

# Modeling ionospheric $f_oF2$ by using empirical orthogonal function analysis

E. A.<sup>1,2</sup>, D.-H. Zhang<sup>1,3</sup>, Z. Xiao<sup>1,3</sup>, Y.-Q. Hao<sup>1</sup>, A. J. Ridley<sup>2</sup>, and M. Moldwin<sup>2</sup>

<sup>1</sup>Institute of Space Physics and Applied Technology, School of Earth and Space Sciences, Peking University, Beijing, China

<sup>2</sup>Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, MI, USA

<sup>3</sup>State Key Laboratory of Space Weather, Chinese Academy of Sciences, Beijing, China

Received: 12 April 2011 – Revised: 11 August 2011 – Accepted: 26 August 2011 – Published: 31 August 2011

**Abstract.** A similar-parameters interpolation method and an empirical orthogonal function analysis are used to construct empirical models for the ionospheric  $f_oF2$  by using the observational data from three ground-based ionosonde stations in Japan which are Wakkanai (Geographic 45.4° N, 141.7° E), Kokubunji (Geographic 35.7° N, 140.1° E) and Yamagawa (Geographic 31.2° N, 130.6° E) during the years of 1971–1987. The impact of different drivers towards ionospheric  $f_oF2$  can be well indicated by choosing appropriate proxies. It is shown that the missing data of original  $f_oF2$  can be optimal refilled using similar-parameters method. The characteristics of base functions and associated coefficients of EOF model are analyzed. The diurnal variation of base functions can reflect the essential nature of ionospheric  $f_oF2$  while the coefficients represent the long-term alteration tendency. The 1st order EOF coefficient  $A_1$  can reflect the feature of the components with solar cycle variation.  $A_1$  also contains an evident semi-annual variation component as well as a relatively weak annual fluctuation component. Both of which are not so obvious as the solar cycle variation. The 2nd order coefficient  $A_2$  contains mainly annual variation components. The 3rd order coefficient  $A_3$  and 4th order coefficient  $A_4$  contain both annual and semi-annual variation components. The seasonal variation, solar rotation oscillation and the small-scale irregularities are also included in the 4th order coefficient  $A_4$ . The amplitude range and developing tendency of all these coefficients depend on the level of solar activity and geomagnetic activity. The reliability and validity of EOF model are verified by comparison with observational data and with International Reference Ionosphere (IRI). The agreement between observations and EOF model is quite well, indicating that the EOF model can reflect the major changes and the temporal distribution characteristics

of the mid-latitude ionosphere of the Sea of Japan region. The error analysis processes imply that there are seasonal anomaly and semi-annual asymmetry phenomena which are consistent with pre-existing ionosphere theory.

**Keywords.** Ionosphere (Modeling and forecasting)

## 1 Introduction

Different types of variability in ionosphere are subject to a number of interconnecting drivers which can be broadly characterized as follows: (a) solar ionizing radiation; (b) geomagnetic activity; and (c) meteorological influences (e.g., Forbes et al., 2000; Rishbeth and Mendillo, 2001; Lei et al., 2008b; Zhang et al., 2011). The ability to model and eventually anticipate the solar cycle, annual, semi-annual and seasonal variations as well as irregularities in ionosphere is of great use for both ionospheric research and space weather applications (Tóth et al., 2005). An ionospheric model can be either a first-principles-based physics model which is developed from a rigorous mathematical analysis of laws of physics and based on numerical solution of the spatial-temporal equations, or an empirical model which refers to any kind of modeling based on empirical observations. The empirical ionospheric model, which usually describes the spatial and temporal variation of electron density, critical frequency, electron temperature and other parameters of ionosphere in the form of various types of functions (e.g., harmonic function, Chapman function), played an important part in extensive practical applications.

Among different empirical models, International Reference Ionosphere (IRI) model (Bilitza et al., 1990; Bilitza, 2001) is the most widely used one. The electron density profile given by IRI is described by special anchor points of characteristic ionospheric parameters including the F2, F1 and E layer peak densities which depend on the critical



Correspondence to: D.-H. Zhang  
(zhangdh@pku.edu.cn)

**Table 1.** Geographic and geomagnetic positions of the ionosondes in Japan.

Station name	Geographic latitude (° N)	Geographic longitude (° E)	Geomagnetic latitude (° N)
Wakkanai	45.4	141.7	35.4
Kokubunji	35.7	139.5	25.6
Yamagawa	31.2	130.6	20.4

frequencies. The relationship of peak density  $NmF2$  (unit:  $m^{-3}$ ) and  $foF2$  (unit: MHz) is as follows.

$$NmF2 = 1.24 \times 10^{10} \times (foF2)^2 \quad (1)$$

Therefore, the F2 layer critical frequency  $foF2$  is one of the most significant ionospheric parameters from which the morphology of topside density profile can be well characterized. IRI provides two choices to describe the  $foF2$ : CCIR (International Radio Consultative Committee) or now known as ITU (International Telecommunication Union) model (CCIR, 1967), and URSI (International Union of Radio Science) model (Rush, 1992). Both cases are based on the observations from the worldwide network of ionosonde stations. The availability of reliable data for the specific region and time determined the accuracy of the model (Bilitza and Reinisch, 2008). The ionosphere in East Asia is an important region where the station density is relatively dense. However, the electro and atmospheric dynamics within middle and low latitude ionosphere of East Asia region which is controlled by the equatorial anomaly phenomena can be very complicated. Several studies have shown that there are relatively large discrepancies between the ionospheric parameters predicted by IRI model and the observational data in equatorial and low latitude regions, especially in East Asia and southern China area (Adeniyi et al., 2003; Liu et al., 2004; Zhang et al., 2004). Zhang et al. (2007) examined that the percentage difference values of  $foF2$  predictions by using URSI coefficients in IRI pattern can reach as large as 30 % around pre-sunrise time, and between  $-5$  % percent and  $-25$  % during most time period of the day. Therefore, it is necessary to update the existing CCIR or URSI  $foF2$  model or build directly a single station or regional model of  $foF2$  among East Asia region.

Several new modeling techniques with respect to different ionospheric parameters have been proposed. Some studies made the temporal and spatial forecasting of ionospheric  $foF2$  and built the model by using neural network analysis (Kumluca et al., 1999; Oyeyemi et al., 2005, 2006; McKinnell and Oyeyemi, 2009, 2010). Of particular intention is concentrated on modeling the ionospheric parameters such as  $foE$ ,  $foF2$ ,  $hmF2$ , and  $M(3000)F2$ , etc. based on empirical orthogonal function analysis (e.g., Dvinskikh, 1988; Liu et al., 2008; Zhang et al., 2009). In this paper, we will focus on constructing single station model of ionospheric  $foF2$

for Wakkanai (Geographic  $45.4^\circ$  N,  $141.7^\circ$  E), Kokubunji (Geographic  $35.7^\circ$  N,  $140.1^\circ$  E) and Yamagawa (Geographic  $31.2^\circ$  N,  $130.6^\circ$  E) during the years of 1971–1987 by using similar-parameters interpolation method and empirical orthogonal analysis, and the results are compared with the observational data and with IRI model.

## 2 Data set for station modeling

The ionospheric F-layer over Japan, which lies near the inner flank of the northern crest of ionospheric equatorial anomaly, is a representative sector of East Asia (geographic longitude range:  $130^\circ$  E– $145^\circ$  E; geographic latitude range:  $30^\circ$  N– $45^\circ$  N; geomagnetic latitude range:  $18^\circ$  N– $35^\circ$  N). Relatively large discrepancies have been measured between IRI and observational values among this sector (Liang, 1990; Adeniyi et al., 2003; Bilitza et al., 2006; Vlasov and Kelley, 2010). Here we use hourly  $foF2$  data observed at three ground-based ionosonde stations in Japan which are Wakkanai, Kokubunji and Yamagawa. These three stations are the oldest established ionosonde sites with long history of reliable data of ionograms (Xu et al., 2004; Lei et al., 2008a). The time period of 1971–1987 are used in the present study because it covers more than one whole solar cycle as long as possible with the maximum data availability. The ionosonde data for 1980, 1981 (high solar activity years) and 1986, 1987 (low solar activity years) are used for data-model comparison in order to assess to what degree the empirical model can represent the observational results. The geographical coordinates and geomagnetic latitudes are listed in Table 1.

## 3 Description of the similar-parameters interpolation method

We use similar-parameters method to refill the missing data for aforementioned three stations during the time period of 1971–1987. Similar-parameters method, which was originally applied to the field of aerodynamics (see NASA website: <http://www.grc.nasa.gov/WWW/k-12/rocket/airsim/>), can be used to interpolate the missing data before constructing the empirical model. In this method, the observational data can be influenced and determined by all possible factors or parameters. If two data sets have the same values for the similarity parameters, the data set which contains the missing value under certain temporal and spatial conditions can be interpolated by using its “control” data set. Here we list several possible drivers of ionospheric variability of F layer in Table 2.

Accordingly we will choose appropriate proxies as parameters from which the influence of aforementioned drivers on ionospheric  $foF2$  can be reflected. For the solar ionizing radiation, we choose solar index  $F_{10.7}$  as proxy because it is an ideal indicator for solar cycle and solar rotation variation.

