and (b) the respective standard deviation of those trials.

provided the best correlations. Furthermore, historical measurements of Kp exist as early as the 1930s, allowing this BF model in conjunction with the VERB code to simulate the radiation belts before satellite observations existed. Provided that Kp exists as an exogenous input, we may derive BF for a prolonged period.

4. Results of NARX-Derived Models

4.1. Results of the NARX Flux Model and Comparison With GOES-15 Observations

Using the parameters derived in Section 3, we trained a NARX geosynchronous flux model with the GOES-15 observations during the period of (June 2013–June 2016). Figure 6a shows the comparison of 0.8 MeV electron flux (#/sr/s/cm²/keV) between the GOES-15 observed electron flux (blue) and the NARX modeled electron flux (maroon, with the bias factor) over the period of July 2016 (after the training period, June 2013–June 2016), which was not included in the training or validation sets, with Figure 6b showing the corresponding Kp values. Overall, the model performs well with maintaining the cyclical nature of the GOES-15 measurements. Throughout this period, the modeled results are considered well correlated with the observations with a CC of ~0.68 and a prediction efficiency (PE) of ~0.39 over the month-long period. While Kp for this period remains relatively low ($Kp \leq 5$), our simulation results for the March 2013 geomagnetic storm ($Kp > 5$) can be found within the Supporting Information S1.
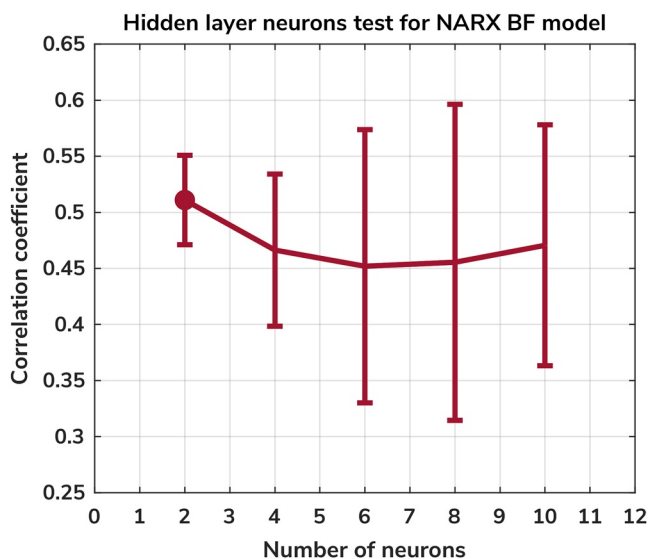
Figure 5. The average correlation coefficients across 20 trials by number of hidden neurons used for the boundary scaling factor model derivation and the standard deviation. The selected number is denoted by the red dot.

4.2. Results of the NARX BF Model, VERB Simulations, and Comparison With CRRES Observations

We have decided to compare our NARX-derived model BF to another period independent of the testing period to examine whether this new BF model can be used in radiation belt simulations independent of contemporaneous (i.e., another solar cycle) *in-situ* observations. We have chosen to examine a period which has been examined with previous VERB simulation studies (e.g., Saikin et al., 2021; Subbotin, Shprits, & Ni, 2011; Subbotin, Shprits, Gkioulidou, et al., 2011), the 1 MeV 85° pitch angle electron CRRES observations between 29 July and 4 November 1990.

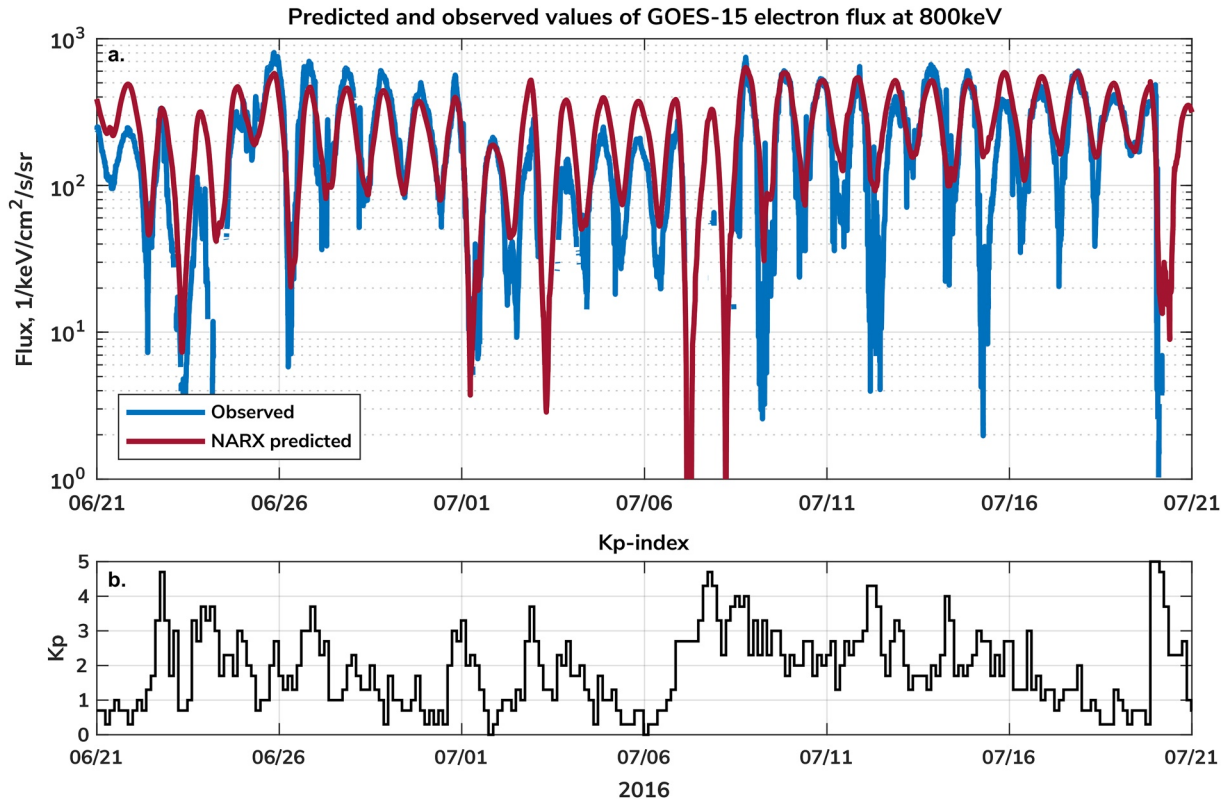**Figure 6.** (a) The logged comparison between the GOES-15 observed 0.8 MeV electrons (#/sr/s/cm^2/keV) (blue) and the corresponding Nonlinear Autoregressive with Exogenous input-derived flux measurements (maroon) during June and July 2016. (b) The corresponding *Kp* index.

