JGR Space Physics

RESEARCH ARTICLE
10.1029/2019JA027703

Special Section:
Long-term changes and trends in the Middle and Upper Atmosphere

New Modes and Mechanisms of Long-Term Ionospheric TEC Variations From Global Ionosphere Maps

Andres Calabia[1,2] **and Shuanggen Jin**[1,2,3]

[1]School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing, China, [2]Jiangsu Engineering Center for Collaborative Navigation/Positioning and Smart Applications, Nanjing, China, [3]Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China

Abstract The ionosphere is very active and complex due to photo-ionization from the solar activity, while traditional empirical models can only give a rough description of its actual variations. Nowadays, global ionosphere maps (GIMs) derived from denser Global Navigation Satellite Systems (GNSS) world-tracking data provide an excellent total electron content (TEC) data set for global ionospheric research and modeling. In this paper, long-tern variations of 16-year (2003–2018) TEC time series from GIMs are investigated by using the principal mode analysis (PCA) technique. We analyze the resulting modes in the time-spectral domain and parameterize the main contributions in terms of solar and magnetospheric forcing, local solar time (LST), and annual variations. The results show that the TEC variability is strongly dependent on the geographical location of the Earth's magnetic field, and the Earth's diurnal rotation modulates its spatial patterns of variability. The latitudinal asymmetry in the global distribution of TEC variations is due to the effects caused by the irregular shape of the Earth's magnetic field along with its diurnal rotation. The analyses of residuals show that periodicities are correlated to the solar wind speed and magnetospheric forcing, especially those located near the southern dip pole at the night side. Furthermore, we found a TEC anomaly at about 15° from the South magnetic dip at the night side, more prominent around 52°S 155°E.

Key Points:
- The solar and the magnetospheric forcing are the main drivers of nonperiodic ionospheric TEC variations
- Main periodic contributions to TEC variations are related to the frequencies of the solar rotation, annual, and subharmonics
- TEC anomaly has been found at about 15° from the South magnetic dip at the night side, more prominent around 52°S 155°E

Supporting Information:
- Supporting Information S1
- Data Set S1

Correspondence to:
S. Jin,
sgjin@nuist.edu.cn;
sg.jin@yahoo.com

Citation:
Calabia, A., & Jin, S. (2020). New modes and mechanisms of long-term ionospheric TEC variations from global ionosphere maps. *Journal of Geophysical Research: Space Physics*, *125*, e2019JA027703. https://doi.org/10.1029/2019JA027703

Received 9 DEC 2019
Accepted 10 APR 2020
Accepted article online 18 May 2020

1. Introduction

The ionosphere is a region of the Earth's upper atmosphere from about 60 to 1,000 km altitude, which is mainly created by photo-ionization from the incoming extreme ultraviolet (EUV) and X radiation from the Sun. This region contains a high concentration of ions and free electrons that influence the propagation of electromagnetic radio waves employed in satellite communication systems, radar remote sensing signals, and Global Navigation Satellite Systems (GNSS) (Jin et al., 2008, 2012). For instance, the presence of ionospheric electron density produces a time delay in the arrival of the modulated modes from GNSS pseudo-range measurements and an advance in the GNSS carrier-phase measurements. These effects decrease the accuracy and reliability of GNSS measurements onboard satellites, aviation, and ground vehicles with errors in the range of 2–50 m for single-frequency receivers (Hofmann-Wellenhof et al., 2008). Fortunately, it is possible to estimate the ionospheric delay by dual-frequency GNSS observations (Jin et al., 2015, 2016, 2017). The resulting estimate is defined as the electron concentration along the path from the receiver to the satellite, namely slant total electron content (STEC), whose unit is TECU (1 TECU = 10^{16} electron/m^2). Furthermore, STEC can be converted into vertical TEC (VTEC) by applying ionospheric mapping functions (Klobuchar, 1987) and used in deriving ionospheric products and models.

Since 1998, the International GNSS Service (IGS) has provided Global Ionospheric Maps (GIMs) VTEC derived from the dual-frequency GNSS data. With more than 350 permanent stations, every single station receives measurements of about 10–30 GNSS satellites every 30 s, resulting in a large volume of worldwide observations per hour. This database is a very valuable source of information, which has been widely employed, among other uses, for ionospheric research and modeling. Empirical models of the ionosphere are usually based on statistical analyses and parameterizations of large data sets that provide the best-fit representation of the real ionospheric state. Currently, there is a wide range of ionospheric empirical models in active use and development designed to give a rough description of the variations in the ionosphere. Well-known empirical models are, for example, the Klobuchar model (Klobuchar, 1987), the Bent Ionospheric Model (Bent & Llewellyn, 1973), the parameterized ionospheric model (PIM) (Daniell

et al., 1995), the SAMI model (Huba et al., 2000), the NeQuick model (Radicella, 2009), and the International Reference Ionosphere (IRI) model (Bilitza et al., 2011). However, the highly variable and complex processes in the ionosphere cause difficulties for accurate modeling, and present empirical models mainly provide monthly averages of the ionospheric behavior for especially magnetically quiet conditions.

During last decades, several new empirical TEC models were developed based on IGS GIMs, mainly through Empirical Orthogonal Function (EOF) analyses (Ercha et al., 2012; Wan et al., 2012) and nonlinear least squares methods (Feng et al., 2019; Jakowski et al., 2011; Mukhtarov et al., 2014). Among them, the nonlinear least squares approach usually presents inconveniences facing the spatial resolution, the complexity of equations, the number of coefficients, and the computing time. On the other hand, EOF analyses are based on the principal component analysis (PCA) technique (Pearson, 1901), which allows transforming a data set of observations of correlated variables into a set of values of linearly uncorrelated variables. In the PCA transformation, the first resulting mode accounts for the largest variance in the observations, and each consecutive mode accounts for the highest possible residual variance under the constraint that it is orthogonal to the preceding mode. The numerous advantages of the PCA technique have made it to become one of the most employed methods to analyze the dominant modes of spatiotemporal data.

Previous authors with the PCA technique for TEC modeling have mapped the data to an orthogonal basis composed of a number of harmonic functions (Ercha et al., 2012; Wan et al., 2012). However, it is clear that besides increasing the complexity of the analysis, the implementation of a spherical harmonic basis can reduce the accuracy of the resulting model. In addition, since the electrons are forced to follow Earth's magnetic field lines, previous authors have adopted a coordinate system expressed by local solar time (LST) and magnetic dip latitude. Though, the irregular shape of the Earth's magnetic field (with the north dip pole close to Earth's rotation axis and the south dip pole far from that axis) under the effects of the Earth's rotation itself might not produce similar variability at all the longitudinal locations (geographic) of the globe, as represented by a LST reference system. Besides, a better abstraction of the global distribution could be interpreted by employing the geographical latitude and longitude coordinate system.

In this paper, instead of attempting to map the observational data to an orthogonal basis composed by a number of spherical harmonics, we employ the full resolution of the IGS GIMs and compute the covariance matrix $R = F^t F$ for the eigenvalue problem, where F is a matrix that represents each location and epoch, by columns and rows, respectively. The resulting time-expansion PCA coefficients are employed to investigate and model the resulting modes independently. Furthermore, we perform analyses in the geographical latitude and longitude coordinates for the 12 different LST epochs given by the IGS TEC GIMs. Then, looking for periodic patterns of variability, we perform a time-domain spectral analysis of the main PCA modes, and correlate the time series looking for the best fit of the most representative solar and geomagnetic indices, including seasonal dependencies. Our results outcome into a new empirical model, which is further assessed with the statistical analysis of the empirical residuals. In the next section, a brief introduction of the IGS TEC estimates, and the space weather proxies employed in this paper are given; section 2 provides the data processing and analysis methods; results of data processing, interpretation of model deficiencies, and accuracy assessment are presented in section 3; finally, conclusions are given in the last section.

2. Observational Data and Methods

2.1. Ionospheric TEC

We employ the final combined solution of IGS GIMs VTEC (Hernández-Pajares et al., 2009; Schaer, 1999) available in IONEX (ionosphere exchange) format at the Crustal Dynamics Data Information System (CDDIS) Goddard Space Flight Center (GSFC) National Aeronautics and Space Administration (NASA) website (https://cddis.nasa.gov/index.html). The gridded data have a temporal resolution of 2 hr and a spatial resolution of 2.5° by 5° in latitude and longitude respectively. In this work, we employ GIMs VTEC data from 2003 to 2018, since these are provided every 2 hr starting from 00:00 hr Universal Time (UT), and the estimates from 1998 to 2002 are provided every 2 hr from 01:00 hr UT.

2.2. Solar and Magnetospheric Indices

The ionosphere is mainly created by photo-ionization from incoming EUV and X radiation from the Sun, with the additional contributions due to geomagnetic activity. Therefore, it is reasonable that predictive

models are based on indices that can represent the solar and magnetospheric forcing. In this study, we employ solar and magnetospheric indices from the Low-Resolution OMNI (LRO) data set of NASA (http://omniweb.gsfc.nasa.gov/form/dx1.html) and the International Service of Geomagnetic Indices (ISGI) website (http://isgi.unistra.fr/data_download.php). We employ the disturbance geomagnetic storm-time index *Dst*, the geomagnetic variation *Am* index, and compute the merging electric field E_m following the indications of Kan and Lee (1979). In addition, solar forcing measurements from Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) satellite are downloaded from the Laboratory for Atmospheric and Space Physics (LASP) University of Colorado Boulder (UCB) website (http://lasp.colorado.edu), and those derived from the National Oceanic and Atmospheric Administration (NOAA) series operational satellites, the Geostationary Operational Environmental Satellite system (GOES), and the NASA/ESA (European Space Agency) Solar and Heliospheric Observatory (SOHO) satellite Solar Extreme-ultraviolet Monitor (SEM) are provided by Bowman et al. (2008).

