Space Weather

RESEARCH ARTICLE
10.1029/2024SW003962

Special Collection:
Recent Advances on Modelling and Observations in Space and Earth Sciences (0)

Predicting Geostationary (GOES) 4.1–30 keV Electron Flux Over All MLT Using LEEMYR Regression Models

L. E. Simms[1,2], N. Y. Ganushkina[1,3], M. van de Kamp[3], and M. W. Liemohn[1]

[1]University of Michigan, Ann Arbor, MI, USA, [2]Department of Physics, Augsburg University, Minneapolis, MN, USA, [3]Finnish Meteorological Institute, Helsinki, Finland

Key Points:

- Stepwise regression picks parsimonious predictive models of 4.1–30 keV geostationary electron flux
- Predictor correlations with 4.1–30 keV flux are lower than those seen with higher energy electrons
- However, reasonably good predictions can be made at geosynchronous orbit over all MLT

Correspondence to:
L. E. Simms,
laurasim@umich.edu

Citation:
Simms, L. E., Ganushkina, N. Y., van de Kamp, M., & Liemohn, M. W. (2024). Predicting geostationary (GOES) 4.1–30 keV electron flux over all MLT using LEEMYR regression models. *Space Weather*, *22*, e2024SW003962. https://doi.org/10.1029/2024SW003962

Received 18 APR 2024
Accepted 6 AUG 2024

Abstract Regression models (LEEMYR: Low Energy Electron MLT geosYnchronous orbit Regression) predict hourly 4.1–30 keV electron flux at geostationary orbit (GOES-16) using solar wind, IMF, and geomagnetic index parameters. Multiplicative interaction and polynomial terms describe synergistic and nonlinear effects. We reduce predictors to an optimal set using stepwise regression, resulting in models with validation comparable to a neural network. Models predict 1, 3, 6, 12, and 24 hr into the future. Validation correlations are as high as 0.78 (4.1 and 11 keV, 1 hr prediction) and Heidke Skill scores (HSS) up to 0.66. A 3 hr ahead prediction is more practical, with slightly lower validation correlation (0.75) and HSS (0.61). The addition of location (MLT: magnetic local time) as a covariate, including multiplicative interaction terms, accounts for location-dependent flux differences and variation of parameter influence, and allows prediction over the full orbit. Adding a substorm index (*SME*) provides minimal increase in validation correlation (0.81) showing that other parameters are good proxies for an unavailable real time substorm index. Prediction intervals on individual values provide more accurate assessments of model quality than confidence intervals on the mean values. An inverse N-weighted least squares approach is impractical as it increases false positive warnings. Physical interpretations are not possible as spurious correlations due to common cycles are not removed. However, *SME*, Bz, Kp, and Dst are the highest correlates of electron flux, with solar wind velocity, density, and pressure, and IMF magnitude being less well correlated.

Plain Language Summary As high levels of electrons in the radiation belts can damage satellites, an accurate forecasting model is needed. Electron levels can be predicted from regression models using data from the solar wind, the interplanetary magnetic field, and indices measuring disturbances in Earth's magnetic field. Dependable predictions can be made using predictors from either an hour or 3 hr before electron changes occur. A 3 hr ahead prediction is more practical, giving time to respond to electron increases, and results in only slightly lower prediction ability.

1. Introduction

In the radiation belts, electrons in the ∼10–50 keV range result in damaging surface charging of satellites (Thomsen et al., 2013), in contrast to >100 keV electrons which generally result in internal charging (e.g., Lam et al., 2012; Loto'aniu et al., 2015). Low energy electron events are detected during even weak to moderate substorm activity (Ganushkina et al., 2021; Matéo-Vélez et al., 2018), but they do not correlate well with geomagnetic storms. This, and their faster reaction time, makes their prediction more difficult than higher energy electrons (Choi et al., 2011; Koons et al., 2000; Matéo-Vélez et al., 2018).

A number of predictive models for electrons >40 keV have been created (e.g., Balikhin et al., 2011; Boynton et al., 2015, 2019; Capman et al., 2019; Chu et al., 2021; Ganushkina et al., 2019; Koons & Gorney, 1991; Li et al., 2001; Ling et al., 2010; Ma et al., 2022; Pakhotin et al., 2014; Rigler et al., 2004; Sakaguchi et al., 2013; Simms, Ganushkina, et al., 2023; Simms et al., 2014; Smirnov et al., 2020), but many fewer for <40 keV electrons (Denton et al., 2016; Swiger et al., 2022). In this paper, we present a regression-based model for prediction of <40 keV electrons at geosynchronous orbit.

The distribution of lower energy (<100 keV) electrons may depend on location in the magnetosphere (Birn et al., 1997; Thomsen et al., 2001), so daily or orbit averaged predictions may not be useful. Data from geosynchronous orbit comes from only one L-shell, but does provide coverage of all MLTs (Magnetic Local Times). Including MLT as a variable in the model would therefore allow predictions over the entire orbit.

© 2024. The Author(s).
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

There are a number of measured parameters that correlate with geosynchronous 1–40 keV electron flux and therefore may be of use in predictions. These include solar wind and IMF (Interplanetary Magnetic Field) descriptors (e.g., Ganushkina et al., 2019; Hartley et al., 2014; Kellerman & Shprits, 2012; Li et al., 2005; Shi et al., 2009; Sillanpää et al., 2017; Simms, Ganushkina, et al., 2022), as well as geomagnetic indices such as *Kp* (Planetarische Kennziffer), *AE* (Auroral Electrojet), or *SME* (SuperMAG Auroral Electrojet) (e.g., Denton et al., 2015, 2016; Freeman, 1974; Korth et al., 1999; Thomsen et al., 2013) (*AE* also correlates with electron dropouts (Boynton, Balikhin, et al., 2016; Ma et al., 2022). While *AE* or *SME* correlate better with the flux-enhancing substorms, neither is available in real time. The 3-hr cadence of *Kp* does not have the capacity to measure rapid changes in geomagnetic activity associated with fast electron enhancements, but its inclusion in a prediction model may be helpful to measure the general background level of disturbance. *Dst* may be helpful to include as it is available in real time and at an hourly cadence. Similarly, it also may be useful to incorporate the solar energy flux (f10.7) even though it changes relatively slowly (Swiger et al., 2022).

At higher energies (>40 keV) combinations of predictors, rather than single factors, generally result in better predictions (Denton et al., 2016; Simms et al., 2014, 2016, 2018a). These combinations may be both additive (adding more predictors to a model) or multiplicative (adding interaction terms that describe how one parameter's impact on flux is affected by other variables). Certain factors may also show nonlinear relationships with electron flux, but this does not preclude using linear modeling techniques (e.g., correlation and regression) because the relationships between variables are usually intrinsically linear. That is, they can be made linear by log transformations of predictor and response variables (Neter et al., 1990). Log transformations or introducing quadratic or cubic terms can also better describe the relationships (Balikhin et al., 2011; Simms, Ganushkina, et al., 2023). In addition, location information can be included by using MLT as a predictor consisting of two variables: $\sin(2\pi/24)$ and $\cos(2\pi/24)$. If these are also included as multiplicative factors—multiplying each other factor by this location information—the possible varying influence of predictors by location will also be incorporated.

Due to the high correlations between these possible predictors, issues arise if all are added simultaneously to a prediction model. High correlation between the predictors (Simms et al., 2014) can lead to overfitting to unique and inconsequential characteristics of the training data set, meaning that the model may not predict well for a novel time period. We need some way to choose only those parameters that make a real contribution. The weighting algorithms used by neural networks are one approach (Chu et al., 2021; Freeman et al., 1998; Katsavrias et al., 2022; Koons & Gorney, 1991; Ling et al., 2010; Ma et al., 2022, 2023; Simms & Engebretson, 2020; Smirnov et al., 2020; Swiger et al., 2022), but stepwise regression can also be used to reduce the number of inputs to a more optimal subset. In fact, we have previously found little difference in the predictive ability of neural network versus regression models, provided the regression models also include multiplicative interaction and quadratic terms, and a time/location variable such as MLT (Simms, Ganushkina, et al., 2023).

We call our current set of regression prediction models the LEEMYR models (Low Energy Electron MLT geosYnchronous orbit Regression).

2. Data and Model Building

To produce our models, we obtained electron data from the SEISS (Space Environment In Situ Suite) MPS (Magnetospheric Particle Sensor) on GOES-16 located at $\sim 6.6 R_E$ (Galica et al., 2016). The sensor consists of MPS-LO (low) and MPS-HI (high) units measuring electrons at energies below and above 40 keV, respectively. The MPS-LO for low energy electrons contains 14 telescopes, which cover 12 15° angular zones that allow determination of pitch-angle distributions from the orbital data. Total omnidirectional electron flux is obtained from the fluxes measured by all the directional telescopes taking into account the pitch angles of each telescope's measured electron flux. Differential electron fluxes are measured in 15 energy bands, logarithmically spaced from 30 eV to 30 keV. We analyze the five highest energy channels with energy band centroids at 4.1, 6.6, 11, 18, and 30 keV (energy band width is 5.8% of the centroid value). Electron flux is measured as spectral density and angular density ($cm^{-2} \cdot s^{-1} \cdot sr^{-1} \cdot keV^{-1}$). We average these 5 min fluxes to 1 hr intervals. Level 1b data for SEISS MPS-LO instrument and Level 2 data for magnetometer MAG are available at https://www.ngdc.noaa.gov/stp/satellite/goes-r.html.

