

Statistical modeling of ionospheric foF2 over Wuhan

Libo Liu, Weixing Wan, and Baiqi Ning

Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China

Received 11 November 2003; revised 16 January 2004; accepted 2 February 2004; published 8 April 2004.

[1] Half-hourly values of the critical frequency of the ionospheric F region, foF2, obtained at Wuhan Ionospheric Observatory (geographic 114.4°E, 30.6°N; 45.2° dip), China, during the period from 1957 to 1991 have been used to investigate the solar activity dependence of the monthly median foF2, and to construct single-station models using Fourier expansion and cubic-B splines approaches. The climatological models incorporate local standard time, month, and solar cycle variations of foF2. Statistical analyses show that, over Wuhan, the monthly median foF2 has a significantly nonlinear dependence on the current and historical solar activities. Introducing the monthly median F107 in the prior month can significantly improve the accuracy of our models. Thus the complex influence of solar activities is approximately expressed by a general nonlinear function. Our models are in good agreement with observations, with standard deviations between about 0.26 to 0.58 MHz. In contrast, the IRI model tends to overestimate foF2 over Wuhan and has a lower accuracy with standard deviations between about 0.5 to 1 MHz. *INDEX TERMS*: 2447 Ionosphere: Modeling and forecasting; 2479 Ionosphere: Solar radiation and cosmic ray effects; 6929 Radio Science: Ionospheric physics (2409); *KEYWORDS*: ionospheric foF2, F107, solar cycle variation, monthly median model

Citation: Liu, L., W. Wan, and B. Ning (2004), Statistical modeling of ionospheric foF2 over Wuhan, *Radio Sci.*, 39, RS2013, doi:10.1029/2003RS003005.

1. Introduction

[2] In both ionospheric researches and applications, developing empirical models is of significant importance to represent the variation patterns seen in the series of ionospheric data. Over decades, great efforts have been made to study the spatial and temporal variations of the Earth ionosphere, thus a large number of station-specific, regional, and global models have been developed greatly based on ionospheric characteristics data. These empirical models are established by performing statistical analysis on long-term measured data. An excellent review of many available empirical models recently has been presented by *Bilitza* [2002]. Among those models, the International Reference Ionosphere (IRI) model is most widely used, and is recognized as the standard specification of ionospheric parameters by the Committee on Space Research (COSPAR) and the International Union of Radio Science (URSI).

[3] Among ionospheric characteristic parameters, the critical frequency of the F region, foF2, is one of the most important ones. At middle and low latitudes, the

primary source of ionization in the F region is the solar extreme ultraviolet (EUV) irradiances. Thus ionospheric empirical models depend on how to describe the relationship between ionospheric characteristics and solar activities. The solar EUV irradiances are now known to show very definitive solar cycle variations. In the absence of long-term direct measurements of solar EUV irradiances, researchers have to rely on the extensive records of longer wavelength solar emissions or sunspot number as solar proxies. Sunspot number and the solar 10.7 cm irradiance index F107 are well-known and widely used indices. Recently the index E107 is suggested as a more accurate proxy compared to F107 for applications including empirical thermospheric models and ionospheric models [*Tobiska*, 2001]. It is designed to offer significant improvement as an index of the energy input to the thermosphere and ionosphere.

[4] The solar activity dependence of ionospheric characteristics has been found in the early routine ionospheric observations. Previous studies have shown that electron density increases with solar activity in a complicated way [e.g., *Balan et al.*, 1994; *Richards*, 2001; *Sethi et al.*, 2002; *Liu et al.*, 2003]. For example, differences on the upward and downgoing parts of a given solar cycle may lead to a so-called 'hysteresis effect' [*Rao and Rao*, 1969; *Triskova and Chum*, 1996]. Thus the physical

nature and dependence of the ionospheric characteristics on solar activity remains a subject for investigation. One objective of this work is to access the solar activity dependence of foF2 over Wuhan.

[5] Because global models may smear out features unique to a particular region [Holt *et al.*, 2002], station-specific models are very useful and widely viewed as the milestone of any ionospheric services [Pancheva and Mukhtarov, 1998]. There is no station-specific model for foF2 available yet based on observations at Wuhan or in the Chinese subcontinent. Another aim of this work is to construct station-specific models for Wuhan foF2 after statistically analyzing the response of foF2 over Wuhan to solar activity. The climatological models are successful in describing the foF2 features and their variations with local time, season and solar cycle. Further efforts will be made toward modeling foF2 with day-to-day variations and during disturbed periods.

2. Data

[6] Since 1957, ionosonde measurements have been routinely made at Wuhan Ionospheric Observatory (geographic 114.4°E, 30.6°N; 45.2° dip), which, located in central China, is just away from the northern crest of equatorial anomaly in East Asia. It has significant values for studying ionospheric dynamics in the equatorial anomaly region. The ionospheric data source for this study is foF2 which has been scaled from the routinely observed ionosonde records from 1957 to 1991. In order to check the ability of our models for the long-term prediction, we use data after the year of 1991 to validate the long-term prediction of the model.

[7] To investigate the foF2 variation, we start from the median patterns extracted from the database. In the present study, half-hourly values of monthly median foF2 are used to represent average conditions during 1957–1991 over Wuhan. Since monthly values should be representative for the ionospheric average behavior, in this analysis, we excluded those data if their median counts are less than 13. The monthly median F107 during 1957–1991 plotted in Figure 1 shows that there is an obviously regular 11-year variation due to the solar cycle variation of solar activity.

