A New Empirical Model of NmF2 Based on CHAMP, GRACE, and COSMIC Radio Occultation

Zhendi Liu 1, Hanxian Fang 1,2,*, M. M. Hoque 3, Libin Weng 1, Shenggao Yang 4 and Ze Gao 1

1 Institute of Meteorology and Oceanography, National University of Defense Technology, Nanjing 211101, China; liuzhendi18@nudt.edu.cn (Z.L.); wenglb@hotmail.com (L.W.); Gaozenudt@hotmail.com (Z.G.)
2 State Key Laboratory of Space Weather, Chinese Academy of Sciences, Beijing 100190, China
3 Institute of Communications and Navigation, German Aerospace Center, Kalkhorstweg 53, 17235 Neustrelitz, Germany; mainul.hoque@dlr.de
4 State Key Laboratory of Astronautic Dynamics, Xi’an Satellite Control Center, Xi’an 710043, China; yangsgg@nudt.edu.cn
* Correspondence: fanghx@hit.edu.cn

Received: 8 March 2019; Accepted: 21 May 2019; Published: 11 June 2019

Abstract: To facilitate F2-layer peak density (NmF2) modeling, a nonlinear polynomial model approach based on global NmF2 observational data from ionospheric radio occultation (IRO) measurements onboard the CHAMP, GRACE, and COSMIC satellites, is presented in this paper. We divided the globe into 63 slices from 80°S to 80°N according to geomagnetic latitude. A Nonlinear Polynomial Peak Density Model (NPPDM) was constructed by a multivariable least squares fitting to NmF2 measurements in each latitude slice and the dependencies of NmF2 on solar activity, geographical longitude, universal time, and day of year were described. The model was designed for quiet and moderate geomagnetic conditions (Ap ≤ 32). Using independent radio occultation data, quantitative analysis was made. The correlation coefficients between NPPDM predictions and IRO data were 0.91 in 2002 and 0.82 in 2005. The results show that NPPDM performs better than IRI2016 and Neustrelitz Peak Density Model (NPDM) under low solar activity, while it undergoes performance degradation under high solar activity. Using data from twelve ionosonde stations, the accuracy of NPPDM was found to be better than that of NPDM and comparable to that of IRI2016. Additionally, NPPDM can well simulate the variations and distributions of NmF2 and describe some ionospheric features, including the equatorial ionization anomaly, the mid-latitude trough, and the wavenumber-four longitudinal structure.

Keywords: F2 peak density; empirical ionospheric model; nonlinear polynomial; radio occultation

1. Introduction

The ionosphere is composed of several layers, namely the D-layer (50–90 km), E-layer (90–140 km), F1-layer (140–210 km), and F2-layer (210–1000 km) [1]. Among these layers, the F2-layer is the most ionized region, having the highest electron density variability. Therefore, the F2-layer is primarily responsible for most of the ionospheric effects, such as signal delay and ray path bending. The F2-layer peak electron density (NmF2) is one of the most important parameters to characterize the ionosphere. It is linearly related to the square of the critical frequency (foF2) of radio wave propagation. The variation of foF2 can result in the interruption or even loss of high frequency signal transmissions. In addition, the research on NmF2 is beneficial to reducing the influence of ionospheric propagation delay on global positioning. Thus, the performance of an NmF2 model is of great significance for ionospheric research and global navigation applications.

NmF2 may vary from about $10^{10}$ to $10^{13}$ m$^{-3}$ during the daytime and decrease during nighttime [2]. The peak density height (hmF2) may range from 350 to 500 km at the equatorial latitudes and from 250
to 350 km at mid-latitudes. The variation of NmF2 is mainly affected by local time, season, sunspot numbers, geographical location, and solar disturbances.

To develop new NmF2 models or improve the existing models, researchers have conducted some interesting studies [3–9]. The International Reference Ionosphere (IRI) [6] and NeQuick model [10,11] are amongst the widely used empirical models. IRI uses CCIR (International Radio Consultative Committee) coefficients [12] and URSI (International Union of Radio Science) coefficients [13] to predict foF2. The CCIR and URSI coefficients were obtained from observations from the existing worldwide network of ionospheric sounding stations. Araujo-Pradere et al. [14] developed an NmF2 model for geomagnetic storms periods. Oyeyemi and McKinnell [15] built an foF2 model that can be used as a replacement for the CCIR and URSI coefficients within the IRI model. Many authors have used empirical orthogonal functions (EOF) to develop models for ionosphere parameter predictions and verified their results [16–18]. A new approach based on the neural network method was also tested for the modeling of foF2 [19–23]. The results showed that it was possible to forecast foF2 in near-real time (5 h in advance) by using this method. Recently, Sai and Tulasi [24] developed the Artificial Neural Network-Based Ionospheric Model (ANNIM) based on COSMIC (Constellation Observing System for Meteorology, Ionosphere and Climate) data to predict NmF2. They used the artificial neural network (ANN) method, a machine learning technique that is capable of solving highly nonlinear problems by systematic learning mechanisms. Tulasi et al. [25] then presented an improved version of ANNIM by including CHAMP (Challenging Minisatellite Payload) and GRACE (Gravity Recovery And Climate Experiment) IRO measurements as well as ionosonde data as data sources. They assessed the performance of the ANNIM by comparing it with the IRI-2016 model, and also showed that the ANNIM well reproduced the spatial and temporal variations of NmF2 and captured the ionospheric anomalies such as the equatorial ionosphere anomaly (EIA) and the mid-latitude summer night-time anomaly (MSNA). Themens et al. [26] used Fourier expansions and a spherical cap harmonic expansion for the quiet time NmF2 model and just a spherical cap harmonic expansion for the perturbed time NmF2 model based on ionosonde and IRO measurements. These models provided NmF2 predictions for a region above 50°N geomagnetic latitude. They demonstrated that their model provided improvements over IRI and was capable of representing auroral enhancements and the midlatitude trough.

