A neural network forecasting of relativistic electron flux at geostationary orbit: solar activity phase dependence

T Yeeram
Department of Physics, Geoinformatics and Space Technology, Mahasarakham University, Kantarawichai, Thailand
thana.y@msu.ac.th

Abstract. Geostationary-satellite anomalies are mainly caused by extreme space environment such as flux enhancements of relativistic electrons. A prediction of the relativistic electrons at geostationary orbit (GEO) is crucial to alert of the flux enhancements before it reaches a harmful condition. This work investigates contribution of solar activities at various phases in the daily electron flux (E > 2 MeV) at GEO measured by GOES satellites during the solar cycles 23-24. An alternative method based on the solar activity and feed forward neural network (NN) scheme with back-propagation learning in Waikato Environment for Knowledge Analysis is adopted to predict the electron fluxes in GEO by sorting the solar activity into maximum, descending, and ascending phases. As the model inputs, the historical daily sum Kp and the > 2 MeV electron fluxes are used. Results indicate that the prediction capability of the NN model is dependent on the solar activity in relation to the sum Kp. Crossing the test over the similar solar phases gives a better prediction result than crossing over different phases. The NN model based on the lagged sum Kp and log-flux input are more suitable for the forecasting during the descending and ascending phases, while the lagged log-flux input is more suitable during the maximum phase.

1. Introduction
Satellites at geostationary orbit (GEO; 35,790 km in altitude) are exposed to space weather hazard mainly caused by relativistic electrons with energy E > 1 MeV [1]. These so-called killer electrons partially reside in the Earth's outer radiation belts (ORBs) which typically lies between L = 3.5 and L = 7 (L is the distance in Earth radii to the equatorial crossing of a given geomagnetic field line). Satellites at GEO, locating at L = 6.6 toward the edge of the ORB, densely populate the orbit; in February 2018 there are over 548 operational satellites in GEO (Wikipedia). These MeV-particles can cause internal charging that is built up over time, leading to arcing or electrostatic discharge which can damage components and even lead to total loss of satellite. The ORB is highly dynamic due to a variety of transport, acceleration, and loss processes in relation to geomagnetic activity driven by the Sun [2]. At GEO, the MeV-energy electron flux varies dramatically between 10 and 10⁵ cm⁻² sr⁻¹ s⁻¹ over a few days. In such an extreme space environment, satellite ground operators should avoid sending critical commands to satellites in order to prevent potentially catastrophic consequences to the spacecrafts. Thus, the prediction of relativistic electrons at GEO is crucial to alert of any flux enhancements before it reaches a harmful condition [3].

Many prediction models of the relativistic electron flux at GEO have long been proposed [3], including an artificial neural network (NN). NN is widely used to model complex and nonlinear relationships between inputs and outputs. Many NN models have been utilized for determining the
relationship between relativistic electron flux and solar wind drivers [4,5,6,7,8,9,10] since Koons et al [11] firstly developed a NN model using the LANL particle data with an input of daily sum $K_p$ over 10 consecutive days that gives prediction of daily averaged > 3 MeV electrons.

Wrenn et al. [12] show that a specific and 27-day recurrent anomaly on a GEO satellite due to internal charging depends on solar cycle (sunspot cycle) and season during 1991-2000. During the descending and ascending phases of solar activity, magnetospheric relativistic electrons are the main cause of the anomalies [12,13]. However, long-term effects of the solar activity on the relativistic electron flux at GEO have not intensively been employed in the model prediction. This work aims to develop a NN prediction model for > 2 MeV daily electron flux variations at GEO based on GOES measurements made during different phases of solar (sunspot) cycles 23-24.

2. Data Sets and Methods

2.1. Data Sets

The relativistic electron flux data at energies of > 2 MeV with resolution of 5 mins were obtained from the Geostationary Operational Environmental Satellite (GOES) observation series in the west direction. The 5 min-data were averaged to daily data for the year 1997, 2000, 2006, 2008, 2011, 2015, and 2017, which were selected during the time interval spanned the solar cycles 23-24. Note the solar cycle is counted from one solar minimum to the next solar minimum; the cycles 23 and 24 began in 1996 and 2008, respectively. The $K_p$ index represents geomagnetic activity in the subauroral region. The $K_p$ index is derived from the $K$ index from each station, which is based on the more disturbed component of the observed horizontal magnetic field after removing the daily quiet variations. After eliminating local time and seasonal effects, concurrent values from the 13 stations are used to create the global $K_p$ index [14]. Note sum $K_p$ is the daily sum of the $K_p$ index.