2.3. Methodology

The aim of a PCA technique (Pearson, 1901) is to determine a new set of bases that capture the largest variance in the data, based on eigenvalue decomposition of the covariance matrix. As mentioned above, our starting data set of TEC is provided as GIMs at 2-hr temporal resolution and a spatial resolution of 2.5° by 5° in geographical latitude and longitude coordinates, respectively. Thus, in order to employ the full resolution of this initial data set, we compute the covariance matrix $R = F^t F$ for the eigenvalue problem where F is a matrix that represents each location (x) by columns and each epoch (t) by rows. This direct method has been presented to investigate conservative force models globally (Calabia & Jin, 2016a) and later employed to extract the principal modes of thermospheric mass density variability (Calabia & Jin, 2016b). In order to simplify the PCA analysis and reduce the size of the covariance matrix, we perform 12 groups, one for each LST epoch (00:00 to 22:00 hr LST every 2 hr). In this scheme, 16-year time series (2003 to 2018) of measurements results in the PCA of 12 matrices of n = 5,183 columns (73 longitudes × 71 latitudes) by m = 5,840 rows (365 days × 16 years). Detailed analyses and the selection of retained modes can be found in Preisendorfer (1988) and Wilks (1995) and a readily computable algorithm in Bjornsson and Venegas (1997). Then, the spatial patterns of TEC variability, their time variation, and the measure of their importance are presented as a low-dimensional space spanned by a set of modes, each of one represented by a pair of time ($\Gamma_{i,l}$) and space ($\Omega_{i,l}$) expansion modes from n locations, x, at m epochs, t, grouped in 12 LST cases, ℓ.

$$TEC(x_n, t_m) = \sum_{i=1}^{i=n} \left[\Gamma_{i,\ell}(t_m) \cdot \Omega_{i,\ell}(x_n)\right] \tag{1}$$

Then, in the next step, making the generalization of the time-expansion PCA components at each LST epoch ($\Gamma_{i,l}$) as Γ_i, we carry out the fitting in terms of the most representative indices and cycles, including solar and magnetospheric indices, and LST and seasonal cycles. In this step, we perform a time-domain spectrum analysis looking for characteristic periodic variations and then a correlation-delay study to obtain the best possible fit. The power spectral density (PSD) estimate (Fourier transform of the biased estimate of the autocorrelation sequence) is employed to determine the main periodicities in the time series, and the Pearson's linear correlation coefficient is estimated sequentially to obtain the time delay between two sequences at maximum correlation. In order to better characterize singular events, we employ the least squares fitting of polynomial and Fourier functions to each time-expansion PCA mode (Γ_i) by minimizing the absolute difference of the residuals (least absolute residuals) in a multivariable parameterization (solar cycle, LST, seasonal, and geomagnetic index). First, the following two-variable polynomial expression is employed to fit the first time-expansion PCA component (Γ_1) as follows:

$$\begin{aligned}\frac{\Gamma_1(t)}{\Upsilon_1(t)} \approx \Psi_1(t) = {} & p_{00} + p_{10} \cdot FLUX(t + \tau_{FLUX}) + p_{01} \cdot MAG(t + \tau_{MAG}) + \\ & + p_{20} \cdot [FLUX(t + \tau_{FLUX})]^2 + p_{11} \cdot FLUX(t + \tau_{FLUX}) \cdot MAG(t + \tau_{MAG})\end{aligned} \tag{2}$$

In this equation, p_{jk} are the fitting coefficients, t is a given epoch to estimate the model, and $FLUX(t + \tau_{FLUX})$ and $MAG(t + \tau_{MAG})$ are the solar and magnetospheric indices evaluated at a given time t, which need to

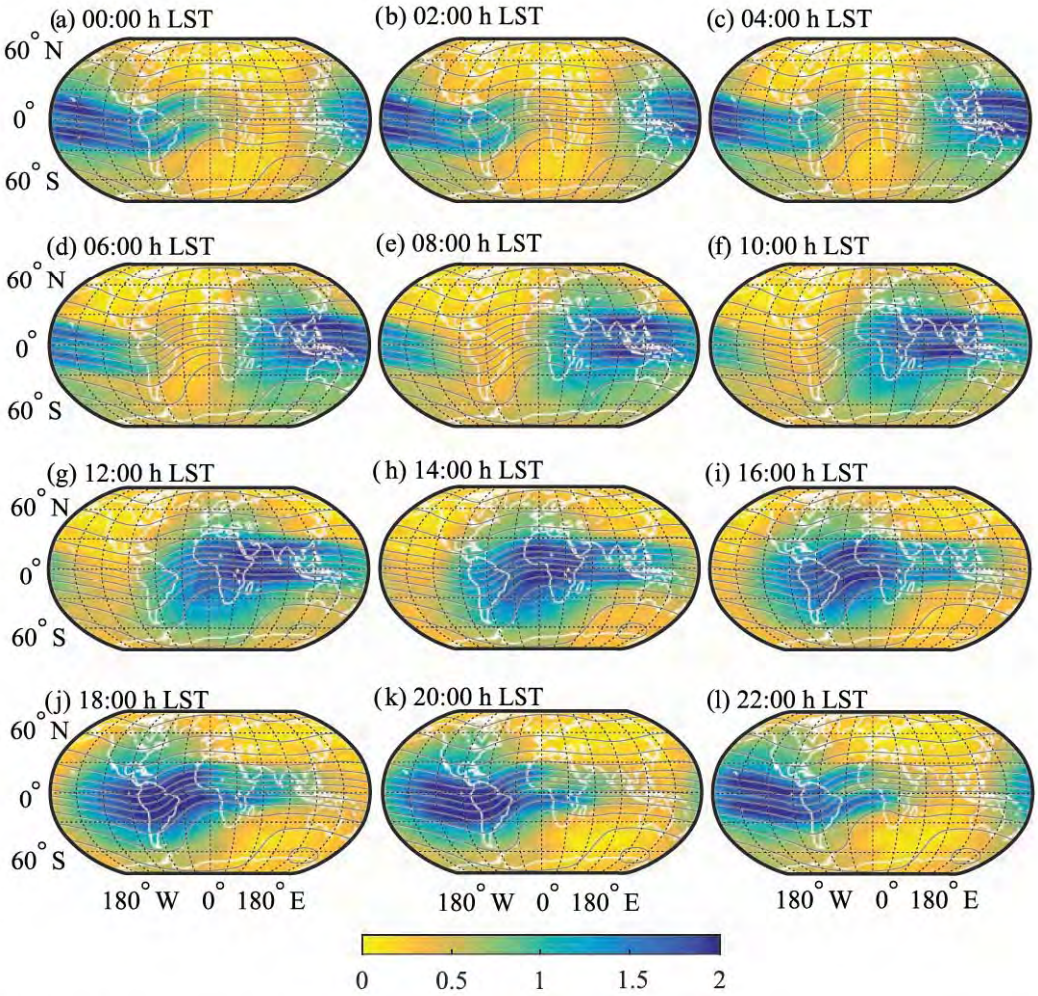

Figure 1. First spatial PCA modes of TEC variability ($\Omega_{1,\ell}$) for each of the 12 LST cases (2-hr resolution) from 2003–2018. This mode explains 75% of the total variance at each LST case. Dip isoclinic lines are plotted in gray to show the alignments. Values are dimensionless.

account for the time delay for each index, τ_{FLUX} and τ_{MAG} respectively. In the first iteration in estimating p_{jk}, Υ_1 is approximated to 1, and then it is updated with a two-variable Fourier expression that fits the annual and LST cycles as follows:

$$\frac{\Gamma_i(t)}{\Psi_i(t)} \approx \Upsilon_i(t) = a_0 + a_{11}\cdot\cos(Sa(t)) + b_{11}\cdot\sin(Sa(t)) + a_{12}\cdot\cos(S1(t)) + b_{12}\cdot\sin(S1(t)) + \\ + a_{21}\cdot\cos(2\cdot Sa(t)) + b_{21}\cdot\sin(2\cdot Sa(t)) + a_{22}\cdot\cos(2\cdot S1(t)) + b_{22}\cdot\sin(2\cdot S1(t)) + \\ + a_{31}\cdot\cos(3\cdot Sa(t)) + b_{31}\cdot\sin(3\cdot Sa(t)) + a_{32}\cdot\cos(3\cdot S1(t)) + b_{32}\cdot\sin(3\cdot S1(t)) \quad (3)$$

In this equation, a_{jk} and b_{jk} are the fitting coefficients, and *Sa* and *S1* are the tidal constituents derived from Doodson's fundamental arguments and corresponding multipliers for the solar annual and the diurnal cycles (Petit & Luzum, 2010). The final expression for each time-expansion mode *I* is given as $\Gamma_i = \Psi_i.\Upsilon_i + \varepsilon_i$, where ε_i is the residual disturbance from each PCA $_i$ mode.