**Table 2.** Possible drivers of ionospheric variability of F layer.

Solar ionizing radiation	Geomagnetic activity	Meteorological influences
Solar cycle variation	Magnetic storms	Solar and lunar tides
Solar rotation variation	Substorms	Acoustic and gravity waves
Seasonal variation	IMF/Solar wind	Planetary waves
Solar flares	Energetic particle precipitation	Lower atmosphere influence

**Table 3.** The settings of the similar parameters.

Parameters	Influence factors	Similar conditions	Remarks
$F_{10.7}$	Solar cycle and Solar rotation variation	Similar solar activity Relative difference <10 %	$F_{10.7} < 150$ : low solar activity $F_{10.7} \geq 150$ : high solar activity
X-ray radiation fluxes	Solar cycle variation	Same flare classes	
Solar zenith angle	Seasonal/Diurnal variation	Identical Solar zenith angle	
Day of year	Seasonal variation	Changes <15 days or >350 days	
Local time	Diurnal variation	Identical local time	
Ap Index	Geomagnetic activity	Similar geomagnetic activity Relative difference < 15 %	Ap < 8: Quiet $8 \leq \text{Ap} < 15$ : Unsettled $15 \leq \text{Ap} < 30$ : Active $30 \leq \text{Ap} < 50$ : Minor storm $50 \leq \text{Ap} < 100$ : Major storm Ap $\geq 100$ : Severe storm

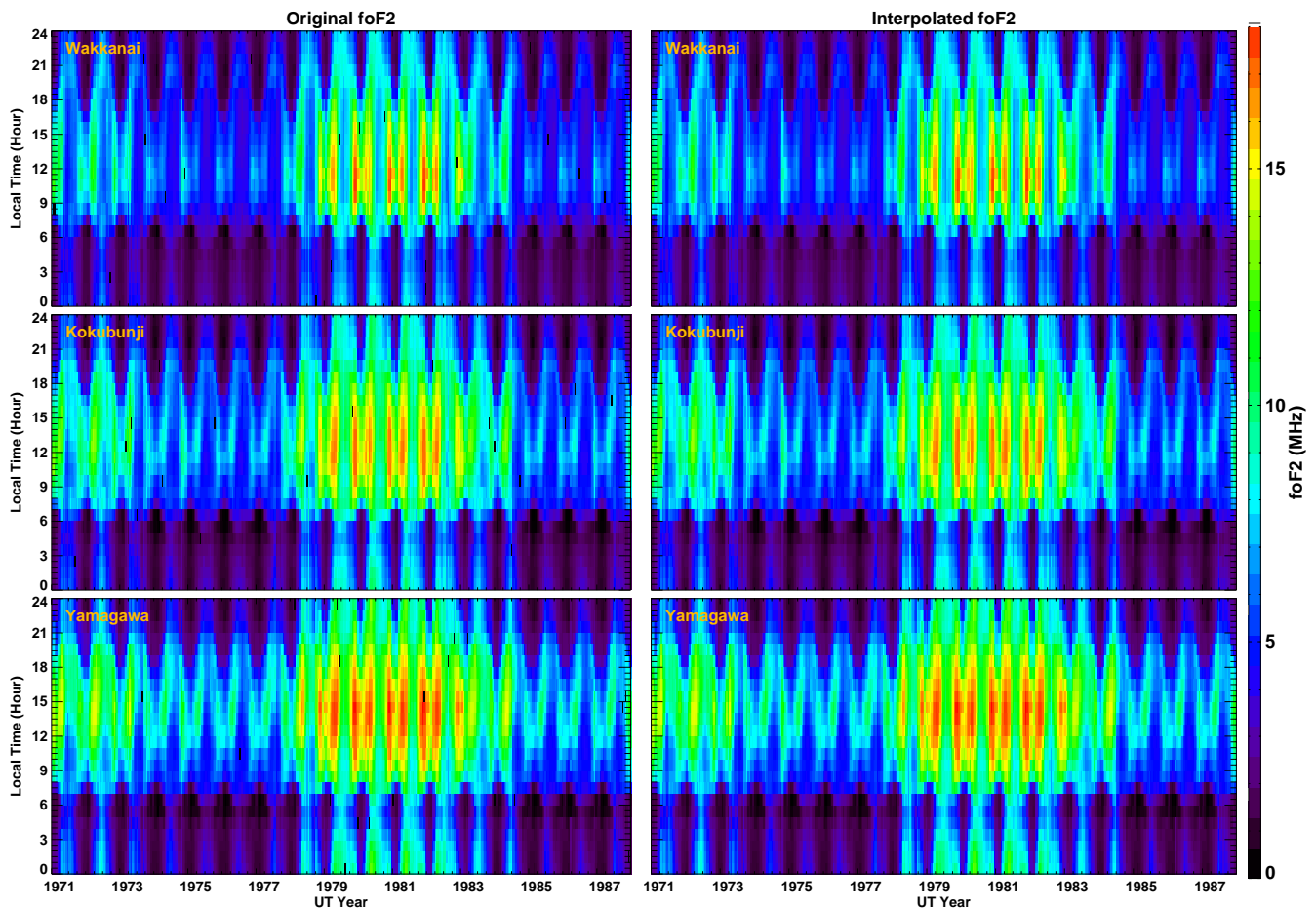
The solar radiation also related to the change in X-ray radiation fluxes which also need to be set as proxy. The seasonal variation is related to change in solar zenith angle which is a function of DOY (day of year), local time and latitude. As the latitude is fixed for single station, here we choose DOY (day of year) and local time as proxies. For the geomagnetic activity, Ap index is a suitable proxy because it is a planetary index for calculating the strength of world-wide geomagnetic disturbances. For the meteorological influences, Mendillo et al. (1998) tried to estimate to what degree the F-layer variability could be attributed to the troposphere and lower stratosphere by using the NCAR (National Center for Atmospheric Research) coupled CCM3/TIME-GCM (Community Climate Model-3/Thermosphere-Ionosphere-Mesosphere-Electrodynamics General Circulation Model; Roble and Ridley, 1994). Mikhailov et al. (2007, 2009) showed that synchronous variation of electron density can be observed during geomagnetic quiet day in both E- and F-region. Such variability is considered to be caused by perturbation originating from lower atmosphere. It is hard to set appropriate proxy for meteorological influences not only because the major influences to the variation of  $f_oF2$  are due to solar ionizing and geomagnetic activity but also since there are so many uncertainties in the climate change as well as in the generation and transmission of the wave. Therefore, the parameters being used here are  $F_{10.7}$ , Ap index, local time and DOY based on aforementioned analysis. At any single station, the missing value can be interpolated by choosing

the median value from the data set which have the “similar” parameters. The settings of the similar parameters are listed in Table 3. One thing worth noting is that the any interpolation method cannot be a good one for interpolating long-gap missing data. However, the data missing for  $f_oF2$  at Wakkanai, Kokubunji, and Yamagawa during the time interval 1971–1987 is only scattered, which makes it possible to implement the interpolation method. Figure 1 displays the comparison of the original and interpolated  $f_oF2$  values. It can be seen from the figure that the similar-parameters interpolation method is fairly feasible for data preprocessing.

## 4 Description of the modeling technique

### 4.1 Brief introduction of EOF analysis method

Empirical Orthogonal Function (EOF) analysis method, which was invented by Pearson (1901), has been widely used to analysis the temporal and spatial variation of the research objects. Dvinskikh (1988) first introduced EOF analysis into the empirical modeling of the ionospheric parameters. It has been confirmed by many researchers that EOF analysis is a powerful method for ionospheric modeling and data analysis (e.g., Singer and Dvinskikh, 1991; Daniell et al., 1995; Matsuo et al., 2002, 2005; Marsh et al., 2004; Zhao et al., 2005; Zapfe et al., 2006; Liu et al., 2008; Mao et al., 2008; Zhang et al., 2009; Matsuo and Forbes, 2010, and more). According to EOF theory, the variation of the ionospheric parameters



**Fig. 1.** Comparison among the original  $foF2$  (left panels) and interpolated  $foF2$  using similar-parameters method during the interval 1971–1987 at three ionosonde stations.

are attributed to different factors which can be extracted and separated in terms of their relative “contribution” with respect to ionospheric parameters. The EOF method can be utilized to decompose and express the variation as a summation of the eigen modes which are not pre-specified artificially, but are calculated according to experimental data itself in the decomposition process. The combination of eigen modes can reproduce the substantive characteristics of the data. The eigen series have rapid convergence velocity and high calculation accuracy. This makes EOF analysis method a highly effective way of empirical modeling not only by greatly reducing the modeling parameters but also by considerably saving the computation time compared with other expansion methods. For further details of EOF decomposition, readers may refer to Dvinskikh (1988), Xu and Kamide (2004), and Zhang et al. (2009).

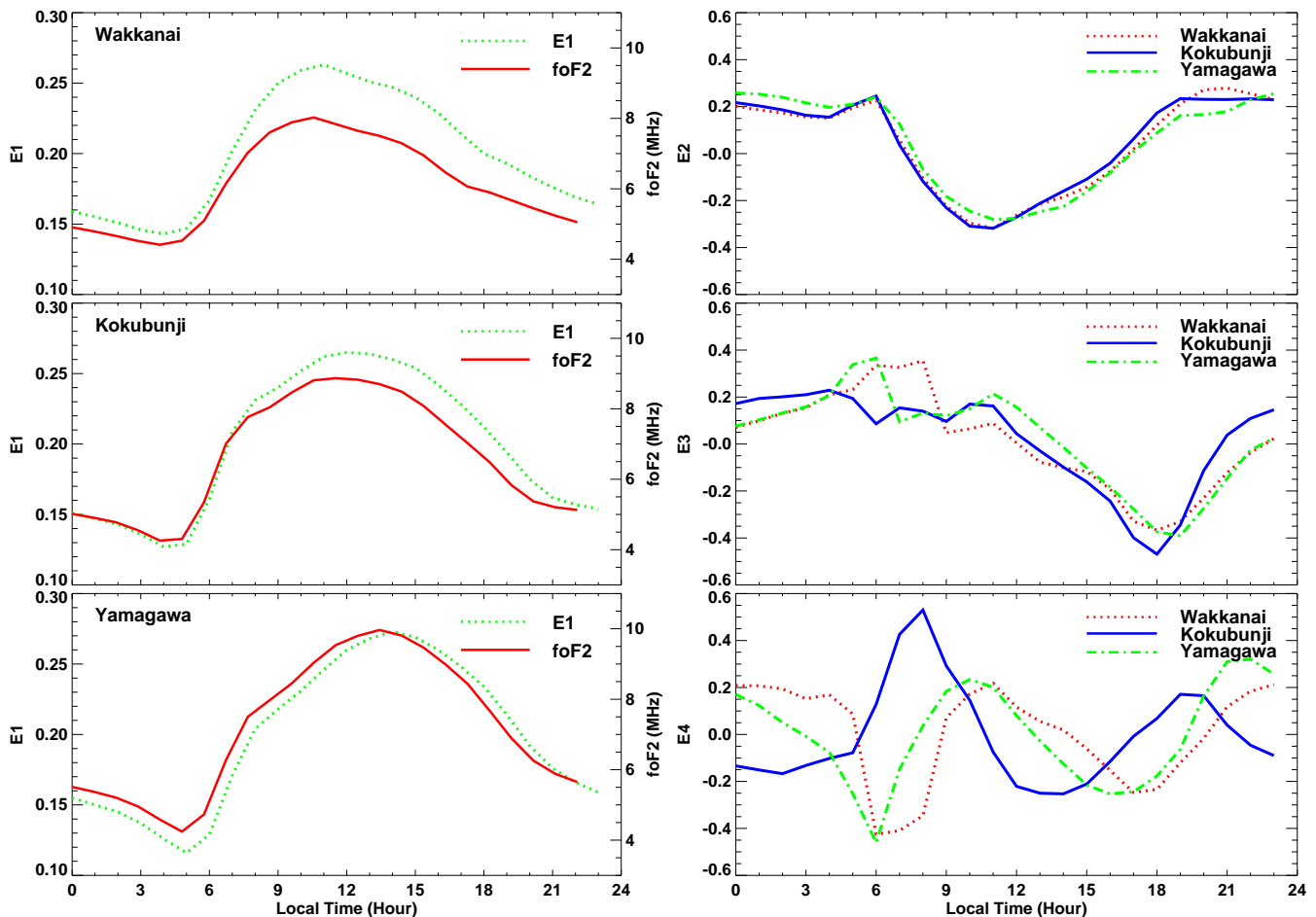
#### 4.2 Data processing with EOF analysis method