2.2. Artificial Neural Network

The NN is trained using a back-propagation algorithm, which is a supervised learning scheme by which a multilayer feed-forward network is trained. The network consists of an input layer, one hidden layer, and one output layer. The inputs consist of $N$ known values of daily sum $K_p$ and log-flux of electrons and the output of one known values of the daily log-flux. The activation function for the hidden nodes is a non-linear sigmoid function (or logistic function). The network training utilizes the error back-propagation algorithm where the actual response $a(n)$ of the network [15], when presenting the $n^{th}$ training example, move closer to the applied desired response $d(n)$ by minimizing the summed squared error

$$\varepsilon = \sum_{n=1}^{N} [d(n) - a(n)]^2,$$  

(1)

where $N$ denotes the total number of examples in the training set.

The learning process adjusts the weights connecting the different layers of the network to minimize $\varepsilon$. The adjustment is made first between the output and hidden layers and then between the hidden and input layers [16]. The weight update is performed after the presentation of all the training examples that constitute an epoch. The weight correction applied to the synaptic weight connection neuron $I$ to neuron $j$ at epoch $s$ is defined by the generalized delta rule:

$$\Delta \omega_{ji}(s) = -\eta \frac{\partial \varepsilon(s)}{\partial \omega_{ji}} + \alpha \Delta \omega_{ji}(s-1).$$  

(2)

The gradient $-\partial \varepsilon(s)/\partial \omega_{ji}(s)$ determines the direction of search in weight space and $\eta$ is the learning rate parameter. To accelerate descent in steady downhill (minus sign) and stabilize the descent in directions that oscillate in sign, a momentum constant $\alpha$ is included and the weight is updated from the previous epoch [16]. The trick of choosing the optimal learning rate is to train a network starting from a low learning rate and increase the learning rate exponentially for every batch.

2
Training, testing, and validating a NN are performed. At the end of the training, the NN can be further tested on an additional independent test set referred to as a validation (prediction) set. To evaluate the amount of the forecast error, we employed the mean error or the mean absolute error (MAE):

$$\text{MAE} = \frac{1}{N} \sum_{n=1}^{N} |d(n) - a(n)|.$$  

(3)

The average of absolute values of the $|d(n) - a(n)|$ is taken to obtain the MAE. We also used the root mean square error (RMSE) that measures the average magnitude of the error:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (d(n) - a(n))^2}.$$  

(4)

2.3. Neural Network Suit and Model Development

The NN scheme used in this study is Waikato Environment for Knowledge Analysis (WEKA) [17], which is a suite of machine learning software written in Java, developed at the University of Waikato, New Zealand. WEKA (≥3.8) has a dedicated time series analysis environment that allows forecasting models to be developed, evaluated and visualized [18]. There are basic and advanced configurations in the WEKA forecasting. In the latter one, the Base learner is available; here multilayer perceptron (MLP) learning algorithm was implemented since MLP is capable to capture high non-linearity of data. The input data set is available in Attributed-Relation File Format (ARFF).

In this study, the forecasting was done with different phases of solar cycles. The yearly data were pruned into three sets based on phases of the solar cycles: descending, ascending, and maximum. Next, the NN inputs were designated to three scenarios: (1) log-electron fluxes only, (2) combination of sum $K_p$ and log-electron flux, and (3) overlaying of sum $K_p$ on log-electron flux. Prior to the analysis, we have removed errors from the data sets. The data are preprocessing to a normalization. To avoid overfitting, which decreases the generalization of the model performance, we used early stopping technique. The early stopping technique works by splitting the whole data set into a training set and a validation set with one-year period for each solar phase. The network was trained using the training set, while testing its performance on the validation set.

3. Neural Network Model Results

In the NN scheme, the output is logarithmic fluxes predicted simultaneously at one day ahead. Here we used $\eta = 0.05$, $\alpha = 0.1$, and the hidden layers are half of the sum of input and output nodes. We found that the best epoch is 100. The previous day’s electron flux is also well correlated with current electron flux. The NN model results are investigated for the selected years both in similar and different solar activity phases.