The LEEMYR models are trained using data over the time interval of 1 January 2020–1 February 2021. To validate the models, we use 28 February 2021–26 August 2021 as the test set. Large data gaps in February 2021 made this period impractical for use. Solar wind parameters (solar wind velocity *V*, number density *N*, pressure *P*, IMF B_y and

Table 1
Means and Standard Deviations for Conversion of log_{10} Electron Flux V, N, P, B, Kp+1, Dst (Unlogged), and SolarFlux to Z-Scores

	Mean	Std Dev
4.1 keV	5.568	0.408
6.6 keV	5.306	0.411
11 keV	4.923	0.408
18 keV	4.614	0.328
30 keV	4.366	0.247
V (*km/s*)	2.578	0.0816
N (#/*cc*)	0.823	0.255
P (*nPa*)	0.226	0.210
B (*nT*)	0.622	0.171
Kp	0.970	0.435
Dst (*nT*)	−5.260	11.403
SolarFlux ($W/m^2/Hz$)	1.873	0.0361

B_z), and magnetic indices (Kp and Dst) were obtained from OMNIWeb web with 1 hr resolution and data time-shifted to the bow shock nose (https://omniweb.gsfc.nasa.gov/form/dx1.html). We use the SuperMAG (https://supermag.jhuapl.edu/) *SME* index (SuperMAG electrojet index) as a proxy for substorm activity for only a subset of the models (Gjerloev, 2012; Newell & Gjerloev, 2011). These models with *SME* were produced to assess how much information we miss by not including this substorm proxy, however these *SME*-models are not useful for actual predictions because *SME* is not available in real time. All variables which are consistently >0 were log_{10}-transformed (this was made possible for Kp by adding 1). Z scores were then calculated by subtracting the mean and dividing by the standard deviation of each variable (Table 1). The Z score (or standard score) is the number of standard deviations an observation lies above or below the mean. It standardizes different variables to a common scale so that (a) variable influences can be directly compared, and (b) variable selection methods are less likely to get stuck in local optima (Hyndman & Athanasopoulos, 2018; Neter et al., 1990). Regressions were then performed on the Z scores of the log transformed data. The transformations were necessary to linearize the relationship between predictor and response variables and to normalize the errors, allowing the use of linear model techniques (such as regression). For presentation in the figures only, we backtransform predictions to log_{10} flux. All statistical analysis and model training was performed in MATLAB (The MathWorks Inc., 2023).

LEEMYR models were built using a variable selection technique: a stepwise regression algorithm which enters variables one by one, eliminating any variable currently in the model if the next addition makes it redundant. We enter each variable if its *p*-value is less than 0.025 and then eliminate any where the *p*-value exceeds 0.05. Assuming a null hypothesis that there is no influence of a variable, the *p*-value is the probability of obtaining a result at least as extreme, or unlikely, as the result observed (Neter et al., 1990). Thus, choosing only those variables with a low *p*-value is equivalent to choosing those with the most influence. Possible predictors include the solar wind, IMF, and geomagnetic variables listed above, as well as each of their square and cubic terms and the multiplicative interaction between each. We convert time to a pair of equations ($\sin(2 \times MLT \times \pi/24)$ and $\cos(2 \times MLT \times \pi/24)$. These sine and cosine variables are also multiplied by the above predictor variables. This gives information on how the effect of each variable may vary with location.

3. Model Validation

Stepwise regression on the full set of obtainable predictor variables (all but *SME* which is not available in real time) gives us a set of coefficients for the LEEMYR models (see Table 2 for the Lag3 coefficients). Over the five energy levels the models vary somewhat both in which variables are included and in the coefficient weightings, but due to the high correlations between predictor variables, these coefficients are not interpretable in a physical sense. Models are created for predicting 1, 3, 6, 12, and 24 hr ahead (Lag1, Lag3, Lag6, Lag12, and Lag24). The models are available for use at https://zenodo.org/records/7520424, including coefficients for Lag1-24 models.

Scatterplots of predicted versus observed log_{10} electron flux for the Lag1 model (prediction 1 hr into the future) at each of the five energies show a reasonably good correspondence (Figure 1), with validation correlations as high as 0.78 for the lower energies. At >11 keV, validation correlations are comparable (0.72–0.76), with standardized RMSE (root mean square error divided by standard deviation) also rising at the higher energies. However, despite the good validation correlation, which is heavily weighted by the bulk of observations in the mid-range values, the lowest flux tends to be overpredicted in value, while the highest flux is underpredicted. This problem is more obvious at the higher energies. This is typical of models using these predictors at higher energies as well (40–150 keV and MeV electrons), no matter the algorithm used to create the model (Simms, Ganushkina, et al., 2023). This is because (a) most observations occur nearer to the mean, with this vast majority of "normal" conditions creating a built in bias drawing all predictions back toward the mean (He & Garcia, 2009), (b) extreme levels of the drivers do not always result in extreme levels of electron flux (there is a high degree of variability in the system), and (c) flux persistence leads to clumping of extreme high values, leading to a high degree of autocorrelation (Bunde, 2023; Simms et al., 2018). However, weighting extreme values in a regression more heavily,

Table 2
Coefficients of the Lag3 Model, Predicting Electron Flux 3 hr Later

Row	4.1 keV	6.6 keV	11 keV	18 keV	30 keV
(Intercept)	0.0196	−0.0208	−0.0420	−0.0782	−0.0729
lag3V	0.1776	0.2211	0.2603	0.2384	0.1477
lag3N	0.1333	0.1488	0.1713	0.1533	
lag3P					0.1220
lag3B	0.0721	0.0648	0.0573	0.0639	0.0641
lag3Kp	0.0947	0.1368	0.1703	0.2018	0.2598
lag3Dst	−0.0175	−0.0492	−0.0602	−0.0529	−0.0650
lag3solar	−0.2331	−0.2507	−0.2705	−0.2729	−0.2677
coshr	0.4772	0.3563	0.2395	0.1490	0.0371
sinhr	0.6672	0.6599	0.6415	0.5753	0.4714
lag3V:lag3Kp		−0.0729			
lag3V:lag3Dst	0.0446	0.0782	0.0455		
lag3V:coshr	0.0961	0.2056	0.1549	0.1244	0.0541
lag3V:sinhr		0.0610	0.1642	0.1558	0.2230
lag3N:lag3Kp	0.0431		0.0369		
lag3N:coshr		0.1279	0.1086	0.0851	
lag3N:sinhr	−0.0955		0.0750		0.1112
lag3P:lag3B		−0.0195			−0.0447
lag3P:coshr	0.0987				
lag3P:sinhr	0.1413	0.0645		0.0696	
lag3B:lag3Dst	−0.0251				
lag3Kp:sinhr	0.1940	0.1920	0.1385	0.1052	0.1044
lag3Dst:coshr			−0.0503	−0.0770	−0.0529
lag3solar:sinhr	−0.0292	−0.0410	−0.0493	−0.0469	−0.0315
coshr:sinhr	0.4136	0.3489	0.2816	0.2160	0.1023
$lag3P^2$	0.0266	0.0612	0.0206	0.0352	0.0455
$lag3Dst^2$			−0.0166		
$lag3solar^2$	0.0608	0.0706	0.0794	0.0810	0.0730

Note. Predictors are log-transformed (when possible) then converted to Z-scores using Table 1 values.

while it may result in more true positive "hits," will also increase false positive "misses" (as we show below) (He & Garcia, 2009). Flux persistence could be added to the model by adding a flux lag term as a predictor, however, this results in predictions that lag behind extreme events, making the predictions useless (Simms, Ganushkina, et al., 2023).

3.1. Predicting Further Than One Hour Ahead

While predicting from 1 hr before might be expected to give the best result, it may not give enough lead time to be of any practical use. We therefore produce models to predict 3, 6, 12, and 24 hr into the future. The validation correlations of these models do drop off (Figure 2), but the correlations from the Lag3 model are fairly close to that of the Lag1. Predictions from 12 to 24 hr previous are not very accurate. Predictive ability is somewhat lower in the higher energies.

It has previously been assumed that a substantial number of previous hours of input data must be available for accurate predictions of electron flux (Chu et al., 2021; Simms & Engebretson, 2020) because certain influences appear to drive flux over many hours or days (Wing et al., 2022). However, when spurious correlations due to trends and cycles are removed, these "long-term" effects tend to disappear (Simms, Engebretson, & Reeves, 2023). Additionally, the persistence of some measured parameters and of their influence means that the predictive ability of inputs over only a short time frame is often sufficient (Simms, Ganushkina, et al., 2023). This is demonstrated by the incorporation of Lag12 inputs into the basic Lag3 model which resulted in little improvement in the validation r compared to the simpler Lag3 alone model (Figure 2).

Lag1 solar wind and IMF parameters reported by the OMNIWeb data site may in fact reach geostationary orbit at the same time as GOES electron flux is measured or even before. It would not make sense, therefore, to interpret correlations in these models as causative. In fact, recent work using neural networks and regression to predict flux, without accounting for confounding cycles and trends, is unable to determine drivers of flux. In future work with this 4.1–30 keV flux GOES-16 data set, we intend to study the relationship between these predictors and flux, using ARMAX (autoregressive-moving average transfer functions) techniques to remove the confounding factors and identify the actual drivers as we have done previously with GOES-13 40–150 keV flux (Simms, Ganushkina, et al., 2022). However, in the present work, we do not use models to identify the drivers of electron flux, merely to predict it. As we are already depending on possible non-causal intercorrelations of variables to create a predictive model, without regard for which correlations are merely the result of common trends and cycles, the only reason to avoid using the Lag1 inputs is the practical matter of lead time before predictions would be useful.

3.2. Is *SME* a Necessary Predictor?

The addition of *SME* (a substorm proxy) to the models does result in somewhat higher validation correlations (Figure 2: 0.743–0.814 for the Lag 1 prediction). If a geomagnetic substorm measure were available for prediction, this would improve the forecasts. However, the other variables do appear to act as reasonable proxies for substorm activity. This means we can produce fairly good predictions even though *SME*, often the most highly correlated variable, is unavailable in real time.