Figure 7 presents the NARX-derived (maroon) and CRRES-derived BF (blue) for the 1 MeV electrons (a) and the corresponding *Kp* activity (b) during the period 29 July–4 November 1990. For this period, the NARX-derived BF mostly overperforms compared to the CRRES-derived BF (between 19 August and 30 September); however, it resembles the variability of the data-driven BF. Figure 7c displays the 1.0 MeV 85° pitch angle (#/sr/s/cm^2/keV) electrons CRRES observed during this period. Here, the CRRES data has been averaged into daily bins for comparison to the VERB simulations. To showcase the performance of the NARX-derived simulation, we have performed two VERB simulations for the 1.0 MeV 85° pitch angle electrons, one which assumes a constant upper boundary condition equal to 1 (Figure 7d), and one which uses the NARX-derived BF (Figure 7e). Note, Figure 7 has been adapted and modified from Saikin et al. (2021).

The constant boundary condition scenario fails to reproduce significant flux enhancements observed by the CRRES observations both in time (ex., 17 August–22 August) and in *L** range (ex., 12 October–21 October). Furthermore, with the constant boundary condition, there are no losses included at the upper *L* shells (i.e., ~5.5–6.6), as this scenario forces the boundary condition to 1, regardless of magnetopause-related losses. Conversely, the NARX-derived BF VERB simulation performs better than the controlled constant upper boundary condition. The NARX VERB simulation appropriately recreates enhancement periods within *L** = 3.6–5.8 (ex., 29 July–5 August; 24 August–26 August; 10 October–24 October). However, both the constant boundary and the NARX VERB code simulation do not accurately predict the decreases in the electron flux observed by CRRES (ex., 2 September–10 September). Lack of electron scattering may be attributed to the NARX-derived BF overperforming the CRRES-derived BF (the NARX BF is higher than the CRRES BF during the September period) or by the need for improvement in the diffusion coefficients related to wave-particle interactions (e.g., hiss, chorus, and VLF). The BF model derived for this paper and used in this simulation (Figure 7) is also the same model used to reconstruct the historical radiation belts used in the companion paper, Saikin et al. (2021).
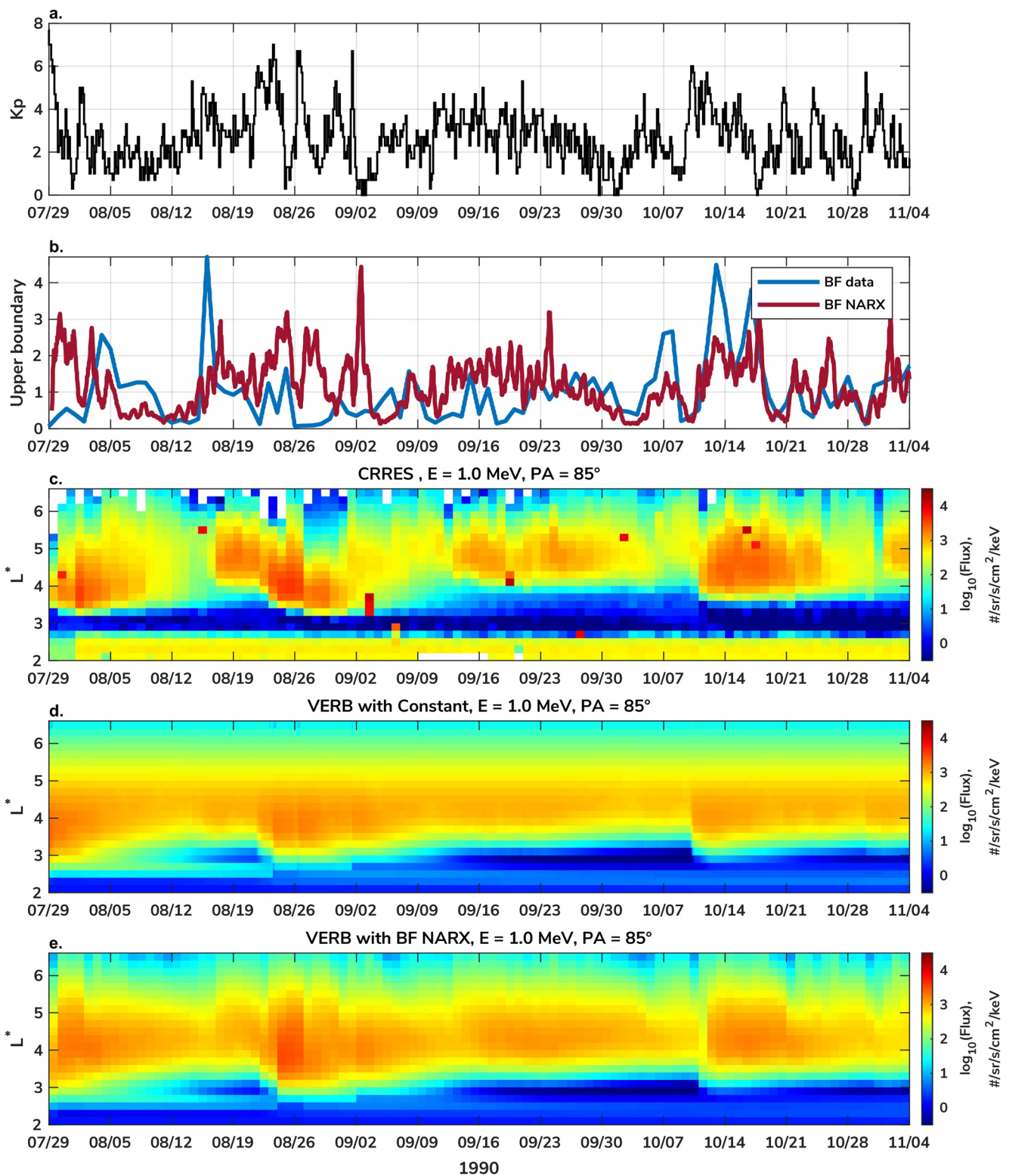
Figure 7. The *Kp* index (a) and the comparison between the normalized boundary scaling factor (BF) (upper boundary condition) index modeled by Nonlinear Autoregressive with Exogenous input (NARX) (blue) and the BF derived from the Combined Release and Radiation Effects Satellite (CRRES; orange) observations between 29 July 1990 and 4 November 1990 (b). Displayed are the corresponding 1.0 MeV 85° pitch angle electrons observed by CRRES (c), produced by the Versatile Electron Radiation Belt (VERB) code, assuming a constant upper boundary condition (set equal to 1) (d), and produced by the VERB code using the NARX-derived boundary condition (e). This figure has been adapted and modified from Saikin et al. (2021).