3. Statistical Relationship Between foF2 and Solar Proxies

[8] We first consider the statistical relationship between foF2 and solar proxies before constructing empirical models. Electron density in the ionospheric F region tends to increase with the increasing of solar activities due to the enhanced photoionization [e.g., Balan *et al.*, 1994; Kouris *et al.*, 1998; Richards, 2001; Sethi *et al.*, 2002; Liu *et al.*, 2003], whereas the daily and monthly

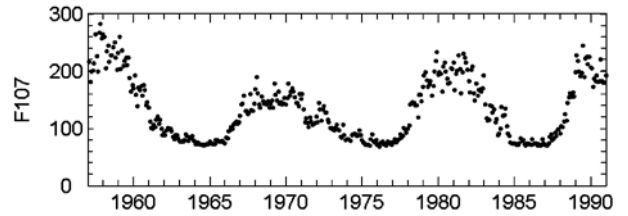


Figure 1. Monthly median F107 index during the interval 1957–1991.

median foF2 increases with solar activities in a rather complicated way [Kouris *et al.*, 1998; Richards, 2001]. Many researchers have investigated the solar-cycle variation of monthly median foF2 using different solar and ionospheric indices [Bilitza, 2000]. Monthly median foF2 linearly increases with the long-term solar activity, but more recent researches have shown evidences that foF2 saturates at extremely high solar epochs [e.g., Liu *et al.*, 2003]. It has also been found that the solar activity dependence of foF2 is related with latitude [Sethi *et al.*, 2002] and historical solar activity [Rao and Rao, 1969; Triskova and Chum, 1996].

[9] The parameters for this statistical analysis are the monthly median values. In the following, variables foF2, F107 and E107 denote the monthly median values for the sake of brevity. We use three regression methods to study the solar activity dependence of foF2 over Wuhan.

[10] The first regression model is a linear approximation to describe the relationship between F107 and foF2:

$$\text{foF2}(t)_m = A_m(t) + B_m(t) \cdot \text{F107}_m. \quad (1)$$

$A_m(t)$ and $B_m(t)$ are coefficients at given local standard time t for different month m . The linear relationship has been applied in many achieved models [e.g., Chen *et al.*, 2002; Gulyaeva, 1999; Holt *et al.*, 2002; Zolesi *et al.*, 1996]. The least squares regression analysis is taken on foF2 at specified local standard times (t) and given months (m) against F107, and sample results at local midnight (00 LST, local standard time over 120°E) and noon (12 LST) are represented with dotted lines in Figures 2 and 3.

[11] Improvement is expected from the second regression model, the quadratic relationship between F107 and foF2:

$$\text{foF2}(t)_m = A_m(t) + B_m(t) \cdot \text{F107}_m + C_m(t) \cdot \text{F107}_m^2. \quad (2)$$

$A_m(t)$, $B_m(t)$ and $C_m(t)$ are coefficients at specified local standard time t and month m . Sample fit results are represented with solid lines in Figures 2 and 3. This relationship has been used by Sethi *et al.* [2002] to investigate the latitudinal dependence of solar cycle