3.3. Validation With the Heidke Skill Score

Validation correlation and RMSE are only one way to compare models. The ability of a model to predict a high flux event may be a more significant capability. This can be assessed using the Heidke Skill Score (HSS)

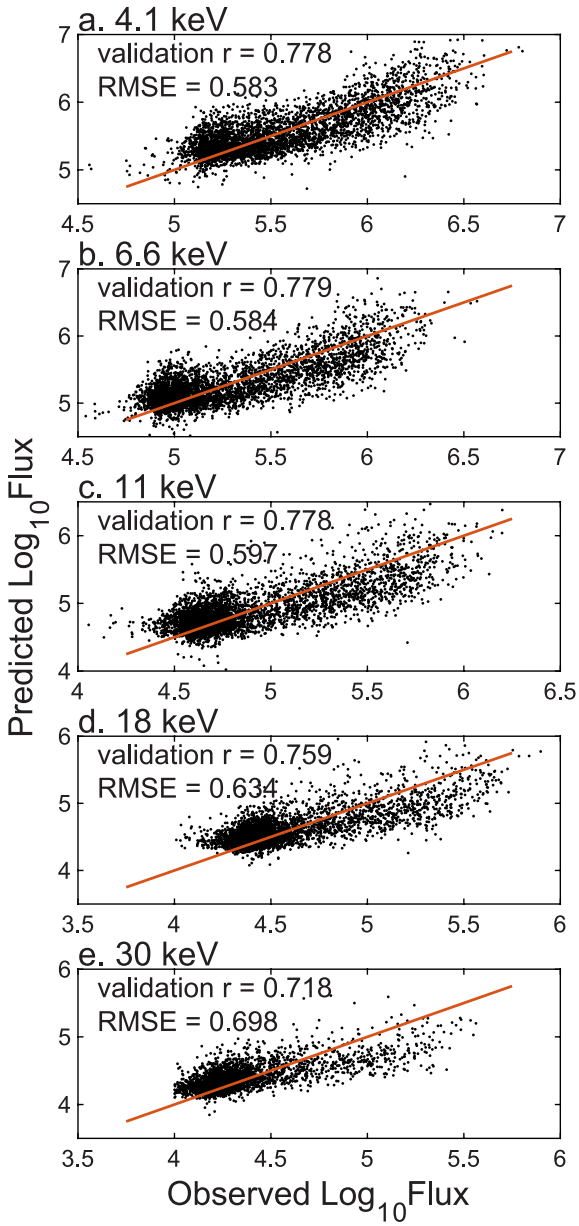


Figure 1. Lag1 LEEMYR model predictions versus observations of $\log_{10}$ electron flux during the test period (28-February-2021 to 26-August-2021). (a) 4.1 keV, (b) 6.6 keV, (c) 11 keV, (d) 18 keV, (e) 30 keV. Validation correlation (r) and standardized RMSE are given.

(Heidke (1926), with calculations in Ganushkina et al. (2019), Simms, Ganushkina, et al. (2023)). Predictions are categorized as correct hits (true positive: TP), incorrect prediction of an event when none occurs (false positive: FP), incorrect missed events where one did occur (false negative: FN), and correct prediction of non-events (true negative: TN). Using these numbers, the HSS calculates a score relative to random predictions. A score below zero will be obtained if the model predicts less well than chance alone. Scores closer to 1 show greater accuracy in prediction, with a score of 1 representing a perfect prediction.

We categorize "events" as above the 75th, 90th, or 97.5th percentile. The models do well in predicting flux occurrences above the 75th percentile, with forecasts 1 and 3 hr in the future having HSS of 0.50–0.659 (Figure 3). Prediction above the 90th percentile is not as good (0.33–0.50 for the Lag1 and Lag3 models). Forecasts of events above the 97.5th percentile (≈2 standard deviations above the mean) are worse (HSS of 0.18–0.24 at Lag 1, 0.02–0.31 at Lag 3), with predictions from >6 hr before being no better than chance. Predictions of the rarest high events, therefore, can only be made 1–3 hr before.

Low electron flux levels are not as well predicted, with the HSS of the lowest 10% of observations ranging from 0.10 to 0.19, and the HSS of the lowest 5% 0 or below (not shown in the figure). This set of solar wind, IMF, and geomagnetic parameters without a substorm measure are much less able to predict low electron fluxes than the high fluxes.

3.4. Over and Under Prediction

A more detailed look at when the models under- and overpredict even within the training set is shown in Figure 4 (29 June–11 July 2020). Over this roughly 2 week period (4.1 keV only), observed flux values are often underpredicted during high flux. This problem becomes more acute if flux is predicted 12 hr into the future (4c).

A 99% confidence interval (gray dotted lines) paints an overly optimistic view of the precision of the prediction, even when the accuracy is low. This tight confidence interval gives the appearance that we have high confidence that the prediciton is true, but even within the training set, this confidence interval often fails to include the actual flux value. This is not surprising, as the confidence interval of a large data set will often be quite narrow. (We use a 99% confidence interval here only so it will be visible on the plot, unlike the narrower 95% confidence interval.) For this reason, we do not recommend relying on confidence intervals to assess model quality.

We interpret a confidence interval as meaning that we are 99% sure the predicted mean lies in this range, but this is not quite what we want. We are more interested in knowing how confident we are of a prediction for a given set of input parameters. This metric is a prediction interval. Prediction intervals include uncertainty both for estimating the mean (as the confidence interval does) but also for the variation due to individual values. For this reason, a prediction interval is wider than a confidence interval and does not converge to the mean prediction as the sample size increases (for calculations see Neter et al., 1990). As expected, the prediction interval of Figure 5 contains the mean prediction at every point and the observed values at nearly every point. In the test period, over a sample period (31 May–12 June 2021), we show predicted flux in dark blue (± a barely visible 95% *confidence* interval), as well as the much wider 95% *prediction* interval (light blue/green outline). Although the prediction interval is wide, it does contain the observed value (red line) most of the time. As we have already noted from the scatterplots, the times when observations do not fall inside the prediction interval tend to be spikes of high electron flux.

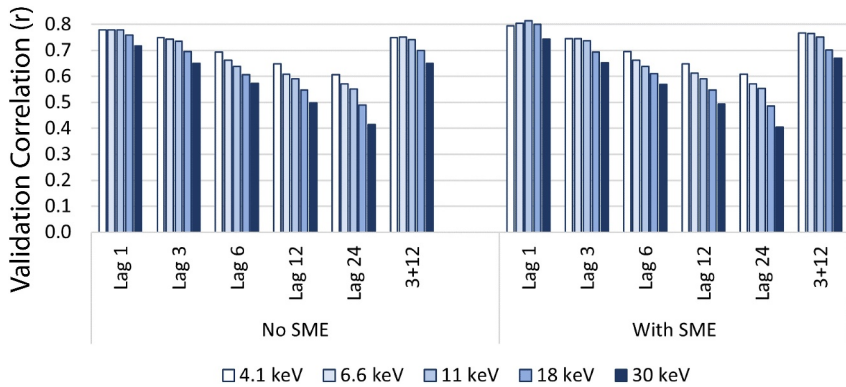


Figure 2. LEEMYR model validation correlations for predictions 1, 3, 6, 12, and 24 hr before the flux observation, with and without *SME* as an added predictor, for each electron energy (4.1, 6.6, 11, 18, and 30 keV).

3.5. Inverse N-Weighted Least Squares

These regression models use ordinary least squares (OLS), which weight observations equally. This means that a non uniform distribution of observations over a range of values will tend to be most influenced by values with larger n. As a result, areas with few observations (e.g., the highest flux values) do not contribute as much to the determination of the models. However, countering this, observations farthest from the mean are naturally weighted more heavily by virtue of their larger residuals. Consequently, the weighting by more common values may not be as extreme as supposed.

The "overweighting" by very common values can be minimized by using an inverse N-weighted least squares approach which assigns more influence to observations in sparse areas and less influence to points in the well-represented regions. This is similar to the imbalanced regression technique used by Chu et al. (2024). Using such an N-weighted approach might be expected to produce a model which more accurately describes rarer events that are of the most interest such as the high flux events. We create 36 bins ($i = 1 - 36$) spread equally over the range of values of the observed, Z score transformed flux, with n_i being the number of observations in each bin. As the bins are determined by the values, the number in each can vary. Using the *lscov* procedure of MATLAB, we create a weighted least squares solution with the weight for observations in any bin (i) being $1/n_i$.

Scatterplots of the inverse N-weighted models may show a better fit, to the eye, in the higher flux ranges (Figures 6 and 7). However, the overall validation correlations are lower at both 4.1 and 30 keV (other energies showed a similar difference). This is due to the much poorer fit in the mid-range values. The HSS values at a given threshold (75th, 90th, or 95th percentile) are similar between the OLS and N-weighted models, with no clear pattern of which is usually higher. At any given threshold, the TPR may be 2–3 times higher in the N-weighted models, which seems promising, however, this comes at the expense of an inflated FPR which can be 5–7 times

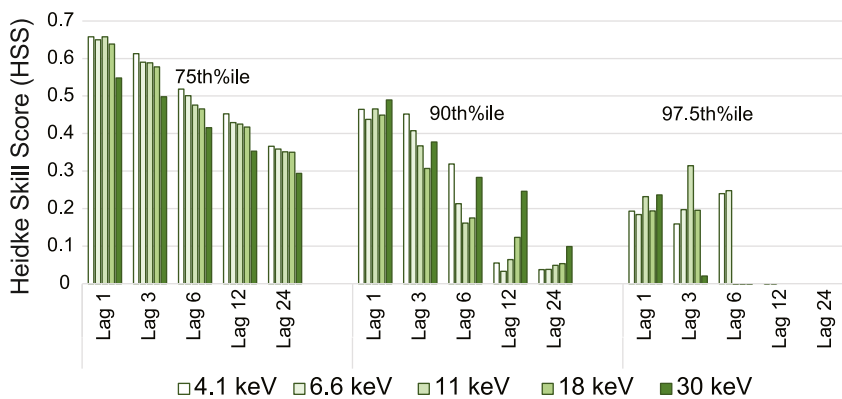


Figure 3. Heidke skill scores for LEEMYR model predictions 1, 3, 6, 12, and 24 hr before the flux observation, at the 75th, 90th, and 97.5th percentiles, for each electron energy (4.1, 6.6, 11, 18, and 30 keV).

