Using Principal Component Analysis of Satellite and Ground Magnetic Data to Model the Equatorial Electrojet and Derive Its Tidal Composition

Gabriel Soares[1,2] [ID], Yosuke Yamazaki[3] [ID], Achim Morschhauser[2] [ID], Jürgen Matzka[2] [ID], Katia J. Pinheiro[1,2], Claudia Stolle[3], Patrick Alken[4,5] [ID], Akimasa Yoshikawa[6] [ID], Kornyanat Hozumi[7] [ID], Atul Kulkarni[8], and Pornchai Supnithi[9]

[1]Observatório Nacional, Rio de Janeiro, Brazil, [2]GFZ German Research Centre for Geosciences, Potsdam, Germany, [3]Leibniz Institute of Atmospheric Physics, University of Rostock, Kühlungsborn, Germany, [4]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA, [5]National Centers for Environmental Information, NOAA, Boulder, CO, USA, [6]International Center for Space Weather Science and Education ICSWSE, Kyushu University, Fukuoka, Japan, [7]Space Environment Laboratory, National Institute of Information and Communications Technology, Koganei, Japan, [8]Indian Institute of Geomagnetism, Navi Mumbai, India, [9]School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

Abstract The intensity of the equatorial electrojet (EEJ) shows temporal and spatial variability that is not yet fully understood nor accurately modeled. Atmospheric solar tides are among the main drivers of this variability but determining different tidal components and their respective time series is challenging. It requires good temporal and spatial coverage with observations, which, previously could only be achieved by accumulating data over many years. Here, we propose a new technique for modeling the EEJ based on principal component analysis (PCA) of a hybrid ground-satellite geomagnetic data set. The proposed PCA-based model (PCEEJ) represents the observed EEJ better than the climatological EEJM-2 model, especially when there is good local time separation among the satellites involved. The amplitudes of various solar tidal modes are determined from PCEEJ based tidal equation fitting. This allows to evaluate interannual and intraannual changes of solar tidal signatures in the EEJ. On average, the obtained time series of migrating and nonmigrating tides agree with the average climatology available from earlier work. A comparison of tidal signatures in the EEJ with tides derived from neutral atmosphere temperature observations show a remarkable correlation for nonmigrating tides such as DE3, DE2, DE4, and SW4. The results indicate that it is possible to obtain a meaningful EEJ spectrum related to solar tides for a relatively short time interval of 70 days.

1. Introduction

An important part of the geomagnetic variations recorded both at ground and at satellite altitudes is related to the dynamics of the Earth's upper atmosphere. The upper atmosphere consists of a combination of neutrals and plasmas. The plasma constituent is produced by the ionization of neutrals due to their photodissociation caused by energetic short-wavelength (<150 nm) solar radiation, giving rise to the so-called ionosphere.

The dayside *E*-region ionosphere is electrically conductive due to the presence of free electrons and ions and an appropriate ion-neutral collision frequency. In the presence of the geomagnetic main field and subjected to tidal winds, electric fields, and currents are generated. This process is referred to as the ionospheric dynamo, which can be described by

$$J = \hat{\sigma} \cdot (E + U \times B) \tag{1}$$

where J is the current density, $\hat{\sigma}$ is the ionospheric conductivity tensor, E is the electric field, U is the neutral wind, and B is the ambient geomagnetic field. This expression shows the connection between ionospheric dynamo currents and driving winds in the neutral atmosphere, including tidal winds. Geomagnetic diurnal variation, which reaches amplitudes of about 200 nT at ground, results from the ionospheric currents generated through the process expressed by Equation 1.