3. Results and Analysis

3.1. Spatial Modes

The first PCA modes at each LST case ($\Omega_{1,l}$) are shown in Figure 1. In this figure, the main variability is represented by the two characteristic enhanced crests located at about ±20° latitude and aligned to the

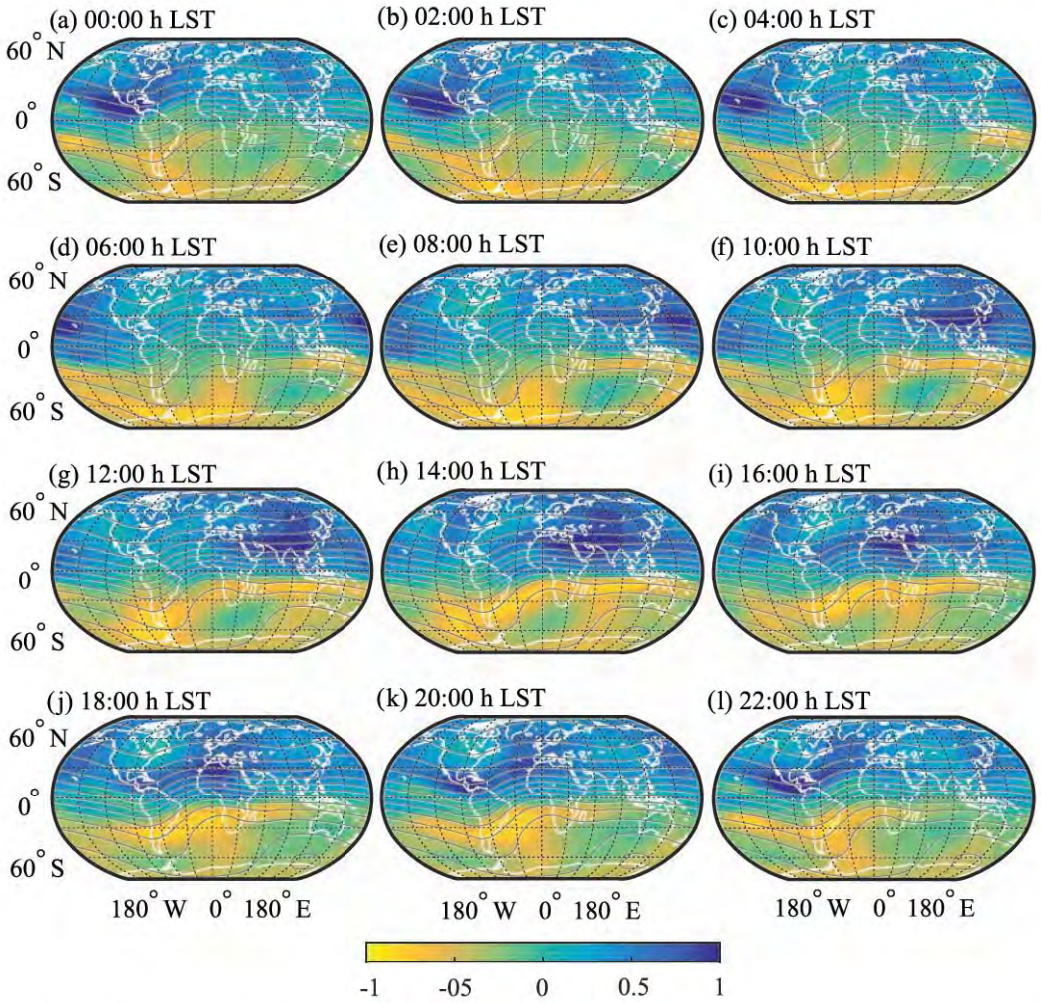

Figure 2. Second spatial PCA modes of TEC variability ($\Omega_{2,\ell}$) for each of the 12 LST cases (2-hr resolution) from 2003–2018. This mode explains 15% of the total variance at each LST case. Dip isoclinic lines are plotted in gray to show the alignments. Values are dimensionless.

magnetic dip equator, namely, equatorial ionospheric anomaly (EIA) or "Appleton anomaly" (Appleton, 1946). The response to solar forcing for these modes shows a time delay of about 2 hr with respect to the location of the subsolar point, for example, maximum values at Greenwich meridian at 14:00 hr LST. As expected in our hypotheses, the different LST cases provide a distinct geographical distribution of the EIA. For instance, long-shaped and well-defined crests at 00:00 hr LST versus a short, abrupt, and not well-defined EIA shape at 18:00 hr LST. Clear disturbances from 12:00 to 20:00 hr LST sectors are seen in the EIA shape, probably caused by the lack of alignment between the magnetic dip equator and the direction of the Earth's rotation. The temporal variation of these modes is mainly driven by the 11-year solar cycle and the seasonal dependence, explained more in detail in the next section.

The second PCA spatial modes at each LST case ($\Omega_{2,l}$) are shown in Figure 2. These modes are mainly driven by the annual variation, as seen from the latitudinal gradient in Figure 2, with a distinct distribution depending on each LST case. By multiplying these second spatial modes ($\Omega_{2,l}$) with the corresponding annual component (Υ_2), that is, Figure 2 multiplied to Figure 4c, the effect is an increase of TEC in local summer and a decrease in local winter. Note also the modulation of the 11-year solar cycle (Ψ_2) as formulated in the previous section. In Figure 2, an enhanced EIA shows a clear alignment with the magnetic field along the magnetic dip equator, with a time delay of about 6 hr with respect to the location of the subsolar point, for

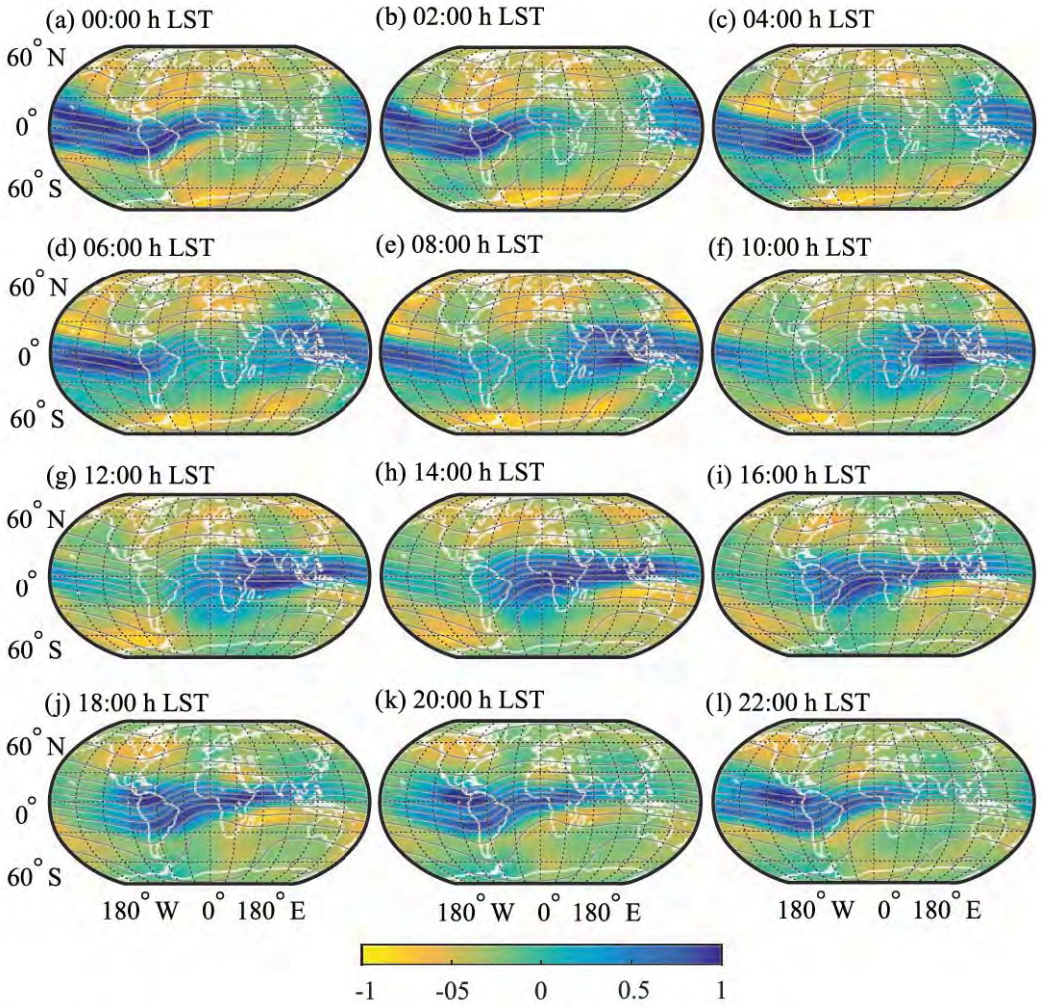

Figure 3. Third spatial PCA modes of TEC variability ($\Omega_{3,\ell}$) for each of the 12 LST cases (2-hr resolution) from 2003–2018. This mode explains 2% of the total variance at each LST case. Dip isoclinic lines are plotted in gray to show the alignments. Values are dimensionless.

example, maximum values at Greenwich meridian at 18:00 hr LST. A persistent contribution is seen from 12:00 to 22:00 hr LST, roughly over 30° South and 50° West, and from 02:00 to 12:00 hr LST in the southern high latitudes between longitudes 30° and 90° West.

The third PCA spatial modes at each LST case ($\Omega_{3,l}$) are shown in Figure 3, whereas each LST case pictures a distinct global distribution. Stronger and bigger modulation of the EIA crests is present from 0:00 to 8:00 hr LST, while, on the contrary, smaller and less pronounced modulation during the rest of the day. As seen from the temporal component (Υ_3 in Figure 4d), the seasonal variation is the main driver of this mode. By multiplying the third PCA spatial modes ($\Omega_{3,l}$) with the corresponding annual component (Υ_3), that is, Figure 3 multiplied to Figure 4d, the resulting effect is an increase of TEC over the EIA around the equinox, and a decrease around the solstice. Note also the modulation of the 11-year solar cycle (Ψ_3) as formulated in the previous section. The contribution over the eastern part of the EIA shows two-crest patterns, while the contribution over the western part of the EIA shows to be located over the magnetic dip equator.