Hoque and Jakowski [27] made an inspiring and successful attempt to develop a nonlinear model (Neustrelitz Peak Density Model—NPDM) with only 13 coefficients and a few empirically fixed parameters, which described the dependencies of NmF2 on local time, geographic/geomagnetic location, and solar irradiance and activity. The NPDM was proved to be a very efficient model with reasonable performance.

In this study, a nonlinear polynomial model approach was also used to develop the Nonlinear Polynomial Peak Density Model (NPPDM). We used the data of ionospheric radio occultation (IRO) measurements as the database, which included CHAMP, GRACE, and COSMIC observations. The model was based on a set of non-linear equations. Unlike the NPDM model, NPPDM included the longitudinal variation of NmF2. Other input variables were latitude, solar activity, day of year, and universal time. The latitudinal dependence of NmF2 was not described by an independent equation in our model. We divided the globe into 63 slices according to the geomagnetic latitude. The coefficients were derived in each slice from a multivariable fit to NmF2 measurements in least squares sense. To evaluate the fitting quality, the residuals of NPPDM were analyzed. Independent observations from IRO and twelve ionosonde stations worldwide were used for accuracy comparisons between NPPDM, IRI2016, and NPDM. In addition, some features of the spatial and temporal variations of NmF2 were simulated by NPPDM for further validations.

2. Data

The IRO technique has the advantages of limb sounding geometry, global coverage, and high vertical resolution in comparison with traditional ionosphere monitoring methods. Under certain assumptions, the total electron content (TEC) derived from the dual-frequency observations along
the GPS ray can be inverted to the electron density profile (EDP) assigned to the tangent points of
the rays by a traditional Abel inversion [28,29]. Investigations by several authors [30,31] showed
that accurate hmF2 and NmF2 of the F2-layer can be obtained from occultation observations through
the Abel transformation. The studies of Lei et al. [32], Angling [33], and Krankowski et al. [34]
showed that the NmF2 and hmF2 derived from IRO technique are consistent with the observations of
ionosonde and incoherent scatter radar. It should be noted that although the retrieved NmF2 by IRO
techniques are generally in good agreement with ionosonde observations at middle latitudes [30,35],
the accuracy of the Abel inverted electron density degrades at low latitudes, especially at the EIA region.
The NmF2 inversion error in the EIA region is characterized by overestimation in the trough region
and underestimation in the crest region due to the smoothing operation during the retrieval process.
The relative error of NmF2 could account for up to 20%, with most amplitudes within ±10% [30].

In this study, the peak electron densities of EDPs were used as the database for NmF2 modeling.
The EDPs from CHAMP and GRACE were provided by the German Aerospace Center (DLR) Neustrelitz
(http://swaciweb.dlr.de/). The EDPs from COSMIC were processed by COSMIC Data Analysis and
Archive Center (CDAAC, http://cosmic-io.cosmic.ucar.edu/cdaac/index.html). We selected profiles
that satisfied the following criteria: (1) The daily geomagnetic activity index (Ap) was less than 32 to
avoid the disturbed NmF2 due to geomagnetic storm events; (2) Observations in the polar regions
(80°–90° N/S) were removed due to data deficiency; (3) A least-squares fitting to a two-layer Chapman
function was performed to each EDP [36]:

\[ Ne(h) = NmF2 \exp \left\{ 0.5\left[ 1 - z - \exp(-z) \right]\right\}, \] (1)
\[ z = (h - hmF2) / H(h), \] (2)

We set the Chapman scale height as \( H(h) = A_1(h - hmF2) + H_m \) for the bottomside and \( H(h) = \)
\( A_2(h - hmF2) + H_m \) for the topside. Here, \( H_m \) is the value of \( H(h) \) at \( hmF2 \), while \( A_1 \) and \( A_2 \)
are coefficients.

After the fitting, the same criteria proposed by Yue et al. [36] was used: if the fitted NmF2 (larger
than \( 10^7 / \text{cm}^3 \)), hmF2 (higher than 600 km or lower than 150 km), Hm (larger than 150 km) values were
unreasonable or the difference between the fitted NmF2/hmF2 and that derived from the EDP directly
by taking the maximum of the profile was larger than 20%, the EDP will be discarded. This procedure
will ensure the integrity of the results in the study because those outliers might mean something
wrong in the profile observation or processing. The final details of the selected database are listed in
Table 1.

