JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 115, A04306, doi:10.1029/2009JA014890, 2010

Solar proxies pertaining to empirical ionospheric total electron content models

Takashi Maruyama[1]

Received 10 September 2009; revised 27 October 2009; accepted 11 November 2009; published 9 April 2010.

[1] Solar proxies and indices exhibiting extreme ultraviolet (EUV) irradiance that affects the ionospheric total electron content (TEC) were examined through training an artificial neural network (ANN). A TEC database was constructed from a dense GPS receiver network over Japan from April 1997 to March 2008, covering an entire 11 year solar activity period. In empirical models of upper atmospheric parameters, such as the International Reference Ionosphere model and the Mass Spectrometer and Incoherent Scatter thermospheric model, the 10.7 cm solar radio flux ($F_{10.7}$) or the sunspot number (R) is used as a proxy for determining the solar activity. In the present study, ANN training for predicting TEC as a target parameter was done by including new solar proxies/indices in the input space that were based on direct measurements of solar EUV/UV flux, SOHO_SEM$_{26–34}$ (the integrated 26–34 nm EUV emission), and Mg II cwr (the core-to-wing ratio of Mg II 280 nm line), as well as the traditional indices $F_{10.7}$ and R. Root mean square errors (RMSEs) of TEC were compared after the training was completed using a variety of combinations of solar proxies. When a single proxy was used, SOHO_SEM$_{26–34}$ yielded the smallest RMSE, or it was the best proxy for modeling ionospheric TEC. Further, general improvements were obtained by combining different types of proxies and short- and long-term means of them. The best combination was the 3 day smoothed daily, 7 day and 27 day backward mean values of Mg II cwr, SOHO_SEM$_{26–34}$, and the 10.7 cm radio flux.

Citation: Maruyama, T. (2010), Solar proxies pertaining to empirical ionospheric total electron content models, *J. Geophys. Res.*, *115*, A04306, doi:10.1029/2009JA014890.

1. Introduction

[2] Total electron content (TEC) is a fundamental parameter of the ionosphere and an important space weather concern because it determines time delays in transionospheric radio propagation that can cause errors in the applications of Global Navigation Satellite Systems such as GPS. There is a requirement for modeling TEC variations by considering solar inputs. The primary factor that controls TEC variations are solar extreme ultraviolet (EUV) radiations with $\lambda \leq$ 102.5 nm that ionize the thermospheric neutral particles. Thus, TEC varies at various time scales such as the 11 year solar activity period, the 27 day solar rotation, and the Earth's rotation (or local time) period, by which incoming solar EUV flux widely varies. Changes in the loss rate of ionospheric plasma are another factor that must be considered in TEC modeling. An important loss process of major ion species, O^+, in the F region is the chemical rearrangement collision with N_2 and O_2 that produces molecular ions, NO^+ and O_2^+. Molecular ions quickly disappear through dissociative recombination. The atmospheric composition represented by [O]/[N_2] is an important factor that affects TEC and depends on the atmospheric temperature, dynamics, and other thermospheric conditions. In the region between 170 and 300 km, the ionization process of neutral particles by the EUV radiation is the primary heat source [*Stolarski et al.*, 1975]. The absorbed EUV energy splits into the photoelectron channel and the chemical energy channel (ion-electron pair) depending on the wavelength. Virtually all photons of $\lambda \leq$ 102.5 nm are absorbed by photoionization [*Richards et al.*, 1994]. Photoionization cross sections of O, N_2, and O_2 broadly peak at wavelengths from 40 to 70 nm. At shorter wavelengths, photoionization cross sections are reduced and the heights of maximum rate of absorption of solar flux become lower. As the ion loss rate increases with lowering the height, contributions to the ionospheric electron density would be reduced. The dissociation process of molecular oxygen (O_2) by the Schumann-Runge continuum radiation (λ = 130 ~ 175 nm) is the primary heat source at heights lower than 170 km. The photodissociation of O_2 also affects the [O]/[N_2] ratio and thus the electron density. Thus solar UV/EUV radiations directly (photo ionization) and indirectly (through thermospheric modification) control TEC variations. Because of the complex relationship between solar spectral irradiance and TEC as described above, we conceptually considered ionospheric effective EUV (IE-EUV) and not any particular wavelength of UV/EUV. Our purpose in this study is to examine how the IE-EUV behaves with various solar proxies.

[1]National Institute of Information and Communications Technology, Koganei, Japan.

0148-0227/10/2009JA014890

[3] As most UV/EUV radiations are absorbed in the upper atmosphere and direct measurements of irradiance by satellites are not always available, several proxies have long been used in the theoretical and empirical modeling of the ionosphere-thermosphere system. The sunspot number (R) and the 10.7 cm radio flux ($F_{10.7}$), which are measured on the ground, have been widely used. Previous empirical atmospheric models such as a series of MSIS (Mass Spectrometer and Incoherent Scatter radar) [*Hedin et al.*, 1977] (see *Picone et al.* [2002] for the latest version) use $F_{10.7}$ and the 81 day centered average of $F_{10.7}$ as proxies of solar UV heating. The International Reference Ionosphere (IRI) model [*Bilitza and Reinisch*, 2008] uses the 12 month average of R as a proxy of solar activity. *Richards et al.* [1994] used the mean of daily $F_{10.7}$ and the 81 day average of $F_{10.7}$ to compose a solar EUV model for aeronomical calculations.

[4] Recently, solar irradiance data sets based on satellite measurements were extensively examined in connection with aeronomical applications. The integrated 26–34 nm solar EUV emission has been measured by the solar extreme ultraviolet monitor (SEM) onboard the Solar and Heliospheric Observatory (SOHO) at the Lagrange Point since its launch in late 1995. These wavelengths overlap with the wavelengths of EUV that ionize the Earth's upper atmosphere unlike the other proxies. *Bowman et al.* [2008a] generate a new index $S_{10.7}$ by converting the integrated 26–34 nm EUV emission through linear regression to implement an empirical thermospheric model. Another index is the ratio of the irradiance in the core of the Mg 280 nm line to the irradiance at neighboring wavelength, referred to as the Mg II core-to-wing ratio (cwr). Mg II cwr is first proposed as a measure of chromospheric solar active region activity by *Heath and Schlesinger* [1986]. The relationship has been examined between Mg II cwr and the UV/EUV irradiance variability at other wavelengths and the 10.7 cm solar radio flux [*Lean et al.*, 1992; *Cebula et al.*, 1992; *Viereck et al.*, 2001; *Thuillier and Bruinsma*, 2001; *Floyd et al.*, 2005]. *Emmert et al.* [2009] demonstrated a correlation between Mg II index and the total EUV energy from 0 to 120 nm measured by using the TIMED/SEE instrument. *Lean et al.* [2001] pointed out that a chromospheric index, either alone or in combination with the coronal 10.7 cm radio flux, offers the possibility of a more accurate representation of EUV irradiance variability. All these studies show that Mg II cwr is a good proxy for the solar activity.

[5] *Floyd et al.* [2005] demonstrated the different long-term behavior among Mg II cwr, $F_{10.7}$, R, and He 1083 solar activity indices. The short-term temporal behavior of these indices over several solar rotation periods was also different [*Donnelly et al.*, 1985; *Floyd et al.*, 2005]. These different behaviors partly originated from the different source regions responsible for the measurements. Regarding the atmospheric modeling, *Tobiska et al.* [2008] pointed out that it is desirable to have solar indices and proxies that vary differently in time.

[6] Another point that we have to pay attention to in the application of solar proxies is the relationship between long-term and short-term variations in measurements. The amplitude of short-term variations induced by solar rotation relative to the long-term variations that follow the 11 year solar activity period varies from one wavelength to another [*Donnelly et al.*, 1986; *Lean*, 1987; *Woods et al.*, 2000]. As a consequence, short-term and long-term variations of proxy are not simply scaled to describe irradiances that we concern. One technique to overcome this in atmospheric models, in which the 10.7 cm solar radio flux is used as a proxy for EUV irradiance, is to use both daily and 81 day mean of radio flux values [e.g., *Hedin et al.*, 1977].

[7] Solar irradiance effects on neutral atmosphere have been widely studied in comparison to those on the ionosphere. The effects are not the same on the thermosphere and the ionosphere, although they are closely connected to each other, because the energy passes are complicated and not exactly the same [*Stolarski et al.*, 1975]. Therefore, it is not quite clear which solar index/proxy and its means over periods are pertaining to describe the solar irradiance effects on the ionosphere. For statistical studies of ionospheric variations caused by solar irradiance changes, an advantage of SOHO_SEM$_{26–34}$ and Mg II cwr is their superior data coverage when compared with other UV/EUV irradiance measurements, and they fulfill our TEC data set from April 1997 to March 2008, covering an entire 11 year solar activity cycle. In this study, we examined SOHO_SEM$_{26–34}$ and Mg II cwr as well as traditional proxies, the sunspot number and the 10.7 cm radio flux, to find a pertaining proxy pattern for IE-EUV. Outputs from an empirical model, SOLAR2000 [*Tobiska et al.*, 2000], were also compared.