3.2. Temporal Modes

The temporal component derived from the first PCA mode (Γ_1) has shown a nature very similar to that of the 11-year solar cycle. Therefore, we employ Equation 2) to fit different solar indices into Γ_1, and include a correlation-delay study for possible delay of the ionospheric response. Thus, we compute Pearson linear

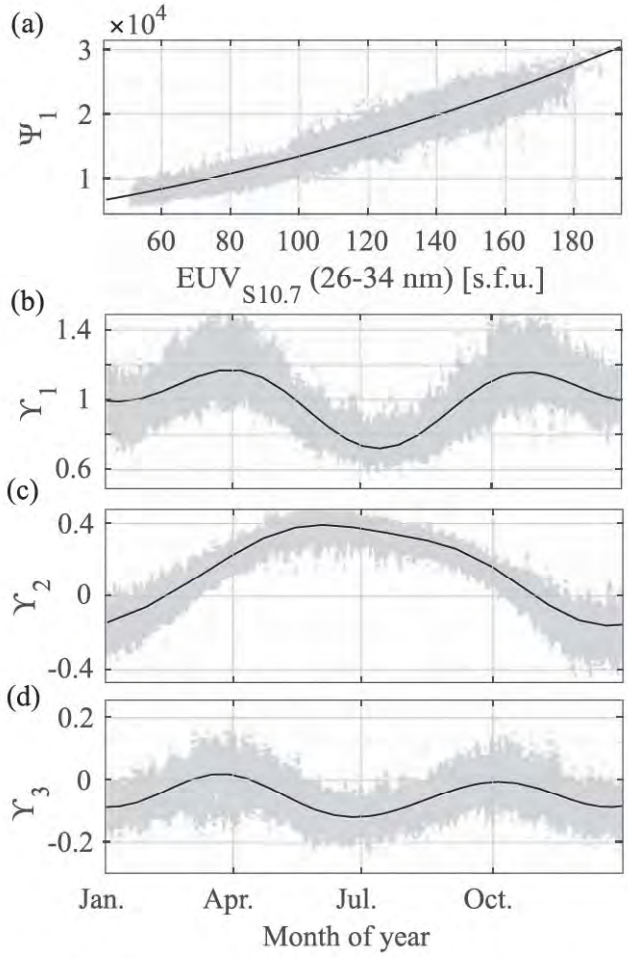

Figure 4. Fit of the three first temporal PCA modes of TEC variability from 2003–2008 (66,800 samples). The solar forcing component of the first mode (Ψ_1) is shown in (a) versus the EUV (26–34 nm) index. In (b), the annual component of the first mode (Υ_1) is shown versus the annual cycle. The second Υ_2 and third Υ_3 components are shown in (c) and (d) versus the annual cycle. Values of the *y*-axes are dimensionless.

Table 1
The Proportion of the Variance (R^2) Explained by Different Solar Forcing Indices in the Regression Model (66,800 samples) of the First Time Expansion PCA Mode Ψ_1. RMSE Values are Dimensionless

Solar forcing input	Source	R^2	RMSE	Correlation	τ_{FLUX} (h)
EUV (26–34 nm)	SOHO	0.95	8.5	97%	18
Mg-II (280 nm)	NOAA	0.88	13.2	96%	8
X-Ray and Lyman-α (0.1–0.8 & 121 nm)	GOES	0.86	15.2	93%	30
F10.7 (10.7 cm)	LRO	0.84	16.0	92%	32
33.5 nm (Fe XVI)	TIMED	0.84	16.0	90%	16
30.4 nm (He II)	TIMED	0.76	21.0	77%	18
36.8 nm (Mg IX)	TIMED	0.66	28.4	65%	18
133.5 nm (C II)	TIMED	0.60	28.6	62%	24
121.5 nm (H I)	TIMED	0.59	29.5	60%	16

correlation coefficients between Ψ_1 and the solar forcing indices using delay times (τ_{FLUX}) within a range of ±48 hr. Analysis results for the 10 solar forcing indices are shown in Table 1, where the highest explained variance (R^2), correlation, and lowest root-mean-square error (RMSE) in the regression model is represented by the EUV index. The resulting time delay for the EUV index is 18 hr, while the "X-Ray and Lyman-α" index provides better scope for prediction (30 hr) with similar accuracy. Figure 4a shows the fit of the first time expansion PCA mode (Ψ_1) in terms of the EUV index.

Furthermore, we employ a similar procedure to investigate the best representation of the magnetospheric forcing by testing the correlation to the disturbance geomagnetic storm-time index *Dst*, the geomagnetic variation *Am* index, and the merging electric field E_m. In order to isolate magnetospheric forcing as much as possible from the first PCA component Ψ_1, we set *MAG* in Equation 2 to a constant value ($Am = 6$) and investigate the corresponding residual disturbances. Due to the noisy nature of these residual disturbances, first we apply a 30-day running-window filter to both residuals and indices to remove the long-term trends. Then the Pearson linear correlation coefficient is calculated first in the correlation-delay analysis using delay times (τ_{MAG}) within a range of ±48 hr and afterward for a 30-day running window. Figure 5 shows the results for *Am*, *Dst*, and E_m indices, revealing correlations up to 75% for *Am* and E_m during the periods of low solar activity (2006–2009 and 2017–2019). Delay times for *Am*, E_m, and *Dst* show at the highest correlation positive values, indicating the ability to predict the future TEC state in about 3.5, 4.5, and 16 hr, respectively.

3.3. Analysis of Residuals

The spectral analyses in the time domain of the five first PCA modes and corresponding residuals (ε_i) are shown in Figure 6. In general, the maximum values of PSD estimates show that main periodicities are at the frequencies of the annual (365 days), solar rotation (27 days), and LST (1 day) cycles; and subharmonics (182 days; 13.5 and 9 days; and 0.5 and 0.25 day; respectively). The annual cycle is very prominent and well modeled in all the components, with the exception of the semiannual variation of the fourth component (Υ_4), which showed difficulties for analytical fitting due to noise in the empirical time series (Figure 6h). On the other hand, although the PCA has been performed separately for each LST case, a residual from the different LST cases is still present in all the components in Figure 6. However, we eliminate these LST residual dependencies by including an additional LST dependency in the fitting (Equation 3), but with the exception of the first mode (Ψ_1), which also showed difficulties to fit due to noise in the empirical time series (Figure 6a). Finally, since the solar rotation periods are included in *MAG* and *FLUX* parameters, clear reductions are seen in Figure 6 at 27, 13.5, and 9 days' periods, and more prominent in the first and third PCA components. Corroborating these periodic variations with the solar and the magnetospheric forcing, the spectral analyses in the time-domain for $EUV_{S10.7}$, solar wind speed, and *Am* are shown in Figure 7. The maximum valued of PSD show that the main signatures belong to the period of the solar rotation and subharmonics (27, 13.5, 9, and 6.75 days). In addition, since Earth dynamics can influence geomagnetic variations, note a diurnal signature in the PSD of the

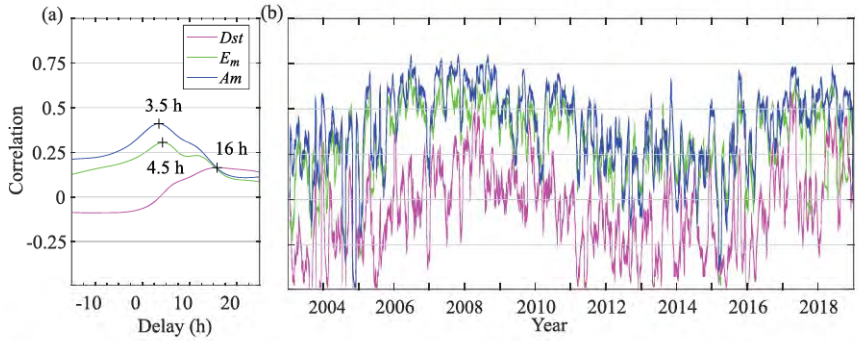

Figure 5. In (a), correlation versus delay between *Am*, *Dst*, and E_m, and the residual disturbances of the first time-expansion PCA fit (i.e., $\varepsilon_1 = \Gamma_1 - \Psi_1 \cdot \Upsilon_1$) when employing constant $MAG \equiv Am = 6$ in Ψ_1. In (b), the 30-day correlation running window for the same cases at a delay of maximum correlation is shown versus time.

Am index. Another interesting periodic signature in the fourth and fifth components (Figures 6g and 6i) can be seen around 14.7 days, which belongs to the solar wind speed (Figure 7c), which is also present in the PSD of the *Am* index (Figure 7e).

We divided the spatial analysis of model residuals in two parts. The first part aims to assess the model for practical applications (e.g., GNSS ionospheric path-delay correction) by studying the location of the major