<table>
<thead>
<tr>
<th>Space Mission</th>
<th>Number of Data Used</th>
<th>Time Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAMP</td>
<td>( 1.7 \times 10^5 )</td>
<td>2001.06–2008.09</td>
</tr>
<tr>
<td>GRACE</td>
<td>( 1.1 \times 10^5 )</td>
<td>2007.02–2013.09</td>
</tr>
<tr>
<td>COSMIC</td>
<td>( 3.8 \times 10^6 )</td>
<td>2006.04–2018.04</td>
</tr>
</tbody>
</table>

Distributions of the database with the solar activity (F10.7 index), day of year, geographical latitude,
geographical longitude, magnetic latitude, and universal time are shown in Figure 1. From Figure 1,
it can be clearly seen that most of the data are distributed in mid and low latitudes and few are
distributed in high latitudes. Similarly, the data under low solar activity are in the majority compared
with that under high solar activity.

Moreover, we also used the NmF2 observations from different ionosonde stations provided by
the Space Physics Interactive Data Resource (SPIDR, http://spidr.ngdc.noaa.gov/spidr/). Since the
ionosonde data from SPIDR were used just for model validation and comparison, the geographical
coordinates of the specific stations and other information of SPIDR data will be shown in Section 4.
3. Method

A set of nonlinear polynomial equations describing different dependencies of NmF2 was used to construct the NmF2 model. The NmF2 was dependent on geographical longitude (Lon), geographical latitude (Lat), solar activity (F10.7), universal time (UT), and season (denoted by DOY, namely day of year). The nonlinear polynomial model was based on the assumption that NmF2 varies with these geophysical elements independently. This assumption has been successfully applied to various empirical models including the topside ionosphere scale height model, the total electron content model, and the thermospheric density model [37–41]. In our model, the functions of NmF2 on various variables were expressed as

\[
NmF2 = F_1(F10.7)F_2(Lon)F_3(UT)F_4(Doy),
\]

\[F_1 = a_1 + a_2F10.7 + a_3F10.7^2,
\]

\[F_2 = b_1 + \sum_{i=1}^{6}\left(b_{2i}\cos(i\cdot\frac{2\pi Lon}{360}) + b_{3i}\sin(i\cdot\frac{2\pi Lon}{360})\right),
\]

\[F_3 = c_1 + \sum_{i=1}^{3}\left(c_{2i}\cos(i\cdot\frac{2\pi UT}{24}) + c_{3i}\sin(i\cdot\frac{2\pi UT}{24})\right),
\]

\[F_4 = d_1 + \sum_{i=1}^{2}\left(d_{2i}\cos(i\cdot\frac{2\pi DOY}{365.25}) + d_{3i}\sin(i\cdot\frac{2\pi DOY}{365.25})\right),
\]

Figure 1. Data distributions with F10.7, day of year, geographical latitude, geographical longitude, magnetic latitude, and universal time.

The longitudinal (F2) was approximated by a 6-order trigonometric function to be able to capture the wave-4 longitudinal variation, which is similar to the study of Liu et al. [41]. A three-order expression was adopted in F3 to represent daily variations including diurnal, semi-diurnal, and terdiurnal terms, which was similar to the study of Hoque and Jakowski [27]. The principal seasonal variations of

Remote Sens. 2019, 11, x FOR PEER REVIEW 4 of 22

Moreover, we also used the NmF2 observations from different ionosonde stations to const... which is similar to the study of Liu et al [41]. A three-order expression was adopted in F3 to represent daily variations including diurnal, semi-diurnal, and terdiurnal terms, which was similar to the study of Hoque and Jakowski [27]. The principal seasonal variations of...
electron density are annual and semiannual components \cite{40,46}. Thus, a two-order trigonometric function of DOY was used in $F_4$.

To improve the computational efficiency and avoid introducing the prior variation with latitude, coefficients were derived in each latitude slice separately. Since the physical processes taking place at low latitudes are controlled mainly by the direction of the magnetic field, and the geographic equator and magnetic equator do not follow each other, we divided the globe into latitude slices from $80^\circ$S to $80^\circ$N with a spatial interval of $5^\circ$ according to geomagnetic latitude (Mlat). In addition, adjacent slices were set to overlap $2.5^\circ$ to avoid the sharp shift between them. Finally, 63 latitude slices were obtained. We considered the middle of each latitudinal slice for the corresponding reference latitude of the derived coefficients. The latitudinal range of each latitude slice and the corresponding reference latitude was obtained by the following equations:

$$\text{Lat}_{\text{min}} = -80 + (i - 1) \times 2.5, \quad (8)$$
$$\text{Lat}_{\text{max}} = -80 + (i + 1) \times 2.5, \quad (9)$$
$$\text{Lat}_{\text{ref}} = -80 + i \times 2.5, \quad (10)$$

$\text{Lat}_{\text{min}}$ and $\text{Lat}_{\text{max}}$ denote the minimum latitude and the maximum latitude of the $i$th latitude slice, respectively. $\text{Lat}_{\text{ref}}$ is the corresponding reference latitude.