Table 1. Prediction accuracy of the NN forecasting relativistic electron flux at GEO in each year. Lagged times are in brackets and the best predicting results are highlighted in bold.

| Year | Phase   | Training (days) | Testing (days) | Log flux $\text{MAE}$ | Log flux $\text{RMSE}$ | Log flux-sum $K_p$ $\text{MAE}$ | Log flux-sum $K_p$ $\text{RMSE}$ | Sum $K_p$ overlay $\text{MAE}$ | Sum $K_p$ overlay $\text{RMSE}$ |
|------|---------|----------------|---------------|------------------------|-------------------------|---------------------------------|-------------------------------|---------------------------------|-------------------------------|
| 2000 | Max     | 250            | 103           | .329(2d)              | .438                    | .334(1d)                       | .447                          | .322(2d)                       | .433                          |
| 2006 | Descend | 209            | 70            | .401(6d)              | .591                    | .426(2d)                       | .590                          | .387(4d)                       | .597                          |
| 2008 | Descend | 215            | 98            | .373(6d)              | .502                    | .369(7d)                       | .512                          | .363(7d)                       | .496                          |
| 2011 | Ascend  | 243            | 110           | .234(6d)              | .288                    | .252(1d)                       | .322                          | .241(2d)                       | .311                          |
| 2015 | Descend | 240            | 109           | .390(2d)              | .564                    | .361(4d)                       | .499                          | .371(2d)                       | .517                          |
| 2017 | Descend | 231            | 104           | .429(4d)              | .564                    | .359(2d)                       | .449                          | .415(4d)                       | .548                          |

3.1. Yearly Data

The best NN model results for the training and testing in yearly data for the three scenarios are tabulated in table 1. The ratio of train to test data was 70:30. For the maximum phase in 2000 the three scenarios give similar good prediction results with time lags of 2-3 days similar to the cross correlation (c.c.) result.
between the log-flux and sum $K_p$ (not shown). In addition, the small MAE values are related to very small seasonal variations in the flux. However, MAE values of the training data are more than 0.4. In 2006, the MAE and RMSE are large and comparable for the 3 cases because of 40 days - data gap and time lags are 2 and 4 days. For 2008, the errors are large because the solar activity is very low at the middle to the end of year. In 2011, the errors are small and the overlay is the best. For 2015 and 2017 the best input is the combination of log-flux and sum $K_p$ with time lag of one day. From these results, the values of MAE and RMSE are dependent on individual factors of the data such as the selected testing interval, difference in the training and testing interval, seasonal variations, and the number of data gaps.

Table 2. Prediction accuracy of the NN forecasting relativistic electron flux at GEO in similar solar phases. Lagged times are in brackets and the best prediction results are highlighted in bold.

| Train-test year | Training (days) | Testing (days) | Log flux MAE | RMSE | Log flux-sum $K_p$ MAE | RMSE | Sum $K_p$ overlay MAE | RMSE |
|-----------------|-----------------|----------------|--------------|------|-------------------------|------|-----------------------|------|
| 2008-2017       | 325             | 347            | .442(4d)     | .622 | .430(3d)                | .580 | .431(4d)              | .596 |
| 2015-2006       | 349             | 289            | .409(6d)     | .574 | .401(5d)                | .526 | .420(3d)              | .580 |
| 2017-2008       | 347             | 148            | .344(4d)     | .509 | .391(1d)                | .575 | .338(1d)              | .536 |
| 2011-1997       | 365             | 149            | .393(1d)     | .538 | .359(1d)                | .499 | .377(1d)              | .535 |

Figure 1. One-day ahead prediction by overlaying sum $K_p$ on log-flux compares to observation for $E > 2.0$ MeV electron fluxes by crossing 2011 over 2006. Dashed line indicates the flux exceeds 1,000 FU used by the NOAA Space Weather Prediction Center for its real-time alerts.