At the magnetic equator, the magnetic field is exactly horizontal, and the zonal electric field sets up a vertical polarization electric field by driving vertical Hall currents. This gives rise to an additional zonal current

Project Administration: Katia J. Pinheiro

Resources: Gabriel Soares, Yosuke Yamazaki, Jürgen Matzka, Katia J. Pinheiro, Claudia Stolle

Software: Gabriel Soares, Achim Morschhauser

Supervision: Yosuke Yamazaki, Jürgen Matzka, Katia J. Pinheiro

Validation: Gabriel Soares, Yosuke Yamazaki

Visualization: Gabriel Soares, Yosuke Yamazaki, Achim Morschhauser

Writing – original draft: Gabriel Soares, Yosuke Yamazaki, Jürgen Matzka, Katia J. Pinheiro, Claudia Stolle, Patrick Alken

Writing – review & editing: Gabriel Soares, Yosuke Yamazaki, Jürgen Matzka, Katia J. Pinheiro, Claudia Stolle, Patrick Alken, Kornyanat Hozumi

superposed to that caused by the original background electric field. This leads to an amplification of the eastward electric current flow at around 110-km height, confined within $\pm 3°$ latitude around the magnetic equator. The resulting enhanced current is known as the equatorial electrojet (EEJ, e.g., Yamazaki & Maute, 2017).

The EEJ was discovered after the installation of the first geomagnetic observatories at magnetic equator latitudes due to its associated abnormally large horizontal component diurnal variation (Chapman, 1951). Since then, many EEJ features were reported by different studies, as its dependence on local time, longitude, season, solar flux, main geomagnetic field, and lunar phase (e.g., Forbes, 1981; Onwumechili, 1997; Yamazaki & Maute, 2017). With magnetic data from dedicated satellite missions and its unprecedented longitudinal coverage, the EEJ spatiotemporal variation became much better described and understood (e.g., Alken & Maus, 2007; Lühr et al., 2004).

The quiet-time EEJ intensity exhibits variations on day-to-day to year-to-year time scales. Year-to-year changes are mainly due to the 11-year solar flux variation (e.g., Matzka et al., 2017), but are also weakly driven by the secular variation of the main field (e.g., Cnossen & Richmond, 2013; Soares et al., 2020) as well as neutral winds (e.g., Yamazaki et al., 2018). Seasonal variations of the EEJ are attributed in part to changes in the solar zenith angle (Chapman & Raja Rao, 1965) but more importantly to neutral winds (Yamazaki et al., 2014b). The EEJ day-to-day changes are primarily due to variable neutral winds (Fang et al., 2013; Miyahara & Ooishi, 1997; Yamazaki et al., 2014a). Large variability of neutral winds can be explained by upward-propagating waves from the regions below the ionosphere (e.g., Liu, 2016). Among various types of waves, atmospheric tides are particularly important for driving the EEJ, as they attain large amplitudes at dynamo region heights (e.g., Oberheide et al., 2011). Tides play a significant role not only for temporal variations of the EEJ but also for the longitudinal structure of the EEJ (e.g., Lühr et al., 2008; Soares et al., 2018).

Atmospheric solar tides are global scale waves that oscillate periodically in time and propagate vertically and zonally in space (Forbes et al., 2008). Different tidal components can be described mathematically as

$$A_{n,s} \cos(n\omega t + s\lambda - \Phi_{n,s}) \tag{2}$$

where $A_{n,s}$ and $\Phi_{n,s}$ are the amplitude and phase, respectively. The variable t is universal time, ω is the rotation rate of the Earth, λ is longitude, n is the subharmonic of a day (where $n = 1, 2, 3, 4$ correspond to oscillations with periods of 24, 12, 8 and 6 hr, called diurnal, semidiurnal, terdiurnal, and quarterdiurnal tides, respectively), and s is the zonal wavenumber ($s < 0$ for waves propagating eastwards and $s > 0$ for waves propagating westwards).

The tides with $n = s$ are called migrating tides as they propagate westwards with the same speed as the apparent motion of the Sun from the perspective of ground observers. All other tides ($n \neq s$) are called nonmigrating tides. Different combinations of n and s represent different tidal modes, which can arise from different excitation mechanisms (e.g., Miyoshi et al., 2017). It is common to use a combination of two letters and one number as a notation to denote specific tidal modes (e.g., DE3, SW6, TW2, …). The first letter stands for the period of oscillation (D = diurnal, S = semidiurnal, T = terdiurnal, Q = quarterdiurnal, …), the second letter refers to the direction of propagation (W = westward or E = eastward) and the number is the zonal wavenumber.