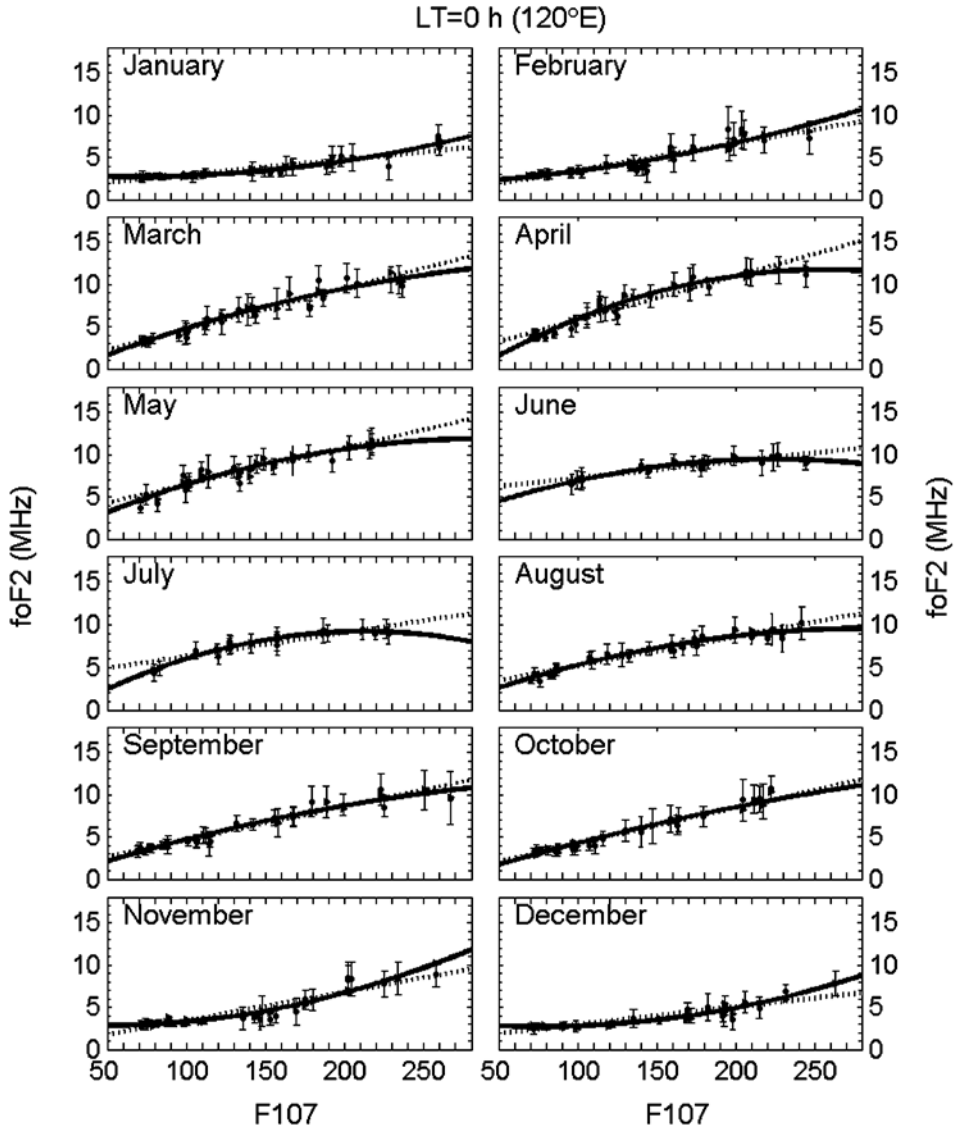


Figure 2. The responses of the monthly median foF2 to solar activities represented by monthly median F107 at local time 0 h (120°E) for the whole interval of 1957–1991 over Wuhan. Observed data are plotted with circle points with error bars. Dotted and solid lines represent the linear and quadratic fits, respectively. Triangle-dot points are for the third fit: $foF2 = A + B \cdot F107 + C \cdot F107^2 + D \cdot F107_p + E \cdot F107_p^2$.

variations of foF2 at stations covering from low- to midlatitudes. Our analysis for Wuhan foF2 supports their conclusion that, for low-latitudes, a second-degree fit makes much improvement.

[12] Further significant improvement is expected from a more strategic use of the parameter Kr, which is introduced by *Pancheva and Mukhtarov* [1998] to describe the tendency of solar activity change in empirical model. However, it is inconvenient in nowcasting/fore-

casting because Kr may be unknown. A new finding of our analyses is that almost identical or even higher accuracy can be achieved by only replacing Kr with F107p, the monthly median F107 in the prior month. Therefore the third regression model can be expressed as

$$foF2(t)_m = A_m(t) + B_m(t) \cdot F107_m + C_m(t) \cdot F107_m^2 + D_m(t) \cdot F107_{p_m} + E_m(t) \cdot F107_{p_m}^2 \quad (3)$$

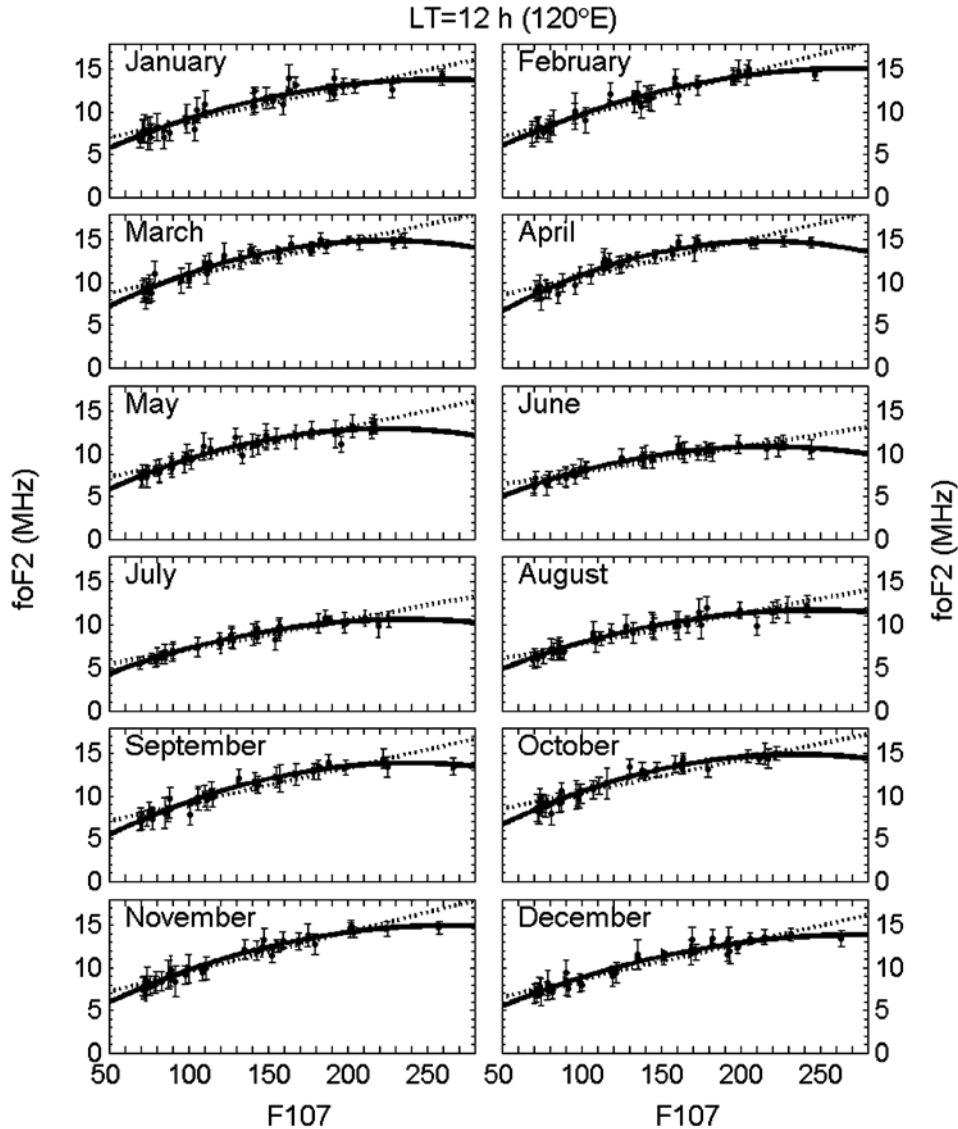


Figure 3. Same as Figure 2, but for local time 12 h (120°E).