The hourly values of  $foF2$  at three stations are decomposed into the EOF base functions  $E_k$ , multiplied by the

corresponding EOF coefficients  $A_k$  using the EOF analysis method:

$$foF2(d, h) = \sum_{k=1}^{24} E_k(h) \times A_k(d) \quad (2)$$

Where  $foF2(d, h)$  is the combination of hourly values of the observational data expressed as a  $6209 \times 24$  array with the rows corresponding to the days ( $d = 1, 2, 3, \dots, 6209$ ) which is calculated from 1 January 1971; The column corresponding to the local time LT ( $h = 0, 1, 2, \dots, 23$ ).  $E_k(h)$  is the  $k$ -th base function of  $foF2$  reflecting the diurnal variation,  $A_k(d)$  is the coefficients of  $E_k(h)$  which represents the long-term variation (solar cycle, annual and seasonal, etc.). Theoretically, all of those 24 order base functions and the associated EOF coefficients are needed to reproduce the variation of the original matrix. However, the EOF decomposition converges fairly quick, which makes it possible to use only a limited number of base functions and the corresponding coefficients to reconstruct the matrix and reflect the principal components of the variation of original data set. In this paper, the first four EOF series, which can reflect the 99.7% information of the



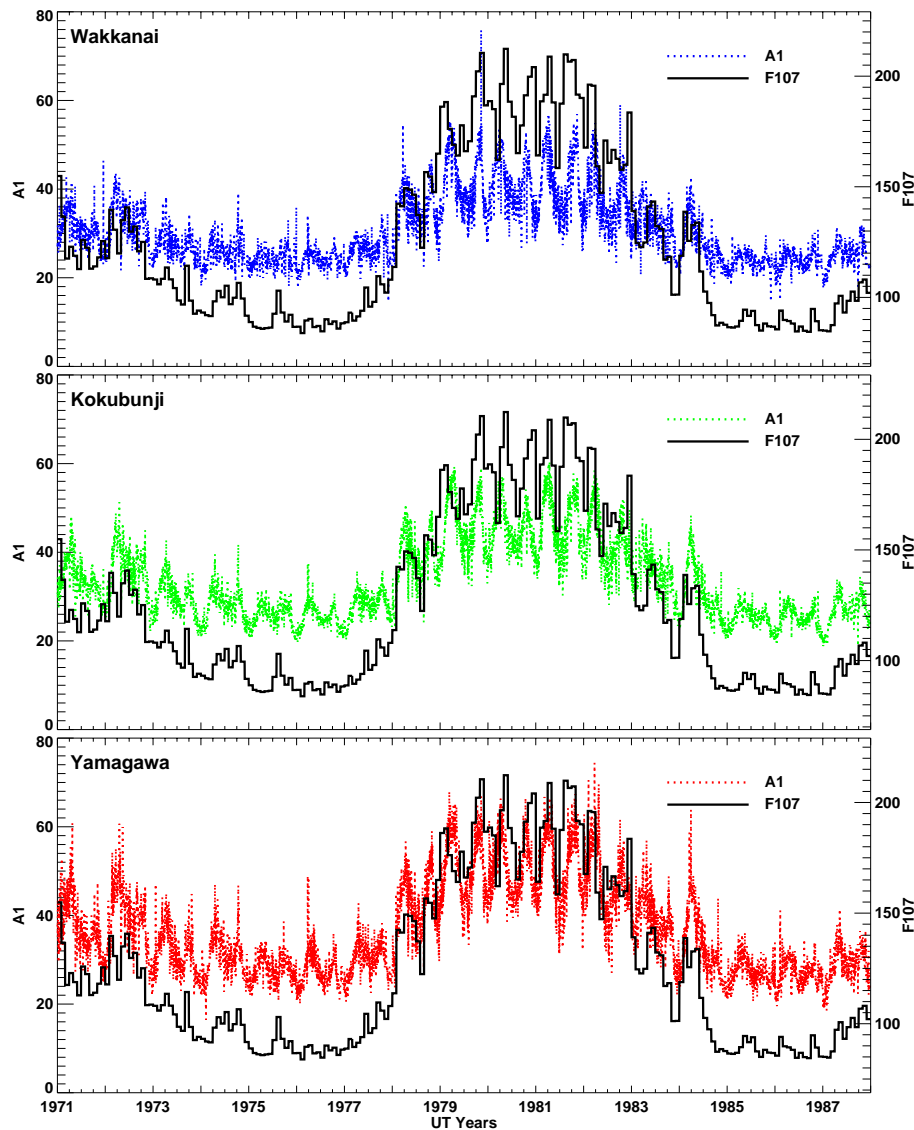
**Fig. 2.** Diurnal variation of the first EOF base functions and the *foF2* (left panels) and 2nd to 4th order base functions (right panels) at three stations.

fluctuation of the original data matrix, are used to reconstruct the *foF2* and build the model. So it is feasible to reduce the number of modeling parameters and to simplify the calculating process to a great extent while the accuracy of data reconstruction being considerably high.

Figure 2 shows the diurnal variation of the first four order EOF base functions and the *foF2* at three stations respectively. It can be clearly seen from the left panels that the diurnal variation of the 1st order base function  $E_1$  and *foF2* are quite similar to each other, the correlation coefficients between  $E_1$  and *foF2* are greater than 0.97 among all those three stations. Therefore  $E_1$  can represent the average diurnal variation trend of *foF2*. The ionospheric *foF2* is also influenced by other factors which including interhemispheric flow, neutral winds and diffusion. Thus one would expect there are small scale disturbances and irregularities superimposed on the diurnal variation due to above influences, which are well represented from the variation of the 2nd, 3rd and 4th order base functions in the right panels. Take the 3rd order base function  $E_3$  as an example.  $E_3$  be-

gins to increase at around 05:00LT and then has a decreasing trend at around 11:00LT and increases again from 17:00LT. These phenomena can be explained by ionospheric sunrise enhancement, bite out phenomena and sunset enhancement respectively (Schunk and Nagy, 2000; Liu et al., 2004).

Figure 3 displays the long-term variations of  $F_{10.7}$  and the first EOF coefficient  $A_1$  from 1971 to 1987. Figure 4 shows the variations of Ap index and EOF coefficients from the 2nd to the 4th order. By comparing these two figures, we can see that the 1st EOF coefficient  $A_1$  has a much larger value than  $A_2$ ,  $A_3$  and  $A_4$  which are the 2nd, 3rd, and 4th order EOF coefficients respectively. The amplitude range of the 1st order coefficient  $A_1$  is about two to three times larger than that of the 2nd order coefficient  $A_2$ , and  $A_2$  is also two times larger than the 3rd order coefficient  $A_3$  while the 4th order coefficient  $A_4$  is the smallest. Referring to the analysis of base function in previous paragraphs, we can deduce that the principal components of variation of *foF2* are reflected by the first-order EOF series which is represented in the form of  $E_1 \times A_1$ .