5. Discussion

In this study, we have utilized a NARX neural network to develop two models: a flux-based model trained on direct GOES-15 observations of 0.8 MeV electrons, and a model based on an upper boundary scaling factor, which can be used in simulations (i.e., VERB) to reconstruct the radiation belts. This marks one of the few attempts in using the NARX neural network to derive a flux model at geosynchronous orbit (Camporeale, 2019) and the first time to derive an upper boundary condition independent of in-situ observations (see the companion paper, Saikin et al. (2021), for more details). For the flux model, Vsw, Dst, MLT, *Kp,* AE, and Pdyn were determined to be effective predictive indicators and used as exogenous inputs. Since the VERB code is *Kp*-driven, the upper boundary scaling factor model used *Kp* as its only input. In this discussion section, we will discuss the interpretation and understanding of our results as well as potential future applications.

5.1. Model Performance, Comparisons, and Future Modifications

Concerning the flux model, the NARX-derived flux was able to successfully replicate the cyclical nature of the 0.8 MeV electrons observed by GOES-15. However, consistently, the model failed to reproduce the high peak electron flux measurements. This is not uncommon when using NARX neural networks as a prediction system. Using two metrics of training sets (a sliding window method and a direct method), Ayala Solares et al. (2016) sought to use NARX neural networks to predict values of *Kp*. Ayala Solares et al. (2016) found that when using the sliding window method (i.e., where the training set is partitioned into time windows of variable hours which are used to predict the next data point; this new data point replaces the oldest data point and is used to train a new neural network), the NARX model had difficulty in matching the lowest ($Kp < \sim 2.3$) and highest ($Kp > 5.3$) *Kp* values by either overpredicting or underpredicting, respectively. The direct method (which uses a singular training data set, like our method) performed much better with predicting the lower *Kp* values, yet still struggled to match the higher *Kp* observations.

Similar results were found using NARX to predict and model AE and Dst. Using their novel Cloud-NARX model, Gu et al. (2019) sought to predict and estimate the AE index using solar wind dynamic pressure, density, speed, and the interplanetary magnetic field as the exogenous inputs. While their model has a CC and PE of ~0.87 and ~0.76, respectively, Cloud-NARX was still unable to consistently predict AE values greater than ~500 nT. Likewise, predictions with NARX were also able to make suitable estimates for Dst except for minimum Dst during geomagnetic storms or quick magnetospheric compressions (i.e., positive Dst values; Wei et al., 2004). Though these studies focused on geomagnetic indices, there is a pattern of NARX-derived neural networks underestimating the higher values of those respective indices (or in this study's case, flux).

While the above comparisons focused on predicting singular geomagnetic indices, other neural network studies have examined the prediction of either GOES-13 or GOES-15 measurements. Shin et al. (2016) derived neural network models for GOES-13 and GOES-15, based on a variable time prediction forecast ranging from 1 hr to 24 hr during March 2015. Their model was trained at several different energy channels (40 keV, 75 keV, 150 keV 275 keV, 475 keV, >0.8 MeV, and >2 MeV) and used numerous input variables (e.g., the respective flux channels, Vsw, solar wind density, IMF [B_x, B_y, and B_z], Sym-H, *Kp*, AE, and the Akasofu parameter). The MLT location of GOES (although incorporated into their model design) was not as heavily weighted as the other input parameters. As such, when examining the PE as a function of MLT, Shin et al. (2016) found a sharp decrease in their model performance for the non-dayside MLTs (especially for the lower energies). Shin et al. (2016) used more inputs overall, so it is possible that the other metrics used may have biased their simulation.

Another neural network flux-derived study, Ukhorskiy et al. (2004), used GOES-7 (1995) and GOES-8 (1996–2000) measurements of >2 MeV electrons. This study used Vsw, the convective electric field, solar wind dynamic pressure, Sym-H, and AsyH as their input variables, and identified that the solar wind speed and Sym-H as the most relevant parameters for geosynchronous MeV electron prediction. Unlike our flux model, the Ukhorskiy et al. (2004) predictive model often overpredicted MeV fluxes, with prediction efficiencies of 0.66 and 0.77, for different time periods. A follow-up work used a variation of the NARX neural network, NARMAX, with the same input parameters as Ukhorskiy et al. (2004) to model >2 MeV electrons (Wei et al., 2011). This NARMAX study incorporated "time lags" deriving 2 models with a 1- and 3-day time lag, respectively, and found that using a longer time delay (i.e., 3 days) yielded a higher PE (0.78–0.91) compared to the 1-day model.

Table 2
Correlation Coefficients (CC) and Prediction Efficiency (PE) Between the Geosynchronous Flux Model, VERB Simulations With the GOES-15, and CRRES Observations (Respectively) at the Respective Energies and L*

	CC	PE	Bias	Precision
Geo. flux model	0.74	0.42	−0.13	0.42
BF model	0.15	0.39	0.11	0.86
CRRES 1.0 MeV, L* = 4	0.86	0.75	0.6	0.4
CRRES 1.0 MeV, L* = 5	0.64	0.74	0.13	0.49
CRRES 1.0 MeV, L* = 6	0.22	0.24	0.32	0.42

NARMAX has also been used to determine daily averaged predictions of >0.8 and > 2.0 MeV electron flux derived from GOES-13 observations (Boynton et al., 2015). Here, solar wind density, velocity, Dst, and IMF B_z were chosen to be the exogenous inputs. The predictions for the >0.8 and > 2.0 MeV models are largely accurate. However, the >0.8 MeV forecast often overpredicts the flux observations, as in Wei et al. (2011). Given the results of the Boynton et al. (2015) study, it may be more accurate and easier to predict daily averages instead of hourly derived flux measurements, like we have done in this work. Similar results were found in Balikhin et al. (2016) (again using NARMAX). In Balikhin et al. (2016), two models (REFM and SNB³GEO) were used to predict geosynchronous flux measurements. Each model could predict either daily (SNB³GEO) or up to 3 days (REFM). However, each model used different input parameters with the SNB³GEO using solar wind velocity, solar wind density, and the percentage of time that IMF remained southward. Conversely, the REFM model used only solar wind velocity.