[8] To model the highly complex responses of the ionosphere-thermosphere system to various proxy patterns, an artificial neural network technique (ANN) was used. The most commonly used ANN is multilayer perceptrons [*Haykin*, 1994] with a target of ionospheric parameter such as f_OF2, $h'F2$, and TEC, and the inputs of parameters affecting the ionospheric variations such as solar irradiance proxies, season, local time, latitude/longitude, and others. To investigate effective solar inputs specifying TEC variations, we used an ANN model similar to that described by *Maruyama* [2007], in which solar inputs and seasons alone are required in the input space. In section 2, the ANN technique is described. Section 3 gives results and section 4 interprets the results. Section 5 provides the conclusion.

2. TEC Database and ANN Modeling

[9] In Japan, a dense GPS receiver network called GPS Earth Observation Network (GEONET) has been developed by the Geographical Survey Institute, Japan, and data from more than 1000 locations have been collected since April 1997, covering the entire solar cycle. About 300 GEONET receivers were chosen for evaluating TEC to ensure uniform coverage over Japan. The major issues in deriving TECs from GPS radio signal observations are the instrumental biases in both satellites and receivers and the conversion process from the observed TECs along the slant path to the vertical ones. In this study, slant TECs were converted to vertical TECs (vTEC) at the piercing point where the ray path crossed a shell at a height of 400 km (a thin shell model). By assuming that the vTECs in a small area 2 × 2° in longitude and latitude surrounding a grid point were the same in a short period of time, we calculated the daily instrumental biases and vTEC in each small area at 15 min intervals, using the algorithm developed by *Ma and Maruyama* [2003]. The vertical TEC obtained in this way is referred to as the grid TEC (gTEC). The grid points are shown in Figure 1.

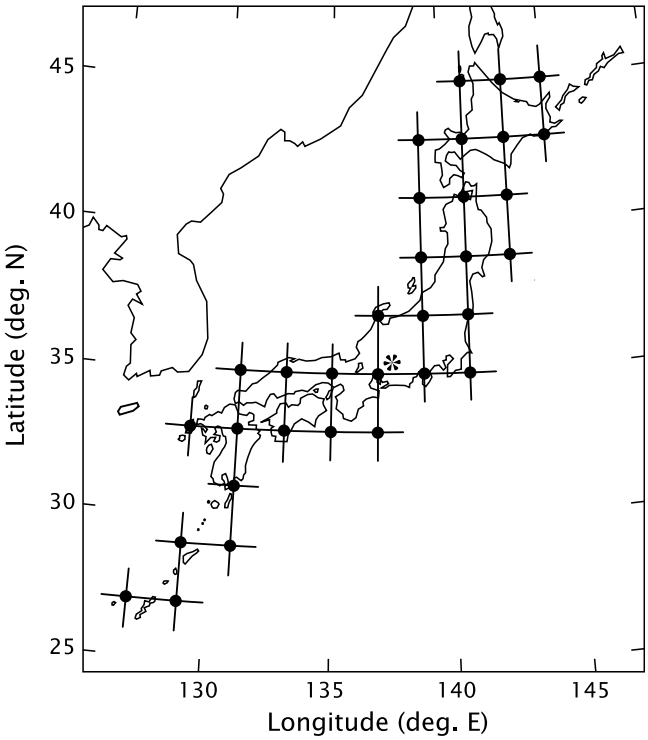

Figure 1. Thirty-two grid points on which TECs were determined at 15 min intervals. Data for the grid point denoted by asterisk (35°N, 137°E) were used for the validation of the ANN.

[10] The major factors that determine the TEC are solar activity, season, local time, and geographic and geomagnetic coordinates. To deal with many factors, a multilayer feed-forward artificial neural network (ANN) [*Rumelhart et al.*, 1986; *Haykin*, 1994] was applied. ANNs have been used to explicitly specify the ionosphere by approximating the relationship between geophysical conditions, as mentioned above (inputs), and the observed ionospheric parameter such as TEC or f_OF2 (a target) [*McKinnell and Poole*, 2004; *Oyeyemi et al.*, 2005; *Nakamura et al.*, 2007]. *Maruyama* [2007], hereinafter referred to as paper 1, developed an ANN in which local time/longitude and latitude were not included in the input space but they were imbedded in the target parameters. That is, TEC variations in time and latitude were expressed as a two-dimensional map; the longitudinal dependence can be assumed to be equivalent to the local mean time (LMT) dependence in a limited longitude area such as Japan's longitudes, as depicted in Figure 1. The two-dimensional TEC map (some examples can be found in paper 1) is the target of the ANN. This technique greatly reduces the computer storage, and the same method was used in this study.

[11] To generate TEC maps from the gTEC, the surface harmonic expansion method based on the associated Legendre's function was used, as shown in (1), by considering the LMT (hour) at each grid point as the azimuth parameter, i.e., $\phi = 2\pi(\mathrm{LMT}/24)$.

$$\mathrm{TEC} = \sum_{m=0}^{M}\sum_{n=m}^{N}(A_{nm}\cos m\phi + B_{nm}\sin m\phi)P_n^m(\cos\theta) \qquad (1)$$

Here, θ is the colatitude and $N = M = 7$. The functional fitting was performed to determine coefficients A_{nm} and B_{nm} in (1). As dummy data were set in the southern hemisphere for mathematical convenience, the entire spherical distribution map is symmetrical with respect to the equator (resultant map data outside the latitude range from 29 to 45° N were disregarded). In other words, A_{nm} and B_{nm} with the odd number of $n + m$ are equal to zero. Thus, our ANN had a total of 36 target parameters that were required to reconstruct diurnal-latitudinal TEC maps. In this study, the TEC map for a given day was constructed by the data gathered from three UT days centered on that day, and the coefficient set was calculated for each day through the entire 11 year solar cycle from 1 April 1997 to 31 March 2008.

[12] The multilayer feed-forward network consisted of an input layer, a hidden layer, and an output layer. The schematic diagram of the network is shown in Figure 2. The input layer had solar proxies and the season which was expressed in sin and cos components of day of the year (DOY); the number of the input nodes varied depending on the combination of solar proxies and the maximum number of solar inputs examined in this study was 16 (4 proxies by 4 types of means). The number of nodes in the hidden layer was chosen to be 100 irrespective of the number of input nodes. The output parameters were the 36 Legendre's coefficients, as described above. Training was divided into two stages. That is, weight updating was performed by the pattern mode in which weights were updated after the presentation of each training example [*Haykin*, 1994] for the first 30 epochs of the back-propagation training. The order of the presentation of training examples was randomized from one epoch to the next. After the weights were coarsely determined, weight updating was continued by the batch mode in which weights were updated after the presentation of all the training examples that constituted an epoch. Both the training and momentum constants [*Haykin*, 1994] were chosen to be 0.9 in the batch mode training. A global error was evaluated at each epoch by taking the summation of square error (square of the difference between the ANN output and the teacher target parameter) over the 36 output

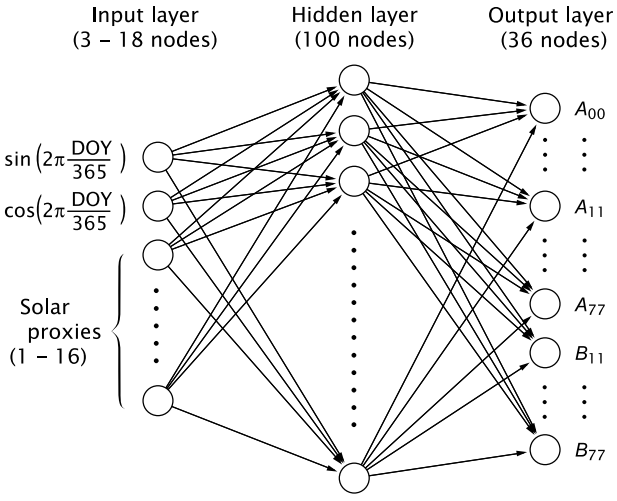

Figure 2. Schematic diagram of the neural network.

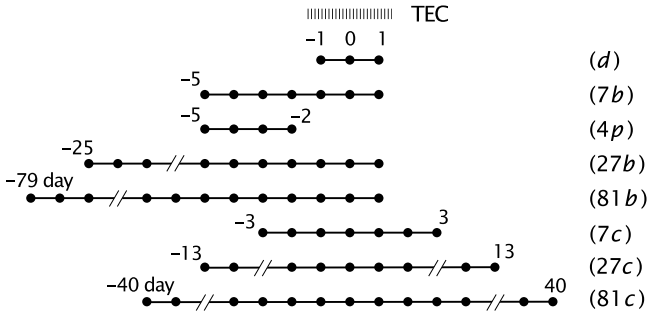

Figure 3. Periods over which proxies are averaged.

nodes and the entire (11 year) training data set. When the absolute rate of changing in the global error per epoch reduced to 10^{-5}, the back-propagation algorithm was considered to have converged (stopping criterion) and the weight updating was terminated. The stopping criterion and training parameters were kept constant through the experiment because our major purpose was to compare the degree of training achievement of the network among various input patterns aimed at finding optimal solar proxies for modeling TEC variations.