3.2. Crossing Over in Similar Phases

It is interesting to examine the solar activity dependence by training the yearly data in a specific solar phase and then testing them in the other similar solar phases. The process was done for most of the possible cases to examine the efficiency of the process. The NN model results for training and testing data in the similar phases are tabulated in table 2. Firstly, during descending phases, the MAE is less than 0.44 for combination input of log-flux and daily sum $K_p$ for 2008- training set and 2017-testing set. The MAE is quite large because from middle to the end of year 2008 the solar activity is minimal, from which the solar phases are different. Conversely, 2017-training set and first 128 days in 2008-testing set give MAE less than 0.34 for the overlay daily sum $K_p$ on the log-flux. This verifies that similarity in phases give a better result and difference in phases give a poorer result. For the 2015-training set and 2006-testing set, the MAE is less than 0.41 for the combination input of log-flux and daily sum $K_p$. The large value may arise from many data gaps in 2006. Secondly, during ascending phases, 2011-training set and 1997-testing set give MAE less than 0.36 for the combination input of log-flux and sum $K_p$ and one-day time lag. Even though the 1997-test set has many data gaps, the prediction accuracy is satisfactory. This suggests that the similar phases give good prediction accuracy. An example of the validation of the test sets in the similar phases is shown in figure 1. The NN prediction model can capture most of the validation set even there are many data gaps in 2006.
3.3. Crossing Over in Different Phases

To further test that forecasting the log-flux is whether dependent on solar phases or not, the NN was trained in yearly data in one solar phase to test for a year in different solar phase. The performance of the model was measured by comparing model outputs with measured fluxes over a year. The NN model results for the different phases are tabulated in table 3. It is found that for ascending phase-training set and maximum phase-testing set such as respective 2011 and 2000, the MAE is less than 0.36 for the overlay of sum $K_p$ with time lag of 10 days. This indicates that $K_p$-related processes in the ascending phase also contribute in the maximum phases in some degrees. In case of descending phase-training set and maximum phase-testing set such as respective 2015 and 2000, the MAE is 0.360 for the log-flux input with time lag of six days. Note the other two input scenarios that using daily $K_p$ give substantially poorer results, indicating that the physical $K_p$-related processes in the descending phase are not significantly contributed to the maximum phases.

Moreover, in case of ascending phase-training set and descending phase-testing set such as respective 2011 and 2006, the MAE is less than 0.39 for the overlay method with time lag of 3 days. The results suggest that physical $K_p$-related processes in the descending phase are also contributed to the ascending phases in some degrees. The model results are consistent with the appearance of similarly prominent solar wind drivers and their geomagnetic effects in the solar phases. Most coronal mass ejection (CME)-driven storms are simultaneous present during solar ascending phase and maximum phase, while most high-speed solar wind (HSS)-driven storms are simultaneous present during solar ascending phase and descending phase [19].

4. Summary and Discussion

The developed NN model to predict the relativistic electron flux based on the solar phases of solar cycles 23-24 indicated some remarkable points. The results will be discussed in the contexts of underlying physics of solar wind-magnetosphere/radiation belt coupling and space weather forecasting.

The NN results indicated that the sum $K_p$ index is more suitable to predict the log-flux in the descending and ascending phases than the maximum phase. One of the reasons for this is the correlation coefficient (c.c.) between the sum $K_p$ and log-flux is higher in the former phases than in the latter phase. The c.c.s of the descending and ascending phases are in the range of 0.37-0.56 with time lags of 0-3 days, whereas the c.c. of the maximum phase is 0.33 with time lag of 4 days. In general, an explanatory time series that shows a high c.c. to the predictor improves the goodness of the model. Previous work found strong nonlinearity of the production function on the $K_p$ for each of 70 flux enhancement intervals as observed in ascending to solar maximum phases (1997-2000) [20]. They found that one of the most important parameters controlling them is 4-day averaged $K_p$ index before the maximal electron flux, which is consistent with the cross correlation and NN results. At GEO the electron flux is more correlated with $K_p$ than $Dst$ [3] since $K_p$ represents global geomagnetic activity in the magnetosphere. The correlation analysis indicates that magnetospheric processes related to the relativistic electron flux and $K_p$ in the maximum phase and descending phase are different. We point here that the best candidate for the difference is substorm events which are most prominent in descending phase or a part of ascending phase. The accumulation processes of substorm activity are longer in CME than in HSS storms as suggested by the time lag analysis of $K_p$ and log-flux in this study. Mark that the $K_p$ increase promotes effective operation of resonant drift mechanisms. At the same time, a significant increase of the $K_p$ leads to intensive electron losses in the outer magnetosphere.