PreMevE (Chen et al., 2019) incorporates Linear Predictive Filters (LPFs) to model MeV and multi-MeV (Pires de Lima et al., 2020; Sinha et al., 2021) electrons within the radiation belts. Each *L* shell (in bin sizes of 0.1, from *L* = 3–6) has its own LPF and outputs predictions in 5-hr, 1 day, and 2-day time steps. This model was developed utilizing MeV electron measurements from RBSP, GEO, and one of NOAA's Polar Observational Environment Satellites (POES). PreMevE is comprised of two submodels (1 and 2). Submodel 1 focuses on timing the arrival of MeV electrons, while Submodel 2 specifies the evolving electron flux. For the purposes of this comparison, we will focus on Submodel 2. Overall, PreMevE performs very well when predicting the onset of all major MeV electron enhancement events. Our NARX derived BF-VERB simulation tends to overproduce electron flux during these enhancement events while PreMevE sometimes underpredicts during enhancement periods at low *L* (e.g., *L* = 3).

Another radiation belt neural network model, SHELLS (Claudepierre & O'Brien, 2020), uses a combination of POES and MagEIS observations in conjunction with both a daily averaged *Kp* and a maximum *Kp* value over the preceding 24 hr. Unlike both models presented in this study, the SHELLS model produces daily averaged predictions. While SHELLS produces reconstructions well in agreement with MagEIS observations of the radiation belts, CC decrease with increasing *L* shell (ranging from r² = 0.1–0.9, per *L* shell and *Kp* index).

To assess the performance of our model predictions, we have employed two metrics: a CC of the logged flux and the PE taken over the entire period. PE has been defined as:

$$PE = 1 - \frac{1}{N} \frac{\sum (y_i - f_i)^2}{Var(X)}$$

where $Var(X) = \frac{1}{N} \sum (y_i - \bar{y})^2$, y_i is the observation, f_i is the prediction, N is the number of data points, and $\bar{y}$ is the sample mean of y. We have also added a calculation of bias and precision for which to judge our model's performance:

$$ME = \bar{M} - \bar{O}$$

where $\bar{M}$ and $\bar{O}$ denote the averaged measurement (or model) and the averaged observation (the GOES-15 and CRRES measurements, respectively). The precision metric is defined as:

$$P_{\sigma, ratio} = \frac{\sigma_M}{\sigma_O}$$

where σ_M and σ_O are the standard deviation of the measurement (model) and the observations (the GOES-15 and CRRES measurements, respectively).

Table 2 shows the results of these calculations, including those for the geosynchronous flux model, the BF model, and the corresponding radiation belt reconstructions at *L** = 4, 5, and 6 an. Our GOES-15 flux correlation, PE, bias, and precision yields values of 0.74, 0.42, −0.13, and 0.04 while, depending on which *L**, the correlations for the radiation belt reconstructions range from 0.22 (*L** = 6) to 0.86 (*L** = 4). PE in the radiation belt

Table 3
Prediction Efficiency (PE) and Coefficient of Determination (r^2) of the Variable Dependent Flux Studies for Comparison

Study	Simulated property	Metric	Score
Claudepierre and O'Brien (2020)[a,b]	Sub MeV electrons	PE	~0.9 (L ~4)
Claudepierre and O'Brien (2020)[a,b]	Sub MeV electrons	PE	~0.9 (L ~5)
Claudepierre and O'Brien (2020)[a,b]	Sub MeV electrons	PE	~0.75 (L ~6)
Claudepierre and O'Brien (2020)[a,b]	Sub MeV electrons	PE	~0.5 (L ~6.6)
Chen et al. (2019)[a,c]	1 MeV electrons	PE	~0.89 (L = 4)
Chen et al. (2019)[a,c]	1 MeV electrons	PE	~0.76 (L = 5)
Chen et al. (2019)[a,c]	1 MeV electrons	PE	−0.55 (L = 6)
Chen et al. (2019)[a,c]	1 MeV electrons	PE	−0.93 (GEO)

[a]*L* dependent. [b]*Kp* dependent. [c]Chen et al. (2019) *Submodel 2, 5-hr time step model.*

reconstructions also vary with L^*, with $L^* = 6$ yielding the lowest PE (0.24), and $L^* = 4$ giving the highest PE (0.75). Despite having low CC and PE scores (0.15 and 0.39, respectively), the BF model still is able to deliver reasonable results for the VERB code radiation belt reconstructions with low bias and precision numbers (see Table 2 for details).

For ease of comparison, we have also included Tables 3 and 4, which showcase the PE, CC, and coefficient of determination (r^2) when appropriate, of the previously mentioned studies.

While we have calculated similar metrics (i.e., CC and PE) to compare to the previous studies, differences remain in the quantity that other studies predict (e.g., >0.8 MeV electrons, > 2.0 MeV electrons, etc.); the satellite missions used (GOES-7 (Ukhorskiy et al., 2004), GOES-8 (Ukhorskiy et al., 2004), GOES-13 (Boynton et al., 2015; Shin et al., 2016, Wei et al., 2011), GOES-15 (Shin et al., 2016), and the Van Allen Probes (Chen et al., 2019; Claudepierre & O'Brien, 2020)); and the time step of the forecast (e.g., 1 day ([Chen et al., 2019; Claudepierre & O'Brien, 2020; Ukhorskiy et al., 2004]) or 2 days [Chen et al., 2019]). Unlike the other models, our geosynchronous flux model predicts in 5-min time steps, specifically for the 800 keV electrons, and only uses the measurements required for the initial input delay (and does not recalibrate itself daily). Once our NARX model begins, it becomes solely driven by just the geomagnetic and solar wind related parameters.

The above-mentioned studies each used their own respective combinations of geomagnetic indices. For our study, we initially began by determining which indices were more predictive of geosynchronous flux, only to determine that excluding none of the parameters yielded the best results. Future work using NARX neural networks to model the radiation belts may wish to consider testing unique combinations of the input parameters (i.e., using only *Kp* and MLT, or Vsw and AE, etc.).

Table 4
Prediction Efficiency (PE) and Correlation Coefficients (CC) of the Other Geosynchronous Flux Studies for Comparison

Study	Simulated property	Metric	Score
Boynton et al. (2015)	>0.8 MeV electrons	PE and CC	0.7 and ~0.85
	>2.0 MeV electrons		~0.79 and 0.89
Wei et al. (2011)	>2.0 MeV electrons	PE	0.78–0.91
Ukhorskiy et al. (2004)[a]	>2.0 MeV electrons	PE	0.66–0.87
		CC	0.80–0.92
Shin et al. (2016)[b]	>0.8 MeV electrons	PE	~0.9 (1 hr)
	>0.8 MeV electrons	PE	~0.5 (24 hr)
Balikhin et al. (2016)	>2.0 MeV electrons	PE and CC	0.63 and 0.82

Note. All studies are daily predictions unless specified otherwise.
[a]Peformed for multiple years. [b]Time dependent.