[13] The TEC data set was divided into two subsets: one was for the training of the network and the other was for the evaluation of the training achievement. The evaluation data set was constructed by extracting a day for ten days each and the rest days were used for the training so that both the training and evaluation data sets uniformly covered various phases of the 11 year solar cycle and the season. The validation of the network training achievement was done by reconstructing TEC maps from the 36 Legendre's coefficients, but not for the coefficients themselves as network outputs. Root mean square errors (RMSEs) were calculated for the hourly values of the observed and network-predicted TECs for the grid point corresponding to (35°N, 137°E), near the center of the GEONET coverage. The initial weights of neurons are given by random numbers in the neural network training, and the achievement of trained network depends not only on the selection of input parameters but also on the initial weights. To overcome this disadvantage in the ANN application to determine pertaining parameters, we run eleven trials of network training and the best result was used for comparison and discussion.

[14] We examined four solar proxies based on measurements and two output indices from an empirical irradiance model. The four measurements were the sunspot number, the 10.7 cm solar radio flux, integrated 26–34 nm irradiance by SOHO_SEM, and Mg II cwr. The daily sunspot number we used is the International Sunspot Number R_i that is provided by the Solar Influences Data Analysis Center (SIDC), Belgium: this number is generated by a network of observing stations and evaluated from a statistical treatment. The daily 10.7 cm solar radio flux, $F_{10.7}$, in sfu (= 10^{-22}W m^{-2} Hz^{-1}) is the value at 20 UT (local noon at Penticton, Canada) observed by the Dominion Radio Astrophysical Observatory, Canada. In the actual application of ANN, all input and output parameters are scaled to take values between 0 and 1. For R_i and $F_{10.7}$, the reported daily values were simply divided by the maximum values in the solar cycle. Mg II cwr is measured twice a day by the NOAA series operational satellites. *Tobiska et al.* [2008] preprocess and convert Mg II cwr to the daily $M_{10.7}$ index that is normalized to the variation of $F_{10.7}$ by performing linear regression. We used the $M_{10.7}$ index divided by the maximum value in the solar cycle. The SOHO_SEM instrument measures the 26–34 nm EUV emission at higher time resolution and daily averages are used to calculate the daily $S_{10.7}$ index [*Tobiska et al.*, 2008] that is also normalized to the variation of $F_{10.7}$. The $S_{10.7}$ index was not available for the periods from 25 June to 25 October 1998 and from 22 December 1998 to 2 February 1999. The periods of data gap are further extended up to 23 April 1999, when 81 day backward means are calculated. Thus, we used the S_{EUV} index, an output of the SOLAR2000 (version 2.35) empirical solar irradiance model as described later, to fill the data gaps by linear regression for the period including 10 solar rotations before and after the periods of the data gap. For this linear regression, the S_{EUV} index values were shifted forward by one day based on the analysis of delayed responses of TEC in this study. We refer to this composite index as $S^*_{10.7}$.

[15] SOLAR2000 (http://www.spacewx.com/solar2000.html) is an empirical solar irradiance model driven by H Lyman-α and the 10.7 cm radio flux [*Tobiska et al.*, 2000]. The integrated irradiance corresponding to SOHO_SEM$_{26–34}$ or S_{EUV} [*Tobiska et al.*, 2008] was examined and used to fill the data gaps of SOHO_SEM$_{26–34}$, as described above. Another index examined is $E_{10.7}$, integrated irradiance over a range of 1–105 nm: this wavelength range of radiations contributes to the ionization of O^+, O_2^+, and N_2^+.

[16] Because the TEC maps were smoothed over 3 UT days, the daily proxies and indices were also averaged over 3 days, and for convenience, we refer to the 3 day means of proxy and indices as 3 day smoothed daily, or simply, daily values (*d*). In addition to the daily values, several mean values were calculated for the periods of 7, 27, and 81 day backward including the 3 days of the TEC map (7*b*, 27*b*, and 81*b*), and for the periods of 7, 27, and 81 day centered on the TEC map (7*c*, 27*c*, and 81*c*). Here, the 27 and 81 day periods correspond to the one and three solar rotation periods, respectively. While the 7 day backward mean was intended to simulate a short term trend of the TEC response or integration effects in the ionosphere/thermosphere system. With this regard, an additional parameter, means over 4 days preceding the 3 day TEC period (4*p*) was examined. These periods are shown in Figure 3. Thus, we have 48 (8 variables for 6 proxies) solar inputs and a variety of combinations of them must be examined. In principle, there are $2^{48} - 1$ ($\sim 10^{14}$) patterns. When the time lag of ionospheric response to solar inputs is considered, the number of cases further diverges. It is impossible and pointless to examine all these combinations of input parameters. We narrowed the cases step by step and about 300 combinations of proxies, periods, and time lags were examined. The maximum number of solar inputs examined was 16 (4 variables for 4 proxies).

3. Results

3.1. Single Proxy Comparisons

[17] In the simplest case, the solar activity is expressed by a single proxy on the same day of the TEC observations. In this study, 3 day smoothed daily proxies were considered

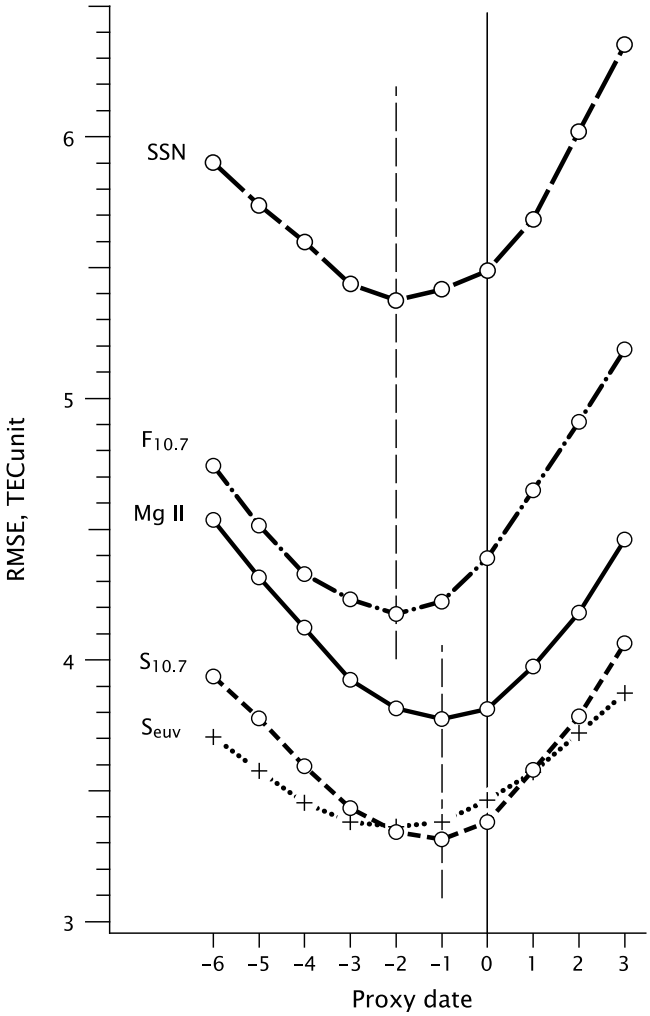

Figure 4. Delayed response of TEC to 3 day smoothed daily solar proxies: sunspot number, 10.7 cm radio flux, Mg II cwr, SOHO_SEM$_{26–34}$, and an output of SOLAR2000 (S_{EUV}). Negative proxy date signifies delayed responses.

corresponding to that the TEC maps were smoothed over 3 days. The comparisons of different proxies are shown in Figure 4. In Figure 4, the horizontal axis represents the day shift of the input proxy with respect to the day of TEC observations. The above mentioned simplest case corresponds to the zero day shift. The RMSE in the vertical axis represents the training achievement of the ANN. As expected from previous studies, RMSE for the sunspot number is the largest and R_i is not a good proxy for IE-EUV. The 10.7 cm radio flux is better than the sunspot number; however, the RMSE is still large. The best result was obtained for SOHO_SEM$_{26–34}$ as expected from the fact that the 26–34 nm wavelengths of SOHO_SEM$_{26–34}$ overlap with the wavelength range of ionization of the Earth's atmosphere. It is to be noted that the data gaps of the $S_{10.7}$ index were not filled and the corresponding data sets were excluded from both training and validation, which may have more or less the effect on the absolute values of RMSE. Mg II cwr based on the measurements at 280 nm, which is outside the wavelength range of ionization of the Earth's atmosphere. The $M_{10.7}$ index was created as a surrogate using the operational Mg II cwr in order to represent the Schumann-Runge Continuum irradiances. These irradiances dissociate O_2 and, in JB2006 thermospheric density model are used with a 5 day lag but in JB2008 are used with a 2 day lag [*Bowman et al.*, 2008b]. The energy that dissociates O_2 is felt with a time lag in higher atmosphere layers. The RMSE for $M_{10.7}$ shows intermediate results between the 10.7 cm radio flux and SOHO_SEM$_{26–34}$. To compare with the $S_{10.7}$ results, S_{EUV} index generated from SOLAR2000 that simulates the SOHO_SEM$_{26–34}$ measurements was examined and the result is shown by the plus sign associated with the dotted line in Figure 4. The periods corresponding to the data gaps of $S_{10.7}$ were excluded in this training and validation. The S_{EUV} index shows a small RMSE comparable to that of $S_{10.7}$; however, the curve for S_{EUV} is somewhat flattened compared with that for $S_{10.7}$.