Note that the coefficients ($a, b, c, \text{and} d$) with subscripts in Equations (4)–(7) are not fixed. Similar to the process presented by Liu et al. \cite{41}, after multiplying $F_1, F_2, F_3$ and $F_4$ together, we obtained $3 \times 13 \times 7 \times 5 = 1365$ coefficients, and they were determined by performing a multivariable least squares fitting to the measurements in each $5^\circ$ latitudinal slice. Thus, a linear fitting approach was used instead of a nonlinear approach. Our goal was to find values of the coefficients that best fitted the data in the sense of minimizing the sum of squares of residual errors in each latitudinal slice. The list of these coefficients is too long to be included in this paper; however, it is available on request. Our model contains more coefficients than NPDM presented by Hoque and Jakowski \cite{27}. Therefore, we expect a better performance of NPPDM, which will be demonstrated in Section 4.

4. Model Performance

For the purpose of validating the NPPDM model, we analyzed the residuals of NPPDM and validated its performance to reproduce the independent observations measured by different instruments and to capture the features of $\text{NmF}_2$. Meanwhile, NPPDM was compared with IRI2016 and NPDM. Note that the resolution of our latitudinal slice was $2.5^\circ$. The $\text{NmF}_2$ value at the reference latitude could be calculated directly using the coefficients. When NPPDM predictions were compared with IRO and ionosonde measurements (Sections 4.2 and 4.3), $\text{NmF}_2$ values within each latitudinal slice range ($2.5^\circ$) were obtained by linear interpolation. The inputs for NPPDM included $F_{10.7}$, day of year, universal time, longitude, and magnetic latitude. Thus, the model was able to directly provide the $\text{NmF}_2$ prediction at a specific location and time. In Section 4.4, the NPPDM predictions were binned into intervals of $1^\circ$ geographic latitude (then transformed into geomagnetic latitude for calculation), $1^\circ$ geographic longitude, and 0.5 h local time to construct the spatial and temporal maps of $\text{NmF}_2$.

4.1. Modeling Results

In order to evaluate the fitting quality, the residuals of NPPDM were analyzed, i.e., the differences between the input database and the NPPDM predictions. The mean relative deviations (Mean) and root mean squared errors (RMSE) between model predictions (mod) and input observations (obs) were calculated as follows,

$$\text{Mean} = \frac{1}{N} \sum_{i=1}^{N} X_{\text{mod}}, \quad (11)$$
\[ X = \frac{NmF_2^{\text{mod}} - NmF_2^{\text{obs}}}{NmF_2^{\text{obs}}}, \quad (12) \]

\[ \text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (NmF_2^{\text{mod}} - NmF_2^{\text{obs}})^2}, \quad (13) \]

Figure 2 shows the statistical result of the relative deviations. For comparison, the residuals of IRI2016 and NPDM are also shown in Figure 2. It can be seen from Figure 2 that the NPPDM histogram peaks at −5–0% and declines steeply in both positive and negative directions. The mean relative deviations of NPPDM, IRI2016, and NPDM are 2.19%, 17.97%, and 16.20%, respectively. The large difference of IRI2016 and NPDM is mainly caused by the systematic deviation between IRO NmF2 and ionosonde NmF2 (used by IRI2016 and NPDM). The distributions of the three models show similar asymmetric patterns: the positive tail is longer than the negative tail. Generally, NPPDM can well reproduce the input database.

The distributions of RMSE with geographical longitude, magnetic latitude, universal time, local time (LT), and day of year are shown in Figure 3. From this figure, it can be seen that the RMSE shows no obvious differences with variations of geographical longitude and universal time. The RMSE is larger in low latitude regions than in middle and high latitude regions and shows two peaks near ±20°. The reason for this may be the EIA, which makes the variation of the F2-layer in low latitude regions a non-linear dynamic process and a result of complex physical systems [47]. It can be seen that the RMSE is larger in equinoxes than in solstices over both the northern and southern hemispheres. This can be explained by the fact that the ionospheric ionization is higher during equinoxes than during other periods. Similarly, RMSE begins to rise from sunrise as a result of the ionization increase [48].
4.2. Independent Test with Occultation Data

In this part, the IRO data for the years 2002 and 2005 were selected for comparisons among the model predictions of NPPDM, IRI2016, and NPDM. It should be pointed out that the NPPDM model we validate here is a reconstructed one, after removing the observational IRO data for 2002 and 2005 from the database. Therefore, the data used here for validation were independent and outside the window of the database. The solar activity was at relatively low and high levels in 2005 and 2002, respectively. Note that the data for these two years cover wide ranges of latitude, longitude, universal time, and day of year. The reproduction results for 2005 and 2002 are shown in Figures 4 and 5, respectively. The diagonal line depicts perfect prediction. Several white lines can be seen in the two figures for the reason that the corresponding values are missing in the data for these two years. The correlation coefficient (R) was calculated for quantitative analysis.