Concerning the NARX-derived BF model, *Kp* was the sole exogenous input used. This was chosen since VERB is a *Kp*-driven code, and historical *Kp* measurements exist as early as the 1930s, unlike other parameters. The derived upper boundary condition was able to predict the enhancement periods observed by the CRRES mission during the test period of July–November 1990. While the peak enhancement periods observed by CRRES were reproduced, the simulation did not scatter enough electrons to match the corresponding dropouts present in the CRRES data. This test simulation period of July–November 2020 is the same as the one used in the Subbotin, Shprits, and Ni (2011); Subbotin, Shprits, Gkioulidou, et al.'s (2011) radiation belt reconstruction. In that study, the authors examined how the inclusion of radial, pitch angle, energy, and mixed diffusion terms impacted VERB code simulations of the radiation belts, all while using a data-derived BF. While the inclusion or exclusion of any diffusion terms impacts the scattering rates produced by the simulation, by using a data derived BF, the Subbotin simulations accurately reproduce the timing of electron dropouts and enhancements. Their most accurate simulation occurs with the inclusion of all diffusion terms in conjunction with the data driven BF. As described in Section 2, we also include all energy, pitch angle, and mixed diffusion terms like Subbotin, Shprits, and Ni (2011); Subbotin, Shprits, Gkioulidou, et al. (2011) (along with updated diffusion coefficients and the contribution of EMIC waves). While Subbotin, Shprits, and Ni (2011); Subbotin, Shprits, Gkioulidou, et al. (2011) had performed stronger in reproducing the CRRES observed radiation belts because of the data-derived BF, this study's goal is to produce an upper boundary condition for historical reconstructions when in-situ observations are unavailable. Therefore, we must address a few reasons why the VERB code produced this result.

First, the NARX-derived BF was, at times, too high, compared to the CRRES-derived BF, thereby causing the VERB code to produce more electron flux. Second, as described in Saikin et al. (2021), current parameterizations of wave-particle interactions may need to be improved. For example, an improved parameterization of plasmaspheric hiss may be required to produce more effective scattering of electrons, especially for the sub-MeV electrons (Malaspina et al., 2020), thereby making post-enhancement periods more consistent with CRRES observations. Finally, our plasmapause location model is based exclusively on *Kp* (Carpenter & Anderson, 1992). Since the plasma density does not evolve with time, the plasma model both inside and outside the plasmasphere remains constant. This will alter the energy and pitch-angle range through which wave-particle interactions may occur, causing electron scattering when appropriate. Alternative models of the plasmasphere exist (e.g., O'Brien & Moldwin, 2003; Zhelavskaya et al., 2017) and should be considered in future simulation studies.

5.2. Possible Applications

The use of neural networks, machine learning, and modeling to develop and to train models has become more common in recent years (Camporeale, 2019). Applications have ranged from the prediction of and modeling of geomagnetic indices (e.g., Ayala Solares et al., 2016; Balikhin et al., 2011; Boynton et al., 2011; Camporeale, 2019; Wei et al., 2004) to real-time measurements (e.g., Ganushkina et al., 2015; Shin et al., 2016; Smirnov et al., 2020; Ukhorskiy et al., 2004; Wei et al., 2011; Zhelavskaya et al., 2017). The work performed in this study pertaining to the training of the upper boundary scaling condition has predictive implications beyond the scope of simply modeling contemporaneous *in-situ* observations.

Previous studies, when using neural networks to model the radiation belts, tended to focus on geostationary orbit flux measurements (e.g., Shin et al., 2016; Ukhorskiy et al., 2004; Wei et al., 2011). These studies would limit themselves to geostationary orbit without any expansion to examine multiple *L* shells. The use of the NARX-derived upper boundary condition in the VERB code allows for the complete simulation and reconstruction of the radiation belts, while providing a reasonable agreement with observations (e.g., CRRES and the Van Allen Probes). Conversely, the assumed constant upper boundary condition simulation failed to reproduce the appropriate flux enhancements observed by CRRES. A previous attempt using the VERB code and a NARMAX-derived upper boundary condition has been performed using GOES-13 measurements and solar wind parameters from L1 as inputs. The 892 keV radiation belt electrons during August, September, and October 2013, were reconstructed and compared against the then-contemporaneous Van Allen Probes observations. The resulting radiation belt reconstructions spanned $L* = 3–7$, and underestimated fluxes compared to the *in-situ* Van Allen Probes (Pakhotin et al., 2014).

The complimentary paper*,* Saikin et al. (2021), performed a long-term reconstruction of radiation belt electrons at three respective energies over a period of 8 Solar Cycles (Solar Cycle 17–24 and 1933–2017), using the techniques pioneered in this study (NARX instead of NARMAX, as in Pakhotin et al. (2014)). Unlike other long-term

simulations studies (e.g., Glauert et al., 2018; Subbotin, Shprits, & Ni, 2011; Subbotin, Shprits, Gkioulidou, et al., 2011, etc.), Saikin et al. (2021) were able to hindcast radiation belt dynamics absent *in-situ* observations, instead of using a NARX-developed upper boundary scaling factor condition trained on GOES-15 measurements. The ability to reconstruct radiation belts absent *in-situ* observations allows us to examine the dynamics of the historical radiation belts, provided there is a guiding geomagnetic index (e.g., *Kp*).

Gray-box models are cases in which physics-based models are supported with machine learning that can provide future advantages for space weather and radiation belt applications (Camporeale, 2019). Our successful use of BF predictions in Saikin et al. (2021) is a compelling example of the importance of gray-box models in the reconstruction of the radiation belts. Furthermore, the derivation of an upper boundary condition showcases the unique ability of machine learning to fill gaps in the parameters of physics-based models.

6. Conclusion

This paper focused on using a NARX neural network to derive a flux model for 0.8 MeV electrons with a 5-min cadence and an upper boundary scaling factor to be used in VERB code simulations. The flux model was trained on GOES-15 measurements from June 2013 to June 2016, with Kp, MLT, Dst, Vsw, AE, and Pdyn being used as the respective exogenous inputs. The NARX flux model reproduces the cyclical nature of the GOES-15 measurements while underperforming the peak flux values by an average of ~46%. The underestimation of peak flux values was consistent with previous neural network forecasts. Unlike previous NARX neural network studies, which focused primarily on predicting singular geomagnetic indices, this study represents one of the early attempts to use NARX to predict and model flux observations.