By fitting Equation 2 to magnetic field measurements from the CHAMP satellite, Lühr and Manoj (2013) determined the average spectrum of the EEJ related to solar tides. They separated 10 years of the CHAMP data into monthly subsets for 2000–2005 (around solar maximum) and 2005–2010 (around solar minimum) periods. The fitting was performed on each of the 24 subsets of data so that the average seasonal dependence of different tidal modes could be determined for different levels of solar activity. The limitation of this approach is that it cannot resolve the tidal variability of the EEJ for individual years. It requires splitting the data into shorter time windows to overcome this limitation. However, doing so would reduce the data available for the fit and the stability of the inverse problem being solved.

Instead of the actual raw observations, an EEJ model could be used in order to determine the EEJ spectrum related to solar tides. There are various types of EEJ models, obtained with different techniques and purposes. Those include, e.g., the modeling of the EEJ ground magnetic effects (Doumouya et al., 2003; Hamid et al., 2015), the EEJ theoretical modeling (Anandarao & Raghavarao, 1987; Richmond, 1973; Sugiura & Poros, 1969), the EEJ morphology modeling (intensity, width, and position— Doumouya et al., 1998; Fambitakoye & Mayaud, 1976; Rigoti et al., 1999) and the EEJ climatology modeling by satellite data (Alken & Maus, 2007). In this work, we propose a new technique aiming to accurately model the EEJ. Improved modeling of the EEJ allows a stable determination of the EEJ spectrum related to solar tides for individual years. To avoid problems during inversion

Table 1
Summary With the Equatorial and Low-Latitude Ground Stations Used to Derive the Equatorial Electrojet (EEJ) Data

Station	Type	Sector	Data repository	Lat.	Lon.
HUA	Equatorial	I	INTERMAGNET	−12.05	−75.33
PIU	Low-latitude	I	LISN	−5.20	−80.60
TTB	Equatorial	II	GFZ Data Services, INTERMAGNET	−1.20	−48.51
KOU	Low-latitude	II	INTERMAGNET	5.21	−52.73
SAM	Equatorial	III	WAMNET	11.39	−5.62
TAM	Low-latitude	III	INTERMAGNET	22.79	5.53
TIR	Equatorial	IV	WDC-Mumbai	8.70	77.80
ABG	Low-latitude	IV	INTERMAGNET	18.64	72.87
PKT	Equatorial	V	NICT Sealion	8.09	98.32
KTB	Low-latitude	V	210° Meridian Chain	−0.20	100.32
BCL	Equatorial	VI	MAGDAS	9.32	105.71
PHU	Low-latitude	VI	INTERMAGNET	21.03	105.95
DAV	Equatorial	VII	MAGDAS	7.00	125.40
MUT	Low-latitude	VII	MAGDAS	14.40	121.00
TND	Low-latitude	VII	WDC-Edinburgh	1.29	124.95
YAP	Equatorial	VIII	MAGDAS	9.30	138.50
GUA	Low-latitude	VIII	INTERMAGNET	13.58	144.87

Note. Their correspondent data repository and geographical coordinates (in degrees) are also shown.

related to the sparsity and quality of the data, we use a principal component analysis (PCA) technique to model an initial EEJ data set. Our approach is based on the combination of ground and satellite EEJ data, since ground data provides very good local time coverage and satellite data provides very good longitudinal coverage, maximizing the spatiotemporal coverage needed when Equation 2 is considered.

2. Data and Model

Data sets from 2000 to 2019 were used and the subsections below detail the type and origin of each data set.

2.1. Ground-Based Magnetic Data

Hourly mean values in units of nanotesla from geomagnetic observatories and magnetometer stations were used. The hourly time resolution is sufficient for resolving the local time variation of the EEJ and it makes the inversion computationally less expensive. Minute mean values were also used during data preprocessing, as explained in Section 3.1. The data were constrained to geomagnetically quiet periods with the geomagnetic activity index Kp (Matzka et al., 2021) being less than or equal to 3.