[18] It has long been known that the upper atmospheric response to solar irradiance changes delays by a couple of days. Figure 4 demonstrates such a delayed response of TEC variations. The proxies are virtually divided into two groups; the sunspot number, radio flux, and SOLAR2000 (S_{EUV}) exhibit the best training achievement with a time lag of 1.5 ~ 2 days, and the proxies based on UV/EUV measurements exhibit a 1 day time lag. The apparent discrepancy in the time lag may be partly caused by the time tag of the proxy, such that the solar radio flux is observed at 20 UT, while the processed $M_{10.7}$ and $S_{10.7}$ indices for Mg II cwr and SOHO_SEM$_{26–34}$ represent values at 12 UT. The S_{EUV} index simulates $S_{10.7}$, but the time lag feature of S_{EUV} is rather close to $F_{10.7}$ than $S_{10.7}$. The S_{EUV} index, while the integrated 26–34 nm irradiance in SOLAR2000, consists of spectral irradiances that are modeled using $F_{10.7}$ and Lyman-α; the $F_{10.7}$ contribution is the reason that S_{EUV} has some similarity to $F_{10.7}$.

3.2. Combination of Means Over Longer Periods

[19] Figure 5a shows improvements of the ANN training results when several means over longer periods were included in the input space of the ANN for the four measurement-based proxies and an index generated from an empirical model. The horizontal axis represents the combination (pattern) of means over several periods. Pattern 1 is the (3 day smoothed) daily proxy. Pattern 2 is the combination of the daily proxy and 7 day backward means, i.e., two solar inputs appear in the input space of the ANN. Pattern 3 has three solar inputs: the daily proxy, 7, and 27 day backward means. Similarly, pattern 4 has four solar inputs: the daily proxy, 7, 27, and 81 day backward means. We had a general improvement of the ANN training results when the input parameter increased. However, the behavior differs from one proxy to the other. For the sunspot number and $F_{10.7}$, better results were observed toward the right-hand side of Figure 5a and the best results were for pattern 4. For Mg II cwr, general improvements were similar to the previous two proxies, except that the best result was obtained for pattern 3. In other words, adding the 81 day mean had no advantage in the ANN training. For the $S^*_{10.7}$ index (composite of SOHO_SEM$_{26–34}$ and S_{EUV} from SOLAR2000) and $E_{10.7}$, the improvements by adding longer-period means were small when compared with the other proxies, and the single daily proxy was fairly good to describe the TEC variations.

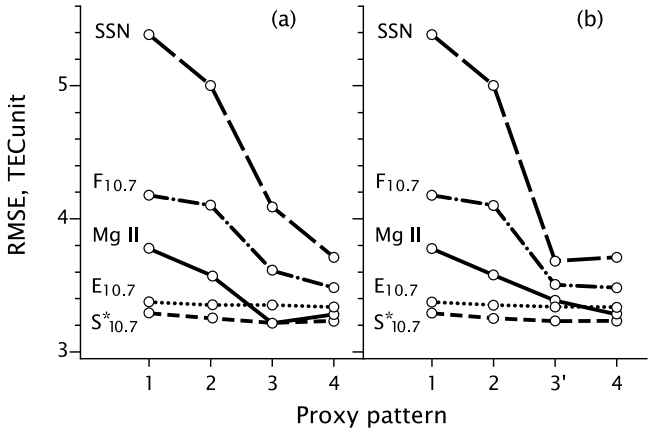

Figure 5. Improvements of the ANN training achievement by combining means over longer period. (a) Pattern 1 is 3 day smoothed daily proxy, pattern 2 is the combination of 3 day smoothed and 7 day backward means of the proxy, pattern 3 is the combination of 3 day smoothed, 7 day, and 27 day backward means of the proxy, and pattern 4 is the combination of 3 day smoothed, 7, 27, and 81 day backward means of the proxy. (b) Same as Figure 5a except that pattern 3′ is the combination of 3 day smoothed, 7 day, and 81 day backward means of the proxy.

[20] Figure 5b is the same as Figure 5a, except that pattern 3′ denotes the combination of the daily proxy, 7, and 81 day means. For the sunspot number, the 81 day mean effectively improved the ANN training result (pattern 3′) and including the 27 day mean (pattern 4) deteriorated the results. The $F_{10.7}$ index behaved similar to the sunspot number, though patterns 3′ and 4 yielded almost identical results. Comparing Figures 5a and 5b, we conclude that adding the 81 day means was essential for the improvement of ANN training results for the sunspot number and $F_{10.7}$ and adding the 27 day means was essential for the improvement of ANN training results for Mg II cwr. On the other hand, adding the longer period means had a small effect for $S^*_{10.7}$ and $E_{10.7}$, although a general improvement similar to $M_{10.7}$ was observed. Among all the proxies, $S^*_{10.7}$ was the best when daily values alone were used, but Mg II cwr yielded a comparable result when 7 and 27 day means were concurrently used.

[21] In the above experiments, all mean values were calculated for the periods backward from the day of concern. However, centered means have been more commonly used in previous empirical models. Table 1 compares centered and backward means for the sunspot number, $F_{10.7}$, Mg II cwr, and $S^*_{10.7}$, where the numbers for the pattern represent periods over which the proxy was averaged, the letters b and c denote backward and centered means, respectively, and the letter d denotes the 3 day smoothed daily proxy. The plus sign represents the concurrent use of those values in the input space of the ANN. The best results (smallest RMSEs) for each proxy are shown in bold. In Table 1, any combination of two solar parameters yielded smaller RMSE than the cases when the daily proxy alone was used. For the sunspot number and $F_{10.7}$, the 81 day means effectively improved the results, and for Mg II cwr, the 27 day mean did, which is consistent with Figure 5. However, except for the sunspot number, there were no differences in the training results between the centered and backward means. For the sunspot number, the 81 day backward mean was better than the 81 day centered mean. For $S^*_{10.7}$, as immediately expected from the results shown in Figure 5, the improvements by any combination of two solar parameters were not much. The 27 and 81 day backward means yielded the best results, and the 27 day centered mean also yielded almost an identical result. Unlike $M_{10.7}$, the 81 day centered mean scarcely improved the RMSE.

3.3. Combination of Two or More Proxies

[22] So far as a single proxy was used, we have had the best result in training ANN using daily, 7, and 27 day backward means of $M_{10.7}$ or $S^*_{10.7}$. However, in a previous work (paper 1) on constructing an ANN-TEC model based on the traditional proxies of solar activity, a combined use of the sunspot number and $F_{10.7}$ improved the model efficiency when compared with the cases in which the sunspot number or $F_{10.7}$ alone was used. Similarly, we compared a variety of combinations of proxies, with each proxy consisting of daily values and 7 and 27 day backward means. The results are summarized in Table 2. We first confirmed the previous result in paper 1 by runs 1-1, 1-2, and 2-6 in Table 2, such that the RMSE for the combination of R_i and $F_{10.7}$ (run 2-6) was smaller than that for runs 1-1 and 1-2. For $M_{10.7}$ (runs 2-1 to 2-3), the best proxy to be combined was $S^*_{10.7}$, and the combination with $F_{10.7}$ also yielded a small improvement. However, the combination with R_i deteriorated the result. For $S^*_{10.7}$ (runs 2-3 to 2-5), the best proxy to be combined was $M_{10.7}$ and the next was $F_{10.7}$, but the combinations with R_i deteriorated the result again. In conclusion, $S^*_{10.7}$ and $M_{10.7}$ were the best combination to improve the result when two different proxies are chosen.

[23] The combinations of three proxies were examined by adding a third proxy to $M_{10.7}$ and $S^*_{10.7}$ (runs 3-1 and 3-2). We noted that $F_{10.7}$ appreciably improved the result, whereas R_i exhibited no effect. Finally, all four proxies were combined in run 4-1, but the result was worse than the case in which R_i was not included (run 3-2).

4. Discussion

4.1. Delayed Response

[24] The most pertaining solar proxy or a combination of several proxies for modeling ionospheric total electron content (TEC) was examined through the training of an artificial neural network (ANN). In this study, we assumed that an input parameter (solar proxy) to the ANN that is better correlated with the target parameter (TEC) yields

Table 1. Comparison of Backward and Centered Means

	Pattern[a]				
Proxy	*d*	*d*+27*b*	*d*+27*c*	*d*+81*b*	*d*+81*c*
R_i	5.38	4.05	4.29	**3.72**	3.94
$F_{10.7}$	4.18	3.54	3.67	**3.48**	**3.48**
$M_{10.7}$	3.77	**3.26**	**3.26**	3.41	3.30
$S^*_{10.7}$	3.28	**3.23**	3.24	**3.23**	3.27

[a]See Figure 3 for the meaning of each variable.