**Fig. 3.** Long-term variation of  $F_{10.7}$  and the 1st EOF coefficients from 1971–1987.

The synchronous solar cycle variation trends both in  $A_1$  and in  $F_{10.7}$  for all three ionosonde stations can reflect a highly positive-correlation relationship. The correlation coefficients for Wakkanai, Kokubunji and Yamagawa are 0.937, 0.945 and 0.928, respectively. This phenomenon illuminated that  $A_1$  can represent the components of solar cycle variation in original foF2 data set. And it can also be seen that there is an evident semi-annual variation and a relatively weak annual-variation component in  $A_1$ . Though they are not as prominent as the solar cycle variation because they are superimposed on the latter one. Figure 4 shows that there are quite obvious annual variation components in  $A_2$ .  $A_2$  also contains relatively weak semi-annual variation components. Both  $A_3$  and  $A_4$  contain mainly the semi-annual variation elements as well as more subtle variations, such as seasonal variations

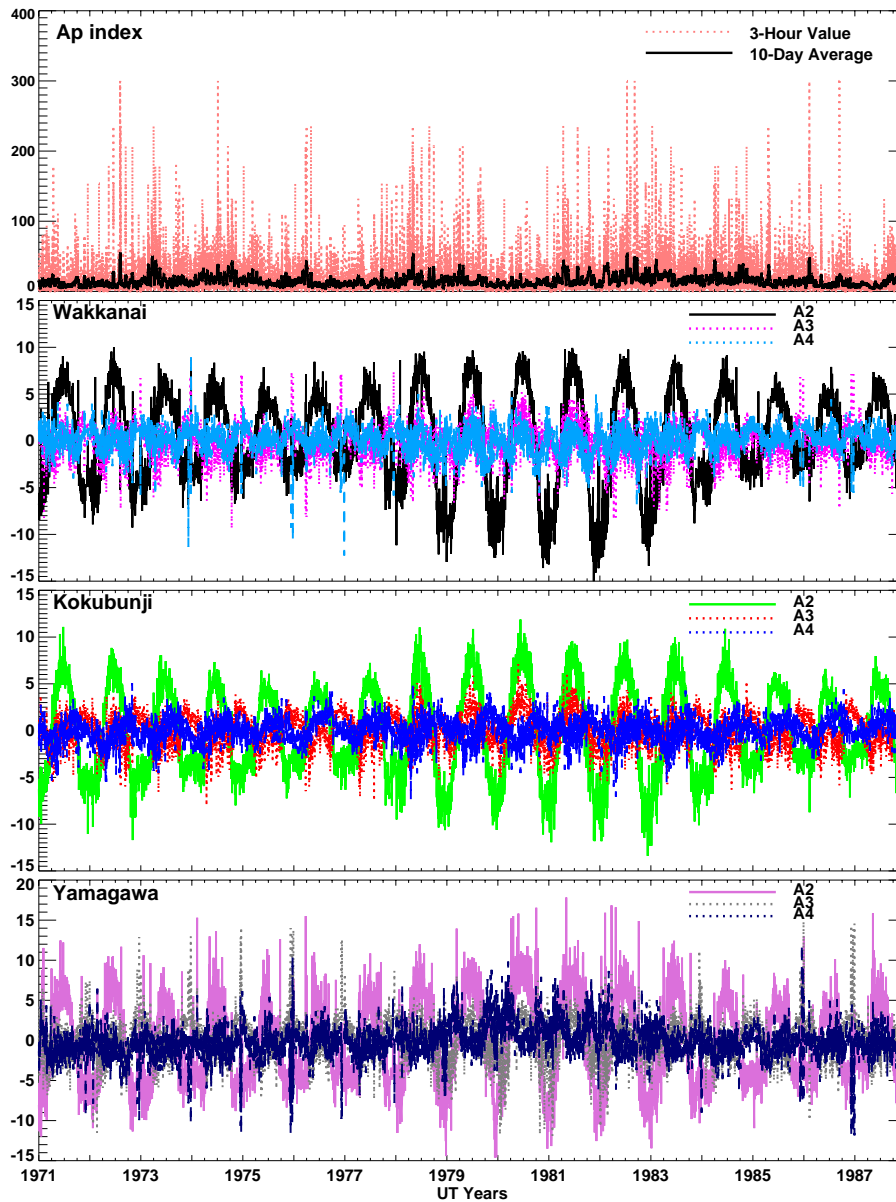
and small-scale irregularities. The semi-annual variation components can be attributed to periodic wave in geomagnetic Ap indices with maxima near equinoxes (Petrukovich and Zakharov, 2007). The amplitudes of the EOF coefficients are influenced by the solar activity represented in the form of  $F_{10.7}$  and by geomagnetic activities represented in the form of Ap index.

From the above analysis, we know that the first four EOF coefficients can reflect the solar cycle variation, annual fluctuation, semi-annual oscillation, and short-term irregularities. So we can use the formal Fourier series to model the first four EOF coefficients  $A_n$  ( $n = 1, 2, 3, 4$ ).

$$A_n(d) = B_{n1}(d) + B_{n2}(d) + B_{n3}(d) + \varepsilon \quad (3)$$

$$B_{n1}(d) = C_{n1} + D_{n1}F_{10.7p}(d) + E_{n1}Ap(d) \quad (4)$$





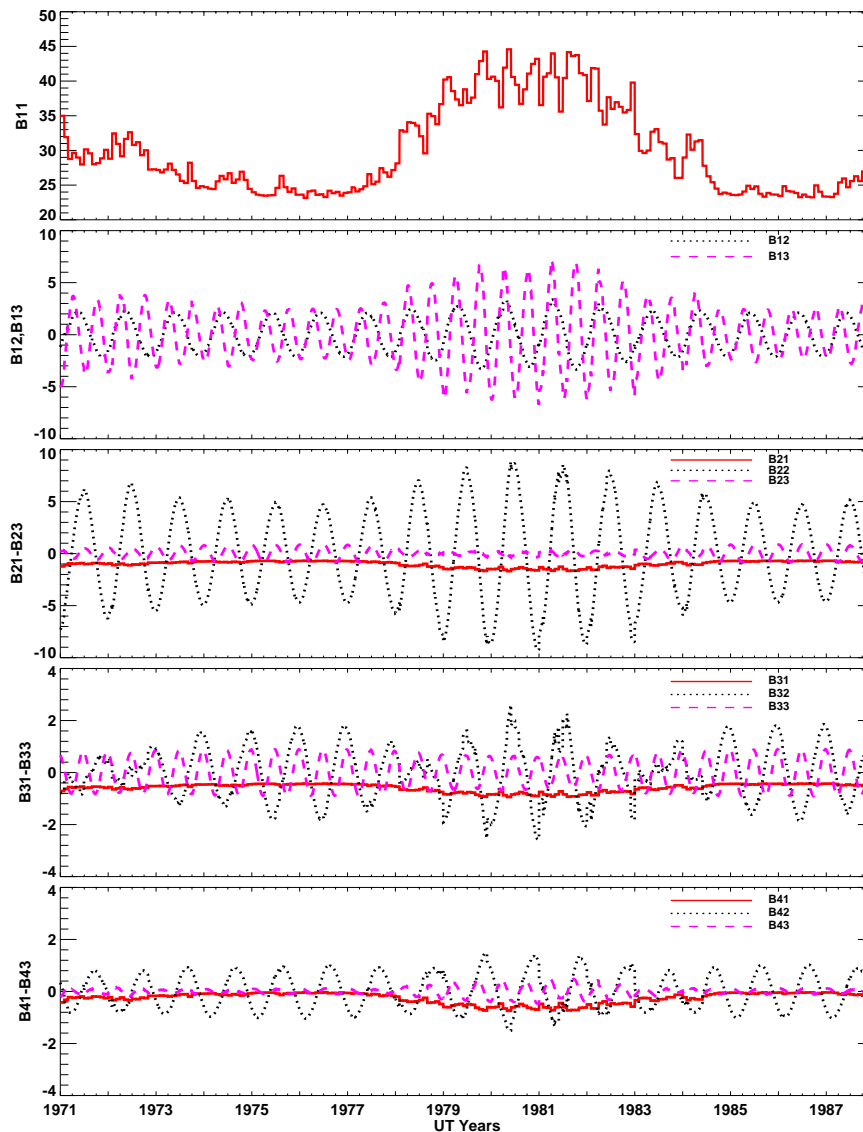
**Fig. 4.** Long-term variation of Ap index and the 2nd to 4th EOF coefficients from 1971–1987.

$$B_{n2}(d) = (C_{n2} + D_{n2}F_{10.7p}(d) + E_{n2}Ap(d))\cos\frac{2\pi d}{365.25} + (F_{n2} + G_{n2}F_{10.7p}(d) + H_{n2}Ap(d))\sin\frac{2\pi d}{365.25} \quad (5)$$

$$B_{n3}(d) = (C_{n3} + D_{n3}F_{10.7p}(d) + E_{n3}Ap(d))\cos\frac{2\pi d}{365.25/2} + (F_{n3} + G_{n3}F_{10.7p}(d) + H_{n3}Ap(d))\sin\frac{2\pi d}{365.25/2} \quad (6)$$

Where *n* is the *n*-th EOF coefficient of base function.  $B_{n1}(d)$ ,  $B_{n2}(d)$ , and  $B_{n3}(d)$  correspond to the solar cycle, annual, and semi-annual variation components in EOF coefficients respectively,  $\varepsilon$  is the residual error.  $F_{10.7p} = (F_{10.7} + F_{10.7A})/2$ , which was established based on

the value of daily  $F_{10.7}$  and its 81-day moving average  $F_{10.7A}$ .  $F_{10.7p}$  has been used in solar irradiance empirical models as a solar EUV proxy (e.g., Hinteregger et al., 1973; Richards et al., 1994). It has been validated that  $F_{10.7p}$  represents quite well the intensity of solar EUV flux, which is considered as a better solar proxy for common use (Liu et al., 2006, 2011). Here we use a linear function of  $F_{10.7p}$  and Ap to express  $B_{n1}(d)$  in which the solar cycle and semi-annual variation components can be represented.  $B_{n2}(d)$  and  $B_{n3}(d)$  are expressed as combination of sinusoidal functions with periods been modulated to one year (denominator = 365.25) and six months (denominator = 365.25/2) respectively in order to represent the different periods of variation components