The NARX upper boundary scaling factor (BF) model was also derived using GOES-15 measurements and only *Kp* as an exogenous input (since it is easier to use for radiation belt predictions). This constructed BF model was used in conjunction with the VERB 3-D code to produce a hindcast simulation of the 1.0 MeV radiation belt electrons during July–November 1990. Unlike previous attempts with modeled BF used with the VERB code (i.e., Pakhotin et al., 2014), this reconstruction was able to simulate flux values on par with the then-contemporaneous CRRES measurements of 1.0 MeV electrons. However, the simulation did provide enough electron scattering to reproduce the observed electron dropouts. The *Kp*-driven hindcast simulation represents one of the first radiation belt reconstructions absent *in-situ* observations and allows for the reconstruction of the historical radiation belts from 1933 to 2017 and Solar Cycles 17–24, with an improved performance over a constant upper boundary condition (Saikin et al., 2021).

Data Availability Statement

Measurements from the CRRES mission were found at https://cdaweb.gsfc.nasa.gov/pub/data/crres/particle_mea/. GOES measurements were obtained from https://satdat.ngdc.noaa.gov/sem/goes/data/avg/. Historical Kp measurements were found at ftp://ftp.gfz-potsdam.de/pub/home/obs/kp-ap/. The data produced and shown (PSD for the electron flux data for Solar Cycles 22 (i.e., the CRRES era), the NARX-derived BF) in this manuscript can be found at this location: https://doi.org/10.25346/S6/HY1DNT. The data relating to the derivation of the geosynchronous flux model can be found at this location: https://doi.org/10.25346/S6/O1ZAZH. This work used computational and storage services associated with the Hoffman2 Shared Cluster provided by the UCLA Institute for Digital Research and Education's Research Technology Group.

References

Abel, B., Thorne, R. M., & Vampola, A. L. (1994). Solar cyclic behavior of trapped energetic electrons in Earth's inner radiation belt. *Journal of Geophysical Research*, *99*, 427–431. https://doi.org/10.1029/94ja01626

Ayala Solares, J. R., Wei, H. L., Boynton, R. J., Walker, S. N., & Billings, S. A. (2016). Modeling and prediction of global magnetic disturbance in near-Earth space: A case study for Kp index using NARX models. *Space Weather*, *14*(10), 899–916. https://doi.org/10.1002/2016SW001463

Baker, D. N., Blake, J. B., Klebesadel, R. W., & Higbie, P. R. (1986). Highly relativistic electrons in the Earth's outer magnetosphere: 1. Lifetimes and temporal history 1979–1984. *Journal of Geophysical Research*, *91*(A4), 4265. https://doi.org/10.1029/ja091ia04p04265

Baker, D. N., Li, X., Turner, N., Allen, J. H., Bargatze, L. F., Blake, J. B., et al. (1997). Recurrent geomagnetic storms and relativistic electron enhancements in the outer magnetosphere: ISTP coordinated measurements. *Journal of Geophysical Research*, *102*, 14141–14148. https://doi.org/10.1029/97ja00565

Balikhin, M. A., Boynton, R. J., Walker, S. N., Borovsky, J. E., Billings, S. A., & Wei, H. L. (2011). Using the NARMAX approach to model the evolution of energetic electrons fluxes at geostationary orbit. *Geophysical Research Letters*, *38*(18), 1–5. https://doi.org/10.1029/2011GL048980

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

Bernoux, G., & Maget, V. (2020). Characterizing extreme geomagnetic storms using extreme value analysis: A discussion on the representativeness of short data sets. *Space Weather*, *18*(6), e02450. https://doi.org/10.1029/2020SW002450

Beutier, T., & Boscher, D. (1995). A three-dimensional analysis of the electron radiation belt by the Salammbo Code. *Journal of Geophysical Research*, *100*, 14853–14861. https://doi.org/10.1029/94ja03066

Bingham, S. T., Mouikis, C. G., Kistler, L. M., Boyd, A. J., Paulson, K., Farrugia, C. J., et al. (2018). The outer radiation belt response to the storm time development of seed electrons and chorus wave activity during CME and CIR driven storms. *Journal of Geophysical Research: Space Physics*, *123*(12), 10139–10157. https://doi.org/10.1029/2018JA025963

Boynton, R. J., Balikhin, M. A., & Billings, S. A. (2015). Online NARMAX model for electron fluxes at GEO. *Annales Geophysicae*, *33*(3), 405–411. https://doi.org/10.5194/angeo-33-405-2015

Boynton, R. J., Balikhin, M. A., Billings, S. A., Wei, H. L., & Ganushkina, N. (2011). Using the NARMAX OLS-ERR algorithm to obtain the most influential coupling functions that affect the evolution of the magnetosphere. *Journal of Geophysical Research*, *116*(5), 1–8. https://doi.org/10.1029/2010JA015505

Brautigam, D. H., & Albert, J. M. (2000). Radial diffusion analysis of outer radiation belt electrons during the October 9, 1990, magnetic storm. *Journal of Geophysical Research*, *105*(A1), 291–309. https://doi.org/10.1029/1999ja900344

Camporeale, E. (2019). The challenge of machine learning in space weather: Nowcasting and forecasting. *Space Weather*, *17*(8), 1166–1207. https://doi.org/10.1029/2018SW002061

Carpenter, D. L., & Anderson, R. R. (1992). An ISEE/whistler model of equatorial electron density in the magnetosphere. *Journal of Geophysical Research*, *97*(A2), 1097–1108. https://doi.org/10.1029/91JA01548

Castillo, A. M., Shprits, Y. Y., Ganushkina, N., Drozdov, A. Y., Aseev, N. A., Wang, D., & Dubyagin, S. (2019). Simulations of the inner magnetospheric energetic electrons using the IMPTAM-VERB coupled model. *Journal of Atmospheric and Solar-Terrestrial Physics*, *191*, 105050. https://doi.org/10.1016/j.jastp.2019.05.014

Cervantes, S., Shprits, Y. Y., Aseev, N. A., Drozdov, A. Y., Castillo, A., & Stolle, C. (2020). Identifying radiation belt electron source and loss processes by assimilating spacecraft data in a three-dimensional diffusion model. *Journal of Geophysical Research*, *125*(1), e27514. https://doi.org/10.1029/2019JA027514

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

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

Craven, J. D. (1966). Temporal variations of electron intensities at low altitudes in the outer radiation zone as observed with satellite Injun 3. *Journal of Geophysical Research*, *71*(23), 5643–5663. https://doi.org/10.1029/jz071i023p05643

Drozdov, A. Y., Aseev, N., Effenberger, F., Turner, D. L., Saikin, A. A., & Shprits, Y. Y. (2019). Storm time depletions of multi-MeV radiation belt electrons observed at different pitch angles. *Journal of Geophysical Research: Space Physics*, *124*(11), 8943–8953. https://doi.org/10.1029/2019JA027332