Table 2. Combination of Different Proxies[a]

Run	Proxy				RMSE
	R_i	$F_{10.7}$	$S^*_{10.7}$	$M_{10.7}$	
1-1	✓				4.09
1-2		✓			3.61
1-3			✓		3.22
1-4				✓	3.21
2-1	✓			✓	3.23
2-2		✓		✓	3.20
2-3			✓	✓	3.16
2-4	✓		✓		3.24
2-5		✓	✓		3.17
2-6	✓	✓			3.52
3-1	✓		✓	✓	3.16
3-2		✓	✓	✓	**3.12**
4-1	✓	✓	✓	✓	3.16

[a]Each proxy consists of 3 day smoothed and 7 and 27 day backward mean values.

smaller RMSE in the training. By analyzing geomagnetic storm responses to solar wind parameters, *Wu and Lundstedt* [1997] showed that both statistical cross-correlation analyses and neural networks yield consistent results. The delayed response of TEC to the solar input, as depicted in Figure 4, can be compared with correlation analyses. The delayed response of the ionosphere/thermosphere to solar irradiance changes has long been known. *Jacchia et al.* [1973] and *Paul et al.* [1974] have found a delay in thermospheric density variations between one and two days after the 27 day variation of $F_{10.7}$ using satellite drag data. *Hedin* [1984] confirmed that the 1 day lag of $F_{10.7}$ was slightly better than the case of no lag or 2 day lag in the N_2 density and temperature modeling. A series of MSIS [*Hedin et al.*, 1977] also uses $F_{10.7}$ on the previous day. *Bowman et al.* [2008a] determined that a time lag of one day was the best to use for $F_{10.7}$ and $S_{10.7}$, while a time lag of five days was the best for $M_{10.7}$, in their thermospheric density model. Similarly, an ionospheric response to solar irradiance changes has been reported [*Jakowski et al.*, 2002; *Afraimovich et al.*, 2008; *Min et al.*, 2009]. *Min et al.* [2009] also demonstrated the correlation amplitudes for neutral density and TEC, in which the diagram (their Figure 6) is quite similar to Figure 4, if it is rotated by 180°. The time lag characteristics of solar inputs to obtain better training results of the ANN shown in Figure 4 are consistent with the previous works. This, in turn, means that the ANN training achievement measured by RMSEs works as well as the commonly used correlation analysis.

[25] Although we have a general consensus of the delayed response of the ionosphere/thermosphere to solar irradiance changes, there were some discrepancies in the amount of delay, not only between the current study and the previous ones but also among the proxies examined in this study. For quantitative discussion, we must consider the time tag (T) of the proxies and TEC data. The 10.7 cm radio flux is measured at 20 UT, SOHO_SEM$_{26–34}$ data are obtained every 5 min, and the daily average is used for the derivation of $S_{10.7}$ (release note on the JB2008 indices file, http://sol.spacenvironment.net); Mg II cwr data are obtained twice daily at 07 and 16 UT (above release note). For the sunspot number, data are gathered from a worldwide network and R_i is statistically generated; however, the time tag is not explicitly described. TEC data are obtained at the longitudes of Japan at around 135°E, and the ionization rate peaks at around local noon or at around 03 UT.

[26] Considering the above time tags, we anticipate an apparent time lag for $F_{10.7}$ (T = 20 UT) longer by 1/3 day (8 hr) than that for $S_{10.7}$ or $M_{10.7}$ (T = 12 UT). However, Figure 4 exhibits an almost a day longer time lag for $F_{10.7}$ when compared with $M_{10.7}$ and $S_{10.7}$. The time lag for R_i is slightly shorter than that for $F_{10.7}$, but still longer than $M_{10.7}$ and $S_{10.7}$. There could be two possible explanations for the time lag: One is the ionosphere/thermosphere response or an integration effect of solar irradiance changes. The other is an earlier evolution of activity region measured at that wavelength when compared with the activity evolution at the wavelengths of IE-EUV. Because the wavelengths of $S_{10.7}$ (26–34 nm) overlap with the wavelengths of a large ionization cross section of the Earth's upper atmosphere, the time lag for $S_{10.7}$ is most probably attributed to the delayed response of the ionosphere/thermosphere system. One possible explanation for that $F_{10.7}$ may have a slightly longer time lag is the fact that it more represents cool corona X-rays than Mg II. The X-rays deposit their energy in the mesosphere/lower thermosphere and the time lag there is 5 days to see its energy pulse at higher altitudes, e.g., behavior of the $Y_{10.7}$ index in JB2008 [*Bowman et al.*, 2008b]. Another explanation for the excess time lag of about one day for R_i and $F_{10.7}$ is the different behavior of the activity region evolution measured by the sunspot number or 10.7 cm radio flux and the activity of IE-EUV. *Donnelly et al.* [1983] and *Floyd et al.* [2005] described the episodes of activity in which the sunspot number and 10.7 cm flux peak occurred earlier than the UV/EUV flux. The cross correlation analysis shows that the 205 nm UV flux during a peak in a series of 27 day peaks is partially related to the values of $F_{10.7}$ or the sunspot number that occurred on the previous rotation [*Donnelly et al.*, 1985]. This asymmetric cross correlation function about the zero time lag between the UV flux and $F_{10.7}$ is ascribed to the longer lifetimes on the solar disc of plage regions relative to the sunspots [*Donnelly et al.*, 1985; *Lean and Repoff*, 1987]. The different behaviors of activity evolutions among the different wavelengths discussed above are over the time scales of solar rotations. On the other hand, the 1 day excess time lag that appeared in the TEC response was derived from the solar rotation modulation of localized activity regions. Thus for quantitative discussion, we require more careful examination of the data in future.

4.2. Improvements Combining Short- and Long-Term Variations

[27] *Hedin* [1984] had found that the slopes of the correlation of EUV with $F_{10.7}$ for the short- and long-term variations are not the same. *Donnelly et al.* [1986] pointed out that the amplitude of long-term variations relative to short-term variations varies from one wavelength to another. Because of these facts, the long-term variations of IE-EUV cannot simply be scaled from short-term variations (and vice versa) using a solar proxy [*Woods et al.*, 2000; *Woods*, 2008]. To compensate the different features of short- and long-term variability in atmospheric models such as a series

of MSIS [*Hedin et al.*, 1977], two-variable (daily and 81 day average) representations of the EUV effects have been used rather than one-variable representations of them. *Bowman et al.* [2008a] compared 54, 81, and 108 day (2, 3, and 4 solar rotations) centered averages for $F_{10.7}$, $M_{10.7}$, and $S_{10.7}$ in thermosphere temperature fitting using the daily and long-term average of solar proxy, and determine that 81 day centered averages are the best long-term average to use.

[28] The current ANN study confirmed the improvements by combining short- and long-term means of solar proxy as summarized in Figure 5. Table 1 compares the effect of the 27 and 81 day means and shows that their effectiveness was not exactly the same from one proxy to the other. A better ANN training result or smaller RMSE was attained both for the sunspot number and the 10.7 cm radio flux by the 81 day mean when compared with the 27 day mean. For the sunspot number, the 81 day backward mean was better than the 81 day centered mean. However, there was no difference between the centered and backward 81 day periods for $F_{10.7}$. On the other hand, the 27 day mean was a better parameter than the 81 day mean for Mg II cwr without difference between the 27 day centered and backward means.

[29] Solar UV irradiance variations are parameterized by a longitudinally dispersed active network and a localized plage area: the short-term variation is due to plages modulated by solar rotation and the long-term variation is due to variations in both plages and active network [*Lean and Skumanich*, 1983]. *Woods et al.* [2000] describe that the active network contribution to the irradiance variability is larger for the transition region emissions than for the chromospheric emissions. Thus, the amplitude of the long-term variations relative to the short-term variations varies from one proxy to other representing different solar atmosphere. Along the evolution of active regions, a plage decays into longitudinally dispersed active network in one to three solar rotations and the active network remains for several solar rotations [*Woods et al.*, 2000]. *Donnelly et al.* [1985] and *Lean and Repoff* [1987] analyzed the power spectra of the temporal variation of the sunspot number, 10.7 cm flux, and UV irradiance and found that both the sunspot number and $F_{10.7}$ have more power at the periods longer than the 27 day peak when compared with the UV irradiance. To compensate the difference between short-term and long-term variations in IE-EUV vs. proxies, the 27 day mean of active regions likely is sufficient to represent long-lasting active network for UV/EUV irradiance; however, 81 days are required to represent it for the more variable sunspot number and 10.7 cm flux.

[30] In Figure 5, a small improvement in RMSE was observed when long-term means were added for the case of SOHO_SEM$_{26–34}$ and the $E_{10.7}$ generated from the SOLAR2000 irradiance model. The wavelengths of these indices overlap with IE-EUV, and the daily proxy values fairly well represent IE-EUV during all phases of long-term changes. In other words, IE-EUV has the similar short-term and long-term variations to SOHO_SEM$_{26–34}$ and the $E_{10.7}$ index.

4.3. Improvements Combining Different Proxies

[31] The ionospheric effective EUV is dominated by emission lines that have components in the chromosphere, transition region, and corona, and their contribution to short-term and long-term variations is different. The emissions throughout the chromosphere, transition region, and corona vary differently mainly because their contrasts for active network and plage components are different [*Woods et al.*, 2000]. Thus, for the better description of the IE-EUV variability, multiple proxies that represent different source regions are essential, as suggested by *Woods et al.* [2000] during the discussion of solar irradiance proxy models. The JB2008 thermospheric density model uses the 10 cm radio flux, SOHO_SEM$_{26–34}$, and Mg II cwr as solar proxies [*Bowman et al.*, 2008a]. The SOLAR2000 empirical irradiance model uses Lyman-α to represent chromospheric emissions and $F_{10.7}$ to represent coronal emissions [*Tobiska et al.*, 2000].