In 2005, the correlation coefficient between NPPDM predictions and IRO measurements is 0.91, higher than that of IRI2016 and NPDM, for which the values are 0.76 and 0.82, respectively. It should be noted that NPPDM used only IRO NmF2 for construction, while IRI2016 used ionosonde NmF2, and NPDM included both in its data source. Therefore, it is not surprising that NPPDM and NPDM show better performances when compared with IRO NmF2. In 2002, NPDM, NPPDM, and IRI2016 show the correlation coefficients of 0.81, 0.82, and 0.72, respectively. However, it can be seen that NPPDM shows a longer tail below the diagonal line than other models and does not distribute as tightly as NPDM. Note that the IRO data in 2002 was included in the database for NPDM, but it is outside the database for NPPDM validated here. The decrease of the accuracy of NPPDM in 2002 may be due to the poor coverage of source data under high solar activity. The COSMIC mission started at 2006 and the peak of the last solar cycle with the highest F10.7 values happened during 2000—2002. Actually, CHAMP mission partly covers this period. However, due to its single satellite, not much data is available.
4.3. Independent test with ionosonde data

We used ionosonde measurements from twelve different stations worldwide to validate NPPDM. The ionosonde observations during low (2007 and 2008) and high (2013 and 2014) solar activity years were selected, covering the months of June, March, September, and December in the temporal range, which represented the four seasons of spring, summer, autumn, and winter, respectively. As for the spatial range, the chosen stations were distributed on both sides of the equator in low, middle, and high latitude regions, covering America, Europe, and Oceania in the eastern and western hemispheres. Table 2 shows the geographical coordinates and the validation periods of all selected stations. A global map showing geographical locations of the selected ionosonde stations is presented in Figure 6.
Table 2. Geographic coordinates and validation periods of the selected stations.

<table>
<thead>
<tr>
<th>Station</th>
<th>Latitude/°N</th>
<th>Longitude/°E</th>
<th>Low solar activity</th>
<th>High solar activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gakona</td>
<td>62</td>
<td>−145</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Juliusruh</td>
<td>55</td>
<td>13</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Boulder</td>
<td>40</td>
<td>−105</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Eglin AFB</td>
<td>30</td>
<td>−87</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>19</td>
<td>−67</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Jicamarca</td>
<td>−12</td>
<td>−77</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Townsville</td>
<td>−20</td>
<td>147</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Louisvale</td>
<td>−29</td>
<td>21</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Camden</td>
<td>−34</td>
<td>151</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Canberra</td>
<td>−35</td>
<td>149</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Port Stanley</td>
<td>−52</td>
<td>−58</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Macquarie Is</td>
<td>−55</td>
<td>159</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

Figure 6. Global map of used ionosonde stations.

The ionosonde measurements were averaged at each UT hour for all days in the selected months. Afterwards, the corresponding hourly means were reproduced by NPPDM, IRI2016, and NPDM, respectively. Analog deviations between ionosonde observations and model predictions were calculated. The analog deviation not only considers the shape analog but also magnitude analog between samples [49]. The analog deviation (C) of two samples can be expressed by

\[ C = \frac{1}{2}(S + D), \]  

where

\[ S = \frac{1}{m} \sum_{k=1}^{m} |X_k - E|, \]  
\[ D = \frac{1}{m} \sum_{k=1}^{m} |X_k|, \]  
\[ X_k = x_{modk} - x_{obsk}, \]  
\[ E = \frac{1}{m} \sum_{k=1}^{m} X_k, \]
Here, $E$ denotes the mean deviation. $D$ is actually the mean value of absolute distance, which is able to accurately reflect the deviation degree in the total means of two samples, and therefore called the magnitude coefficient. $S$ reflects the discrete degree of the element difference $X_k$ from $E$ of the two samples. Obviously, only when each $X_k$ equals $E$, then $S = 0$, and the shapes of the two sample curves are completely similar; conversely, the larger the discrete degree of each $X_k$ from $E$, the more dissimilar the two sample curves are. It is known from the above that $S$ is able to better reflect the analog degree in shapes of two samples, and is thus called the shape coefficient. Expression (19) was used for data normalization in this study before the analog deviation was calculated,

$$x'_k = \frac{(x_k - x_{min})}{(x_{max} - x_{min})}, k = 1, 2, \ldots, m,$$

where $x_k$ is the original value, $x'_k$ is the normalized value, $x_{min}$ and $x_{max}$ are the minimum and maximum values of the original data, respectively.

It can be proven that $0 \leq C \leq 1$. The less the value of $C$ is, the more analogous the two samples are.