Drozdov, A. Y., Shprits, Y. Y., Orlova, K. G., Kellerman, A. C., Subbotin, D. A., Baker, D. N., et al. (2015). Energetic, relativistic, and ultrarelativistic electrons: Comparison of long-term VERB code simulations with Van Allen Probes measurements. *Journal of Geophysical Research: Space Physics*, *120*, 3574–3587. https://doi.org/10.1002/2014JA020637.Received

Drozdov, A. Y., Shprits, Y. Y., Usanova, M. E., Aseev, N. A., Kellerman, A. C., & Zhu, H. (2017). EMIC wave parameterization in the long-term VERB code simulation. *Journal of Geophysical Research: Space Physics*, *122*(8), 8488–8501. https://doi.org/10.1002/2017JA024389

Fennell, J. F., Claudepierre, S. G., Blake, J. B., O'Brien, T. P., Clemmons, J. H., Baker, D. N., et al. (2015). Van allen Probes show that the inner radiation zone contains no MeV electrons: ECT/MagEIS data. *Geophysical Research Letters*, *42*(5), 1283–1289. https://doi.org/10.1002/2014GL062874

Ganushkina, N. Y., Amariutei, O. A., 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

Glauert, S. A., Horne, R. B., & Meredith, N. P. (2014). Simulating the Earth's radiation belts: Internal acceleration and continuous losses to the magnetopause. *Journal of Geophysical Research: Space Physics*, *119*(9), 7444–7463. https://doi.org/10.1002/2014JA020092

Glauert, S. A., Horne, R. B., & Meredith, N. P. (2018). A 30-year simulation of the outer electron radiation belt. *Space Weather*, *16*(10), 1498–1522. https://doi.org/10.1029/2018SW001981

Gu, Y., Wei, H. L., Boynton, R. J., Walker, S. N., & Balikhin, M. A. (2019). System identification and data-driven forecasting of AE index and prediction uncertainty analysis using a new cloud-NARX model. *Journal of Geophysical Research: Space Physics*, *124*(1), 248–263. https://doi.org/10.1029/2018JA025957

Jordanova, V. K., Albert, J. M., & Miyoshi, Y. (2008). Relativistic electron precipitation by EMIC waves from self-consistent global simulations. *Journal of Geophysical Research*, *113*, A00A10–A00A21. https://doi.org/10.1029/2008JA013239

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. *Journal of Geophysical Research*, *116*(10), A10214. https://doi.org/10.1029/2011JA016684

Lyons, L. R., Thorne, R. M., & Kennel, C. F. (1972). Pitch-angle diffusion of radiation belt electrons within the plasmasphere. *Journal of Atmospheric and Solar-Terrestrial Physics*, *77*(19), 3455–3474. https://doi.org/10.1029/JA077i019p03455

Ma, Q., Li, W., Thorne, R. M., Nishimura, Y., Zhang, X. J., Reeves, G. D., et al. (2016). Simulation of energy-dependent electron diffusion processes in the Earth's outer radiation belt. *Journal of Geophysical Research: Space Physics*, *121*(5), 4217–4231. https://doi.org/10.1002/2016JA022507

Maget, V., Bourdarie, S., Boscher, D., & Friedel, R. H. W. (2006). Data assimilation of LANL satellite data into the Salammbô electron code over a complete solar cycle by direct insertion. *Space Weather*, *4*(4), 19–31. https://doi.org/10.1029/2007SW000322

Malaspina, D. M., Zhu, H., & Drozdov, A. Y. (2020). A wave model and diffusion coefficients for plasmaspheric hiss parameterized by plasmapause location. *Journal of Geophysical Research: Space Physics*, 1–23. https://doi.org/10.1029/2019ja027415

O'Brien, T. P., & Moldwin, M. B. (2003). Empirical plasmapause models from magnetic indices. *Geophysical Research Letters*, *30*(4), 1152. https://doi.org/10.1029/2002GL016007

Orlova, K., & Shprits, Y. Y. (2011). On the bounce-averaging of scattering rates and the calculation of bounce period. *Physics of Plasmas*, *18*(9), 092904. https://doi.org/10.1063/1.3638137

Orlova, K., Shprits, Y. Y., & Spasojevic, M. (2016). New global loss model of energetic and relativistic electrons based on Van Allen Probes measurements. *Journal of Geophysical Research: Space Physics*, *121*(2), 1308–1314. https://doi.org/10.1002/2015ja021878

Pakhotin, I. P., Drozdov, A. Y., Shprits, Y. Y., Boynton, R. J., Subbotin, D. A., & Balikhin, M. A. (2014). Simulation of high-energy radiation belt electron fluxes using NARMAX-VERB coupled codes. *Journal of Geophysical Research: Space Physics*, *119*(10), 8073–8086. https://doi.org/10.1002/2014ja020238

Pires de Lima, R., Chen, Y., & Lin, Y. (2020). Forecasting megaelectron-volt electrons inside Earth's outer radiation belt: PreMevE 2.0 based on supervised machine learning algorithms. *Space Weather*, *18*(2), 1–23. https://doi.org/10.1029/2019SW002399

Reeves, G. D., Chen, Y., Cunningham, G. S., Friedel, R. W. H., Henderson, M. G., Jordanova, V. K., et al. (2012). Dynamic radiation environment assimilation model: DREAM. *Space Weather*, *10*(3), 1–25. https://doi.org/10.1029/2011SW000729

Reeves, G. D., Friedel, R. H. W., Belian, R. D., Meier, M. M., Henderson, M. G., Onsager, T., et al. (1998). The relativistic electron response at geosynchronous orbit during the January 1997 magnetic storm. *Journal of Geophysical Research*, *103*(A8), 17559–17570. https://doi.org/10.1029/97ja03236

Reeves, G. D., McAdams, K. L., Friedel, R. H. W., & O'Brien, T. P. (2003). Acceleration and loss of relativistic electrons during geomagnetic storms. *Geophysical Research Letters*, *30*(10), 1529. https://doi.org/10.1029/2002gl016513

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

Roederer, J. L. (1970). *Dynamics of geomagnetically trapped radiation*. Springer US.