[32] The variation of the chromospheric emissions shows the greatest similarity to the variation of Mg II cwr [e.g., *Woods et al.*, 2000], and Mg II cwr is considered to be a good proxy for the chromospheric activity. The major contribution to SOHO_SEM$_{26–34}$ is the transition region He II line at 30.4 nm with the secondary contribution of coronal Fe XV line at 28.4 nm. The 10.7 cm radio flux originates primarily in the transition region and corona [*Lean*, 1987], while *Tapping* [1987] describes that the 10.7 cm flux is probably neither purely chromospheric or coronal in origin, and the relative importance of the chromopsheric and coronal contribution is likely to vary from source to source.

[33] We do not intend to discuss the source regions of the proxies in detail. Rather, *Tobiska et al.* [2008] note that it is desirable to have solar indices and proxies that vary differently in time if one wants to capture the greatest range of energies from solar temperature regions. From this view point, we observed a different behavior between the proxies about improvements in the ANN training by a concurrent use of long-term means. The 81 day mean was the best long-term mean to compensate for the difference between short-term and long-term variations in the 10.7 cm radio flux and IE-EUV, the 27 day mean was the best in the case of Mg II cwr, and the long term means did not show significant improvements for SOHO$_{26–34}$. In our ANN model, the best training result was obtained when the above three proxies were concurrently used, as shown in Table 2, which is consistent with the general comment given by *Tobiska et al.* [2008].

4.4. Role of 7 Day Means

[34] Delayed responses of Earth's upper atmospheric parameters to solar irradiance changes have long been known [*Jacchia et al.*, 1973; *Paul et al.*, 1974; *Jakowski et al.*, 1991, 2002; *Rich et al.*, 2003; *Bowman et al.*, 2008a; *Afraimovich et al.*, 2008; *Min et al.*, 2009] and they were confirmed by the ANN approach in this study, as shown in Figure 4, although the apparent delay time differs with the proxy. The delayed response is mostly caused by an accumulated or integrated effect of the solar input that is modulated with a period of ~27 days. If a short-term mean over a period shorter than one solar rotation (but longer than 3 days) backward (7 days in this study) is included in the input space of ANN, the ANN is expected to learn this effect. The second points from the left in Figures 5a and 5b may include the integration effect. However, the 7 day smoothed temporal variation curves are similar to the original daily proxy curve, but the amplitude of the 27 day modulation is reduced. If the

Table 3. Role of Short-Term Mean

Run	Proxy Pattern[a]					RMSE
	d	4*p*	7*b*	7*c*	27*b*	
3-2	✓		✓		✓	3.12
5-1	✓	✓			✓	3.12
5-2	✓				✓	3.19
5-3	✓			✓	✓	3.14
5-4	✓	✓		✓	✓	3.12

[a]Proxy patterns are common to $F_{10.7}$, $S^*_{10.7}$, and $M_{10.7}$.

amplitude of the 27 day variation for a given proxy is larger than that for IE-EUV, the 7 day smoothed curve with the reduced amplitude may result in an apparent improvement of the ANN training irrespective of the integration effect.

[35] To avoid the above uncertainty, we performed an additional experiment, in which 7 day backward means were replaced with the means over the 4 preceding days to the 3 day smoothed daily proxy, but the other inputs were the same as run 3-2. The new run (run 5-1) yielded an identical result with run 3-2, as shown in Table 3, and both the 7 day backward and 4 preceding day means express the integration effect and improved the training results as compared with the run without the short-term mean (run 5-2).

[36] Another consideration regarding the short-term variation less than one solar rotation is limb darkening/brightening. *Donnelly et al.* [1986] discussed the dependence of solar active region radiation on the region's solar central meridian distance (CMD). The center-to-limb behavior differs with wavelength [*Donnelly et al.*, 1986; *Crane et al.*, 2004; *Woods et al.*, 2005], and the flux of optically thick lines has narrow CMD dependence. If a proxy is based on optically thick lines, the actual IE-EUV flux is larger than that predicted from that proxy when the active region is close to the limb. Thus, the information on CMD of the active region is expected to improve the ANN model. If the short-term variation in irradiance is caused by the solar rotation alone, the 7 day backward means are smaller than the daily proxy value when the active region is near the east limb and they are reversed when the active region is near the west limb. Therefore, the limb darkening/brightening effect might be canceled after the training of the ANN with the entire data set is completed. In an additional experiment (run 5-3), 7 day centered means were used instead of 7 day backward means. In this pattern, when the active region is near the central meridian, 7 day centered means are smaller than the 3 day smoothed daily proxy, and when the active region is on the other side, 7 day centered means are larger than the 3 day smoothed daily proxy. Thus using 7 day centered means, the center to limb information can be put to the ANN (run 5-3). The result was improved, as expected, when compared with the run without short-term means (run 5-2), as shown in Table 3. In the final experiment, we added both 4 day preceding (short-term trend), and 7 day centered means (center to limb characteristics) in the input space. The result is shown at the bottom line of Table 3 (run 5-4). Although improvement was observed when compared with run 5-3, RMSE was no better than the case in which 7 day centered means were excluded (run 5-1). Thus, we did not have any conclusive evidence of the compensation of center to limb characteristics by the comparison of RMSEs of ANN training. For the discussion of subtle effects such as the short-term variation less than the solar rotation period, we need to eliminate nonsolar origin factors that affect TEC variations as described in section 4.5, because RMSE values are likely to be saturated.

4.5. Comparison of TECs on an Individual Day

[37] In Figure 6a, the observed (3 day smoothed values) and ANN predicted TECs at 03 UT (12 LT) at the latitude of 35°N are compared for 30 solar rotations or approximately 2 years and 3 months around the solar maximum. The solar inputs are 3 day smoothed, 7, and 27 day backward means of $M_{10.7}$, $S^*_{10.7}$, and $F_{10.7}$, i.e., the combination that yielded the best results in Table 2. Figure 6a (top) is the observed and ANN predicted TECs and their difference (ΔTEC). Figure 6a (middle) is the three solar proxies (3 day smoothed values only), and Figure 6a (bottom) is the daily geomagnetic activity index, Ap. The seasonal trend of summer minimum and the 27 day modulations of TEC are generally reproduced. However, relatively large errors related or unrelated with magnetic disturbances are noticed occasionally. The largest magnetic disturbance was on 31 March 2001, and several moderate to weak disturbances occurred before and after it within one month. A weak positive ionospheric disturbance occurred in association with the largest storm, and several intense negative ionospheric storms occurred associated with the other magnetic disturbances [*Maruyama and Nakamura*, 2007]. These ionospheric storms are recognized in the ΔTEC curve, and the ANN prediction errors are ascribed to the magnetic activity. The second largest magnetic disturbance on 6 November 2001 caused a strong TEC enhancement [*Maruyama et al.*, 2004], and preceding moderate to weak magnetic disturbances also caused positive ΔTECs. However, from 7 to 24 November, the persistent negative ΔTEC were not related with magnetic activities. The TEC response to the magnetic storm on 24 November (third largest one) is not quite clear. The large positive and negative ΔTECs in January and February 2002 are not related with magnetic activities.

[38] Figure 6b shows the next 30 solar rotations for a moderate solar activity period. The magnetic disturbance from 29 to 31 October 2003 caused a negative ΔTEC disturbance. This is the famous Halloween super storm, and during such severe disturbances, ionospheric heights are often significantly increased. Thus, the assumption of a constant value of the thin-shell approximated ionospheric height may cause an error in evaluating vertical TECs, depending on the geometry of the radio propagation path. However, the negative ΔTECs around 20 November 2003 persisted for more than 10 days even after the magnetic storm and it was not simply connected with the magnetic activity. The strong positive ΔTECs at the end of December 2003 are also difficult to relate with magnetic activities. For the second largest magnetic activity on 27 July 2004, the TEC response was not so significant. On 8 November 2004, the TEC response was a large positive storm and the storm enhanced density (SED) was followed after sunset [*Maruyama*, 2006], although it is smoothed in the ΔTEC plots.