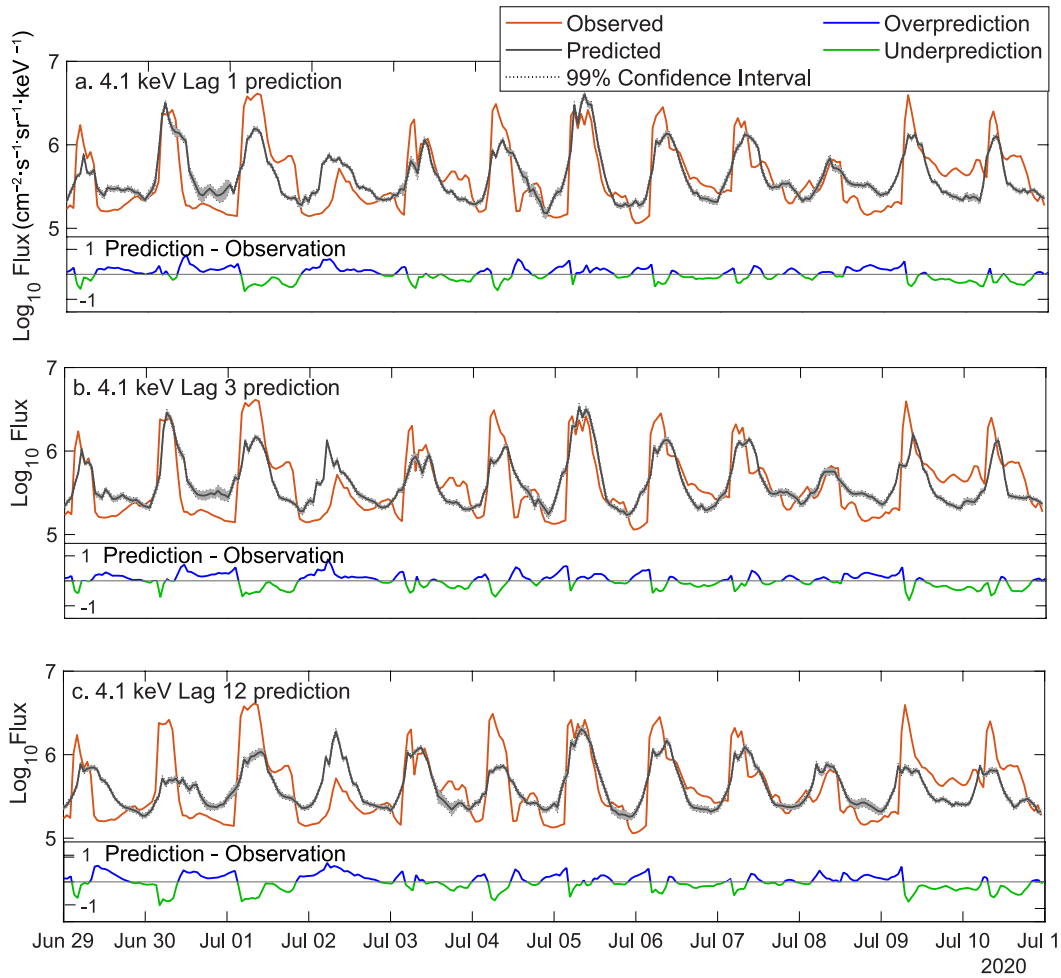


Figure 4. Prediction versus observation time plots (29 June–10 July 2020; 4.1 keV). Predictions (gray) often underpredict observations (orange) at the flux highs (green prediction–observation values). Overprediction (blue) is more common at the flux lows. Prediction 1 or 3 hr ahead (a, b) show similar accuracy, while 12 hr predictions (c) are somewhat less accurate. A 99% confidence interval around the model output (gray dotted line) lies very close to the predicted values and often does not include the test set observations.

higher in the N-weighted models. This can be seen graphically in the "bump" of midrange values wrongly classified into the FP region of the N-weighted regression scatterplot.

Rare observations far from the mean already are weighted more heavily by regression analysis (Neter et al., 1990), so further manipulation of the weighting may not be as useful as it may appear. Rather than manipulating the model to weight more heavily on the rare high flux events, a better approach may be to move the threshold for classifying an event.

4. The Models in Use

As MLT is included as a parameter, we can predict, for given solar wind, IMF, geomagnetic, and solar flux observations, electron flux levels over the entire orbit, however, a comparison observation is not at a single point in time but must be gathered over a full orbit. Figure 8 is the observed data, over a full orbit, gathered over the period centered at 20 March 2021 11:00:00 (UT). For each MLT, in the outer ring, electron flux above the 90th percentile is marked in yellow, above the 95th percentile in red, and above the 99th percentile in purple. There is an observed event (flux >95th percentile) in the 4.1 keV electrons. (Note that the viewing geometry is from above the South Pole.)

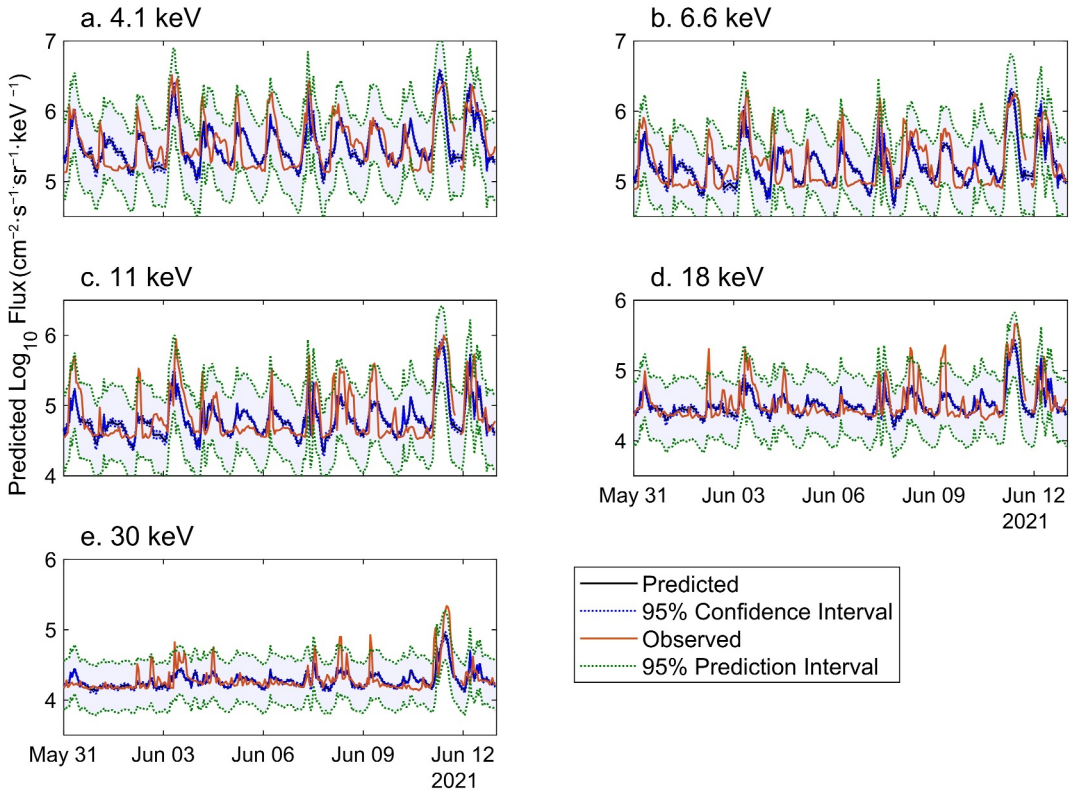


Figure 5. Prediction interval versus confidence interval for the Lag1 LEEMYR model (prediction 1 hr ahead) over 31 May–12 June 2021. (a) 4.1 keV, (b) 6.6 keV, (c) 11 keV, (d) 18 keV, (e) 30 keV. 95% confidence interval of the model (blue dashed line) is much smaller than the single observation 95% prediction interval (shaded blue with green outline). However, the single observation prediction interval is much more likely to contain the observed values.

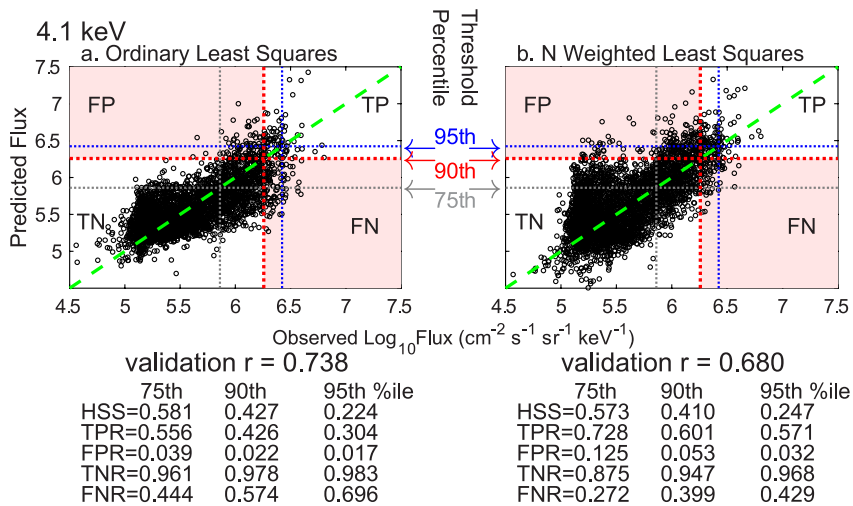


Figure 6. Observation versus prediction of 4.1 keV flux using (a). The Lag3 LEEMYR model (prediction 3 hr ahead, ordinary least squares) and (b). a Lag3 N-weighted model. True positive (TP) and false positive (FP) numbers can be more usefully manipulated by moving the "event" thresholds in the OLS model (75th, 90th or 95th percentiles) than by using an n-weighted model.

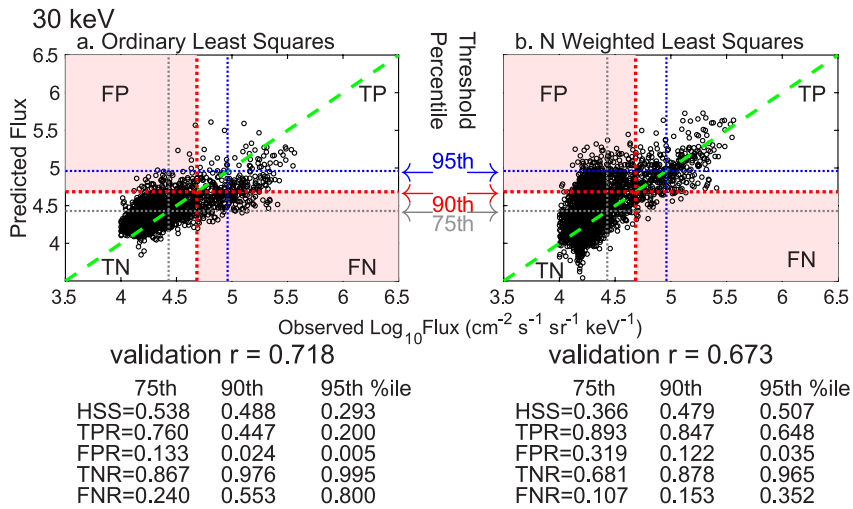


Figure 7. Observation versus prediction of 30 keV flux using (a). The Lag3 LEEMYR model (prediction 3 hr ahead, ordinary least squares) and (b). A Lag3 N-weighted model. True positive (TP) and false positive (FP) numbers can be more usefully manipulated by moving the "event" thresholds in the OLS model (75th, 90th or 95th percentiles) than by using an n-weighted model.