To obtain EEJ data from the ground-based magnetometer data, we used the 2-station method described in Soares et al. (2018). To extract the EEJ signal, we calculated the difference between the H component measured at an equatorial station and at a low-latitude station with a similar longitude, but outside the influence of the EEJ. Then, the nighttime quiet level is defined for each longitude by calculating the average for each night by using an interval of 4 hr around local midnight and, then, linearly interpolating a baseline between successive nighttime averages. After subtracting the nighttime baseline, the final EEJ signal at the longitude of the equatorial station is obtained, being hereinafter referred to as ΔH.

Table 1 lists all the ground observatories and stations used to derive the EEJ signal, their type (equatorial or low-latitude), sector (identifier of longitudinal sector), source of data, geographic latitude, and longitude. A total of eight longitudinal sectors (I up to VIII) are used as sources of ground-based data in the analysis. In Table 1, the sector VII shows one equatorial station (DAV) and two low-latitude stations (MUT and TND) because TND
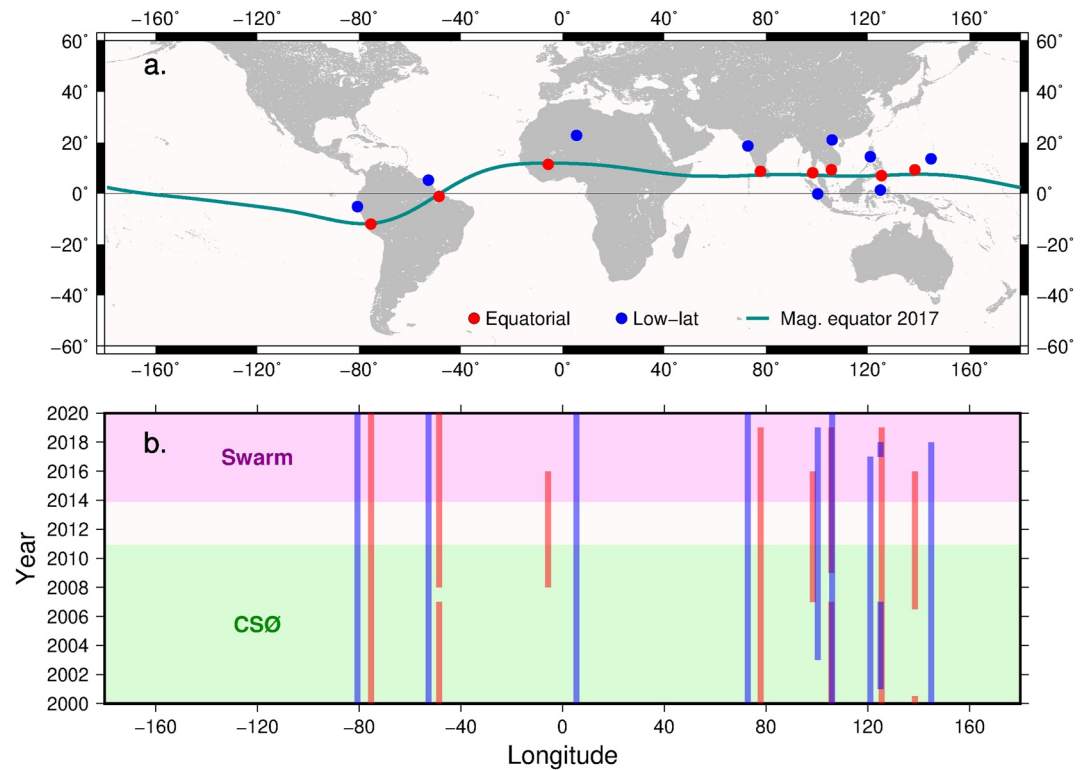
Figure 1. (a) Map with the equatorial (red circles) and low-latitude (blue circles) stations used to obtain the ground-based equatorial electrojet (EEJ) data. The magnetic equator for 2017 is shown as a green line. (b) Longitude versus time plot indicating the data availability of each ground station shown in the panel a (red and blue lines) and of the satellite missions (Swarm as purple area and CHAMP/SAC-C/Ørsted as green area with CSØ label), for the period from 2000 to 2019.

is used for 2017, when MUT data are not available. The data providers and data repositories used are listed in Table 1 and in the data availability statement section.