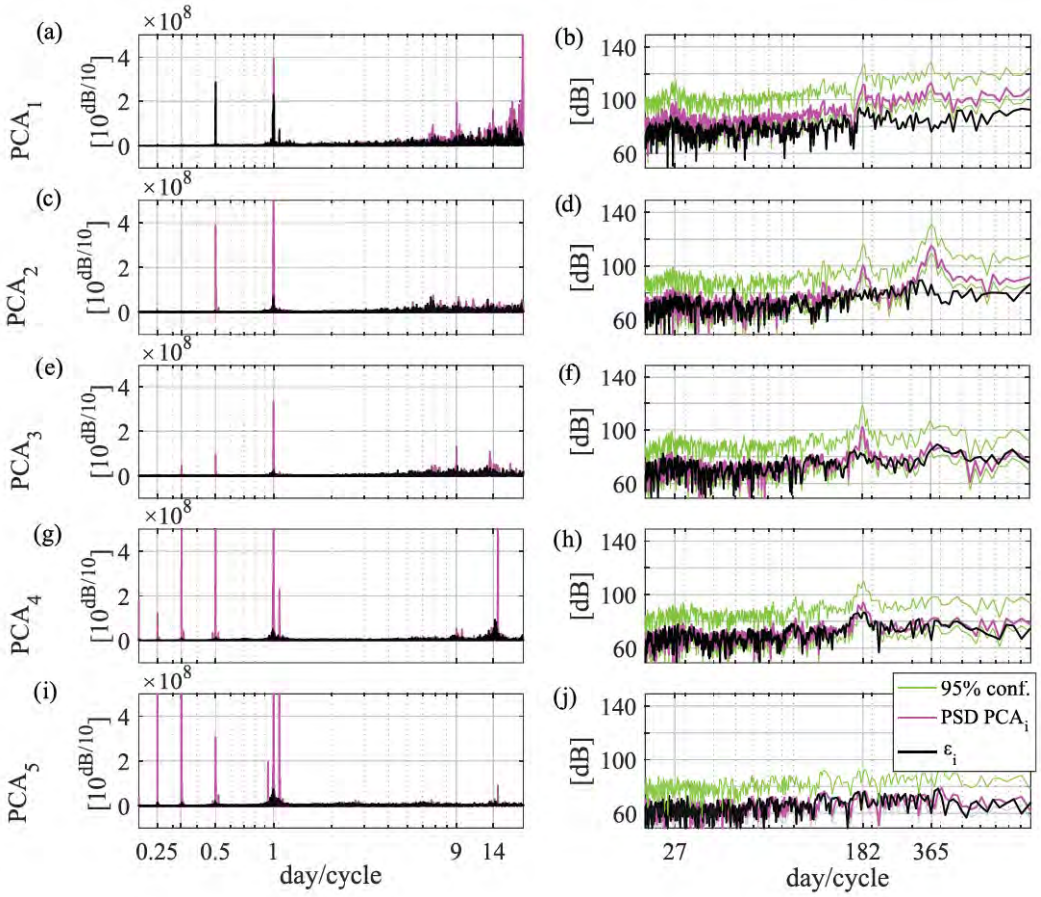

Figure 6. In magenta PDS estimate of PCA time expansion modes, and, in black solid line the corresponding residuals of the fitting model (i.e., $\varepsilon_i = PCA_i - \Psi_i \cdot \Upsilon_i$). Gray dash-dotted line represents the 95% confidence bounds.

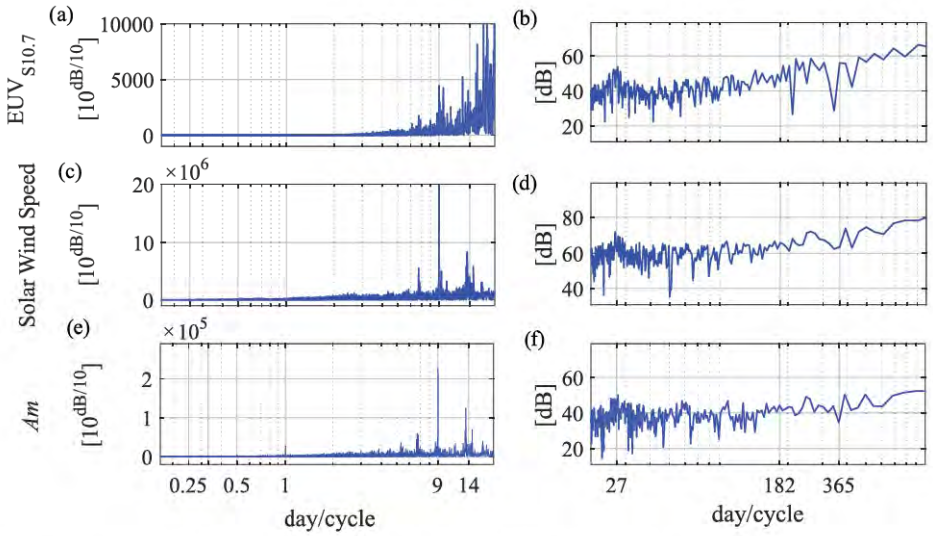

Figure 7. The PSD estimate of (a,b) $EUV_{S10.7}$, (c,d) solar wind speed, and (e,f) *Am*.

absolute residuals (Figure 8). For the second part, the assessment is addressed to enhance the understanding of ionospheric processes by studying local features that are not sensitive to the PCA. In this second part, we uncover and investigate features derived from the relative residuals with respect to the background density (Figure 9). Moreover, we subdivide the analysis for the low solar activity period (2006–2010) and for the high solar activity period (2011–2016).

In Figures 8a and 8d, the median average of absolute residuals (i.e., $TEC_{IGS} - \Sigma[\Psi \cdot \Upsilon \cdot \Omega]$) shows a positive pattern (model underestimation) between the front sides of the EIA crests (following the subsolar point), more pronounced from 00:00 to 12:00 hr LST during high solar activity. On the other hand, under similar conditions, a negative pattern (model overestimation) is present at the tails of the EIA crests. The peak values of median absolute residuals for these regions range below 10 TECU. As for the standard deviation, the absolute residuals show larger variability at the EIA crests. Similar patterns are found when comparing low and high solar activity periods, showing that the standard deviation of residuals along the EIA crests is about 10 TECU and 3 TECU, and away from the EIA crests is about 1–3 TECU and 1 TECU, respectively, for the high and the low solar activity periods. Concerning the relative residuals with respect to the background density, the corresponding median average shows higher values during low solar activity periods (Figure 9b) than during high solar activity periods (Figure 9a). In general, positive patterns (model underestimation) are located at middle and high latitudes, and negative patterns follow the subsolar point at the tails of the EIA crests.

In Figure 9, a prominent deviation with negative values is visible from 10:00 to 18:00 hr LST at 52°S 155°E in the southern high latitude, more pronounced during low solar activity. The standard deviation of relative residuals shows above 100% variability at this same location from 08:00 to 16:00 hr LST. The northern high latitude, on the contrary, lacks of such a large mean deviation or variability. We further investigate this anomaly in the southern high latitude by changing the projection of Figure 9 and centering the map in the Geographic South Pole (90°S). We separate the different LST cases in two figures, one from 00:00 to 10:00 hr LST in Figure 10 and the other from 12:00 to 22:00 hr LST in Figure 11. Analogous to Figure 9, we subdivide the analysis for high and low solar activity periods. In these figures, we confirm the location the negative anomaly at about 52°S 155°E, at about 15° from the magnetic dip pole, and with a minimal displacement from 08:00 to 16:00 hr LST for the low solar activity period (the location of the subsolar point is shown with a red asterisk). For the high activity period, the anomaly is aligned to the dip isoclinic lines at about 15° from the magnetic dip pole. It seems that the course of the subsolar point along the LST passage is reflected as a TEC decrease anomaly at about 15° from the South magnetic dip pole, only in the night side. Figure 12 shows the minimum and maximum values of median and standard deviation of relative residuals, respectively, for all the LST cases. In this figure, minimum values of the median average of relative residuals

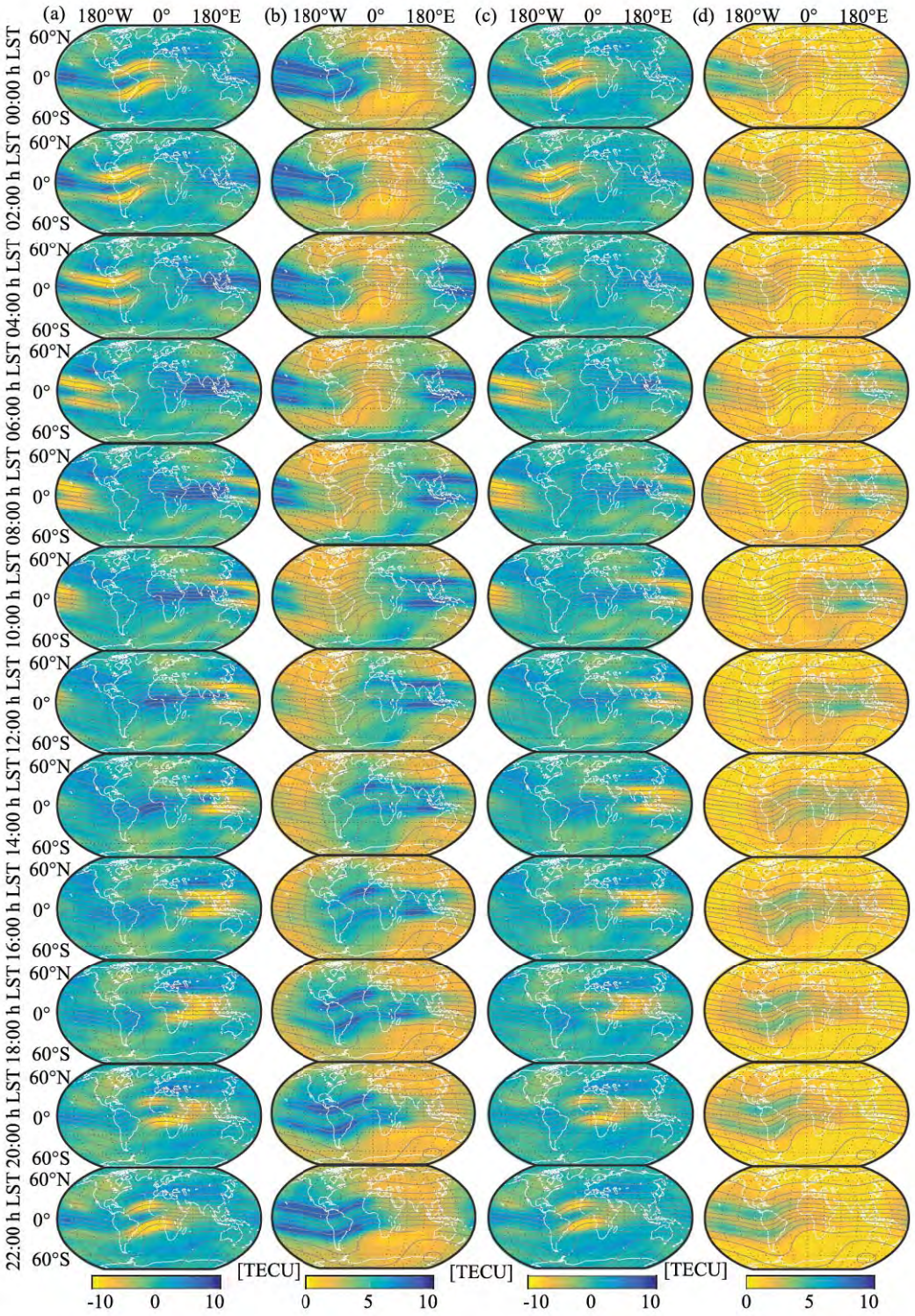

Figure 8. The (a–c) median average and (b–d) standard deviation of absolute residuals (i.e., "$TEC_{IGS} - \Sigma[\Psi \cdot \Upsilon \cdot \Omega]$"), for the (a,b) high and (c,d) low solar activity periods of 2011–2016 and 2006–2010, respectively. LST from top to bottom. Dip isoclinic lines are plotted in gray to show the alignments.