[13] Triangular points in Figures 2 and 3 illustrated the sample results with the third fit model. Evidently the third fit is the best. The agreement between modeled and observed foF2 is quite remarkable when the parameter F107p are taken into consideration. In other words, it has shown that the historical solar activity may have an important influence on foF2 over Wuhan, but its effect depends on month (Figure 4).

[14] Figure 4 illustrates the diurnal variations of standard deviations for these regression fitted models from the observed foF2 for 1957–1991. Over Wuhan the solar cycle dependence of foF2 is obviously nonlinear, because Figures 2–4 show that, in general, a second degree fit gives a significantly better correlation and a much lower

standard deviation than the linear fit. The most significant decrease in standard deviation occurs in October from around 1.35 MHz to 0.5 MHz. Over Wuhan, foF2 generally shows either saturation or a decrease for high solar epoch. But in some months (like winter months in Figure 2) nighttime foF2 tends to increase with increasing solar activities. This feature has not been reported yet. However, it needs further confirmation due to too sparse data at extremely high solar epoch. Introducing F107p generally further improves the description of solar cycle dependence of foF2. Figure 4 reveals that, over Wuhan, the influence of historical solar activity on foF2 is most distinct in February and March, and less obvious at most time in July and October.

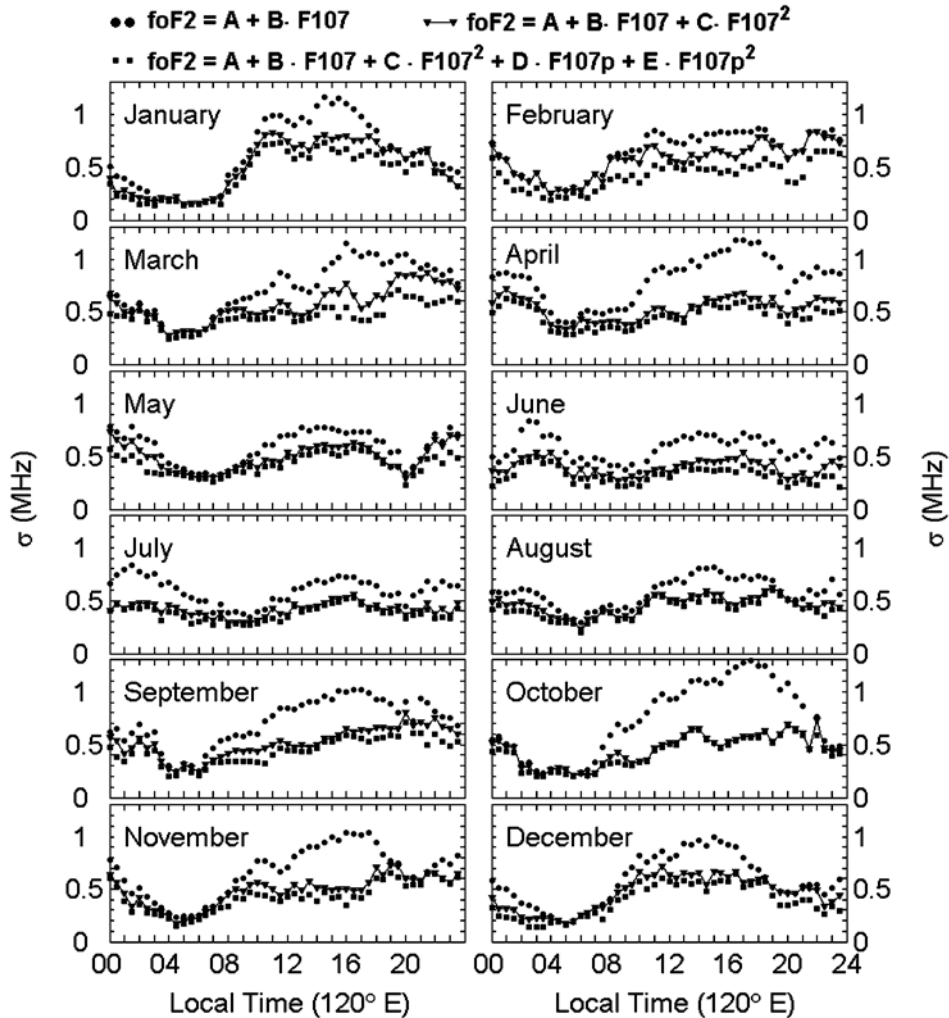


Figure 4. Diurnal variation of standard deviations of the values of three regression fits from observed foF2 for the whole interval of 1957–1991 over Wuhan.

[15] We also made analogous analyses on foF2 over Wuhan against E107 (figures not shown). Identical features can be found regardless of taking F107 or E107 as solar proxies, indicating no improvements with introducing E107 for empirical modeling long-term foF2 data over Wuhan. The analogous conclusion has made for the Millstone Hill incoherent scatter data by *Zhang and Holt* [2002]. Therefore, in our models, the monthly median F107 is still used as the indicator for the level of solar activity.

4. Models

[16] Empirical models use appropriate mathematical functions to reproduce the characteristic variations

seen in long-term data records. The complex influence of solar activity on foF2 over Wuhan is expressed as a general nonlinear approximation (equation (3) in Section 3).