The comparisons between model predictions and ionosonde observations are shown in Figures 7 and 8, corresponding to low and high solar activity, respectively. It can be seen from Figures 7 and 8 that NPPDM could not only reproduce actual measurements with small differences in values, but also describe the diurnal variation trend with small discrepancies. In Figures 7 and 8, NPDM shows larger analog deviations than NPPDM over majority of the selected stations, and this is especially true when the shape analog is considered. Note that the NPDM model used data until 2010, so its performance may be relative inferior when compared with data after 2010. The models with the smallest analog deviations in low ($0^\circ$–$30^\circ$), middle ($30^\circ$–$50^\circ$), and middle-high (>50°) latitude stations under low and high solar activity (as displayed in Figures 7 and 8) are listed in Tables 3 and 4, respectively. It can be seen from Tables 3 and 4 that NPPDM, IRI2016, and NPDM show the smallest analog deviations in 24, 25, and 2 station-month validations (as displayed in each subplot of Figures 7 and 8), respectively. Therefore, we can find that the performance of NPPDM is comparable to that of IRI2016. It is expected that the accuracy of NPPDM has not been improved significantly compared with IRI2016 when ionosonde data were used for comparisons, because ionosonde data were used in the IRI model development, but were not included in the database of NPPDM.

Table 3. Models with the smallest analog deviations in the selected stations under low solar activity.

<table>
<thead>
<tr>
<th>Station Region</th>
<th>Model</th>
<th>Northern Hemisphere</th>
<th>Southern Hemisphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low latitude (0°–30°)</td>
<td>NPPDM</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>IRI2016</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>NPDM</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Middle latitude (30°–50°)</td>
<td>NPPDM</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>IRI2016</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>NPDM</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Middle-high latitude (&gt;50°)</td>
<td>NPPDM</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>IRI2016</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>NPDM</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Figure 7. Variations of hourly mean of NPPDM (red solid lines), IRI2016 (black dotted lines) and NPDM (red dotted lines) as a function of UT in comparisons with ionosonde observation (dots) under low solar activity. C1, C2, and C3 are the analog deviations between NPPDM, IRI2016, and NPDM and ionosonde observations, respectively.
Figure 8. Same as Figure 7, but for high solar activity.
### Table 4. Same as Table 3, but for high solar activity.

<table>
<thead>
<tr>
<th>Station Region</th>
<th>Model</th>
<th>Northern Hemisphere</th>
<th>Southern Hemisphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low latitude (0°–30°)</td>
<td>NPPDM</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>IRI2016</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>NPDM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle latitude (30°–50°)</td>
<td>NPPDM</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>IRI2016</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>NPDM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle-high latitude (&gt;50°)</td>
<td>NPPDM</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>IRI2016</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>NPDM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 4.4. Ionospheric Features Simulation

Hoque and Jakowski [27] used the similar approach and developed an NmF2 model (Neustrelitz Peak Density Model—NPDM) with 13 model coefficients. In their model, F10.7, DOY, LT, Lat and Mlat were chosen as input elements to describe the dependencies of NmF2. They included longitude behavior by virtue of the local time term. This local time coordinate, however, did not allow them to model non-migrating tides or other non-sun-fixed features. Figures 9 and 10 show the global constant local time maps of NmF2 predicted by NPDM. Note that the local time is fixed, while the universal time changes along the horizontal axis in each submap of Figures 9 and 10. As shown in Figures 9 and 10, the well-known longitudinal structure in the low-latitude ionosphere (wavenumber-4, WN4) cannot be displayed by NPDM. According to the previous studies, the WN4 patterns in the F2 region mainly result from the influence of the lower atmospheric non-migrating tide through either interaction with the neutral background and the ionosphere itself or modulating the E region dynamo [50–52]. Therefore, it is necessary to do a preliminary verification of the NPPDM’s capability to capture ionospheric features.

![Figure 9](image9.png)

**Figure 9.** Global distribution of NPDM NmF2 (cm$^{-3}$) at the September equinox under low (F10.7 = 80 sfu) solar activity.
Figure 10. Same as Figure 9, but for high (F10.7 = 150 sfu) solar activity.

To validate the capability of NPPDM to present the expected latitudinal features, the global distributions of NmF2 at 08 UT and 20 UT under moderate solar activity (F10.7 = 120 sfu) produced by NPPDM at the September equinox are shown in Figure 11. In the day sector of Figure 11, when the ionosphere is irradiated, a pronounced double peak centers around ±20–±40° dip on either side of the magnetic dip equator (black curves), which is known as the EIA. This phenomenon is induced by the well-known fountain effect, which lifts the equatorial plasma through $\mathbf{E} \times \mathbf{B}$ to a height of several hundreds of kilometers [53,54]. The plasma diffuses downward along magnetic field lines under the influence of gravity and pressure gradient forces after the plasma is lifted to greater heights. In Figure 11, the EIA asymmetry is also shown: the denser NmF2 is in either the northern or the southern hemisphere, where neutral winds are equatorward [55].

Figure 11. Global distributions of NPPDM NmF2 (cm$^{-3}$) at the September equinox under middle (F10.7 = 120 sfu) solar activity.

Figure 12 shows the longitudinal mean LT-Lat maps of NmF2 reproduced by NPPDM in equinoxes and solstices. Results are considered under both low (F10.7 = 80 sfu) and high (F10.7 = 150 sfu) solar activities. The EIA is also shown in each submap in Figure 12. In the daytime, a trough around the geomagnetic equator and two crests in either side of the equator can be found. After the sunset, the two crests on both sides of the geomagnetic equator gradually recede and merge into one, and finally disappear. In addition, some other features of NmF2 can be seen in Figure 12. NmF2 increases due to the solar activity enhancement. NmF2 is found to be higher in equinoxes than in solstices.
This semi-annual variation is attributed to the neutral O/N\textsubscript{2} concentration ratio variation according to the previous study [56]. The hemispheric asymmetry is also illustrated in Figure 12: NmF2 is higher in the summer hemisphere than in the winter hemisphere, and the EIA daily NmF2 peaks appear earlier in the winter hemisphere than in the summer hemisphere. The hemispheric asymmetry presented here agrees well with previous results [57,58].