Rothwell, P., & Mcilwain, C. E. (1960). Magnetic Storms and the Van Allen Radiation Belts- Observations from Satellite 1958e (Explorer IV). *Journal of Geophysical Research*, *65*(3), 799–806. https://doi.org/10.1029/JZ065i003p00799

Saikin, A. A., Shprits, Y. Y., Drozdov, A. Y., Landis, D. A., Zhelavskaya, I. S., & Cervantes, S. (2021). Reconstruction of the radiation belts for solar cycles 17–24 (1933–2017). *Space Weather*, *19*, e2020SW002524. https://doi.org/10.1029/2020sw002524

Saikin, A. A., Zhang, J.-C., Allen, R. C., Smith, C. W., Kistler, L. M., Spence, H. E., et al. (2015). The occurrence and wave properties of H+-, He+-, and O+-band EMIC waves observed by the Van Allen Probes. *Journal of Geophysical Research: Space Physics*, *120*, 1–16. https://doi.org/10.1002/2015JA021358

Saikin, A. A., Zhang, J.-C., Smith, C. W., Spence, H. E., Torbert, R. B., & Kletzing, C. A. (2016). The dependence on geomagnetic conditions and solar wind dynamic pressure of the spatial distributions of EMIC waves observed by the Van Allen Probes. *Journal of Geophysical Research: Space Physics*, *121*, 4362–4377. https://doi.org/10.1002/2016JA022523

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. Y., Li, W., & Thorne, R. M. (2006). Controlling effect of the pitch angle scattering rates near the edge of the loss cone on electron lifetimes. *Journal of Geophysical Research*, *111*(12), 1–7. https://doi.org/10.1029/2006JA011758

Shprits, Y. Y., & Ni, B. (2009). Dependence of the quasi-linear scattering rates on the wave normal distribution of chorus waves. *Journal of Geophysical Research*, *114*(11), 1–10. https://doi.org/10.1029/2009JA014223

Shprits, Y. Y., Subbotin, D. A., Meredith, N. P., & Elkington, S. R. (2008). Review of modeling of losses and sources of relativistic electrons in the outer radiation belt I: Radial transport. *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*(14), 1694–1713. https://doi.org/10.1016/j.jastp.2008.06.014

Sinha, S., Chen, Y., Lin, Y., & Pires de Lima, R. (2021). PreMevE update: Forecasting ultra-relativistic electrons inside Earth's outer radiation belt. *Space Weather*, *19*(9), 1–23. https://doi.org/10.1029/2021SW002773

Smirnov, A. G., Berrendorf, M., Shprits, Y. Y., Kronberg, E. A., Allison, H. J., Aseev, N. A., et al. (2020). Medium energy electron flux in Earth's outer radiation belt (MERLIN): A machine learning model. *Space Weather*, *18*, 1–20. https://doi.org/10.1029/2020SW002532

Subbotin, D. A., & Shprits, Y. Y. (2009). Three-dimensional modeling of the radiation belts using the versatile electron radiation belt (verb) code. *Space Weather*, *7*(10), 1–15. https://doi.org/10.1029/2008SW000452

Subbotin, D. A., & Shprits, Y. Y. (2012). Three-dimensional radiation belt simulations in terms of adiabatic invariants using a single numerical grid. *Journal of Geophysical Research*, *117*(5), 1–12. https://doi.org/10.1029/2011JA017467

Subbotin, D. A., Shprits, Y. Y., Gkioulidou, M., Lyons, L. R., Ni, B., Merkin, V. G., et al. (2011). Simulation of the acceleration of relativistic electrons in the inner magnetosphere using RCM - VERB coupled codes. *Journal of Geophysical Research*, *116*, 1–12. https://doi.org/10.1029/2010JA016350

Subbotin, D. A., Shprits, Y. Y., & Ni, B. (2011). Long-term radiation belt simulation with the VERB 3-D code: Comparison with CRRES observations. *Journal of Geophysical Research*, *116*(12), 1–14. https://doi.org/10.1029/2011JA017019

Summers, D., Ni, B., & Meredith, N. P. (2007a). Timescales for radiation belt electron acceleration and loss due to resonant wave-particle interactions. *Journal of Geophysical Research*, *112*, 1–11. https://doi.org/10.1029/2006JA011801

Summers, D., Ni, B., & Meredith, N. P. (2007b). Timescales for radiation belt electron acceleration and loss due to resonant wave-particle interactions: 2. Evaluation for VLF chorus, ELF hiss, and electromagnetic ion cyclotron waves. *Journal of Geophysical Research*, *112*, 1–21. https://doi.org/10.1029/2006JA011993

Summers, D., & Throne, R. M. (2003). Relativistic electron pitch-angle scattering by electromagnetic ion cyclotron waves during geomagnetic storms. *Journal of Geophysical Research*, *108*(1143), 1–12. https://doi.org/10.1029/2002JA009489

Tsyganenko, N. A. (1989). A magnetospheric magnetic field model with a warped tail current sheet. *Planetary and Space Science*, *37*(1), 5–20. https://doi.org/10.1016/0032-0633(89)90066-4

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

Turner, D. L., Kilpua, E. K. J., Hietala, H., Claudepierre, S. G., O'Brien, T. P., Fennell, J. F., et al. (2019). The response of Earth's electron radiation belts to geomagnetic storms: Statistics from the Van Allen Probes era including effects from different storm drivers. *Journal of Geophysical Research: Space Physics*, *124*, 1013–1034. https://doi.org/10.1029/2018JA026066

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

Vampola, A. L., Osborn, J. V., & Johnson, B. M. (1992). CRRES magnetic electron spectrometer AFGL-701-5A (MEA). *Journal of Spacecraft and Rockets*, *29*(4), 592–595. https://doi.org/10.2514/3.25504

Van Allen, J. A., & Frank, L. A. (1959). Radiation around the Earth to a radial distance of 107,400 km. *Nature*, *183*(183), 430–434. https://doi.org/10.1038/183430a0

Wei, H. L., Billings, S. A., & Balikhin, M. (2004). Prediction of the Dst index using multiresolution wavelet models. *Journal of Geophysical Research*, *109*(A7), 1–12. https://doi.org/10.1029/2003JA010332

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

Zhang, J.-C., Halford, A. J., Saikin, A. A., Huang, C.-L., Spence, H. E., Larsen, B. A., et al. (2016). EMIC waves and associated relativistic electron precipitation on 25-26 January 2013. *Journal of Geophysical Research: Space Physics*, *121*, 1–15. https://doi.org/10.1002/2016ja022918

Zhelavskaya, I. S., Shprits, Y. Y., & Spasojević, M. (2017). Empirical modeling of the plasmasphere dynamics using neural networks. *Journal of Geophysical Research: Space Physics*, *122*(11), 11227–11244. https://doi.org/10.1002/2017JA024406

Zhu, H., Shprits, Y. Y., Spasojevic, M., & Drozdov, A. Y. (2019). New hiss and chorus waves diffusion coefficient parameterizations from the Van Allen Probes and their effect on long-term relativistic electron radiation-belt VERB simulations. *Journal of Atmospheric and Solar-Terrestrial Physics*, *193*, 105090. https://doi.org/10.1016/j.jastp.2019.105090