For this point in time, a prediction is made over the entire orbit (Figure 9). Alongside the visual prediction, the model outputs a $\log_{10}$ flux prediction in the accompanying text (± the 95% prediction interval). An event is predicted for energies ≤11 keV, but not for higher energies. This roughly approximates the observed flux at this point in time, where an event occurs in the 4.1 keV flux and there is a rise in flux at 6.6 and 11 keV.

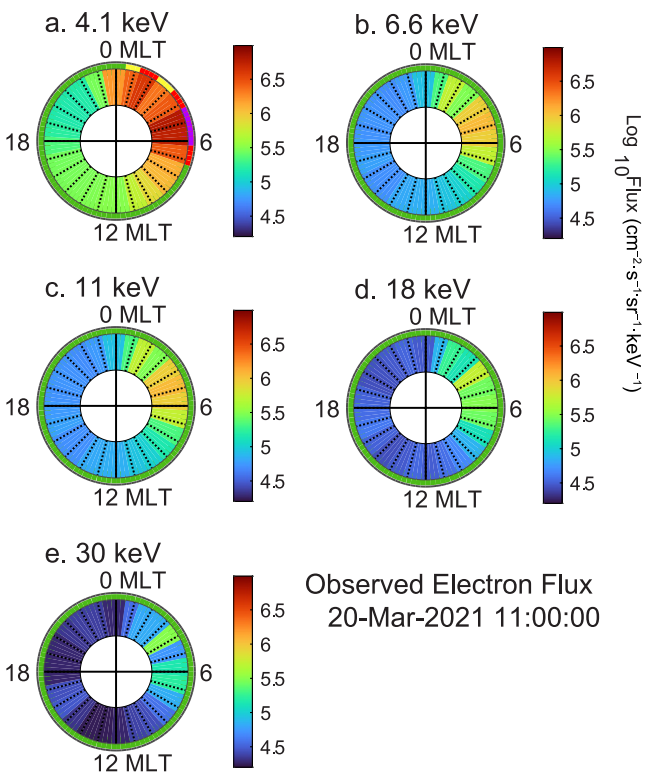


Figure 8. Observed flux over all MLT centered at 20 March 2021 11:00 UT. Observations above the 99th/95th/90th flux percentile are marked in purple/red/yellow in the outer ring. (a) 4.1 keV, (b) 6.6 keV, (c) 11 keV, (d) 18 keV, (e) 30 keV. This is an idealized representation, with viewing geometry from above the South Pole.

At this one point in time, we then explore how well flux is predicted from earlier hours (Figure 10). The observed event (flux >95th percentile) in the 4.1 keV flux at 0–6 MLT is labeled as being at the >99th percentile by the 1 and 3 hr ahead predictions. The 6 hr ahead prediction also flags the event, but the 12 hr ahead prediction warns only of a >90th percentile event and the 24 hr ahead prediction misses it entirely. While the validation correlations and HSS numbers do show that >3 hr ahead predictions are less reliable, this single event does suggest that earlier predictions could be monitored for signs of upcoming increases in flux.

5. Individual Correlations

The use of solar wind, IMF, and geomagnetic parameters in a predictive model is based on the assumption that these are correlated with electron flux. We test that assumption with cross correlations of each parameter over each flux energy (Figure 11). These show an apparent influence of some of the parameters in the 24 hr leading up to the flux measurement, with this correlation decreasing the further back in time we look. Correlations can be as high as 0.51 (*SME*). Although *SME* cannot be included in an actual prediction because it is not available in real time, several other variables show peak correlations that are high enough to be useful in flux prediction (e.g., Kp: 0.40, Bz: −0.39). Importantly, for prediction purposes, these correlations are often nearly as high at a 3 hr lag as they are at 0 or 1 hr lags. As no one variable (not even the unavailable *SME*) is tightly correlated with electron flux, it is useful to include all these variables in the stepwise regression to choose a set of parameters that best predict flux. Even though it may be tempting to manually drop such lower correlating parameters as V, N, P, and B to make a

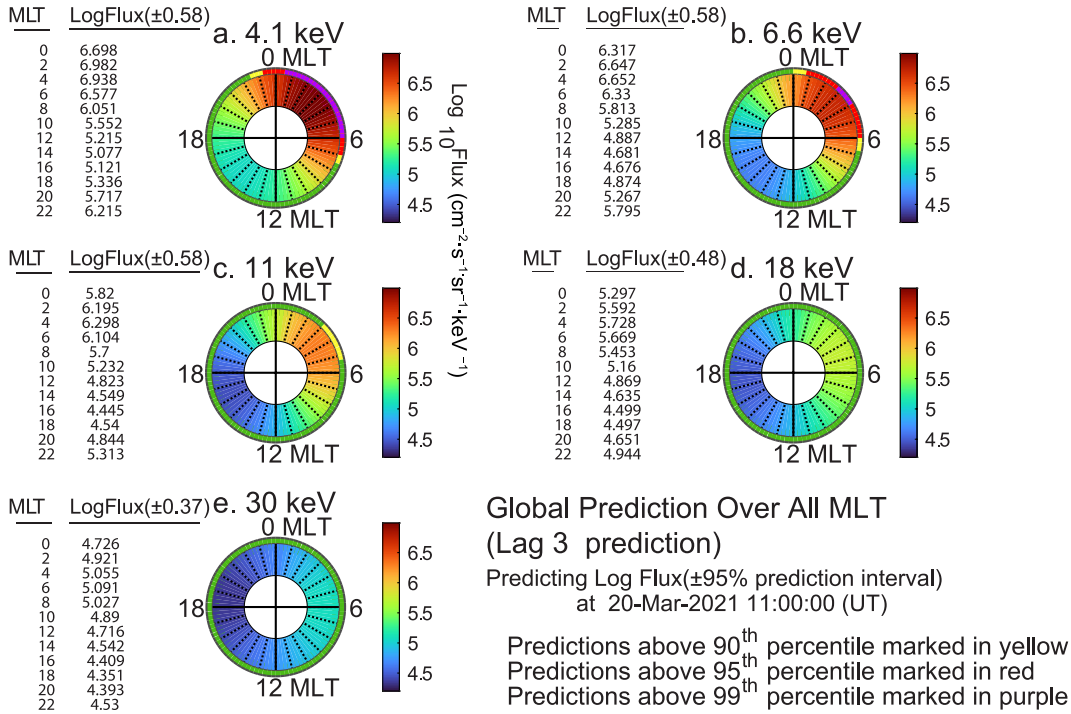


Figure 9. Example prediction from LEEMYR model over all MLT, centered at 20 March 2021 11:00 UT (3 hr ahead prediction; predictors measured at 08:00). Tables show $\log_{10}$ electron flux prediction at each MLT, at each energy, ± the 95% prediction interval. Predictions above the 99th/95th/90th flux percentile are marked in purple/red/yellow, respectively, in the outer ring. (a) 4.1 keV, (b) 6.6 keV, (c) 11 keV, (d) 18 keV, (e) 30 keV.

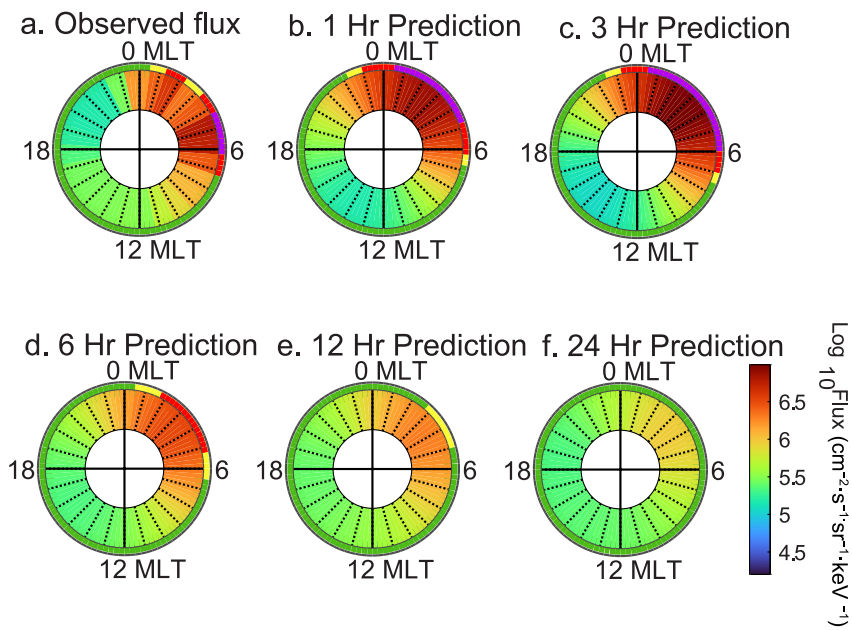


Figure 10. Example observation and prediction at 4.1 keV over all MLT, centered at 20 March 2021 11:00 UT. Predictions made from 1, 3, 6, 12, and 24 hr before. Predictions above the 99th/95th/90th flux percentile are marked in purple/red/yellow, respectively, in the outer ring. (a) Observed 4.1 keV flux in the 24 hr period centered on the prediction hour, (b) Prediction from 1 hr before, (c) From 3 hr before, (d) From 6 hr before, (e) From 12 hr before, (f) From 24 hr before.