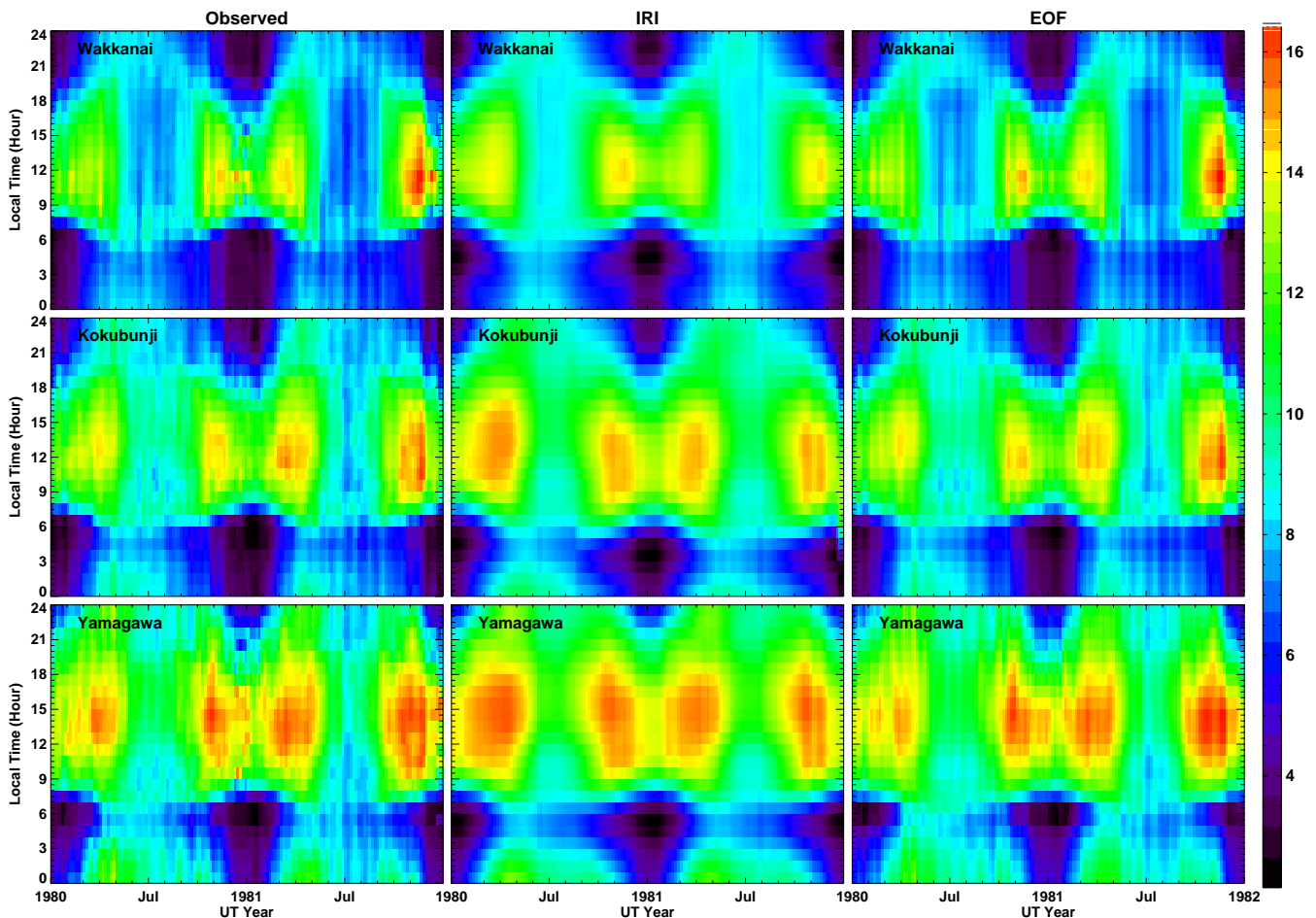


**Fig. 5.** Variation components of fitted EOF coefficients with different period.

in EOF coefficients. The amplitudes of those trigonometric functions can also be expressed in the form of linear functions of  $F_{10.7p}$  and  $A_p$  index to show their dependence on the level of solar activity as well as geomagnetic activity.  $C$ ,  $D$ ,  $E$ ,  $F$ ,  $G$  and  $H$  are coefficients of various parts in above equations. Those coefficients can be calculated by using linear regression analysis method, and thus the EOF coefficients  $A_n(d)$  can be acquired by using Eqs. (3)–(6) with those determined coefficients. We can still add shorter periods variation components into the EOF coefficients (e.g., the seasonal variation components, the solar rotational components with the period of 27-days, and 16-days solar oscillation, etc.) for the accuracy of the fitted coefficients. The modeled values of  $foF2$  at single stations can be acquired using Eq. (2) with the EOF base functions multiplied by the coefficients calculated with aforementioned linear regression method.

Figure 5 shows the solar cycle, annual, and semi-annual variation components of the fitted EOF coefficients at Kokubunji respectively. We can see from the figure that: (1) The contribution to solar cycle variation is mainly made by the 1st order component which has a much larger value than the others. (2) Annual variation can be traced in all of those 4-order components, but the major contribution comes from the 2nd order component. (3) The semi-annual variation is mainly concentrated in the 1st, 3rd and 4th order components whereas contribution from the 2nd order components is smaller. (4) The influence due to solar activity and geomagnetic activity can be indicated from these components.





**Fig. 6.** Comparison among the observational  $f_oF2$ , IRI model value and EOF model value during 1980–1982.

### 4.3 Data-model comparison and discussion

The model constructed using EOF method combined with linear regression analysis need to be verified via the data-model comparison from which the accuracy of model can be evaluated. The data among period of high solar activity years (1980, 1981) and low solar activity years (1986, 1987) are selected from the observational  $f_oF2$  value as validating samples. One thing worth noting is that the chosen data for data-model comparison is excluded from the original data set when building the model. In other words, the data among the selecting time period is not included in the data set to generate the EOF coefficients, which makes the testing data independent for model validation.

Figures 6 and 7 show the comparison of the daily variation of the observational  $f_oF2$  values, the values given by IRI model, and the EOF modeled values during years with high solar activity (1980, 1981) as well as years with low solar activity (1986, 1987). It can be seen from the figure that the IRI modeled values have a relatively smooth boundary from which the details of variation in original data set can

not be well expressed. This illustrated that some variation with short periods and small scale irregularities/disturbances are beyond the level that can be well represented by IRI. On the contrary, the EOF modeled values can reflect quite well different scales of variation in original  $f_oF2$ , which is mainly attributed to that the EOF coefficients are directly obtained from the decomposition of original data set. The quick convergence of EOF expansion made it possible to use limited number of base functions and the corresponding fitted coefficients generated by linear regression analysis to reproduce the observational data set.

Figure 8 shows the scatter plots of IRI and EOF model versus the observational  $f_oF2$ . The left two panels are for high solar activity years (1980, 1981), and the right two panels are for low solar activity years (1986, 1987). One phenomenon worth noting is that the correlation coefficients between the modeled values and the observational values are larger in high solar activity years for both IRI and EOF model. For IRI model, some studies have noted major shortcoming for low electron densities conditions during which the discrepancies between the model and measurement can be consistently

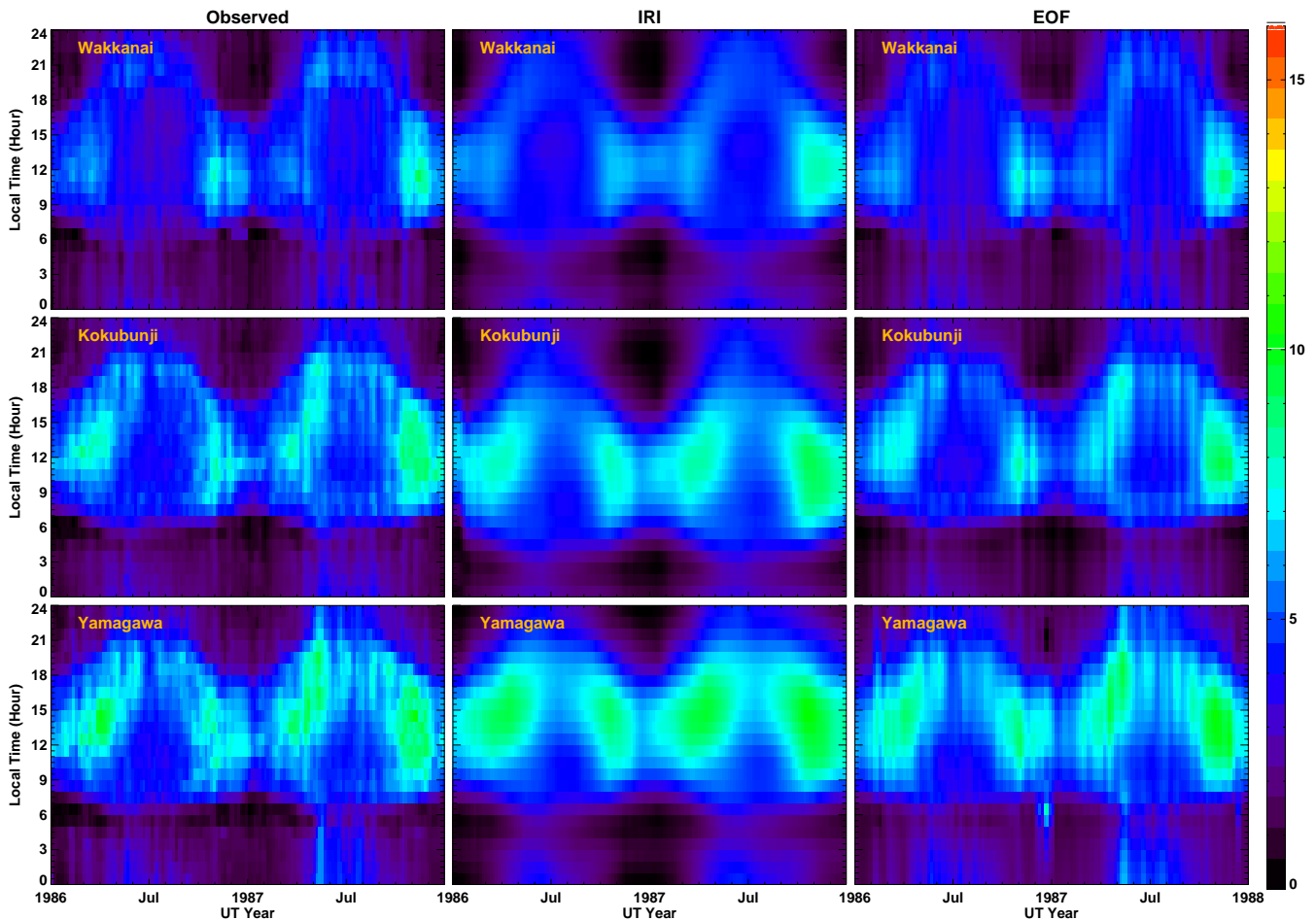


Fig. 7. Comparison among the observational foF2, IRI model value and EOF model value during 1986–1988.