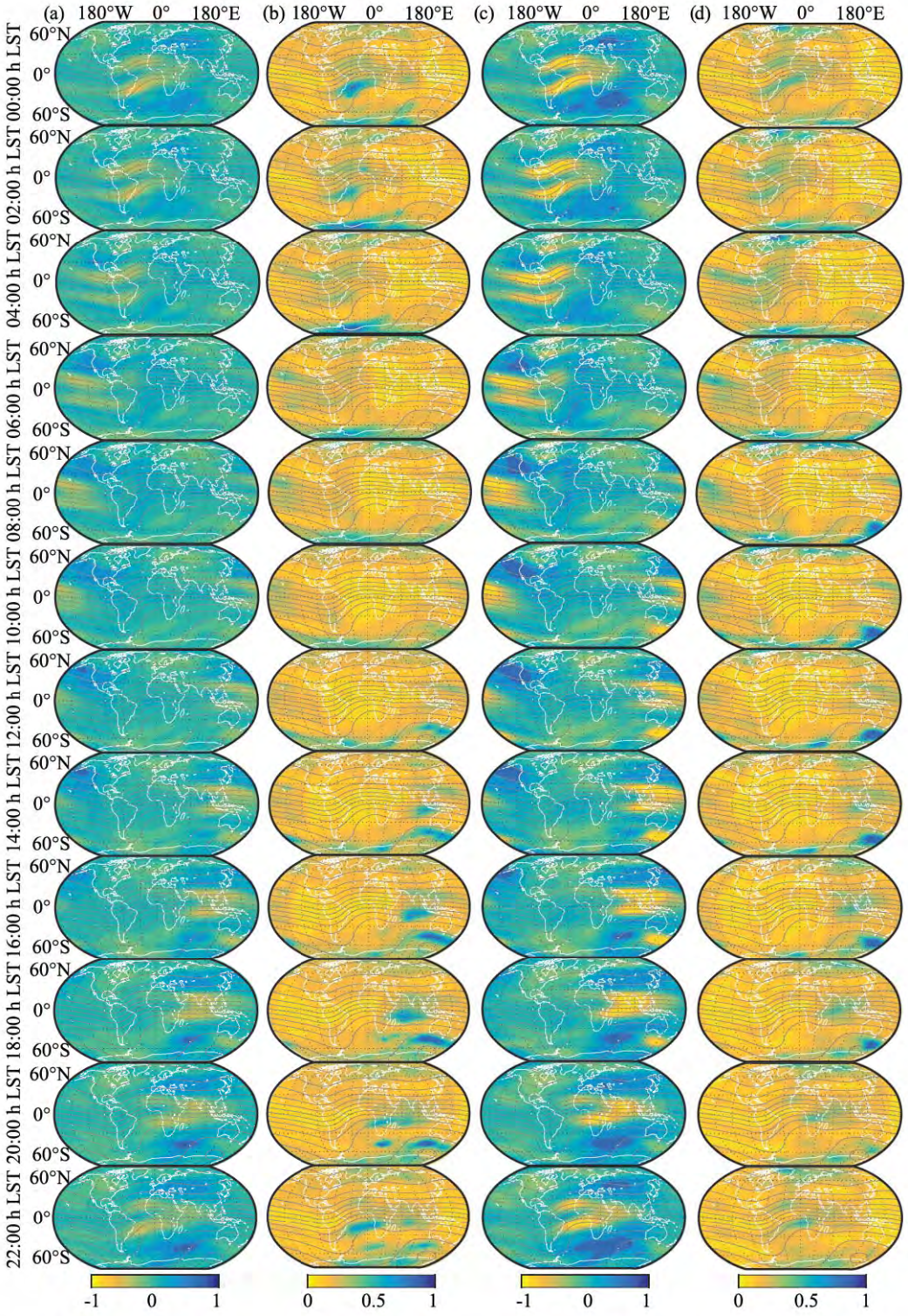

Figure 9. The (a–c) median average and (b–d) standard deviation of relative residuals (i.e., "($TEC_{IGS} - \Sigma[\Psi \cdot \Upsilon \cdot \Omega]$)/$TEC_{IGS}$"), for the (a,b) high and (c,d) low solar activity periods of 2011–2016 and 2006–2010, respectively. LST from top to bottom. Dip isoclinic lines are plotted in gray to show the alignments. Values are dimensionless.

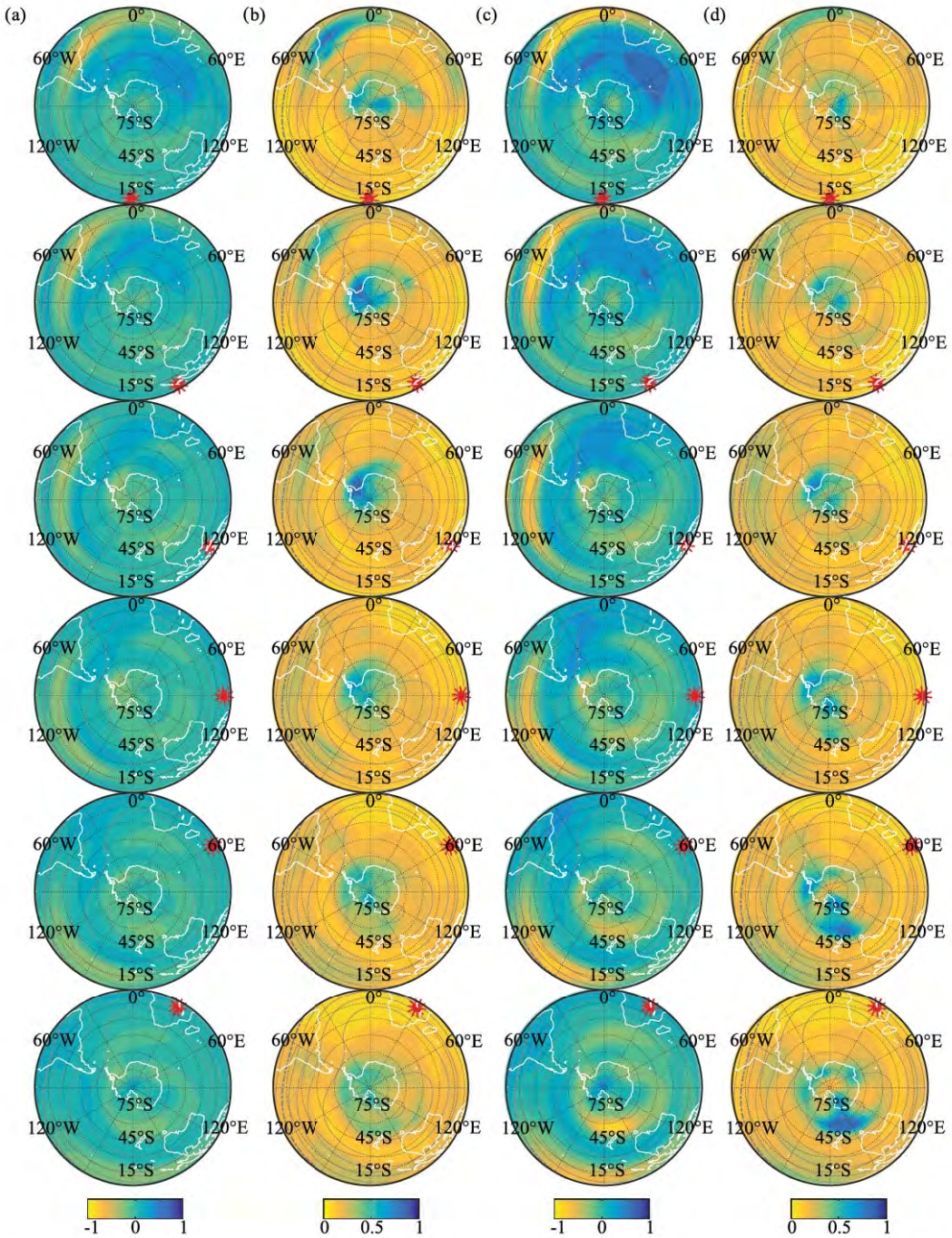

Figure 10. The (a–c) median average and (b–d) standard deviation of relative residuals (i.e., "$(TEC_{IGS} - \Sigma[\Psi \cdot \Upsilon \cdot \Omega])/TEC_{IGS}$"), for the (a,b) high and (c,d) low solar activity periods of 2011–2016 and 2006–2010, respectively. Southern Hemisphere from 0:00 to 10:00 hr LST, rows from top to bottom (red asterisks indicate the subsolar point). Dip isoclinic lines are plotted in gray to show the alignments. Values are dimensionless.

(Figures 12a and 12c) show to decease at all the LST locations at about 15° from the South magnetic dip pole, more prominent at 52°S 155°E. The northern high latitudes lack of such deviations. Figures 12a and 12c show a similar pattern for the maximum values of standard deviation.