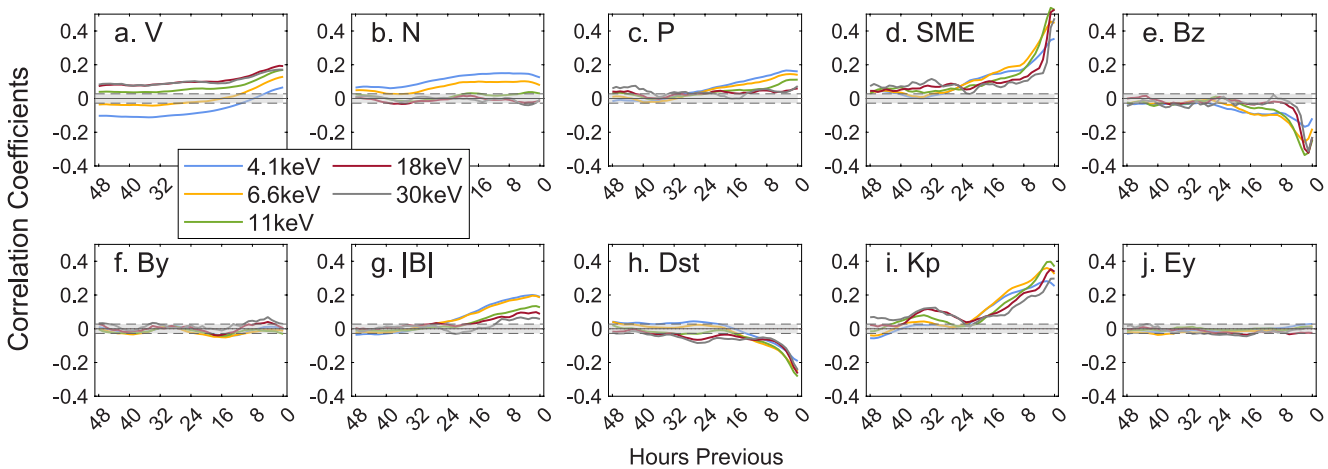


Figure 11. Cross correlations of each parameter with 4.1–30 keV flux over 48 hr previous to the flux observation (1 January 2020–26 August 2021). Statistically significant correlations lie outside the gray bars.

more parsimonious model, it is better, to increase predictive ability, to leave this determination to the stepwise procedure. As these lower correlating variables are often chosen by the stepwise procedure, this does indicate that there is some independent information contained in them that is worth retaining. In fact, in our models here, Bz was never chosen, even though it correlates more highly with electron flux indvidually than some other variables. This suggests (a) we can often get a better model if we allow the inclusion of less correlated variables and (b) inclusion in the model by a stepwise procedure is not a guarantee that the variables chosen are, individually, highly correlated with the response. As with other machine learning models, the algorithm is not intended to always include the most highly correlated variables, but instead the combination of parameters that best describe the variation in the response. As a consequence, the best predictive model is often not the most explanatory.

While many of these correlations are statistically significant ($p < 0.05$), this does not allow us to conclude that any of these variables drive electron flux. Possible confounding correlations with unmeasured variables, common trends in both flux and predictors, and autocorrelations in time series data all mean that these correlations may be due to factors other than direct physical relationships. However, even if these correlations were corrected for these sources of error, this series of cross correlations over time does not show evidence of either long term effects or cumulative influences because each variable is correlated with itself over these short time periods. Similarly, although several variables show up in the stepwise regression coefficients with apparently strong quadratic influences, we are unable to make the case, from this uncorrected analysis, that the physical relationships are nonlinear.

6. Discussion

1. To predict 4.1–30 keV electron flux at geosynchronous orbit, we develop the LEEMYR regression-based models, reduced to a more optimal set of predictors by a stepwise variable selection procedure, from GOES-16 data. As we have previously found that regression, ARMAX models, and neural networks all produce models that predict electrons with about the same level of accuracy (in the 40–150 keV range, Simms, Ganushkina et al. (2023)), we choose the simplest method here for developing models for lower energy electrons. It is unsurprising that the algorithm for model development would make little difference in outcome because (a) the time behavior of flux that can be modeled separately by the autoregressive and moving average terms of ARMAX models or with the addition of more past behavior (e.g., in a recurrent neural network) is already inherent in the time series data and therefore modeled by the regression (even more so if MLT is included as a parameter), and (b) interactive and polynomial terms describing nonlinear relationships that would be included in a neural network can also be easily added in a regression. To reduce overfitting to the specific training set, and to produce a more optimal set of predictors, the stepwise procedure of variable addition and elimination can be used, much as a neural network procedure would choose the most influential variables.

As the data in this energy range covers only 2020–2021, in the rising portion of the beginning of Solar Cycle 25, the model may be limited to periods of low solar activity. It is hoped that future work will explore how this model behaves in more active periods, perhaps with updates to the coefficients for periods of more intense solar activity.

2. When predicting 1 hr into the future, the LEEMYR models achieve a validation correlation as high as 0.778 for the 4.1 keV flux (0.779, 0.778, 0.759, and 0.718 for 6.6, 11, 18, and 30 keV flux respectively). Validation correlations for the Lag3 model (predicting 3 hr ahead) are somewhat lower (as high as 0.750 in the 4.1 keV energy range). These $R_E \sim 6.6$ (geosynchronous orbit) validation correlations are similar to those obtained for electrons of this energy in the plasmasheet region (18 and 06 MLT, $R_E 6 - 12$) of 0.74–0.76 (Swiger et al., 2022). The LEEMYR model is also comparable to other models using less similar energy ranges. Prediction-observation correlations from within the training sets used to build regression models of plasmasheet electrons range from $r = 0.65 - 0.81$ (Stepanov et al., 2021) to 0.82 (Dubyagin et al., 2016), but these are higher because they are not true validation correlations performed on a withheld test set. The RMSE for the LEEMYR models (0.58–0.70) exceeds that of the 32 keV Denton et al. (2016) model (0.34–0.36), however, these are not comparable because (a) the Denton and LEEMYR RMSEs are normalized in different ways, and (b) the Denton RMSE covers the training period of the model. A validation correlation with a withheld test set is not provided for the Denton model. The Boynton et al. (2011) 40 keV autoregressive model accounted for spatial differences by creating separate models for each MLT and thereby achieved an average validation correlation of 0.67, ranging from 0.51 to 0.79. The LEEMYR model validation correlation is comparable or better (0.72–0.78 over all MLT) because many of the MLTs show lower validation in the Boynton model. The Boynton model describes each MLT region separately so as to isolate the dynamics that may vary by location. However, we were able to incorporate both this location information (MLT represented by paired cosine and sine variables) as well as the differing dynamics in each region by allowing the models to choose MLT X input interaction terms if appropriate. In so doing, the LEEMYR model not only accounts for the different dynamics at various MLTs, it also achieves more predictive ability for the entire MLT range. At geosynchronous orbit, the Ma et al. (2022) neural network ORIENT-M model at 54 keV (the closest to the LEEMYR energies) achieves a validation correlation of 0.67, which is lower than the LEEMYR r for lower energies. However, the difference is even greater, as the ORIENT-M model incorporates *AE* or AL as an input. We do not use *AE* or *SME* in the final models because these are not available in real time and thus would make the model useless for prediction. When we do include *SME* as an input, for comparison, the validation correlation is as high as 0.81 (Figure 2).
3. The validation correlations are weighted heavily by the most common mid-range values, which are well predicted. As has been found with 40–150 keV electrons, the highest flux values tend to be underpredicted (Simms, Ganushkina, et al., 2023). To measure how well these models predict at the high flux events that are of most interest, the Heidke Skill score may be more informative. The HSS will be >0 if classification is better than chance, up to a maximum of 1 which indicates perfect sorting above and below the threshold. The HSS at the 75th percentile (0.548–0.659) is similar to that found for 40 keV electrons with a similar regression model (0.625, Simms, Ganushkina, et al., 2023). It is higher than that found for 40 keV electrons by Ganushkina et al. (2014, 2019) (as high as 0.17), probably because our current regression models incorporate both the nonlinear aspects of flux response (using polynomial and interaction terms) and the variation of predictor influence by MLT. At the 90th percentile these scores range from 0.449 to 0.496. Practically speaking, for hourly flux predictions, a dip below the median (50th percentile) occurs in almost every orbit and the 75th percentile is not much above this threshold. The most extreme events (>97.5th percentile or roughly 2 standard deviations above the mean) cannot be predicted more than 3 hr beforehand, and predictions from 1 to 3 hr ahead are less reliable. Thus, flagging increases above the 90th percentile may be more useful.
4. Both these metrics are slightly improved by the addition of *SME* to the model (validation r of 0.743–0.814; HSS at the 75th percentile of 0.564–0.724 and at the 90th percentile, 0.463–0.584). This is of no use for an actual predictive model, as *SME* is not available in real time, but it is notable that there is not more improvement. While substorm activity is related to low energy electron increases (Arnoldy & Chan, 1969; Birn et al., 1997; Thomsen et al., 2001; Turner et al., 2015), the high correlation of other parameters to substorms means that a specific measure of substorm activity is not needed in these models. Other parameters act as a proxy and further manipulation of them into an index is not needed.