[17] Because of the rather complicated nature of the ionosphere, there have been numerous approaches to ionospheric modeling over decades. We construct station-specific models for foF2 over Wuhan using two algorithms, which are based on Fourier analysis and cubic-B splines methods, respectively. Four input parameters of our models are t , m , F107 and F107p, which represent the diurnal, season-month, and solar cycle variations and the influence of historical solar activity on foF2. Here t is local standard time over 120°E, m is the given month, and F107 and F107p are the monthly

median values of F107 in the specified and previous months, respectively.

4.1. Fourier Model

[18] Since there are diurnal periodicities inherent in all the geophysical data, Fourier expansion generally has the highest priority in constructing empirical models. The diurnal variations can be well expressed by a Fourier expansion of cosine and sine functions with periods of 24 hours and higher harmonics. The formulation of the Fourier model for foF2 can be generally expressed as

$$foF2(t)_m = c_{0,m} + \sum_{n=1}^N \left(c_{n,m} \cos \frac{2\pi nt}{T} + s_{n,m} \sin \frac{2\pi nt}{T} \right), \quad (4)$$

where $n(= 0, 1, \dots, N)$ is the harmonic number and T is equal to 24 hours. The spectral coefficients, $c_{n,m}$ and $s_{n,m}$, are a function of F107 and F107p. In the model, these coefficients are approximately expressed as equation (3), and can be easily estimated with a least squares regression analysis for all the foF2 data set in a specified month. It is found that only the first three harmonic components are most significant in foF2 over Wuhan, but we take the highest harmonic number, N , up to 5 in our model to insure our model precision.

4.2. Cubic-B Splines Model

[19] The second approach is the cubic-B splines method (called splines model for short). The cubic-B splines analysis has been successfully used by *Scherliess and Fejer* [1999] in developing a model of global equatorial F region vertical drifts, and by *Fejer and Scherliess* [1997] in the empirical models of storm time equatorial zonal electric fields.

[20] In the splines model, the foF2 data can be expressed with the univariate normalized cubic-B splines of order four, $N_{i,4}(t)$, as

$$foF2(t)_m = \sum_{n=1}^N \Gamma_{n,m} N_{i,4}(t). \quad (5)$$

The basis functions $N_{i,4}(t)$ add up to unity at all local standard time, t , and are nonvanishing over limited time intervals. Here again, the coefficients $\Gamma_{n,m}$ in equation (5) are approximately expressed as equation (3). These coefficients can be easily estimated with a least squares regression analysis for all the measured foF2 data set in the given month, m . The analysis is constrained to make the fit periodic in 24 hours.

[21] Figure 5 illustrates the functions $N_{i,4}(t)$ and the node positions. The node positions should be properly selected according to the diurnal variation of what be fitted, as suggested by *Scherliess and Fejer* [1999].

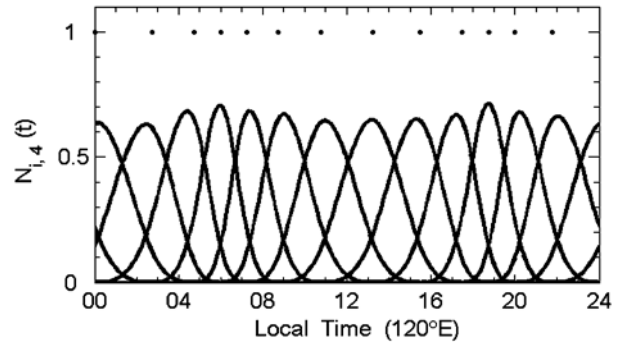


Figure 5. The nodes (circle points) and basis cubic-B splines functions.

Readers may refer to the work of *Scherliess and Fejer* [1999] for more information on the cubic-B splines method. To better reproduce the diurnal pattern of foF2 over Wuhan, we take 13 nodes in our splines model and place those nodes at local standard times 0, 2.75, 4.75, 6, 7.25, 8.75, 10.75, 13.25, 15.5, 17.5, 18.75, 20, and 21.75.

[22] In essence, the analyses mentioned in Section 3 can be used to predict the half-hourly values of foF2 at given months. The approaches we described in this section may reduce the number of model coefficients and predict values at any time.

5. Verification of the Models

[23] Sample values of the observed (circles) and modeled (lines) foF2 are demonstrated in Figure 6, left panels for the splines model, and right for the IRI model. Years for high solar activity (1958), low solar activity (1965), the descending part of solar cycle (1973) and the ascending part of solar cycle (1978) were chosen. Results of the Fourier model are not plotted in Figure 6 for brevity because they are little different from the splines model. It should be emphasized that, when we predict foF2 with our model, the observed data in the specified year are not included in the database used for the model coefficients generation. We also found that the model result almost does not change whether data in the year is included or not. foF2 in 2001 over Wuhan are also plotted in Figure 6 to illustrate the model ability in long-term prediction.