| Train-test year | Training (days) | Testing (days) | Log flux MAE | Log flux RMSE | Log flux-sum $K_p$ MAE | Log flux-sum $K_p$ RMSE | Sum $K_p$ overlay MAE | Sum $K_p$ overlay RMSE |
|-----------------|-----------------|----------------|--------------|--------------|------------------------|------------------------|------------------------|------------------------|
| 2017-2006       | 347             | 289            | 0.426(3d)    | 0.583        | 0.367(3d)              | 0.530                  | 0.388(3d)              | 0.557                  |
| 2011-2006       | 365             | 289            | 0.428(3d)    | 0.619        | 0.394(2d)              | 0.558                  | 0.386(3d)              | 0.590                  |
| 2011-2000       | 365             | 353            | 0.368(10d)   | 0.512        | 0.379(4d)              | 0.512                  | 0.358(10d)             | 0.499                  |
| 2015-2000       | 349             | 353            | **0.360(6d)**| 0.493        | 0.431(1d)              | 0.571                  | 0.418(7d)              | 0.567                  |

Table 3. Prediction accuracy of the NN forecasting relativistic electron flux at GEO in different solar phases. Lagged times are in brackets and the best prediction results are highlighted in bold.
5. Conclusions
This study investigated the solar phase dependence on the prediction of relativistic electron \((E > 2 \text{ MeV})\) flux at GEO in different phases of the solar cycles 23-24. We adopted an alternative method based on the feed forward NN scheme with back-propagation learning for the prediction. Results indicate that the NN model based on sum \(K_p\) and log-flux input is more suitable for forecasting relativistic electron flux at GEO during ascending and/or ascending phases, while the log-flux input is more suitable during maximum phase. Crossing the test over the similar phases gives a better prediction result, while crossing over different phases gives a poorer result. The \(K_p\)-related processes (substorms) are similar for the ascending-maximum phases and for the descending-ascending phases, while they are different for the descending-maximum phases. The results are consistent with the appearance of similarly prominent solar wind drivers in each solar phase. This study can be useful for the forecasting of relativistic electrons in each solar phase. Typical training sets based on solar activity can be prepared for the interested testing sets of similar activity. To improve the forecast capability, the seasonal effects as well as more inputs of feature will be focused in future work.

Acknowledgements
This research was financially supported by Mahasarakham University 2018, the national research council of Thailand (NRCT), and Mahasarakham University Development Fund. The author is grateful to the GOES mission. The OMNI data were obtained from http://omniweb.gsfc.nasa.gov/.

References
[1] Wrenn G L 1995 J. Spacecraft and Rockets 32 514
[2] Thorne R M 2010 Geophys. Res. Lett. 37 L22107
[3] Sakaguchi K, Nagatsuma T, Reeves G D and Spence H E 2015 Space Weather 13 853
[4] Stringer G A, Heuten I, Salazar C and Stokes B 1996 Washington DC Amer. Geophys. Union Geophys. Mono. Ser. 97 291
[5] Freeman J W, O’Brien T P, Chan A A and Wolf R A 1998 J. Geophys. Res. 103 26251
[6] Fukata M, Taguchi S, Okuzawa T and Obara T 2002 Ann. Geophys. 20 947
[7] Ling A G, Ginet G P, Hilmer R V and Perry K L 2010 Space Weather 8 S09003
[8] Kitamura K, Nakamura Y, Tokumitsu M, Ishida Y and Watari S 2011 Artif. Life Robot. 16 389
[9] Shin D K, Lee D Y, Kim K C, Hwang J and Kim J 2016 Space Weather 14 313
[10] Wei L, Zhong Q, Lin R, Wang J, Liu S and Cao Y 2018 Space Weather 16 903
[11] Koons H C and Gorney D J 1991 J. Geophys. Res. 96 5549
[12] Wrenn G L, Rodgers D J and Ryden K A 2002 Ann. Geophys. 20 953
[13] Pilipenko V, Yagova N, Romanova N and Allen J 2006 Adv. in Space Res. 37 1192
[14] Menvielle M and Berthelier A 1991 Rev. Geophys. 29 415
[15] Boberg F, Wintoft P and Lundstedt H 2000 Phys. Chem. of the Earth, Part C: Solar, Terr. Planet. Sci. 25 275
[16] Haykin S 1999 Neural Networks: A Comprehensive Foundation Prentice Hall
[17] Witten I H, Frank E, Hall M A and Pal C J 2016 Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques, Morgan Kaufmann Publishers Inc.
[18] Hall M A, Frank E, Holmes G, Pfahringer B, Reutemann P and Witten I H 2009 SIGKDD Explorations 11 10
[19] Richardson I G, Cane H V and Cliver E W 2002 J. Geophys. Res. 107 1187
[20] Dmitriev A V and Chao J K 2003 J. Geophys. Res 108 1388