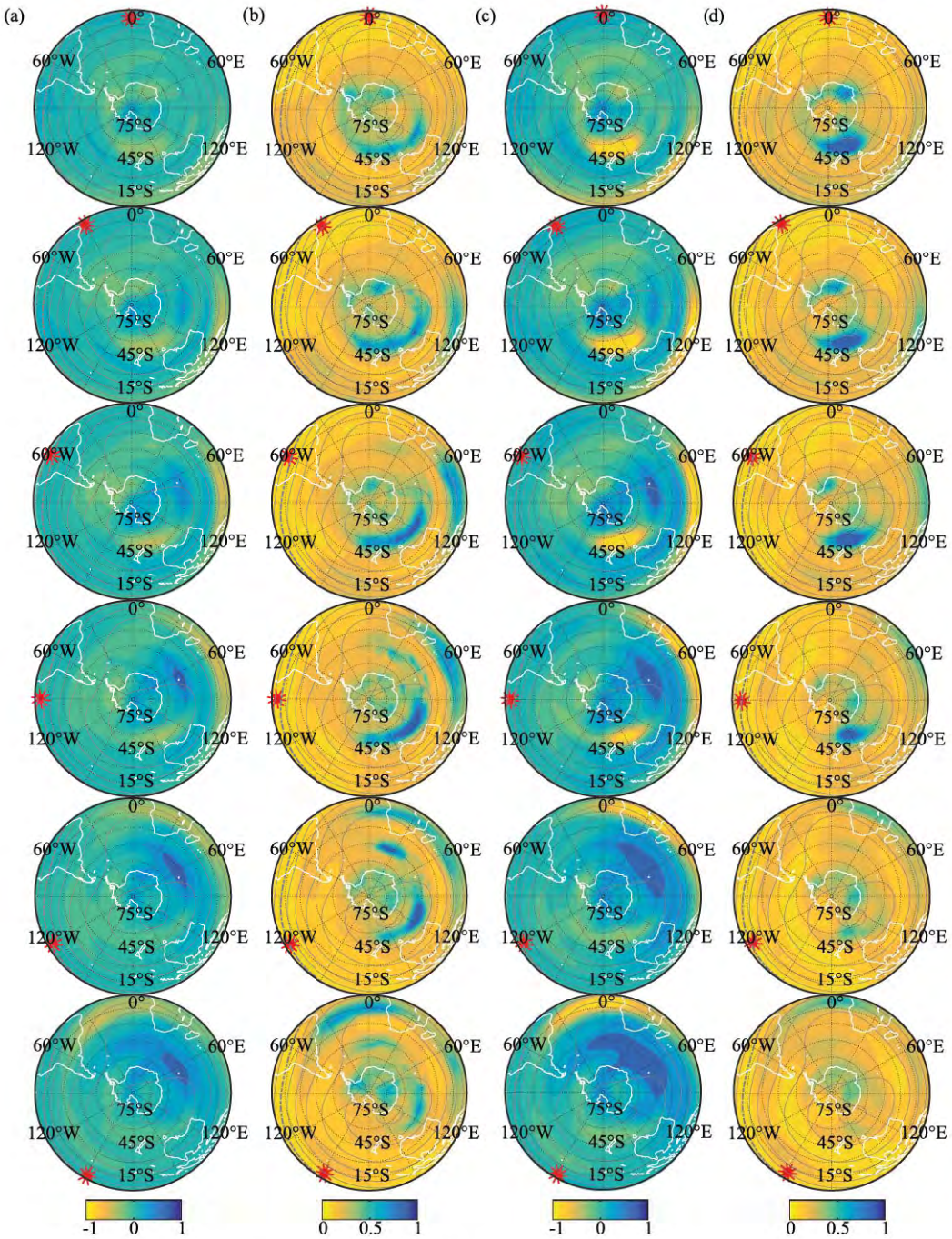

Figure 11. The (a–c) median average and (b–d) standard deviation of relative residuals (i.e., "$(TEC_{IGS} - \Sigma[\Psi \cdot \Upsilon \cdot \Omega])/TEC_{IGS}$"), for the (a,b) high and (c,d) low solar activity periods of 2011–2016 and 2006–2010, respectively. Southern Hemisphere from 12:00 to 22:00 hr LST, rows from top to bottom (red asterisks indicate the subsolar point). Dip isoclinic lines are plotted in gray to show the alignments. Values are dimensionless.

3.4. Parameterized PCA Modes

TEC observables from 2003 to 2018 have been reduced to a lower-dimensional representation through the PCA for each of the 12 LST cases provided by the IGS, that is, every 2 hr since 00:00 hr LST. In each LST case, the three leading modes resulting from the analysis together account for 93% of the total variance and, individually, explain 75%, 15%, and 2% of the total variability (Table 2). The high values of explained variance

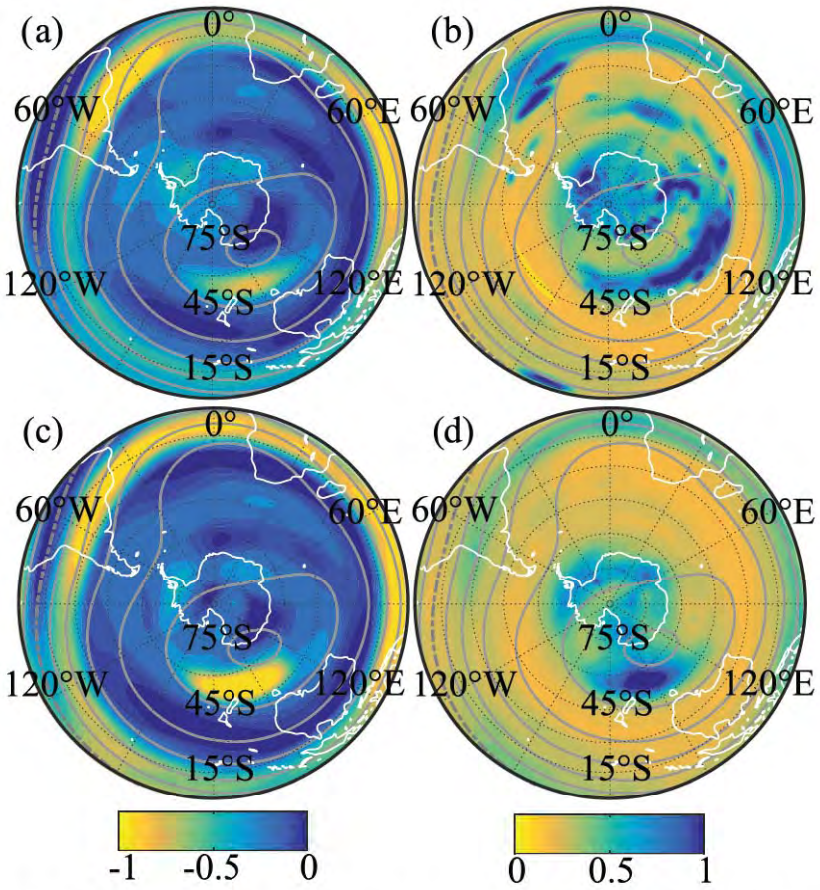

Figure 12. The (a–c) median average and (b–d) standard deviation of relative residuals (i.e., "$(TEC_{IGS} - \Sigma[\Psi \cdot \Upsilon \cdot \Omega])/TEC_{IGS}$"), for the (a,b) high and (c,d) low solar activity periods of 2011–2016 and 2006–2010, respectively. Minimum (a–c) and maximum (b–d) values from all LST cases at the Southern Hemisphere. Dip isoclinic lines are plotted in gray to show the alignments. Values are dimensionless.

given by the first modes for each LST case indicate marked patterns of variability. The resulting time-expansion PCA modes have been parameterized in terms of solar and magnetospheric forcing, annual, and LST cycles, and the resulting correlation coefficients, coefficients of determination (R^2), and RMSE of the parameterizations indicate high accuracy in the model with high correlation and R^2 and low RMSE. These coefficients are summarized in Table 2, and the model is provided in an SI file, and it is also available in Calabia and Jin (2019).

Table 2
Variance Explained by each PCA Mode at each LST Epoch and Fitting Statistics. Values are Dimensionless

Mode	Variance explained (per LST)	Data versus model (66,894 samples)		
		Correlation	R^2	RMSE
PCA 1	75%	98%	0.97	902
PCA 2	15%	97%	0.98	454
PCA 3	2%	71%	0.73	413
PCA 4	1%	43%	0.43	534
PCA 5	1%	64%	0.51	528

The analysis of absolute residuals shows positive patterns (model underestimation) located between the front sides of the EIA crests, more pronounced from 00:00 to 12:00 hr LST and during high solar activity. On the other hand, under similar conditions, a negative pattern (model overestimation) is present at the tails of the EIA crests. The standard deviation of absolute residuals along the EIA crests are about 10 TECU and 3 TECU, and away from the EIA crests are about 1–3 TECU and 1 TECU, respectively, for the high and the low solar activity periods. The analysis of relative residuals shows positive patterns (model underestimation) located at middle and high latitudes and negative patterns following the subsolar point at the tails of the EIA and at about 15° from the South magnetic dip at the night side, more prominent at 52°S 155°E.

4. Conclusions

Monitoring and modeling Earth's ionospheric variations and processes are a key fundamental to better understand how space weather responds to variable solar and magnetospheric conditions and to forecast the detrimental effects on life and society. Unfortunately, the negative effects of miss-modeled variations in ionospheric electron density result in uncertainty errors far to meet the operational requirements, and the negative repercussions and concerns in the space industry and space situational awareness force to demand better modeling of the upper atmospheric mass density distribution and variations. Therefore, new models, estimates, and differences between them are very important because it can help to better understand the complex processes in the upper atmosphere system and calibrate physics-based models. In this paper, we have investigated variations and characteristics of 16-year IGS GIMs TEC time series using the PCA technique and parameterized the main modes in terms of solar and magnetospheric forcing, diurnal, and annual variations. The main conclusions derived from this study can be summarized as follows:

1. TEC variability is strongly dependent on the geographical location of the Earth's magnetic field, and the Earth's diurnal rotation modulates its spatial patterns of variability.
2. The response of the EIA crests to solar forcing shows a time delay of about 2 hr with respect to the location of the subsolar point, and the EUV index (26–34 nm) is the best proxy for prediction, with a time scope of 18 hr.
3. The annual variation has a distinct distribution depending on each LST case, showing an enhanced balance of the EIA crests with a time delay of about 6 hr with respect to the location of the subsolar point.
4. The disturbances due to magnetospheric forcing have shown the best correlation with the Am, and E_m indices (up to 0.75 during low solar activity) with a time scope for prediction of 3.5 and 4.5 hr, respectively. On the other hand, the Dst, index has shown worse correlation but a better time scope for prediction (16 hr).
5. The residual TECs show a high latitude anomalous enhancement only at the Southern Hemisphere, specifically at about 15° from the South magnetic dip at the night side, more prominent around 52°S 155°E. The latitudinal asymmetry in the global TEC distribution is due to the effects caused by the irregular shape of the Earth's magnetic field (with the north dip pole close to Earth's rotation axis and the south dip pole far from that axis) along its diurnal rotation. The coverage of GNSS signal ionosphere piercing points over this position allocates small range of error induced by artifacts in the interpolation process carried by the IGS Ionospheric Associate Analysis Centers.