5. However, for low flux values (<10th percentile), this model does not work well. An HSS of 0.10–0.19 is lower than that achieved for the >90th percentile fluxes. This set of solar wind, IMF, and geomagnetic parameters is of less use in predicting when fluxes will be low. This may be due to a stronger dependence of low flux on the *AE* or *SME* index (Boynton, Mourenas, & Balikhin, 2016; Ma et al., 2022). However, as we hope to create a model for real time predictions, the *SME* or *AE* substorm index cannot be used.
6. Because these models may be more heavily weighted on the much more numerous mid-range flux values, rather than the rarer high values that are of most interest, we produce several inverse N-weighted models that draw most influence from the less common but more important high and low values. However, although there is a modest increase in the number of true positive identifications of high flux events, it comes at the cost of a large increase in false positive identifications. Besides this, small populations of observations far from the mean already have an increased influence on the regression slopes (Neter et al., 1990), making this manipulation of the weightings less useful than it might at first appear. Rather than manipulating the model to weight more heavily on the rare high flux events, a better approach is to simply move the threshold for event classification. Moving the threshold increases the number of true positive identifications, but without an outsize increase in the number of false positives.
7. The lower predictive ability in the 18–30 keV range is also seen in the 40–75 keV electrons (Simms, Ganushkina, et al., 2023). Better predictive models, using these parameters, can be made for electrons <18 keV and >75 keV. It may be that we are not measuring the right parameters, or that the midrange of electron flux is simply more variable or less accurately measured. Whether this is an indication that there are processes we are simply not measuring, or if there is higher inherent stochasticity in these energy ranges is unknown.
8. We have found that reasonably accurate predictions can be made 1 hr into the future. However, a 1 hr warning may not be enough for practicality. Predictions 3 hr ahead are nearly as accurate and may be of more use. The accuracy is lower the further out predictions are made. However, a continuous process of predicting 24, 12, and 6 hr into the future before the 1 and 3 hr predictions would be helpful in notifying when an event might occur.
9. Although processes may act at different time scales, we have found that a Lag3 alone prediction is nearly as good as the Lag1 prediction, and that adding previous hours (e.g., Lag12) to the Lag3 model results in little improvement. Persistence of predictors or their influence may account for this, but previous work with MeV electrons suggests that, when spurious correlations due to cycles are eliminated, much of the influence is within only a few hours (Simms, Engebretson, & Reeves, 2023).
10. There is an advantage to using data from geosynchronous orbit. Although we are not able to forecast into other L shells, we are able to produce a full MLT forecast for any point in time.
11. Our models report prediction intervals rather than confidence intervals as we are more interested in knowing how confident we are of a prediction for a given set of input parameters. Prediction intervals include uncertainty both for estimating the mean (as a confidence interval does) but also for the variation due to individual values. A prediction interval is wider than a confidence interval but better describes the uncertainty in the prediction.
12. Although electron flux in this period shows some correlations with various predictors (most notably, *SME*, Bz, Dst, and Kp: Figure 11), these may be "spurious" (resulting from common correlation to unmeasured factors) rather than explanatory. In space weather data, these unmeasured variables are often common diurnal cycles and trends associated with long term variation in solar output or satellite measurements (Simms, Engebretson, & Reeves, 2022). While we may find correlations that produce a successful predictive model, this may say little about the physical relationships between parameters and electron flux. Loadings from regression models uncorrected for cycles and trends cannot be interpreted as explanatory. Similarly, the determination of neural network loadings (via the Shapley method, for example (Ma et al., 2023; Swiger et al., 2022)), do nothing to distinguish which of the correlation is due to physical relationships and which to spurious correlation. On the other hand, that Bz was not chosen in these stepwise regressions for the predictive models does not mean Bz has no physical influence, merely that other, correlated variables are already describing the physical effects of Bz. Neither the stepwise regression algorithm, nor other machine learning techniques, is optimized to necessarily include the most individually correlated variables, but instead the combination of parameters that best describe the variation in the response. As a consequence, the best predictive model may or may not be particularly explanatory in a physical sense. We therefore warn against ascribing physical interpretations to purely predictive models. In future work, we intend to remove cycles and trends to determine the physical drivers of electron flux in this keV energy range.

Data Availability Statement

OMNIWeb data is available at https://omniweb.gsfc.nasa.gov/form/dx1.html, the SuperMAG *SME* index at (https://supermag.jhuapl.edu/), and GOES-16 electron flux data at https://www.ngdc.noaa.gov/stp/satellite/goes-r.html. The Lag1-24 models and coefficients are available for use at https://zenodo.org/records/7520424.

Acknowledgments
The work at the University of Michigan was partly funded by National Aeronautics and Space Administration Grants NNX17AI48G, 80NSSC20K0353, and National Science Foundation Grants 1663770 and 2246912. The contributions by M. van de Kamp and N. Ganushkina were also partly supported by the Academy of Finland (Grant 339329). We gratefully acknowledge the SuperMAG collaborators (https://supermag.jhuapl.edu/info/?page=acknowledgement) for the *SME* index data, and two anonymous reviewers for their helpful comments.

References

Arnoldy, R. L., & Chan, K. W. (1969). Particle substorms observed at geostationary orbit. *Journal of Geophysical Research*, *74*(21), 5019–5028. https://doi.org/10.1029/JA074i021p05019

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). https://doi.org/10.1029/2011GL048980

Birn, J., Thomsen, M. F., Borovsky, J. E., Reeves, G. D., McComas, D. J., & Belian, R. D. (1997). Characteristic plasma properties of dispersionless substorm injections at geosynchronous orbit. *Journal of Geophysical Research*, *102*, 2309. https://doi.org/10.1029/96JA02870

Boynton, R. J., Amariutei, O. A., Shprits, Y. Y., & Balikhin, M. A. (2019). The system science development of local time-dependent 40-keV electron flux models for geostationary orbit. *Space Weather*, *17*(6), 894–906. https://doi.org/10.1029/2018SW002128

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*(A5). https://doi.org/10.1029/2010JA015505

Boynton, R. J., Balikhin, M. A., Sibeck, D. G., Walker, S. N., Billings, S. A., & Ganushkina, N. (2016). Electron flux models for different energies at geostationary orbit. *Space Weather*, *14*(10), 846–860. https://doi.org/10.1002/2016SW001506

Boynton, R. J., Mourenas, D., & Balikhin, M. A. (2016). Electron flux dropouts at geostationary earth orbit: Occurrences, magnitudes, and main driving factors. *Journal of Geophysical Research: Space Physics*, *121*(9), 8448–8461. https://doi.org/10.1002/2016JA022916

Bunde, A. (2023). The different types of noise and how they effect data analysis. *Chemie Ingenieur Technik*, *95*(11), 1758–1767. https://doi.org/10.1002/cite.202300031

Capman, N. S. S., Simms, L. E., Engebretson, M. J., Clilverd, M. A., Rodger, C. J., Reeves, G. D., et al. (2019). Comparison of multiple and logistic regression analyses of relativistic electron flux enhancement at geosynchronous orbit following storms. *Journal of Geophysical Research: Space Physics*, *124*(12), 10246–10256. https://doi.org/10.1029/2019JA027132

Choi, H. S., Lee, J., Cho, K. S., Kwak, Y. S., Cho, I. H., Park, Y. D., et al. (2011). Analysis of GEO spacecraft anomalies: Space weather relationships. *Space Weather*, *9*(5), 1–12. https://doi.org/10.1029/2010SW000597

Chu, X., Bortnik, J., Li, W., Shen, X., Ma, Q., Ma, D., et al. (2024). Distribution and evolution of chorus waves modeled by a neural network: The importance of imbalanced regression. *IEEE Xplore*. https://doi.org/10.23919/USNC-URSINRSM60317.2024.10464461

Chu, X., Ma, D., Bortnik, J., Tobiska, A., Cruz, W. K., Bouwer, S. D., et al. (2021). Relativistic electron model in the outer radiation belt using a neural network approach. *Space Weather*, *19*(12), e2021SW002808. https://doi.org/10.1029/2021SW002808

Denton, M. H., Henderson, M. G., Jordanova, V. K., Thomsen, M. F., Borovsky, J. E., Woodroffe, J., et al. (2016). An improved empirical model of electron and ion fluxes at geosynchronous orbit based on upstream solar wind conditions. *Space Weather*, *14*(7), 511–523. https://doi.org/10.1002/2016SW001409

Denton, M. H., Thomsen, M. F., Jordanova, V. K., Henderson, M. G., Borovsky, J. E., Denton, J. S., et al. (2015). An empirical model of electron and ion fluxes derived from observations at geosynchronous orbit. *Space Weather*, *13*(4), 233–249. https://doi.org/10.1002/2015SW001168

Dubyagin, S., Ganushkina, N. Y., Sillanpää, I., & Runov, A. (2016). Solar wind-driven variations of electron plasma sheet densities and temperatures beyond geostationary orbit during storm times. *Journal of Geophysical Research: Space Physics*, *121*(9), 8343–8360. https://doi.org/10.1002/2016JA022947

Freeman, J. W. (1974). Kp dependence of plasma sheet boundary. *Journal of Geophysical Research*, *79*(28), 4315–4317. https://doi.org/10.1029/ja079i028p04315

Freeman, J. W., O'Brien, T. P., Chan, A. A., & Wolf, R. A. (1998). Energetic electrons at geostationary orbit during the November 3-4, 1993 storm: Spatial/temporal morphology, characterization by a power law spectrum and, representation by an artificial neural network. *Journal of Geophysical Research*, *103*(A11), 26251–26260. https://doi.org/10.1029/97JA03268

Galica, G. E., Dichter, B. K., Tsui, S., Golightly, M. J., Lopate, C., & Connell, J. J. (2016). GOES-R space environment in-situ suite: Instruments overview, calibration results, and data processing algorithms, and expected on-orbit performance. In X. J. Xiong, S. A. Kuriakose, & T. Kimura (Eds.), *Earth observing missions and sensors: Development, implementation, and characterization IV* (Vol. 9881, pp. 237–251). SPIE. https://doi.org/10.1117/12.2228537

Ganushkina, N. Y., Liemohn, M. W., Amariutei, O. A., & Pitchford, D. (2014). Low-energy electrons (5–50 keV) in the inner magnetosphere. *Journal of Geophysical Research: Space Physics*, *119*(1), 246–259. https://doi.org/10.1002/2013JA019304

Ganushkina, N. Y., Sillanpää, I., Welling, D., Haiducek, J., Liemohn, M., Dubyagin, S., & Rodriguez, J. V. (2019). Validation of Inner Magnetosphere Particle Transport and Acceleration Model (IMPTAM) with long-term GOES MAGED measurements of keV electron fluxes at geostationary orbit. *Space Weather*, *17*(5), 687–708. https://doi.org/10.1029/2018SW002028

Ganushkina, N. Y., Swiger, B., Dubyagin, S., Matéo-Vélez, J.-C., Liemohn, M. W., Sicard, A., & Payan, D. (2021). Worst-case severe environments for surface charging observed at LANL satellites as dependent on solar wind and geomagnetic conditions. *Space Weather*, *19*(9), e2021SW002732. https://doi.org/10.1029/2021SW002732

Gjerloev, J. W. (2012). The SuperMAG data processing technique. *Journal of Geophysical Research*, *117*(A9), A09213. https://doi.org/10.1029/2012JA017683

Hartley, D. P., Denton, M. H., & Rodriguez, J. V. (2014). Electron number density, temperature, and energy density at GEO and links to the solar wind: A simple predictive capability. *Journal of Geophysical Research: Space Physics*, *119*(6), 4556–4571. https://doi.org/10.1002/2014JA019779