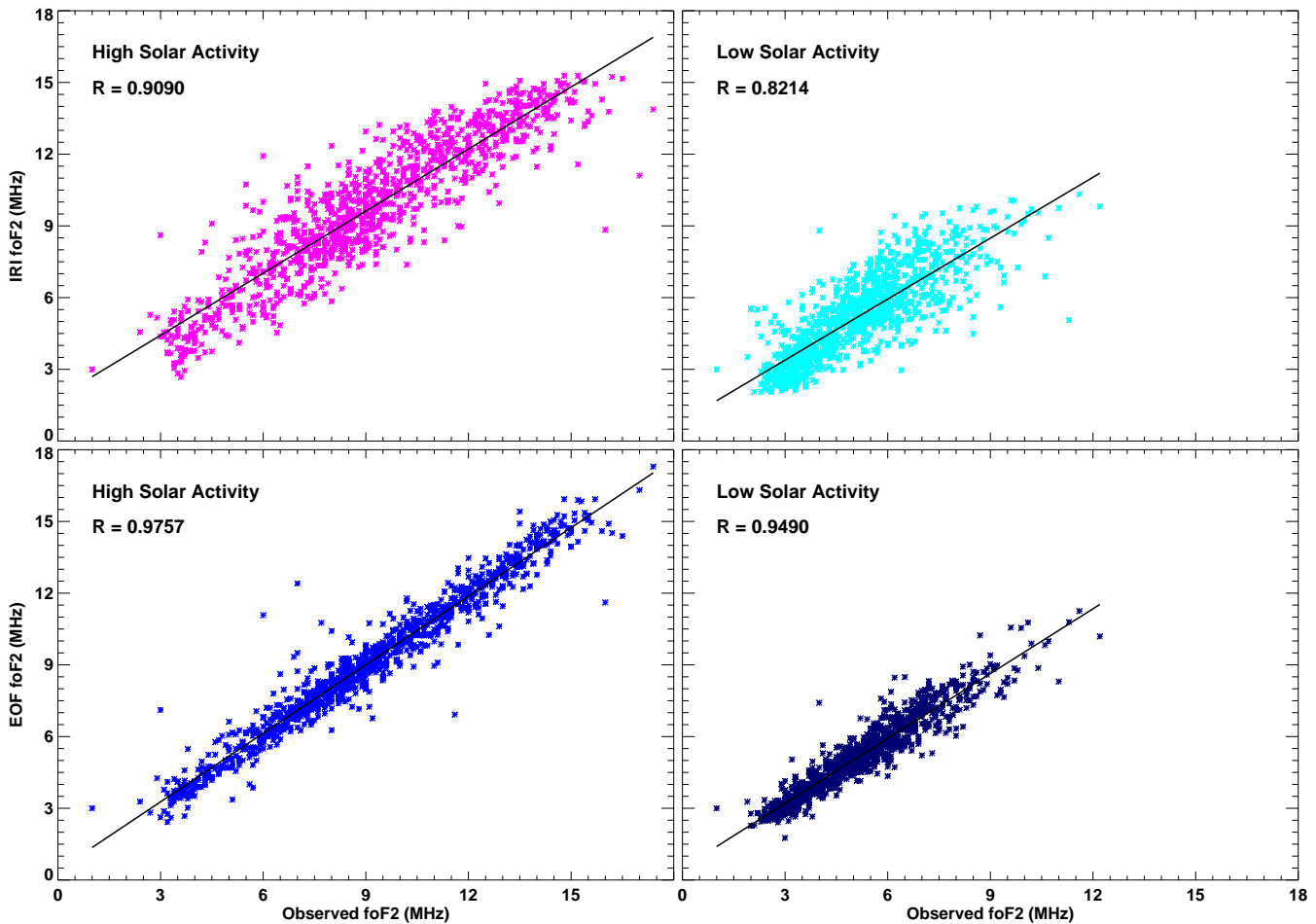
large (e.g., Bilitza et al., 1998, 2006; Bilitza and Williamson, 2000). For EOF model, the cumulative percentage variances of the first 4-order coefficients associated with the principal components of EOF analysis during solar high activity years are higher than that in solar low activity years. No matter for high or low solar activity years, the correlation coefficients between the EOF modeled values and the observational values, which are 0.9757 and 0.9490 respectively, are higher than that for IRI model. So the EOF model can reproduce the original data quite well by better reflecting the temporal distribution characteristics.

As the discrepancies between model and observational values varied with the temporal distribution of foF2, the model-measurements deviation, rather than the model result itself, seems better suited for investigating the accuracy degree of the model. Here we use the relative error (RE) and the relative mean square error (RMS) to represent the deviation. The RE and RMS are calculated by the following formulas.

$$RE = \frac{1}{N} \sum_{i=1}^N \left( \frac{foF2_{model} - foF2_{obs}}{foF2_{obs}} \right) \times 100\% \quad (7)$$

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{foF2_{model} - foF2_{obs}}{foF2_{obs}} \right)^2} \times 100\% \quad (8)$$

Figure 9 shows the diurnal variation of relative errors and relative mean square errors between the model and the measurements in spring (March, April and May), summer (June, July and August), autumn (September, October and November), and winter (December, January and February) of high and low solar activity years. It can be seen that the EOF modeled values agree well with the observational foF2 than that for IRI modeled values due to low relative errors and low relative mean square errors in each season. One important feature worth noting is that during high solar activity years, the deviation error of EOF model in winter is greater than that in summer despite smaller solar zenith angle in summer. This phenomenon, which is similar with ionospheric seasonal anomaly, might be attributed to the strong summer-to-winter neutral circulation in high solar activity years. The circulation will result in fluctuation in the O/N<sub>2</sub> ratio, electron density and foF2 profile, which could decelerate the velocity of convergence of EOF decomposition, thereby reduce



**Fig. 8.** Scatter plots of IRI and EOF model values versus observational data of  $f_oF2$ .

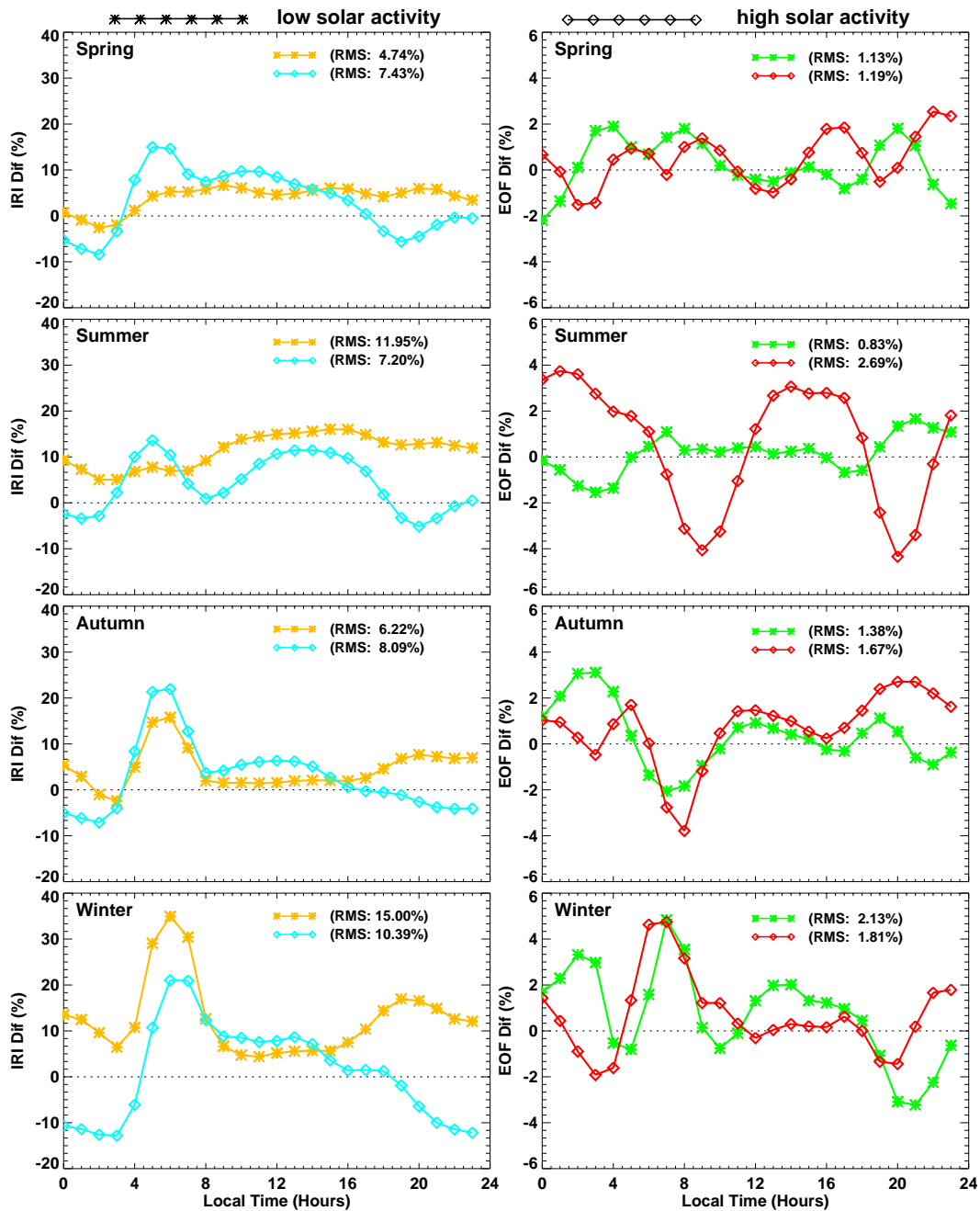
the accumulated percentages of first 4-order coefficients of EOF base function.