![Figure 12](image)

**Figure 12.** Longitudinal mean variations of NPPDM as functions of latitude and local time under both low solar activity (upper panel) and high solar activity (lower panel) (cm\(^{-3}\)).

Latitude versus longitude distributions of NmF2 at 02 LT and 14 LT produced by NPPDM at the September equinox are displayed in Figures 13 and 14, corresponding to low (F10.7 = 80 sfu) and high (F10.7 = 150 sfu) solar activity, respectively. We can see from the figures that WN4 patterns are shown at 14LT: the EIA crests are enhanced over West Africa, Southeast Asia, Central Pacific Ocean, and South America. The satellite observations of different ionospheric parameters, such as total electron content (TEC) and hmF2 also present WN4 longitude structures [59–62]. However, probably because the wave-like longitude structures were observed by quasi-sun-synchronous satellites, they have been researched only at fixed local time (LT) mostly before. Using NPPDM, the development and propagation of the wave-like patterns can be further investigated. Beyond this, it can be seen that a large-scale electron density depletion structure appears near ±60° latitudes at 02 LT. This ionospheric midlatitude trough is also described by NPPDM, which forms at the interface between the midlatitude ionosphere and the high-latitude auroral region as a result of the complex interplay between different geophysical processes [63]. We have noticed the enhancement of NmF2 at 02 LT at high latitudes, which is caused by the disappearance of photoionization at middle latitudes and auroral particle precipitation at high latitudes [64].
Recently, Sai and Tulasi [24] developed ANNIM to predict global NmF2 based on COSMIC data. Afterwards, Tulasi et al. [25] presented an improved version of ANNIM by assimilating more IRO measurements as well as ionosonde data using a modified spatial gridding approach based on the magnetic dip latitudes. They used an Artificial Neural Networks technique, which enabled the estimation of the coefficients without knowing the base function of independent variables in advance. More variables were selected as inputs in their model than NPPDM, including DOY, UT, latitude, longitude, F10.7, Kp, declination, inclination, dip latitude, zonal wind, and meridional wind. The Kp input enabled ANNIM to operate under disturbed geomagnetic conditions and to be more applicable than NPPDM, which is only effective under quiet and moderate geomagnetic conditions [25]. As declared by Sai and Tulasi [24], ANNIM showed the regression coefficient of 0.89 when compared with the GRACE data. It can be seen from Figures 4 and 5 (in this paper) that NPPDM is well linearly correlated with the independent IRO data. For comparison, the regression coefficients in Figures 4 and 5 for 2005 and 2002 were calculated and show values of 0.96 and 0.77, respectively. The relative low value of the regression coefficient in 2002 can be due to the poor coverage of source data under high solar activity, as we mentioned above in Section 4.2. Therefore, we think that the performance of NPPDM is generally comparable with that of ANNIM. However, the validation data set (15% of total data), which was prepared for the validation process to avoid the overfitting of ANNIM parameters.
and to ensure the errors were within the permissible range, was outside the database and was not used to derive the non-linear relationship or the coefficients of the empirical model. Thus, it is possible that the ANNIM has a limitation over high latitude areas, where the data distribution is relatively sparse. As mentioned by Sai and Tulasi [24], ANNIM underestimated the NmF2 at high latitudes and the mean relative deviation increased to ~20–25% when compared with IRO data. We have calculated the mean relative deviation of NPPDM based on the same measurements from the four high latitude ionosonde stations (Gakona, Juliusruh, Macquarie Is, and Port Stanley) in Figures 7 and 8. The result shows a mean relative deviation of 14%. Note that the deviation for NPPDM may be even smaller when compared with IRO data due to the systematic deviation between ionosonde NmF2 and occultation NmF2 (used as data source by NPPDM). In addition, the performance of ANNIM in higher latitudes (>60°) has not been reviewed in their studies, while it can be seen from this work that NPPDM shows a reasonable performance there and well presents the midlatitude trough near ±60° latitudes (Figures 13 and 14).