Figure 1a indicates the geographical positions of the equatorial (red circles) and low-latitude (blue circles) stations, as well as the magnetic equator for 2017. Figure 1b is a longitude versus time plot indicating the data availability for each ground station from 2000 to 2019 (red and blue lines). Note that if a data gap occurs either at the equatorial or at the low-latitude station, there will be a correspondent data gap in the EEJ data set. Some station pairs provide quite continuous records, namely HUA-PIU (Peru), TTB-KOU (Brazil), and TIR-ABG (India).

2.2. Satellite Magnetic Data

Geomagnetic data from the Ørsted, CHAMP, SAC-C, and Swarm satellite missions were used. Ørsted operated from February 1999 to January 2014 in a near polar orbit with an inclination of 96.5°, an apogee around 865 km, and a perigee around 650 km, drifting slowly in local time by −0.88 min per day (Neubert et al., 2001). CHAMP operated from July 2000 to September 2010 with a local time drift rate of 5.44 min per day, starting its operation at a height of 454 km and progressively decayed until 200 km in 2010 (Reigber et al., 2002). SAC-C operated from November 2000 to August 2013 (data after 2010 is not used in this study to be consistent with CHAMP data availability) in a polar circular orbit of altitude 702 km with an inclination of 98.2° (Colomb et al., 2004). Its orbit is sun-synchronous, and it remains sampling at a fixed local time of around 10:25 a.m. Swarm is a constellation of three satellites launched in November 2013 (Friis-Christensen et al., 2006, 2008) and still under operation. Swarm A and C fly at an altitude of around 450 km, while Swarm B flies at an altitude of around 530 km. The Swarm satellites drift faster in local time than the other missions, at an average rate of 10.5 min per day. The satellite data can be divided into two periods, according to its availability through time: the CHAMP/SAC-C/Ørsted (CSØ) period from 2000 to 2010 (indicated as green area in Figure 1b) and the Swarm period from 2014 to 2019

Table 2
Summary With the Linear Transformation Coefficients c_1 and c_2 for Each Satellite

Satellite	c_1	c_2
Swarm A and C	0.8511	1.3254
Swarm B	0.8600	−1.5380
CHAMP	0.9353	−2.0449e−5
SAC-C	0.7987	−7.6188
Oersted	0.7073	2.0909e−4

(magenta area in Figure 1b). There is a data gap between the end of the CSØ period and beginning of the Swarm period, as also indicated in Figure 1b.

The satellite EEJ data used in this work are given as electric current intensity values in mA/m. The EEJ electric current intensity is obtained by inverting the observed satellite magnetic field data based on an EEJ current sheet model. To perform the inversion, first, it is necessary to remove the core (by the CHAOS-6 model, Finlay et al., 2016), lithospheric (by the MF7 model, Maus et al., 2008), magnetospheric (by the POMME-6 model, Maus & Lühr, 2005), and Sq (by fitting a low-degree spherical harmonic field model to the higher-latitude data) magnetic fields from the original 1 Hz magnetic field data that comes from a scalar magnetometer. After removing these contributions, the residual data represents the latitudinal magnetic signature of the EEJ current for every orbit on the dayside. Then, the EEJ magnetic signature along each track is inverted for an estimate of its height integrated current. To do this, we considered a simple sheet current model of line currents spaced at 0.5°, flowing longitudinally eastward along lines of constant quasi-dipole latitude, and at an altitude of 110 km. In this work, only the peak EEJ value at the magnetic equator is used (i.e., no latitudinal averaging is performed). This approach is presented and explained in detail in Alken et al. (2013). Like the ground-based data, the Kp index ≤3 criteria was also used to constrain the satellite data to geomagnetically quiet periods.