[39] As examined above, even for large magnetic disturbances the TEC errors were comparable with those not related with magnetic activities. Further, there are quasi-periodic increases in errors in Figure 6, which are difficult

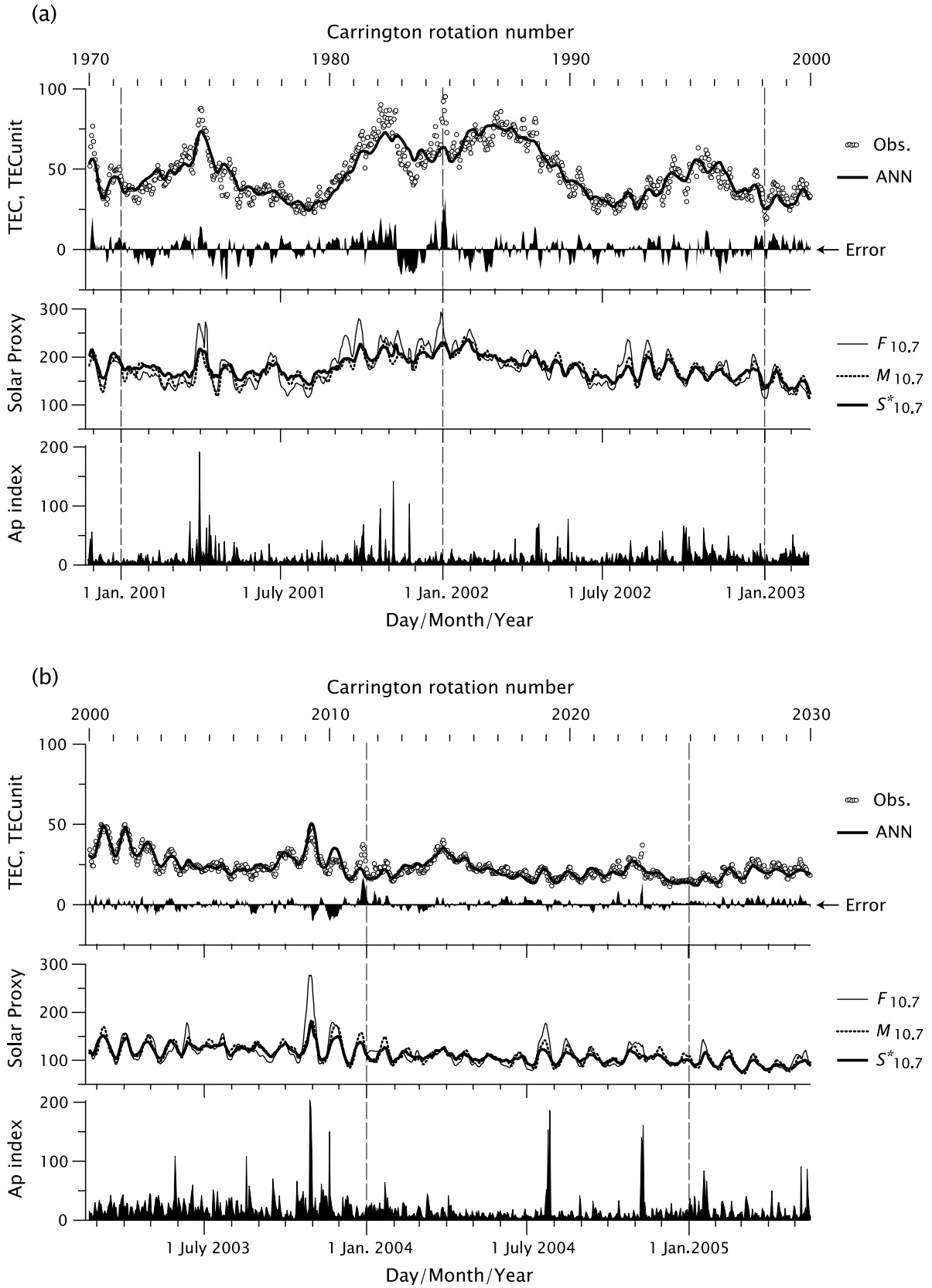

Figure 6. (a) (top) Observed and ANN modeled daily TEC for 30 solar rotations (approximately 2 years and 3 months) and their difference, (middle) 3 day smoothed $F_{10.7}$, $M_{10.7}$, and $S^*_{10.7}$, and (bottom) planetary Ap index for the solar maximum. (b) Same as Figure 6a but for the moderate to low solar activity period.

to distinguish whether they are related with magnetic activities or not. The TEC errors unrelated with magnetic activities should be ascribed to other origin such as forcing below the ionosphere including coupling with planetary wave activities [e.g., *Laštovička*, 2006; *Borries et al.*, 2007], although an incomplete representation of solar inputs by the proxies cannot be ruled out. From the other point of view, the proxy model benefits to separate solar effects from the ionospheric variations or to provide a background state for ionospheric storm studies. Such applications are for future issues.

5. Conclusions

[40] By using the artificial neural network (ANN) mapping technique, variations in ionospheric total electron content (TEC) in response to solar activity changes were modeled. Various proxies and their means over several periods were used to represent the solar activity. After the ANN training was completed, root mean square errors between the observed and the modeled TEC were compared to find the most pertaining proxies for the TEC prediction. The proxies examined are the international sunspot number (R_i), the 10.7 cm solar radio flux ($F_{10.7}$), the Mg II core-to-wing ratio ($M_{10.7}$), and the integrated 26–34 nm emissions by SOHO_SEM ($S_{10.7}$). For the latter two indices, measurements are preprocessed and normalized to the variation of $F_{10.7}$ by linear regression. Data gaps of SOHO_SEM$_{26–34}$ were filled by the output (S_{EUV}) of an empirical irradiance model SOLAR2000 using linear regression ($S^*_{10.7}$). Further, two outputs from SOLAR2000, $E_{10.7}$ and S_{EUV} were compared. The major points we observed are as follows.

[41] 1. When a single daily proxy alone was used, proxies from poor to good were in the order of R_i, $F_{10.7}$, $M_{10.7}$, $E_{10.7}$, and $S^*_{10.7}$.

[42] 2. The delayed response of TEC to the solar activity was recognized for all the proxies and its amount was approximately two days for R_i, $F_{10.7}$, and S_{EUV}, and approximately one day for $M_{10.7}$ and $S^*_{10.7}$.

[43] 3. The combination of long-term means improved the modeling for all the proxies. The best long-term mean to use was 81 day backward for R_i and $F_{10.7}$, and 27 day backward for $M_{10.7}$ and $S^*_{10.7}$.

[44] 4. The combination of different proxies yielded further improvement. The best combination was 3 day smoothed daily values and 7 and 27 day backward means of $F_{10.7}$, $M_{10.7}$, and $S^*_{10.7}$; 9 solar inputs in total.

[45] These results were interpreted on the basis of a number of published works on the solar irradiance behavior.

[46] 1. The relatively longer delay time of the TEC response to R_i and $F_{10.7}$ variations is ascribed to the relatively rapid growth and decay of the active region corresponding to those indices compared with that measured by UV/EUV emissions.

[47] 2. Different lengths of the period (27 days or 81 days), over which mean values are calculated, are effective to compensate the different features of short- and long-term variability between the proxy and IE-EUV. These different lengths were attributed to the different feature of power spectra of the temporal variation of the proxies; both the sunspot number and $F_{10.7}$ have more power at the periods longer than the 27 day peak when compared with the UV irradiance.

[48] 3. The combination of 3 proxies, $M_{10.7}$, $S^*_{10.7}$, and $F_{10.7}$ specify well the behavior of IE-EUV because those proxies exhibit temporally different behaviors and thus represent activities in different solar atmosphere regions.

[49] **Acknowledgments.** The international sunspot number is supplied by the Solar Influences Data Analysis Center, WDC for the Sunspot Index, Royal Observatory of Belgium. The 10.7 cm solar radio flux data are observed by the Dominion Radio Astrophysical Observatory, Canada, and these data are supplied by the National Geophysical Data Center. The Solar Irradiance Platform (SIP) provides $M_{10.7}$, $S_{10.7}$, S_{EUV}, and $E_{10.7}$, and they are downloaded at the SIP quick link http://spacewx.com.

[50] Zuyin Pu thanks W. Kent Tobiska and Manuel Hernandez-Pajares for their assistance in evaluating this paper.

References

Afraimovich, E. L., E. I. Astafyeva, A. V. Oinats, Y. V. Yasukevich, and I. V. Zhivetiev (2008), Global electron content: A new conception to track solar activity, *Ann. Geophys.*, *26*, 335–344.

Bilitza, D., and B. W. Reinisch (2008), International Reference Ionosphere 2007: Improvements and new parameters, *Adv. Space Res.*, *42*, 599–609.

Borries, C., N. Jakowski, C. Jacobi, P. Hoffmann, and A. Pogoreltsev (2007), Spectral analysis of planetary waves seen in ionospheric total electron content (TEC): First results using GPS differential TEC and stratospheric reanalyses, *J. Atmos. Sol. Terr. Phys.*, *69*, 2442–2451.

Bowman, B. R., W. K. Tobiska, F. A. Marcos, and C. Valladares (2008a), The JB2006 empirical thermospheric density model, *J. Atmos. Sol. Terr. Phys.*, *70*, 774–793.

Bowman, B. R., W. K. Tobiska, F. A. Marcos, C. Y. Huang, C. S. Lin, and W. J. Burke (2008b), A new empirical thermospheric density model JB2008 using new solar and geomagnetic indices, paper AIAA 2008-6438 presented at AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Am. Inst. of Aeronaut. and Astronaut., Honolulu.

Cebula, R. P., M. T. DeLand, and B. M. Schlesinger (1992), Estimates of solar variability using the solar backscatter ultraviolet (SBUV) 2 Mg II index from the NOAA 9 satellite, *J. Geophys. Res.*, *97*(D11), 11,613–11,620.

Crane, P. C., L. E. Floyd, J. W. Cook, L. C. Herring, E. H. Avrett, and D. K. Prinz (2004), The center-to-limb behavior of solar active regions at ultraviolet wavelengths, *Astron. Astrophys.*, *419*, 735–746, doi:10.1051/0004-6361:20040012.