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284. https://doi.org/10.1109/TKDE.2008.239

Heidke, P. (1926). Measures of success and goodness of wind force forecasts by the gale-warning service. *Geografiska Annaler*, *8*(4), 301–349. https://doi.org/10.1080/20014422.1926.11881138

Hyndman, R., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts.
Katsavrias, C., Aminalragia-Giamini, S., Papadimitriou, C., Daglis, I. A., Sandberg, I., & Jiggens, P. (2022). Radiation belt model including semi-annual variation and solar driving (Sentinel). *Space Weather*, *20*(1), e2021SW002936. https://doi.org/10.1029/2021SW002936
Kellerman, A. C., & Shprits, Y. Y. (2012). On the influence of solar wind conditions on the outer-electron radiation belt. *Journal of Geophysical Research*, *117*(A5). https://doi.org/10.1029/2011JA017253
Koons, H. C., & Gorney, D. J. (1991). A neural network model of the relativistic electron flux at geosynchronous orbit. *Journal of Geophysical Research*, *96*(A4), 5549–5556. https://doi.org/10.1029/90JA02380
Koons, H. C., Mazur, J. E., Selesnick, R. S., Blake, J. B., Fennell, J. F., Roeder, J. L., & Anderson, P. C. (2000). The impact of the space environment on space systems. AFRL-VS-TR-20001578.
Korth, H., Thomsen, M. F., Borovsky, J. E., & McComas, D. J. (1999). Plasma sheet access to geosynchronous orbit. *Journal of Geophysical Research*, *104*(A11), 25047–25061. https://doi.org/10.1029/1999JA900292
Lam, H.-L., Boteler, D. H., Burlton, B., & Evans, J. (2012). Anik-E1 and E2 satellite failures of January 1994 revisited. *Space Weather*, *10*(10). https://doi.org/10.1029/2012SW000811
Li, X., Baker, D. N., Temerin, M., Reeves, G., Friedel, R., & Shen, C. (2005). Energetic electrons, 50 keV to 6 MeV, at geosynchronous orbit: Their responses to solar wind variations. *Space Weather*, *3*(4). https://doi.org/10.1029/2004SW000105
Li, X., Temerin, M., Baker, D. N., Reeves, G. D., & Larson, D. (2001). Quantitative prediction of radiation belt electrons at geostationary orbit based on solar wind measurements. *Geophysical Research Letters*, *28*(9), 1887–1890. https://doi.org/10.1029/2000GL012681
Ling, A. G., Ginet, G. P., Hilmer, R. V., & Perry, K. L. (2010). A neural network-based geosynchronous relativistic electron flux forecasting model. *Space Weather*, *8*(9). https://doi.org/10.1029/2010SW000576
Loto'aniu, T. M., Singer, H. J., Rodriguez, J. V., Green, J., Denig, W., Biesecker, D., & Angelopoulos, V. (2015). Space weather conditions during the Galaxy 15 spacecraft anomaly. *Space Weather*, *13*(8), 484–502. https://doi.org/10.1002/2015SW001239
Ma, D., Bortnik, J., Chu, X., Claudepierre, S. G., Ma, Q., & Kellerman, A. (2023). Opening the black box of the radiation belt machine learning model. *Space Weather*, *21*(4), e2022SW003339. https://doi.org/10.1029/2022SW003339
Ma, D., Chu, X., Bortnik, J., Claudepierre, S. G., Tobiska, W. K., Cruz, A., et al. (2022). Modeling the dynamic variability of sub-relativistic outer radiation belt electron fluxes using machine learning. *Space Weather*, *20*(8), e2022SW003079. https://doi.org/10.1029/2022SW003079
Matéo-Vélez, J.-C., Sicard, A., Payan, D., Ganushkina, N., Meredith, N. P., & Sillanpäa, I. (2018). Spacecraft surface charging induced by severe environments at geosynchronous orbit. *Space Weather*, *16*(1), 89–106. https://doi.org/10.1002/2017SW001689
Neter, J., Kutner, M. H., & Wassermann, W. (1990). *Applied linear statistical models* (3rd ed.). Irwin.
Newell, P., & Gjerloev, J. W. (2011). Evaluation of SuperMAG auroral electrojet indices as indicators of substorms and auroral power. *Journal of Geophysical Research*, *116*(A12), A12211. https://doi.org/10.1029/2011JA016779
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
Rigler, E. J., Baker, D. N., Weigel, R. S., Vassiliadis, D., & Klimas, A. J. (2004). Adaptive linear prediction of radiation belt electrons using the Kalman filter. *Space Weather*, *2*(3). https://doi.org/10.1029/2003SW000036
Sakaguchi, K., Miyoshi, Y., Saito, S., Nagatsuma, T., Seki, K., & Murata, K. T. (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
Shi, Y., Zesta, E., & Lyons, L. R. (2009). Features of energetic particle radial profiles inferred from geosynchronous responses to solar wind dynamic pressure enhancements. *Annales Geophysicae*, *27*(2), 851–859. https://doi.org/10.5194/angeo-27-851-2009
Sillanpää, I., Ganushkina, N. Y., Dubyagin, S., & Rodriguez, J. V. (2017). Electron fluxes at geostationary orbit from GOES MAGED data. *Space Weather*, *15*(12), 1602–1614. https://doi.org/10.1002/2017SW001698
Simms, L. E., & Engebretson, M. (2020). Classifier neural network models predict relativistic electron events at geosynchronous orbit better than multiple regression or ARMAX models. *Journal of Geophysical Research: Space Physics*, *125*(5), e2019JA027357. https://doi.org/10.1029/2019JA027357
Simms, L. E., Engebretson, M., Clilverd, M., Rodger, C., Lessard, M., Gjerloev, J., & Reeves, G. (2018). A distributed lag autoregressive model of geostationary relativistic electron fluxes: Comparing the influences of waves, seed and source electrons, and solar wind inputs. *Journal of Geophysical Research: Space Physics*, *123*(5), 3646–3671. https://doi.org/10.1029/2017JA025002
Simms, L. E., Engebretson, M., & Reeves, G. (2022). Removing diurnal signals and longer term trends from electron flux and ULF correlations: A comparison of spectral subtraction, simple differencing, and ARIMAX models. *Journal of Geophysical Research*, *127*(2), e2021JA030021. https://doi.org/10.1029/2021JA030021
Simms, L. E., Engebretson, M. J., Pilipenko, V., Reeves, G. D., & Clilverd, M. (2016). Empirical predictive models of daily relativistic electron flux at geostationary orbit: Multiple regression analysis. *Journal of Geophysical Research: Space Physics*, *121*(4), 3181–3197. https://doi.org/10.1002/2016JA022414
Simms, L. E., Engebretson, M. J., & Reeves, G. D. (2023). Determining the timing of driver influences on 1.8–3.5 mev electron flux at geosynchronous orbit using ARMAX methodology and stepwise regression. *Journal of Geophysical Research: Space Physics*, *128*, e2022JA030963. https://doi.org/10.1029/2022JA030963
Simms, L. E., Ganushkina, N. Y., der Kamp, M. V., Balikhin, M., & Liemohn, M. W. (2023). Predicting geostationary 40-150 kev electron flux using ARMAX (an autoregressive moving average transfer function), RNN (a recurrent neural network), and logistic regression: A comparison of models. *Space Weather*, *21*(5), e2022SW003263. https://doi.org/10.1029/2022SW003263
Simms, L. E., Ganushkina, N. Y., van de Kamp, M., Liemohn, M. W., & Dubyagin, S. (2022). Using ARMAX models to determine the drivers of 40–150 keV GOES electron fluxes. *Journal of Geophysical Research*, *127*(9), e2022JA030538. https://doi.org/10.1029/2022JA030538
Simms, L. E., Pilipenko, V., Engebretson, M. J., Reeves, G. D., Smith, A. J., & Clilverd, M. (2014). Prediction of relativistic electron flux at geostationary orbit following storms: Multiple regression analysis. *Journal of Geophysical Research: Space Physics*, *119*(9), 7297–7318. https://doi.org/10.1002/2014JA019955
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*(11), e2020SW002532. https://doi.org/10.1029/2020SW002532
Stepanov, N. A., Sergeev, V. A., Sormakov, D. A., Andreeva, V. A., Dubyagin, S. V., Ganushkina, N., et al. (2021). Superthermal proton and electron fluxes in the plasma sheet transition region and their dependence on solar wind parameters. *Journal of Geophysical Research: Space Physics*, *126*(4), e2020JA028580. https://doi.org/10.1029/2020JA028580
Swiger, B. M., Liemohn, M. W., Ganushkina, N. Y., & Dubyagin, S. (2022). Energetic electron flux predictions in the near-earth plasma sheet from solar wind driving. *Space Weather*, *20*(11), e2022SW003150. https://doi.org/10.1029/2022SW003150
The MathWorks Inc. (2023). MATLAB version: 23.2.0.2599560 (R2023b) Update 8. The MathWorks Inc.

Thomsen, M. F., Birn, J., Borovsky, J. E., Morzinski, K., McComas, D. J., & Reeves, G. D. (2001). Two-satellite observations of substorm injections at geosynchronous orbit. *Journal of Geophysical Research*, *106*(A5), 8405–8416. https://doi.org/10.1029/2000JA000080

Thomsen, M. F., Henderson, M. G., & Jordanova, V. K. (2013). Statistical properties of the surface-charging environment at geosynchronous orbit. *Space Weather*, *11*(5), 237–244. https://doi.org/10.1002/swe.20049

Turner, D. L., Claudepierre, S. G., Fennell, J. F., O'Brien, T. P., Blake, J. B., Lemon, C., et al. (2015). Energetic electron injections deep into the inner magnetosphere associated with substorm activity. *Geophysical Research Letters*, *42*(7), 2079–2087. https://doi.org/10.1002/2015GL063225

Wing, S., Johnson, J. R., Turner, D. L., Ukhorskiy, A. Y., & Boyd, A. J. (2022). Untangling the solar wind and magnetospheric drivers of the radiation belt electrons. *Journal of Geophysical Research: Space Physics*, *127*(4). https://doi.org/10.1029/2021JA030246