To guarantee that the EEJ variation is represented similarly by the different satellite data sets, we performed an intercalibration of Swarm, CHAMP, Ørsted, and SAC-C data by using a common reference data set. The EEJM-2 model (Alken & Maus, 2007; detailed in Section 2.4) was used as the reference data set. We calibrated each satellite data set by minimizing its differences to the EEJM-2 reference values. This minimization was achieved by finding linear transformation coefficients c_1 and c_2 in the form $c_1 \cdot data + c_2$, determined separately for each satellite data set (listed in Table 2). This procedure was used to minimize any bias or sensitivity differences between the satellite data sets. The effect of the different satellite orbital altitudes was already addressed during the EEJ current modeling by the Bio-Savart law (Alken et al., 2013). Therefore, no further altitude correction is necessary, as the current intensities are always estimated at the same reference of 110-km height.

2.3. SABER Temperature Satellite Data

Tidal signatures in the EEJ are compared with tides in the neutral atmosphere. The tides that affect the EEJ current involve temperature perturbations in the same height interval, and therefore similarity can be expected in temporal variations of tidal signatures in the EEJ and temperature.

We use atmospheric temperature data from the NASA TIMED (Thermosphere Ionosphere Mesosphere Energetics Dynamics) satellite (Kusnierkiewicz, 2003). The TIMED temperature data are recorded by its SABER (Sounding of the Atmosphere using Broadband Emission Radiometry) instrument (Russell et al., 1999). The SABER instrument performs global measurements of the atmosphere using a 10-channel broadband limb-scanning infrared radiometer covering the spectral range from 1.27 to 17 μm. SABER observes infrared emissions from CO_2, O_2, H_2O, NO, O_3, and OH. The measured thermal infrared radiance values are then mathematically inverted to obtain some of the mission data products (Russell et al., 1999), including atmospheric temperature that is used in this work. SABER temperature data have been widely used for studying tides in the mesosphere and lower thermosphere region (e.g., Forbes et al., 2008; Oberheide et al., 2009; Zhang et al., 2006).

2.4. EEJM-2 Climatological Model

The EEJM-2 is an empirical climatological model of the EEJ based on CSØ satellite data (Alken & Maus, 2007). Besides its use in the satellite data intercalibration, the model is also used to evaluate the effectiveness of our PCA model in representing the EEJ variations. The EEJM-2 model was chosen as a reference for comparison due to two main reasons. First, like our PCA proposed model, the EEJM-2 is an empirical model which takes advantage of important longitudinal coverage from satellite data. Second, unlike the other aforementioned EEJ models, the EEJM-2 is publicly available (https://geomag.colorado.edu/equatorial-electrojet-model-eejm2.html).

EEJM-2 uses different basis functions to represent the EEJ longitudinal, local time, season, and solar flux dependence. With our PCA approach, we aim at obtaining different basis functions to capture the EEJ variations.

3. Method

Our strategy can be divided into two parts: first, performing an improved EEJ modeling and, second, obtaining its spectrum related to solar tides.

The EEJ modeling consists of three main steps. First, the ground and satellite EEJ data are combined to form a hybrid data set with common unit of measure. Second, EEJ basis functions are derived from multiyear satellite observations using the PCA method. Then, in the last step, the EEJ for a specific time of an individual year is modeled by fitting the PCA basis functions to the combined ground-satellite data. The obtained PCA model will be hereinafter called as PCEEJ model.

The tidal analysis part is performed after the PCEEJ model is obtained. In this stage, tidal components are fitted to the PCEEJ model by using Equation 2. This fit provides the importance of each tide in the modeled EEJ. Lastly, comparisons between geomagnetic EEJ data analysis and SABER temperature data analysis are performed to evaluate their level of similarity. The EEJ and SABER tidal analyses are useful to confirm whether the PCEEJ has a realistic tidal composition or not.

3.1. Combining Ground and Satellite EEJ Data

In order to gain the best possible local time and longitudinal coverage for EEJ observations, we combine estimates of the EEJ from ground-based data (expressed as ΔH in nT) and from satellite data (expressed as peak height integrated current density ICD, in mA/m). To keep all observations with the same unit of measure prior to running the inversion, we converted the ΔH data given in nT to mA/m. We determine by linear regression the relationship between ΔH and ICD, that should be linear for the same longitude (Manoj et al., 2006). This was done by selecting ground and satellite data from the same temporal interval in a longitude interval defined as the longitude of the ground data $\pm 5°$. For this purpose, we used ground data with the temporal resolution of 1 min, and the analysis was made separately for different local times to take into account the local time dependence of the ΔH and ICD linear relation.