[24] To verify the quality of our models described above and compare with the IRI model, we calculate half-hourly values of foF2 with the Fourier model, the splines model and IRI model over Wuhan in each month for the whole interval 1957–1991. Figure 7 depicts the error distributions and standard deviations of relative deviations of model predictions against observations

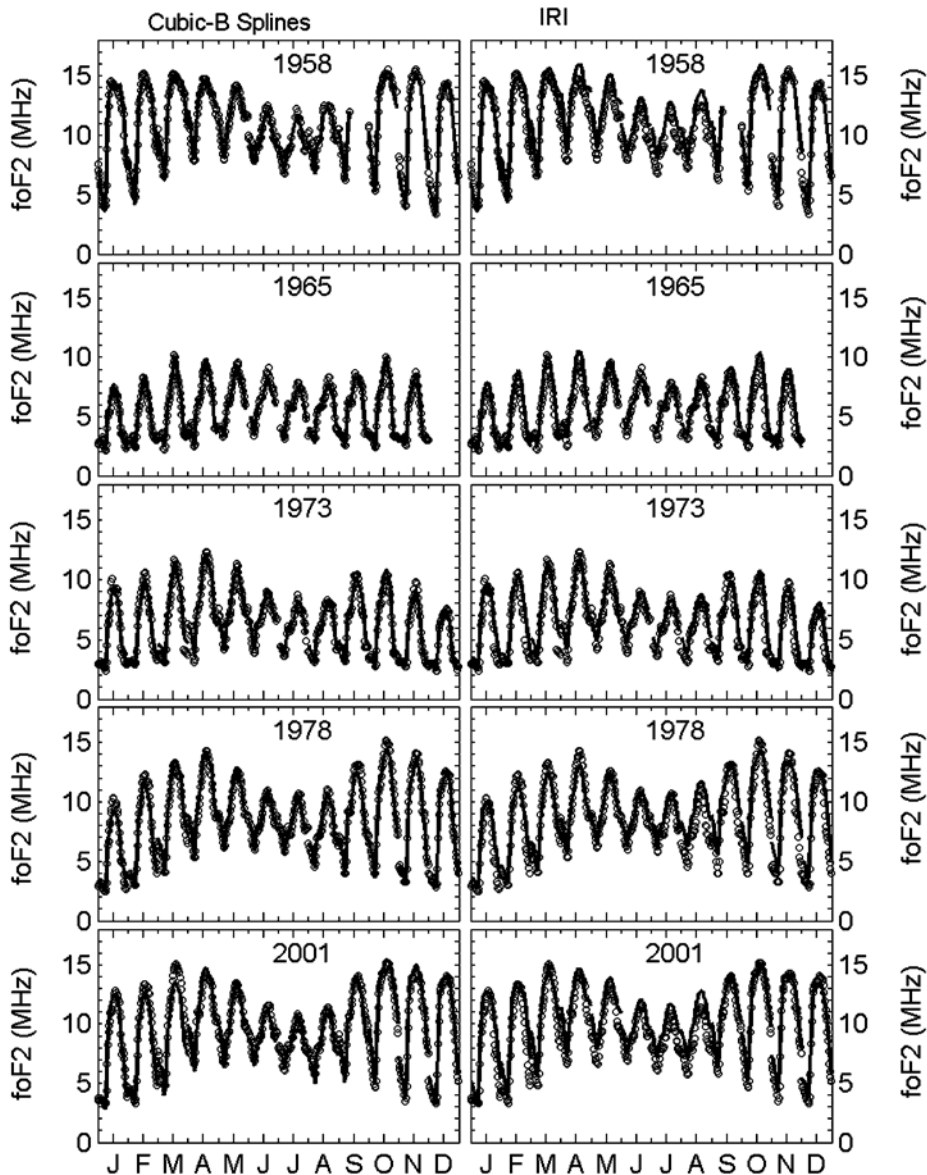


Figure 6. Monthly medians of half-hourly values of observed (open circles) and modeled (lines) foF2 for high solar activity (1958, 2001), low solar activity (1965), the descending part of solar cycle (1973), and the ascending part of solar cycle (1978).

for all local standard time during the whole interval of 1957–1991 in given months. Corresponding diurnal and yearly variations of standard deviations are illustrated in Figure 8.

[25] When validating models' performance illustrated in Figures 7 and 8, it should be noted that the Fourier model and the splines model are of 13 and 11 coefficients, respectively. It is found that the accuracy of the splines model (middle panels in Figure 7) is slightly higher than

that of the Fourier model (left panels in Figure 7), when the splines nodes are properly placed and the numbers of model coefficients are taken as the same. Errors for our models distribute from about -2 to 2 MHz, and their standard deviations lie around 0.26 to 0.58 MHz.

[26] In contrast, the IRI model has a lower accuracy. The error distributions of the IRI model shift to the positive side, which means that the IRI model tends to overestimate foF2 over Wuhan. At the same time,