In the future, more studies should consider the potential effects of the solar wind and magnetospheric forcing carefully, for example, the Dungey cycle effects, when the magnetic tail reconnection occurs explosively in a substorm, producing intense auroras in the night side (Dungey, 1961). The resulting field-guided currents and energetic particle precipitation may induce to decrease TEC by chemical reactions with the night side ion outflow, including H^+, O^+, He^+, and NOx species. The accurate modeling of the ionospheric variability has a practical importance for real operational needs such as predictions to aid communication, navigation, and for satellite trajectories in low Earth orbit. This study has proven the predictive capability of a new global ionosphere model to estimate TEC variability, which can model the errors in electromagnetic signals passing through the ionosphere, such as satellite communication systems, radar signals, and GNSS, as well as calibrate physics-based models.

Acknowledgments

This work was supported by the National Natural Science Foundation of China-German Science Foundation (NSFC-DFG) Project (grant no. 41761134092), the Startup Foundation for Introducing Talent of NUIST (grant no. 2243141801036), and the Talent Start-Up Funding project of NUIST (grant no. 1411041901010). Great appreciation is extended to IGS (https://cddis.nasa.gov/index.html) for providing the TEC data set and to NASA/ESA SOHO (http://omniweb.gsfc.nasa.gov/form/dx1.html) and ISGI (http://isgi.unistra.fr/data_download.php) for the solar and magnetospheric indices. There is no conflict of interest regarding the publication of this paper. Supporting data are included in an SI file, and it is also available at the data repository http://doi.org/10.5281/zenodo.3563463; any additional data may be obtained from A. Calabia (email: andres@calabia.com).

References

Appleton, E. V. (1946). Two anomalies in the ionosphere. *Nature*, *157*(3995), 691–691. https://doi.org/10.1038/157691a0

Bent, R. B., & Llewellyn, S. K. (1973). *Documentation and description of the bent ionospheric model*. Los Angeles, CA, USA: Space and Missile Organisation.

Bilitza, D., McKinnell, L. A., Reinisch, B., & Fuller-Rowell, T. (2011). The international reference ionosphere (IRI) today and in the future. *Journal of Geodesy*, *85*(12), 909–920. https://doi.org/10.1007/s00190-010-0427-x

Bjornsson, H., & Venegas, S.A. (1997). A manual for EOF and SVD analyses of climatic data. McGill Univ., CCGCR report no. 97-1, Montréal, Québec, 52 pp.

Bowman, B. R., et al. (2008). A new empirical thermospheric density model JB2008 using new solar and geomagnetic indices. In *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 18-21 August 2008, Honolulu, Hawaii, n. AIAA 2008-6438.

Calabia, A., & Jin, S. G. (2016a). Assessment of conservative force models from GRACE accelerometers and precise orbit determination. *Aerospace Science and Technology*, *49*, 80–87. https://doi.org/10.1016/j.ast.2015.11.034

Calabia, A., & Jin, S. G. (2016b). New modes and mechanisms of thermospheric mass density variations from GRACE accelerometers. *Journal of Geophysical Research: Space Physics*, *121*, 11,191–11,212. https://doi.org/10.1002/2016JA022594

Calabia, A., & Jin, S. G. (2019). Supporting information for "new modes and mechanisms of long-term TEC variations from global ionosphere maps". *Zenodo*. http://doi.org/10.5281/zenodo.3563463

Daniell, R. E., Brown, L. D., Anderson, D. N., Fox, M. W., Doherty, P. H., Decker, D. T., et al. (1995). Parameterized ionospheric model: A global ionospheric parameterization based on first principles models. *Radio Science*, *30*(5), 1499–1510. https://doi.org/10.1029/95RS01826

Dungey, J. W. (1961). Interplanetary magnetic fields and the auroral zones. *Physical Review Letters*, *6*(2), 47–48. https://doi.org/10.1103/PhysRevLett.6.47

Ercha, A., Zhang, D., Ridley, A. J., Xiao, Z., & Hao, Y. (2012). A global model: Empirical orthogonal function analysis of total electron content 1999-2009 data. *Journal of Geophysical Research*, *117*, A03328. https://doi.org/10.1029/2011JA017238

Feng, J., Han, B., Zhao, Z., & Wang, Z. (2019). A new global Total Electron content empirical model. *Remote Sensing*, *11*(6), 706. https://doi.org/10.3390/rs11060706

IIernández-Pajares, M., Juan, J. M., Sanz, J., Orus, R., Garcia-Rigo, A., Feltens, J., et al. (2009). The IGS VTEC maps: A reliable source of ionospheric information since 1998. *Journal of Geodesy*, *83*(3-4), 263–275. https://doi.org/10.1007/s00190-80-266-1

Hofmann-Wellenhof, B., Lichtenegger, H., & Wasle, E. (2008). *GNSS—Global Navigation Satellite Systems, GPS, GLONASS, Galileo and more*, (1st ed.p. 518). Wien, Austria: Springer-Verlag. https://doi.org/10.1007/978-3-211-73017-1

Huba, J. D., Joyce, G., & Fedder, J. A. (2000). Sami2 is another model of the ionosphere (SAMI2): A new low-latitude ionosphere model. *Journal of Geophysical Research*, *105*(A10), 23,035–23,053. https://doi.org/10.1029/2000JA000035

Jakowski, N., Hoque, M. M., & Mayer, C. (2011). A new global TEC model for estimating transionospheric radio wave propagation errors. *Journal of Geodesy*, *85*(12), 965–974. https://doi.org/10.1007/s00190-011-0455-1

Jin, R., Jin, S. G., & Feng, G. P. (2012). M_DCB: Matlab code for estimating GNSS satellite and receiver differential code biases. *GPS Solutions*, *16*(4), 541–548. https://doi.org/10.1007/s10291-012-0279-3

Jin, S. G., Jin, R., & Kutoglu, H. (2017). Positive and negative ionospheric responses to the march 2015 geomagnetic storm from BDS observations. *Journal of Geodesy*, *91*(6), 613–626. https://doi.org/10.1007/s00190-016-0988-4

Jin, S. G., Jin, R., & Li, D. (2016). Assessment of BeiDou differential code bias variations from multi-GNSS network observations. *Annals of Geophysics*, *34*(2), 259–269. https://doi.org/10.5194/angeo-34-259-2016

Jin, S. G., Luo, O., & Park, P. (2008). GPS observations of the ionospheric F2-layer behavior during the 20th November 2003 geomagnetic storm over South Korea. *Journal of Geodesy*, *82*(12), 883–892. https://doi.org/10.1007/s00190-008-0217-x

Jin, S. G., Occhipinti, G., & Jin, R. (2015). GNSS ionospheric seismology: Recent observation evidences and characteristics. *Earth Science Reviews*, *147*, 54–64. https://doi.org/10.1016/j.earscirev.2015.05.003

Kan, J. K., & Lee, L. C. (1979). Energy coupling function and solar wind-magnetosphere dynamo. *Geophysical Research Letters*, *6*(7), 577–580. https://doi.org/10.1029/GL006i007p00577

Klobuchar, J. A. (1987). Ionospheric time-delay algorithm for single-frequency GPS users. *IEEE Transactions on Aerospace and Electronic Systems*, *AES-23*(3), 325–331. https://doi.org/10.1109/TAES.1987.310829

Mukhtarov, P., Pancheva, D., & Andonov, B. (2014). Hybrid model for long-term prediction of the ionospheric global TEC. *Journal of Atmospheric and Solar-Terrestrial Physics*, *119*, 1–10. https://doi.org/10.1016/j.jastp.2014.05.009

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, *2*(11), 559–572. https://doi.org/10.1080/14786440109462720

Petit, G., & Luzum, B. (2010). *IERS Conventions (2010)*, IERS Tech. Note, vol 36. Verlag des Bundesamts für Kartogr. und Geod., Frankfurt am.

Preisendorfer, R. (1988). In D. Curtis (Ed.), *Principal mode analysis in meteorology and oceanography, Mobley*. Amsterdam: Elsevier.

Radicella, S. (2009). The NeQuick model genesis, uses and evolution. *Annales de Geophysique*, *52*(3–4), 417–422. https://doi.org/10.1109/ISAPE.2010.5696491

Schaer, S. (1999). Mapping and predicting the earth's ionosphere using the global positioning system, Ph.D. Dissertation, Astronomical Institute, University of Berne, Berne, Switzerland, 25 March.

Wan, W. X., Ding, F., Ren, Z. P., Zhang, M. L., Liu, L. B., & Ning, B. Q. (2012). Modeling the global ionospheric total electron content with empirical orthogonal function analysis. *Science China Technological Sciences*, *55*(5), 1161–1168. https://doi.org/10.1007/s11431-012-4823-8

Wilks, D. S. (1995). *Statistical methods in the atmospheric sciences*. San Diego, Calif., USA: Academic Press.