Figure 2 shows scatter plots of ICD versus ΔH for eight longitudinal sectors that have ground data available during the Swarm data period (see Table 1). In Figure 2 panels, data from all local times are shown together, with no distinction, to facilitate the visualization. The scatter plots confirm that a linear relation exists for all longitudinal sectors and a linear fit for each sector is shown at each panel of Figure 2, together with the corresponding linear fit slope (s) and intercept (i) coefficients.

3.2. Principal Component Analysis

Like the Fourier analysis, the PCA uses a series of orthogonal functions, but it does not use a fixed set of basis functions. The purpose of PCA is to find linear combinations of the data which are uncorrelated with each other and maximize the variance explained in the data, as mentioned, e.g., in Alken et al. (2017), which also applied PCA technique to data of ionospheric currents. It indicates which parts of a data set provide redundant information or noise that are not useful for understanding a given system. In practice, it also acts as a filter and reduces the dimensionality of a complex data set. This is achieved after constructing a covariance matrix of all the data and deriving the associated eigenvalues and eigenvectors. The eigenvalues represent the amount of variance explained by each eigenvector, and the eigenvectors represent the principal components (PCs; basis functions) that describe the initial data set. Based on cumulative variance analysis, the most important PCs are used to model the data (Alken et al., 2017).

Fitting selected PCA basis functions to the initially sparse EEJ data set yields a densified EEJ data. This is important because the usage of modeled and densified EEJ data helps to stabilize the inversion when extracting solar tidal signatures from the data set in a later step. For our PCA analysis, the idea is to use a very large satellite data set, with good longitude and local time coverage, to derive the basis functions. Thus, we have used a total of 17 years of satellite data: 11 years from CSØ period, and 6 years from Swarm period. Data from SAC-C and ground were not used in this step because they are limited to specific local time and longitudes, which could lead to certain level of bias in the basis functions.