Figure 10 shows the long-term variation of relative errors and relative mean square errors between the model and the observational data in day (6 a.m.–6 p.m.) and night (6 p.m.–6 a.m.). The IRI model results are obviously inferior to that of EOF model results, and there are some other features worth noting. First, there are distinct semi-annual variation components in the relative errors of EOF model. Second, the wave crest and the wave trough of the variation of relative errors are asymmetric for day and night. In the daytime, the relative errors have positive values (near wave crest) or increasing trends around solstice and the negative values (near wave trough) or decreasing trends around equinox. However, the night time conditions are to the contrary. Positive error appears near equinox and negative error appears near solstice. For the seasonal asymmetry, three different explanations are suggested: (1) neutral wind hypothesis: the large-scale interhemispheric circulation induced by asymmetric heat distribution can cause asymmetric density variation (Johnson and Gottlieb, 1970; Fuller-Rowell, 1998). The circulation

may also exert influence on  $f_oF2$  and even on deviation with small scales due to the proportional relation between  $f_oF2$  and electron density. (2) Equinoctial hypothesis: the variation of ionospheric parameters are related to the angle between solar wind flow and geomagnetic dipole field and to the solar asymmetric illumination (McIntosh, 1959; Lyatsky et al., 2001). (3) Axial hypothesis: the increase of heliographic latitude around equinoxes can place Earth closer to fast solar wind streams from coronal holes (Bohlin, 1977).

## 5 Summary and conclusions

Ionospheric  $f_oF2$  measurements in three ionosonde stations of Japan during 1971–1987 with maximum data availability offer the opportunity to construct the empirical model in order to provide the basis for updating the existing CCIR or URSI  $f_oF2$  model. In the present paper, the original  $f_oF2$  data is interpolated with similar-parameters method, then an empirical model is constructed with EOF decomposition combined with regression analysis. We made the following observations and conclusions:



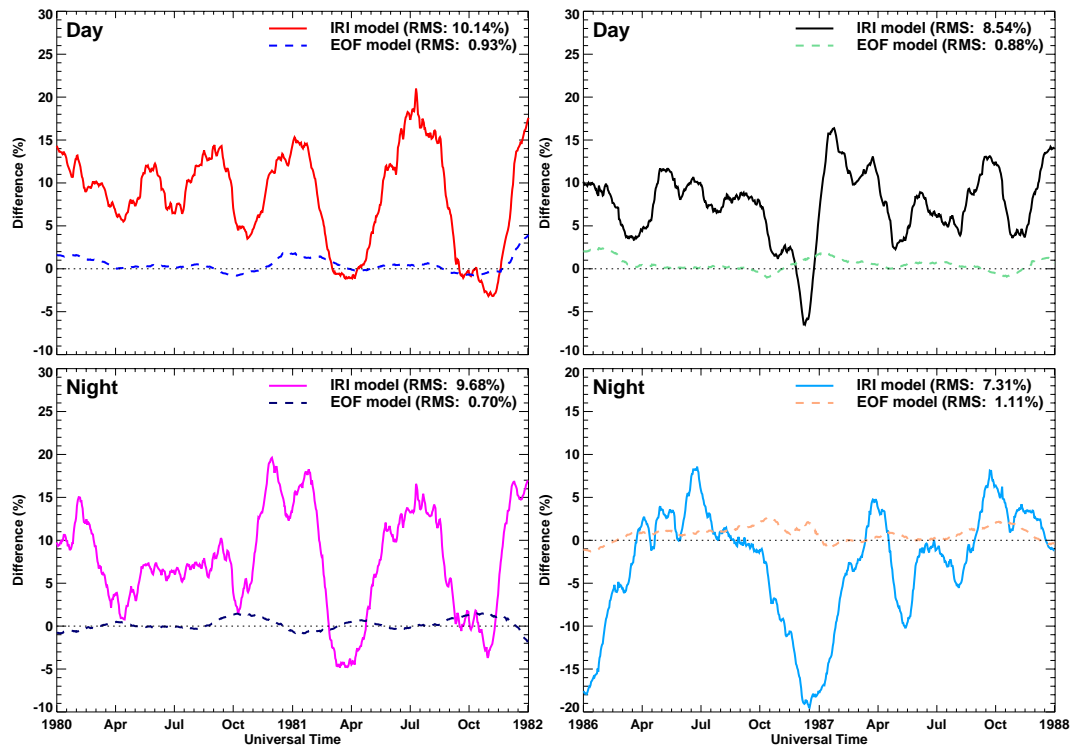
**Fig. 9.** Diurnal variation of relative errors and relative mean square errors between model and the observational data.

First, ionospheric  $f_oF_2$  is subject to a number of drivers which can be broadly divided into three categories: solar ionizing radiation, geomagnetic activity and meteorological influences. It is feasible to use  $F_{10.7}$ , X-ray fluxes, solar zenith angle and  $A_p$  index as appropriate proxies to reflect the aforementioned influences and to refill the missing value of original  $f_oF_2$  with similar-parameters method.

Second, the EOF model can reproduce quite well the original data sets of  $f_oF_2$  by utilizing only the first 4-order base

functions as well as corresponding coefficients. The base functions can express the diurnal variation of the original  $f_oF_2$  data while the corresponding coefficients can represent the long-term variation (solar cycle, annual, semi-annual, seasonal, solar rotation, and irregularities, etc.).

Third, comparisons between the EOF model and the observational data show higher linear correlation coefficients and more accurate degree than that of IRI model. The EOF model, which agrees quite well with the observational data,



**Fig. 10.** Long-term variation of relative errors and relative mean square errors between model and the observational data.

can reflect the major change tendencies and the temporal distribution characteristics of the mid-latitude ionosphere of the Sea of Japan region.

Finally, the error analysis reveals that there are seasonal anomaly and semi-annual variation phenomena. These results can be attributed to neutral wind hypothesis, equinoctial hypothesis, and axial hypothesis.

*Acknowledgements.* This research was jointly supported by the National Natural Science Foundation of China (41174134 and 40904036), National Important basic Research Project of China (2011CB811405), the State Key Laboratory of Space Weather, and the China Scholarship Council (2009601221). The  $F_{10.7}$  index data, Geomagnetic Ap index data and ionospheric  $f_oF_2$  data for Wakkanai, Kokubunji and Yamagawa were downloaded from the SPIDR (Space Physics Interactive Data Resource) website: <http://spidr.ngdc.noaa.gov/spidr/home.do>.

Topical Editor K. Kauristie thanks two anonymous referees for their help in evaluating this paper.

## References

- Adeniyi, J. O., Bilitza, D., Radicella, S. M., and Willoughby, A. A.: Equatorial  $F_2$ -peak parameters in the IRI model, *Adv. Space Res.*, 31, 507–512, doi:10.1016/S0273-1177(03)00039-5, 2003.
- Bilitza, D.: International Reference Ionosphere 2000, *Radio Sci.*, 36, 261–275, doi:10.1029/2000rs002432, 2001.
- Bilitza, D. and Reinisch, B. W.: International Reference Ionosphere 2007: Improvements and new parameters, *Adv. Space Res.*, 42, 599–609, doi:10.1016/j.asr.2007.07.048, 2008.
- Bilitza, D. and Williamson, R.: Towards a better representation of the IRI topside based on ISIS and Alouette data, *Adv. Space Res.*, 25, 149–152, doi:10.1016/S0273-1177(99)00912-6, 2000.
- Bilitza, D., Rawer, K., Bossy, L., Kutiev, I., Oyama, K., Leitinger, R., and Kazimirovsky, E.: International Reference ionosphere 1990, Nat. Space Sci. Data Cent., Report 90-22, Greenbelt, Md., 1990.
- Bilitza, D., Koblinsky, C., Williamson, R., and Bhardwaj, S.: Improving the topside electron density model for IRI, *Adv. Space Res.*, 22, 777–787, doi:10.1016/S0273-1177(98)00098-2, 1998.
- Bilitza, D., Reinisch, B. W., Radicella, S. M., Pulnits, S., Gulyaeva, T., and Triskova, L.: Improvements of the International Reference Ionosphere model for the topside electron density profile, *Radio Sci.*, 41, RS5S15, doi:10.1029/2005RS003370, 2006.
- Bohlin, J. D.: Extreme-ultraviolet observations of coronal holes., *Sol. Phys.*, 51, 377–398, doi:10.1007/BF00216373, 1977.
- CCIR: Comité Consultatif International des Radiocommunications, Reports 340, 340-2 and later supplements, Geneva, 1967.
- Daniell, R. E., Brown, L. D., Anderson, D. N., Fox, M. W., Doherty, P. H., Decker, D. T., Sojka, J. J., and Schunk, R. W.: Parameterized ionospheric model: A global ionospheric parameterization based on first principles models, *Radio Sci.*, 30, 1499–1510, doi:10.1029/95RS01826, 1995.
- Dvinskikh, N. I.: Expansion of ionospheric characteristics fields in empirical orthogonal functions, *Adv. Space Res.*, 8, 179–187, doi:10.1016/0273-1177(88)90238-4, 1988.