It is known that geomagnetic storms may profoundly affect the variation of NmF2. However, NPPDM was designed for quiet and moderate geomagnetic conditions and the IRO measurements under disturbed geomagnetic conditions (Ap > 32) were removed from the database when the model was developed. Therefore, it should be noted that our model may show large errors during storm events. Jin et al. [65] reported a fairly large increase of NmF2 during the super geomagnetic storm (20 November 2003) at Anyang station (37.4°N, 127.0°E). Figure 15 shows the NmF2 hourly variations of ionosonde observations and model predictions at Anyang station for a 2-day period (20–21 November 2003). The variation of Kp index during this period can be seen in Jin’s paper (Figure 4). Some ionosonde measurements are missing due to the absorption during the storm. The predictions using the STORM version of the IRI2016 are presented in Figure 15 (black dotted lines). The STORM model [14] was included in IRI and was designed to be dependent on the intensity of the storm (Ap index over the 33 previous hours), latitude, and season. In Figure 15, the abnormal variations of ionosonde NmF2 during 10 to 14 UT (20, November) and 02 to 07 UT (21, November) may be associated with the variations of geomagnetic activity. The limitation of NPPDM during storm can be seen in Figure 15. We can see that IRI2016 (STORM on) shows the best performance between the three models, while NPPDM and NPDM mainly overestimate the NmF2 during this storm event. Joshua et al. [66] reported that NmF2 was more perturbed than TEC, which further stressed the importance of NmF2 in ionospheric modeling, particularly during disturbed geophysical conditions. Therefore, further research on solving this limitation of NPPDM is needed.

![Figure 15. Variations of hourly mean of NPPDM (red solid lines), IRI2016 (black dotted lines) and NPDM (red dotted lines) as a function of UT in comparisons with ionosonde observation (dots) during 20-21 November 2003.](image-url)
The F2-layer peak density height hmF2 is another important parameter in empirical ionospheric modeling. Hoque and Jakowski [67] followed the same basic approach and developed a nonlinear hmF2 model in another paper. Therefore, the approach we used in this paper may have the potential application in hmF2 modeling. This possibility requires more investigations. Since the hmF2 modeling is beyond the scope of this paper, further research will be performed in the future.

6. Conclusions

In this study, an empirical model NPPDM was developed based on CHAMP, GRACE, and COSMIC radio occultation data. The model was designed as a simple pattern consisting of nonlinear polynomial equations to describe the dependences of NmF2 on universal time, day of year, solar activity, and longitude. We divided the globe into 63 slices according to the geomagnetic latitude. The coefficients were derived by a multivariable fitting in each latitude slice. To validate the model, the predictions of NPPDM were compared with IRO data and ionosonde data, as well as the predictions of the IRI2016 and NPDM models. In addition, we simulated some ionospheric climatological features using the new model. The simulation shows good agreement with the previous observational conclusions. The main results can be summarized as follows:

1. Using NPPDM to reproduce the input data, the mean relative deviation between the input data and the new model is 2.19%. NPPDM shows larger RMSE in the EIA region than in other regions.
2. The predictions of NPPDM were compared with independent IRO measurements, IRI2016 predictions, and NPDM predictions. The correlation coefficients between model predictions and actual measurements show that the NPPDM performs better than IRI2016 and NPDM under low solar activity, while it undergoes performance degradation under high solar activity as a result of the poor database coverage.
3. Ionosonde data from twelve stations were selected to validate the accuracy of NPPDM. The analog deviations between ionosonde observations and model predictions indicate that the precision of NPPDM is better than that of NPDM and comparable to that of IRI2016.
4. The variations and distributions of NmF2 simulated by NPPDM agree well with previous research. NPPDM can capture some ionospheric features including the EIA, the longitudinal wavenumber-4 structure, and the midlatitude trough.

Supplement of the database from upcoming satellite missions, especially under high solar activity, is required to further improve the model.


Acknowledgments: The authors are grateful to sponsors and operators of the CHAMP and GRACE missions; Deutsches GeoForschungs Zentrum (GFZ) Potsdam, German Aerospace Center (DLR), and the US National Aeronautics and Space Administration (NASA). The authors would like to give thanks to sponsors and operators of the FORMOSAT-3/COSMIC mission; Taiwan’s National Science Council and National Space Organization (NSPO), the US National Science Foundation (NSF), National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA) and the University Corporation for Atmospheric Research (UCAR). The authors are also grateful to NGDC for disseminating historical ionosonde data via SPIDR.

Conflicts of Interest: The authors declare no conflict of interest.

References
3. Halcrow, B.W.; Nisbet, J.S. A model of F2 peak electron densities in the main trough region of the ionosphere. Radio Sci. 1977, 12, 815–820. [CrossRef]


43. Liu, H.; Stolle, C.; Förster, M.; Watanabe, S. Solar activity dependence of the electron density in the equatorial anomaly regions observed by CHAMP. *J. Geophys. Res.* 2007, 112, A11311. [CrossRef]


52. Wan, W.; Liu, L.; Pi, X.; Zhang, M.; Ning, B.; Xiong, J.; Ding, F. Wavenumber 4 patterns of the total electron content over the low latitude ionosphere. *Geophys. Res. Lett.* **2008**, *35*, L12104. [CrossRef]


56. Hanson, W.B.; Moffett, R.J. Ionization transport effects in the equatorial F region. *J. Geophys. Res.* **1966**, *71*, 5559–5572. [CrossRef]


65. Jin, S.; Luo, O.F.; Park, P. GPS observations of the ionospheric F2-layer behavior during the 20th November 2003 geomagnetic storm over South Korea. *J. Geod.* **2008**, *82*, 883–892. [CrossRef]


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).