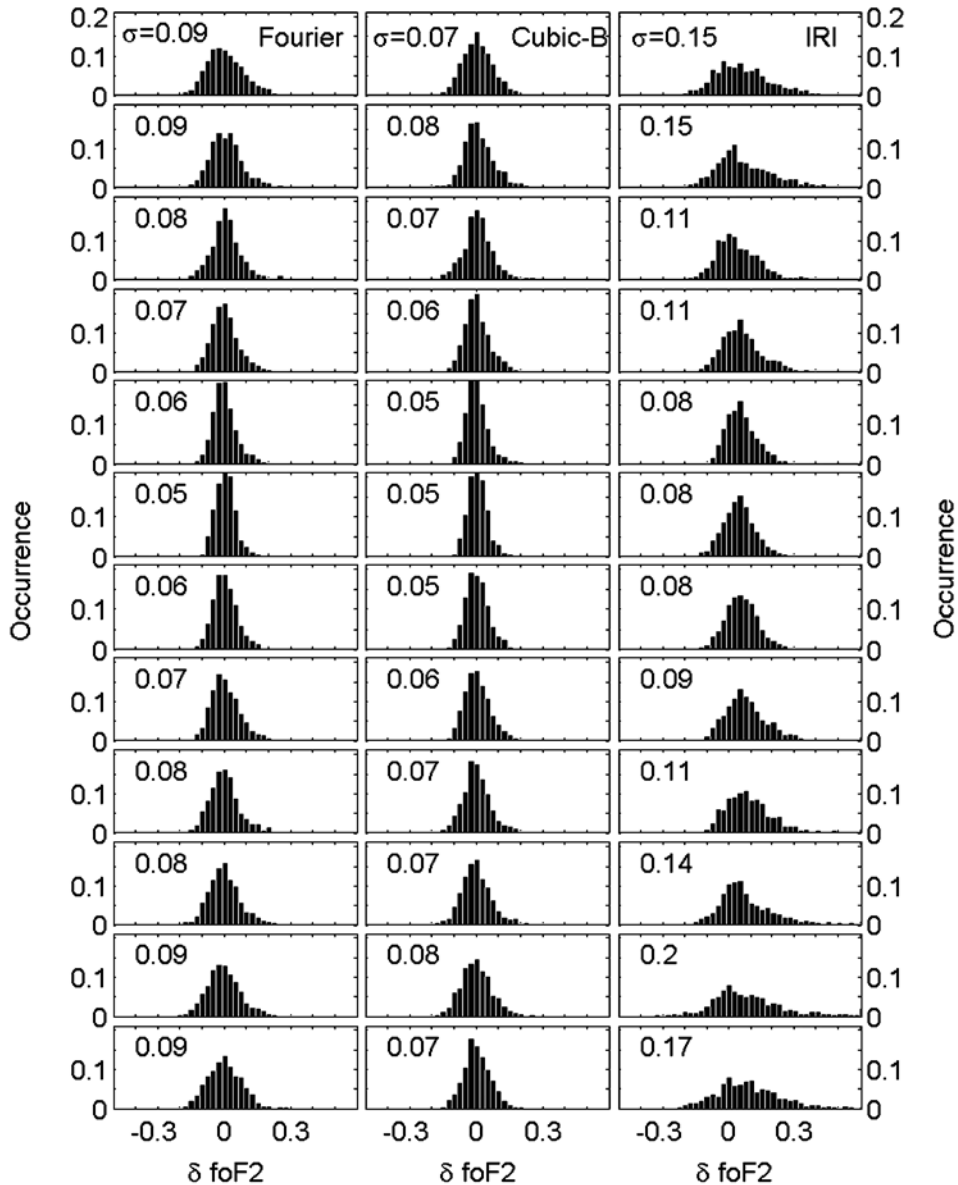


Figure 7. The distributions and standard deviations of relative deviations of model values from observed foF2 over Wuhan during the whole interval of 1957–1991 in months. The results for the Fourier, Cubic-B splines, and IRI models are illustrated in the left, middle, and right panels, respectively. Panels from top to bottom are for months from January to December, respectively.

standard deviations of the IRI model are also higher, lying from around 0.5 to 1 MHz. It is interesting that the IRI model also overestimates TEC over Wuhan [Chen et al., 2002].

[27] As a global empirical model, the IRI model is successful at many locations. However, due to the rather complicated nature of the ionosphere, sparse Chinese data used in the IRI model [Wu et al., 1996], and the limitation

of global empirical models, our models provide a higher accuracy than the IRI model. Figures 6–8 have basically demonstrated the advantage of our models.

6. Summary

[28] Long-term routine ionosonde observation over Wuhan and lack of available models of foF2 based on

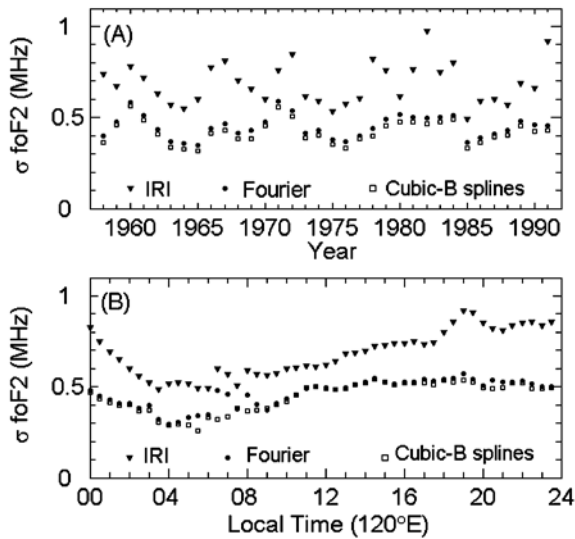


Figure 8. The (a) yearly and (b) diurnal variation of standard deviations of model values from observed foF2 for the whole interval of 1957–1991 over Wuhan. Results for the Fourier, Cubic-B splines, and IRI models are illustrated with solid circles, open squares, and solid diamonds, respectively.

initial observations activate us to develop a single station foF2 model over Wuhan. The models incorporate local standard time, month and solar cycle variations of foF2. The following conclusions can be drawn from the present investigation:

[29] 1. There is a nonlinear relationship between monthly median F107 and foF2 over Wuhan. A second-degree fit gives a much better correlation than the linear fit. Similar features can be found by analogously analyzing the relationship between foF2 and E107, indicating no improvements on empirical modeling long-term foF2 data over Wuhan replacing F107 with E107 as an indicator of the level of solar activity.

[30] 2. Historical level of solar activity may significantly influence monthly median foF2 over Wuhan. Introducing monthly median F107 in the prior month can significantly improve the description of the solar cycle dependence of foF2. This adjustment also makes the model prediction better than the Kr characteristic [Pancheva and Mukhtarov, 1998], at least for Wuhan station.