- Forbes, J. M., Palo, S. E., and Zhang, X.: Variability of the ionosphere, *J. Atmos. Sol. Terr. Phys.*, 62, 685–693, doi:10.1016/S1364-6826(00)00029-8, 2000.
- Fuller-Rowell, T. J.: The “thermospheric spoon”: A mechanism for the semiannual density variation, *J. Geophys. Res.*, 103, 3951–3956, doi:10.1029/97JA03335, 1998.
- Hinteregger, H. E., Bedo, D. E., and Manson, J. E.: The EUV spectroheliometer on Atmosphere Explorer, *Radio Sci.*, 8, 349–359, doi:10.1029/RS008i004p00349, 1973.
- Johnson, F. S. and Gottlieb, B.: Eddy mixing and circulation at ionospheric levels, *Planet. Space Sci.*, 18, 1707–1718, doi:10.1016/0032-0633(70)90004-8, 1970.
- Kumluca, A., Tulunay, E., Topalli, I., and Tulunay, Y.: Temporal and spatial forecasting of ionospheric critical frequency using neural networks, *Radio Sci.*, 34, 1497–1506, doi:10.1029/1999RS900070, 1999.
- Lei, J., Burns, A. G., Tsugawa, T., Wang, W., Solomon, S. C., and Wiltberger, M.: Observations and simulations of quasiperiodic ionospheric oscillations and large-scale traveling ionospheric disturbances during the December 2006 geomagnetic storm, *J. Geophys. Res.*, 113, A06310, doi:10.1029/2008JA013090, 2008a.
- Lei, J., Thayer, J. P., Forbes, J. M., Wu, Q., She, C., Wan, W., and Wang, W.: Ionosphere response to solar wind high-speed streams, *Geophys. Res. Lett.*, 35, L19105, doi:10.1029/2008GL035208, 2008b.
- Liang, S.: The Deviation of IRI from the Ionospheric Parameters Observed Over East Asia, *Publications of the Yunnan Observatory*, 128, 1–+, 1990.
- Liu, C., Zhang, M., Wan, W., Liu, L., and Ning, B.: Modeling M(3000)F2 based on empirical orthogonal function analysis method, *Radio Sci.*, 43, RS1003, doi:10.1029/2007RS003694, 2008.
- Liu, L., Wan, W., and Ning, B.: Statistical modeling of ionospheric foF2 over Wuhan, *Radio Sci.*, 39, RS2013, doi:10.1029/2003RS003005, 2004.
- Liu, L., Wan, W., Ning, B., Pirog, O. M., and Kurkin, V. I.: Solar activity variations of the ionospheric peak electron density, *J. Geophys. Res.*, 111, A08304, doi:10.1029/2006JA011598, 2006.
- Liu, L., Wan, W., and Le, H.: Solar activity effects of the ionosphere: A brief review, *Chinese Science Bulletin*, 56, 1202–1211, doi:10.1007/s11434-010-4226-9, 2011.
- Lyatsky, W., Newell, P. T., and Hamza, A.: Solar illumination as cause of the equinoctial preference for geomagnetic activity, *Geophys. Res. Lett.*, 28, 2353–2356, doi:10.1029/2000GL012803, 2001.
- Mao, T., Wan, W., Yue, X., Sun, L., Zhao, B., and Guo, J.: An empirical orthogonal function model of total electron content over China, *Radio Sci.*, 43, RS2009, doi:10.1029/2007RS003629, 2008.
- Marsh, D. R., Solomon, S. C., and Reynolds, A. E.: Empirical model of nitric oxide in the lower thermosphere, *J. Geophys. Res.*, 109, A07301, doi:10.1029/2003JA010199, 2004.
- Matsuo, T. and Forbes, J. M.: Principal modes of thermospheric density variability: Empirical orthogonal function analysis of CHAMP 2001–2008 data, *J. Geophys. Res.*, 115, A07309, doi:10.1029/2009JA015109, 2010.
- Matsuo, T., Richmond, A. D., and Nychka, D. W.: Modes of high-latitude electric field variability derived from DE-2 measurements: Empirical Orthogonal Function (EOF) analysis, *Geophys. Res. Lett.*, 29, 1107, doi:10.1029/2001GL014077, 2002.
- Matsuo, T., Richmond, A. D., and Lu, G.: Optimal interpolation analysis of high-latitude ionospheric electrodynamic using empirical orthogonal functions: Estimation of dominant modes of variability and temporal scales of large-scale electric fields, *J. Geophys. Res.*, 110, A06301, doi:10.1029/2004JA010531, 2005.
- McIntosh, D. H.: On the Annual Variation of Magnetic Disturbance, *Philos. Trans. R. Soc. London. Ser. A.*, 251, 525–552, doi:10.1098/rsta.1959.0010, 1959.
- McKinnell, L. and Oyeyemi, E.: Progress towards a new global foF2 model for the International Reference Ionosphere (IRI), *Adv. Space Res.*, 43, 1770–1775, doi:10.1016/j.asr.2008.09.035, 2009.
- McKinnell, L. and Oyeyemi, E.: Equatorial predictions from a new neural network based global foF2 model, *Adv. Space Res.*, 46, 1016–1023, doi:10.1016/j.asr.2010.06.003, 2010.
- Mendillo, M., Rishbeth, H., Roble, R., Damboise, E., and Wroton, J.: Ionospheric variability originating from tropospheric and stratospheric sources, *EOS*, 79, 238, 1998.
- Mikhailov, A. V., Depuev, V. H., and Depueva, A. H.: Synchronous NmF2 and NmE daytime variations as a key to the mechanism of quiet-time F2-layer disturbances, *Ann. Geophys.*, 25, 483–493, doi:10.5194/angeo-25-483-2007, 2007.
- Mikhailov, A. V., Depueva, A. H., and Depuev, V. H.: Quiet time F2-layer disturbances: seasonal variations of the occurrence in the daytime sector, *Ann. Geophys.*, 27, 329–337, doi:10.5194/angeo-27-329-2009, 2009.
- Oyeyemi, E. O., Poole, A. W. V., and McKinnell, L. A.: On the global model for foF2 using neural networks, *Radio Sci.*, 40, RS6011, doi:10.1029/2004RS003223, 2005.
- Oyeyemi, E. O., McKinnell, L., and Poole, A.: Near-real time foF2 predictions using neural networks, *J. Atmos. Sol. Terr. Phys.*, 68, 1807–1818, doi:10.1016/j.jastp.2006.07.002, 2006.
- Pearson, K.: On Lines and Planes of Closest Fit to Systems of Points in Space, *Philosophical Magazine*, 2, 559–572, 1901.
- Petrukovich, A. A. and Zakharov, M. Y.:  $a_p$ -index solar wind driving function and its semiannual variations, *Ann. Geophys.*, 25, 1465–1469, doi:10.5194/angeo-25-1465-2007, 2007.
- Richards, P. G., Fennelly, J. A., and Torr, D. G.: EUVAC: a solar EUV flux model for aeronomic calculations., *J. Geophys. Res.*, 99(A5), 8981–8992, doi:10.1029/94JA00518, 1994.
- Rishbeth, H. and Mendillo, M.: Patterns of F2-layer variability, *J. Atmos. Sol. Terr. Phys.*, 63, 1661–1680, doi:10.1016/S1364-6826(01)00036-0, 2001.
- Roble, R. G. and Ridley, E. C.: A thermosphere-ionosphere-mesosphere-electrodynamics general circulation model (time-GCM): Equinox solar cycle minimum simulations (30–500 km), *Geophys. Res. Lett.*, 21, 417–420, doi:10.1029/93GL03391, 1994.
- Rush, C. M.: URSI foF2 model maps (1988), *Planet. Space Sci.*, 40, 546–547, doi:10.1016/0032-0633(92)90181-M, 1992.
- Schunk, R. and Nagy, A. (Eds.): *Ionospheres: Physics, Plasma Physics and Chemistry*, Cambridge University Press, 2000.
- Singer, W. and Dvinskikh, N. I.: Comparison of empirical models of ionospheric characteristics developed by means of different mapping methods, *Adv. Space Res.*, 11, 3–6, doi:10.1016/0273-1177(91)90311-7, 1991.
- Tóth, G., Sokolov, I. V., Gombosi, T. I., Chesney, D. R., Clauer,



- C. R., De Zeeuw, D. L., Hansen, K. C., Kane, K. J., Manchester, W. B., Oehmke, R. C., Powell, K. G., Ridley, A. J., Roussev, I. I., Stout, Q. F., Volberg, O., Wolf, R. A., Sazykin, S., Chan, A., Yu, B., and Kóta, J.: Space Weather Modeling Framework: A new tool for the space science community, *J. Geophys. Res.*, 110, A12226, doi:10.1029/2005JA011126, 2005.
- Vlasov, M. N. and Kelley, M. C.: Crucial discrepancy in the balance between extreme ultraviolet solar radiation and ion densities given by the international reference ionosphere model, *J. Geophys. Res.*, 115, A08317, doi:10.1029/2009JA015103, 2010.
- Xu, W. and Kamide, Y.: Decomposition of daily geomagnetic variations by using method of natural orthogonal component, *J. Geophys. Res.*, 109, A05218, doi:10.1029/2003JA010216, 2004.
- Xu, Z., Wu, J., Igarashi, K., Kato, H., and Wu, Z.: Long-term ionospheric trends based on ground-based ionosonde observations at Kokubunji, Japan, *J. Geophys. Res.*, 109, A09307, doi:10.1029/2004JA010572, 2004.
- Zapfe, B., Materassi, M., Mitchell, C., and Spalla, P.: Imaging of the equatorial ionospheric anomaly over South America—A simulation study of total electron content, *J. Atmos. Sol. Terr. Phys.*, 68, 1819–1833, doi:10.1016/j.jastp.2006.05.025, 2006.
- Zhang, D. H., Mo, X. H., Cai, L., Zhang, W., Feng, M., Hao, Y. Q., and Xiao, Z.: Impact factor for the ionospheric total electron content response to solar flare irradiation, *J. Geophys. Res.*, 116, A04311, doi:10.1029/2010JA016089, 2011.
- Zhang, M. L., Shi, J. K., Wang, X., Wu, S. Z., and Zhang, S. R.: Comparative study of ionospheric characteristic parameters obtained by DPS-4 digisonde with IRI2000 for low latitude station in China, *Adv. Space Res.*, 33, 869–873, doi:10.1016/j.asr.2003.07.013, 2004.
- Zhang, M., Shi, J., Wang, X., Shang, S., and Wu, S.: Ionospheric behavior of the F2 peak parameters *foF2* and *hmF2* at Hainan and comparisons with IRI model predictions, *Adv. Space Res.*, 39, 661–667, doi:10.1016/j.asr.2006.03.047, 2007.
- Zhang, M.-L., Liu, C., Wan, W., Liu, L., and Ning, B.: A global model of the ionospheric F2 peak height based on EOF analysis, *Ann. Geophys.*, 27, 3203–3212, doi:10.5194/angeo-27-3203-2009, 2009.
- Zhao, B., Wan, W., Liu, L., Yue, X., and Venkatraman, S.: Statistical characteristics of the total ion density in the topside ionosphere during the period 1996–2004 using empirical orthogonal function (EOF) analysis, *Ann. Geophys.*, 23, 3615–3631, doi:10.5194/angeo-23-3615-2005, 2005.