Donnelly, R. F., D. F. Heath, J. L. Lean, and G. J. Rottman (1983), Differences in the temporal variations of solar UV flux, 10.7-cm solar radio flux, sunspot number, and Ca-K plage data caused by solar rotation and active region evolution, *J. Geophys. Res.*, *88*(A12), 9883–9888.

Donnelly, R. F., J. W. Harvey, D. F. Heath, and T. P. Repoff (1985), Temporal characteristics of the solar UV flux and He I line at 1083 nm, *J. Geophys. Res.*, *90*(A7), 6267–6273.

Donnelly, R. F., H. E. Hinteregger, and D. F. Heath (1986), Temporal variations of solar EUV, UV, and 10,830-Å radiations, *J. Geophys. Res.*, *91*(A5), 5567–5578.

Emmert, J. T., J. L. Lean, and J. M. Picone (2009), Comment on "Oscillations of global mean TEC" by K. Hocke, *J. Geophys. Res.*, *114*, A01309, doi:10.1029/2008JA013679.

Floyd, L., J. Newmark, J. Cook, L. Herring, and D. McMullin (2005), Solar EUV and UV spectral irradiances and solar indices, *J. Atmos. Sol. Terr. Phys.*, *67*, 3–15.

Haykin, S. (1994), *Neural Networks: A Comprehensive Foundation*, Macmillan, New York.

Heath, D. F., and B. M. Schlesinger (1986), The Mg 280-nm doublet as a monitor of changes in solar ultraviolet irradiance, *J. Geophys. Res.*, *91*(D8), 8672–8682.

Hedin, A. E. (1984), Correlations between thermospheric density and temperature, solar EUV flux, and 10.7-cm flux variations, *J. Geophys. Res.*, *89*(A11), 9828–9834.

Hedin, A. E., et al. (1977), A global thermospheric model based on mass spectrometer and incoherent scatter data, MSIS: 1. N_2 density and temperature, *J. Geophys. Res.*, *82*, 2139–2147.

Jacchia, L. G., J. W. Slowey, and I. G. Campbell (1973), An analysis of the solar-activity effects in the upper atmosphere, *Planet. Space Sci.*, *21*, 1835–1842.

Jakowski, N., B. Fichtelmann, and A. Jungstand (1991), Solar activity control of ionospheric and thermospheric processes, *J. Atmos. Terr. Phys.*, *53*, 1125–1130.

Jakowski, N., S. Heise, A. Wehrenpfennig, S. Schlüter, and R. Reimer (2002), GPS/GLONASS-based TEC measurements as a contributor for space weather forecast, *J. Atmos. Sol. Terr. Phys.*, *64*, 729–735.

Laštovička, J. (2006), Forcing of the ionosphere by waves from below, *J. Atmos. Sol. Terr. Phys.*, *68*, 479–497.

Lean, J. (1987), Solar ultraviolet irradiance variations: A review, *J. Geophys. Res.*, *92*(D1), 839–868.

Lean, J. L., and T. P. Repoff (1987), A statistical analysis of solar flux variations over time scales of solar rotation: 1978–1982, *J. Geophys. Res.*, *92*(D5), 5555–5563.

Lean, J. L., and A. Skumanich (1983), Variability of the Lyman alpha flux with solar activity, *J. Geophys. Res.*, *88*(A7), 5751–5759.

Lean, J., M. VanHoosier, G. Brueckner, D. Prinz, L. Floyd, and K. Edlow (1992), SUSIM/UARS observations of the 120 to 300 nm flux variations during the maximum of the solar cycle: Inferences for the 11-year cycle, *Geophys. Res. Lett.*, *19*, 2203–2206.

Lean, J. L., O. R. White, W. C. Livingston, and J. M. Picone (2001), Variability of a composite chromospheric irradiance index during the 11-year activity cycle and over longer time periods, *J. Geophys. Res.*, *106*(A6), 10,645–10,658.

Ma, G., and T. Maruyama (2003), Derivation of TEC and estimation of instrumental biases from GEONET in Japan, *Ann. Geophys.*, *21*, 2083–2093.

Maruyama, T. (2006), Extreme enhancement in total electron content after sunset on 8 November 2004 and its connection with storm enhanced density, *Geophys. Res. Lett.*, *33*, L20111, doi:10.1029/2006GL027367.

Maruyama, T. (2007), Regional reference total electron content model over Japan based on neural network mapping techniques, *Ann. Geophys.*, *25*, 2609–2614.

Maruyama, T., and M. Nakamura (2007), Conditions for intense ionospheric storms expanding to lower midlatitudes, *J. Geophys. Res.*, *112*, A05310, doi:10.1029/2006JA012226.

Maruyama, T., G. Ma, and M. Nakamura (2004), Signature of TEC storm on 6 November 2001 derived from dense GPS receiver network and ionosonde chain over Japan, *J. Geophys. Res.*, *109*, A10302, doi:10.1029/2004JA010451.

McKinnell, L.-A., and A. W. V. Poole (2004), Predicting the ionospheric *F* layer using neural networks, *J. Geophys. Res.*, *109*, A08308, doi:10.1029/2004JA010445.

Min, K., J. Park, H. Kim, V. Kim, H. Kil, J. Lee, S. Rentz, H. Lühr, and L. Paxton (2009), The 27-day modulation of the low-latitude ionosphere during a solar maximum, *J. Geophys. Res.*, *114*, A04317, doi:10.1029/2008JA013881.

Nakamura, M., T. Maruyama, and Y. Shidama (2007), Using a neural network to make operational forecasts of ionospheric variations and storms at Kokubunji, Japan, *Earth Planets Space*, *59*, 1231–1239.

Oyeyemi, E. O., A. W. V. Poole, and L. A. McKinnell (2005), On the global model for f_oF_2 using neural networks, *Radio Sci.*, *40*, RS6011, doi:10.1029/2004RS003223.

Paul, G., H. Volland, and M. Roemer (1974), A study of the time lag between the 27-day variations of thermospheric density and 10.7 cm solar radiation, *Space Res.*, *14*, 189–193.

Picone, J. M., A. E. Hedin, and D. P. Drob (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *J. Geophys. Res.*, *107*(A12), 1468, doi:10.1029/2002JA009430.

Rich, F. J., P. J. Sultan, and W. J. Burke (2003), The 27-day variations of plasma densities and temperatures in the topside ionosphere, *J. Geophys. Res.*, *108*(A7), 1297, doi:10.1029/2002JA009731.

Richards, P. G., J. A. Fennelly, and D. G. Torr (1994), EUVAC: A solar EUV flux model for aeronomic calculations, *J. Geophys. Res.*, *99*(A5), 8981–8992.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1986), Learning representations by back-propagating errors, *Nature*, *323*(9), 533–536.

Stolarski, R. S., P. B. Hays, and R. G. Roble (1975), Atmospheric heating by solar EUV radiation, *J. Geophys. Res.*, *80*, 2266–2276.

Tapping, K. F. (1987), Recent solar radio astronomy at centimeter wavelengths: The temporal variability of the 10.7-cm flux, *J. Geophys. Res.*, *92*(D1), 829–838.

Thuillier, G., and S. Bruinsma (2001), The Mg II index for upper atmosphere modelling, *Ann. Geophys.*, *19*, 219–228.

Tobiska, W. K., T. Woods, F. Eparvier, R. Viereck, L. Floyd, D. Bouwer, G. Rottman, and O. R. White (2000), The SOLAR2000 empirical solar irradiance model and forecast tool, *J. Atmos. Sol. Terr. Phys.*, *62*, 1233–1250.

Tobiska, W. K., S. D. Bouwer, and B. R. Bowman (2008), The development of new solar indices for use in thermospheric density modeling, *J. Atmos. Sol. Terr. Phys.*, *70*, 803–819.

Viereck, R., L. Puga, D. McMullin, D. Judge, M. Weber, and W. K. Tobiska (2001), The Mg II index: A proxy for solar EUV, *Geophys. Res. Lett.*, *28*, 1343–1346.

Woods, T. N. (2008), Recent advances in observations and modeling of the solar ultraviolet and X-ray spectral irradiance, *Adv. Space Res.*, *42*, 895–902.

Woods, T. N., W. K. Tobiska, G. J. Rottman, and J. R. Worden (2000), Improved solar Lyman α irradiance modeling from 1947 through 1999 based on UARS observations, *J. Geopys. Res.*, *105*(A12), 27,195–27,215.

Woods, T. N., F. G. Eparvier, S. M. Bailey, P. C. Chamberlin, J. Lean, G. J. Rottman, S. C. Solomon, W. K. Tobiska, and D. L. Woodraska (2005), Solar EUV Experiment (SEE): Mission overview and first results, *J. Geophys. Res.*, *110*, A01312, doi:10.1029/2004JA010765.

Wu, J.-G., and H. Lundstedt (1997), Geomagnetic storm predictions from solar wind data with the use of dynamic neural networks, *J. Geophys. Res.*, *102*(A7), 14,255–14,268.

T. Maruyama, National Institute of Information and Communications Technology, 2-1 Nukuikita 4-chome, Koganei, Tokyo, Japan. (tmaru@nict.go.jp)