[31] 3. Two models are successfully constructed using the Fourier expansion and cubic-B splines methods respectively, which are verified with standard deviations between about 0.26 to 0.58 MHz. In contrast, the IRI model has a lower accuracy with standard deviations of about 0.5 to 1 MHz and tends to overestimate foF2 over Wuhan.

[32] Good agreements have been found between the observed monthly median foF2 and our models. The complex influence of the solar cycle characteristics is expressed by a general nonlinear approximation. We recommend the approximation can be applied in other models to improve their predictions. Our model can also be used as a suitable tool to study the basic physical processes controlling the *F* region of the Earth's ionosphere and as a reference for short-term and disturbance predictions. In future, the cubic-B splines functions will be applied to construct empirical models for storm-time foF2 over Wuhan.

[33] **Acknowledgments.** The codes of the IRI models are provided by the World Data Center-A. The Ap and F107 indices are downloaded from the SPIDR web site <http://spidr.ngdc.noaa.gov/>. The E107 index is downloaded from the web site <http://www.SpaceWx.com/>. This research was supported by the National Natural Science Foundation of China (40274054, 40134020) and Important Basis Research Project of China (G2000078407).

References

- Balan, N., G. J. Bailey, B. Jenkins, P. B. Rao, and R. J. Moffett (1994), Variations of ionospheric ionization and related solar fluxes during an intense solar cycle, *J. Geophys. Res.*, *99*(A2), 2243–2253.
- Bilitza, D. (2000), The importance of EUV indices for the International Reference Ionosphere, *Phys. Chem. Earth*, *25*(5–6), 515–521.
- Bilitza, D. (2002), Ionospheric models for radio propagation studies, in *The Review of Radio Science 1999–2002*, edited by W. R. Stone, pp. 625–679, IEEE Press, Piscataway, N. J.
- Chen, Y., W. Wan, L. Liu, and L. Li (2002), A statistical TEC model based on the observation at Wuhan Ionospheric Observatory, *Chin. J. Space Sci.*, *34*(1), 27–35.
- Fejer, B. G., and L. Scherliess (1997), Empirical models of storm-time equatorial zonal electric fields, *J. Geophys. Res.*, *102*, 24,047–24,056.
- Gulyaeva, T. L. (1999), Regional analytical model of ionospheric total electron content: Monthly mean and standard deviation, *Radio Sci.*, *34*(6), 1507–1512.
- Holt, J. M., S.-R. Zhang, and M. J. Buonsanto (2002), Regional and local ionospheric models based on Millstone Hill incoherent scatter radar data, *Geophys. Res. Lett.*, *29*(8), 1207, doi:10.1029/2002GL014678.
- Kouris, S. S., P. A. Bradley, and P. Dominici (1998), Solar-cycle variation of the daily foF2 and M(3000)F2, *Ann. Geophys.*, *16*, 1039–1042.
- Liu, J. Y., Y. I. Chen, and J. S. Lin (2003), Statistical investigation of the saturation effect in the ionospheric foF2 versus sunspot, solar radio noise, and solar EUV radiation, *J. Geophys. Res.*, *108*(A2), 1067, doi:10.1029/2001JA007543.

- Pancheva, D., and P. Mukhtarov (1998), A single-station spectral model of the monthly median foF2 and M(3000)F2, *Stud. Geophys. Geod.*, *42*, 183–196.
- Rao, M. S. J. G., and R. S. Rao (1969), The hysteresis variation in F2-layer parameters, *J. Atmos. Terr. Phys.*, *31*, 1119–1125.
- Richards, P. G. (2001), Seasonal and solar cycle variations of the ionospheric peak electron density: Comparison of measurement and models, *J. Geophys. Res.*, *106*(A12), 12,803–12,819.
- Scherliess, L., and B. G. Fejer (1999), Radar and satellite global equatorial *F* region vertical drift model, *J. Geophys. Res.*, *104*(A4), 6829–6842.
- Sethi, N. K., M. K. Goel, and K. K. Mahajan (2002), Solar cycle variations of foF2 from IGY to 1990, *Ann. Geophys.*, *20*, 1677–1685.
- Tobiska, W. K. (2001), Validating the solar EUV proxy, E107, *J. Geophys. Res.*, *106*(A12), 29,969–29,978.
- Triskova, L., and J. Chum (1996), Hysteresis in dependence of foF2 on solar indices, *Adv. Space Res.*, *18*(6), 145–148.
- Wu, J., K. Quan, K. Dai, F. Luo, X. Sun, Z. Li, C. Cao, R. Liu, and C. Shen (1996), Progress in the study of the Chinese reference ionosphere, *Adv. Space Res.*, *18*(6), 187–190.
- Zhang, S. R., and J. M. Holt (2002), Empirical modeling ionospheric electron density variations using F107, E107 and MgII indices based on long-term incoherent scatter radar observations over Millstone Hill, paper presented at 4th (Virtual) Thermospheric/Ionospheric Geospheric Research (TIGER) Symposium, Internet, 10–14 June.
- Zolesi, B., L. R. Cander, and G. D. Franeschi (1996), On the potential applicability of the simplified ionospheric regional model to different midlatitude areas, *Radio Sci.*, *31*(3), 547–552.

L. Liu, B. Ning, and W. Wan, Institute of Geology and Geophysics, Chinese Academy of Sciences, 100029 Beijing, China. (liul@mail.igcas.ac.cn; lliu@wipm.ac.cn; nbq@mail.igcas.ac.cn; wanw@mail.igcas.ac.cn)