The first step in PCA is to grid the data set according to longitude and local time bins. We found that increments of 10° in longitude and 1 hr in local time provide an optimum grid configuration, ensuring the presence of a
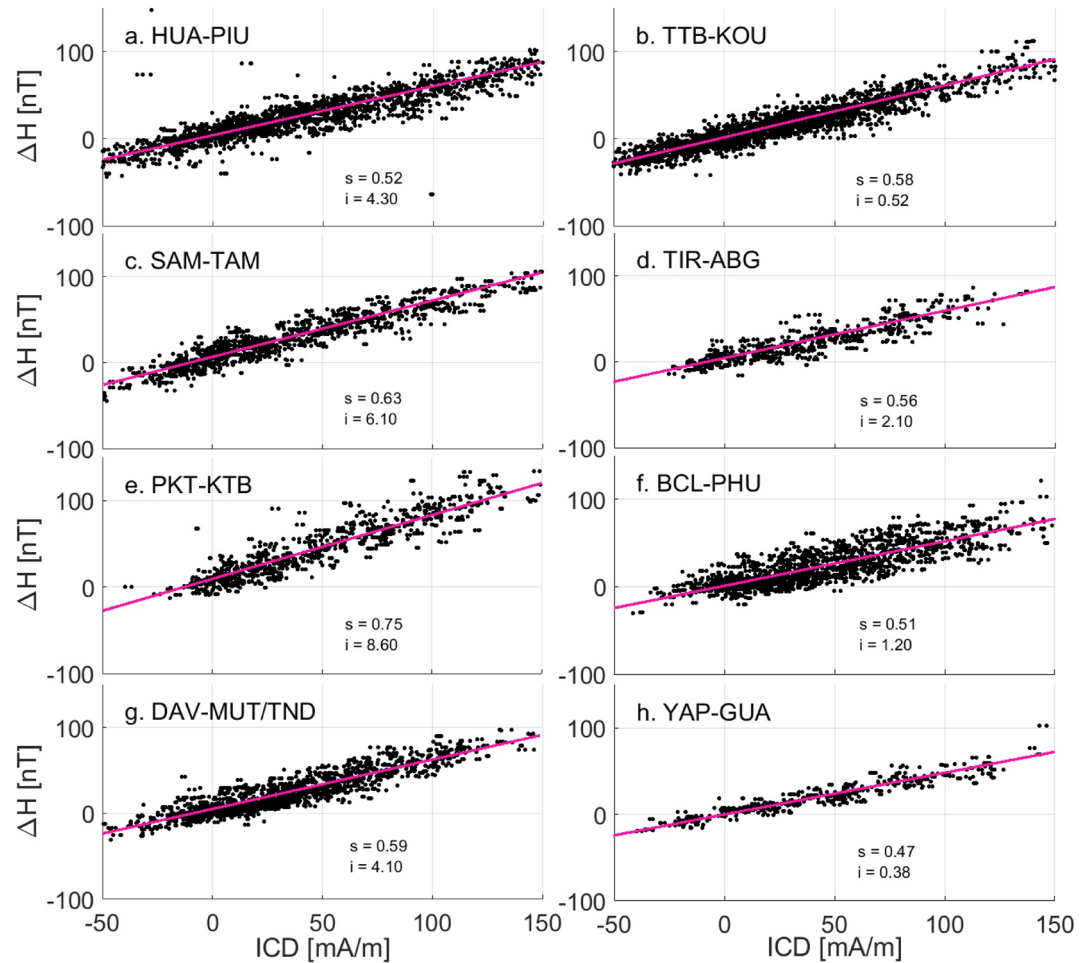
Figure 2. Relation between ground-based (ΔH, in nT) and satellite (ICD, in mA/m) equatorial electrojet (EEJ) data for eight considered longitudinal sectors, represented by the ground stations of HUA-PIU (a), TTB-KOU (b), SAM-TAM (c), TIR-ABG (d), PKT-KTB (e), BCL-PHU (f), DAV-MUT/TND (g), and YAP-GUA (h). Satellite data are from Swarm A and B satellites. Pink lines show a linear fit for each longitudinal sector. The obtained slope s and intercept i coefficient values of the linear fit in the form of $\Delta H = s*$ICD $+ i$ are shown in each panel. Data from all local times are considered in the panels.

substantial amount of data points in each bin. Figure 3 shows this grid, with dimensions of 36×12, resulting in 432 bins.
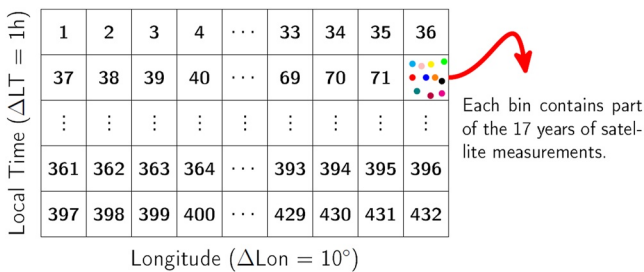
Figure 3. Sketch of the initial longitude versus local time data grid to be applied in the principal component analysis (PCA). Here, the optimum 36×12 grid is shown, with a total number of 432 bins. The numbering of each bin is shown, starting at bin number 1 and ending at bin number 432.

The second step is to construct an EEJ time series matrix X. As indicated in Figure 3 (red arrow), each grid bin contains some of the 17 years worth of EEJ satellite measurements. The time series matrix X reorganizes this information by splitting the data according to day of year (DoY). X has the dimension 432×365, with 432 lines from the longitude and local time binning and 365 columns representing the 365 days of the year. Each element of X is given as an average value, since more than one sample can be found with the same longitude, local time, and DoY.

The third step is the calculation of the covariance matrix COV related to X, as given by Equation 3

$$COV = \frac{1}{N} X X^T \tag{3}$$

where N (=365) is the number of samples for each bin. The eigenvalues and associated eigenvectors of the covariance matrix COV are computed
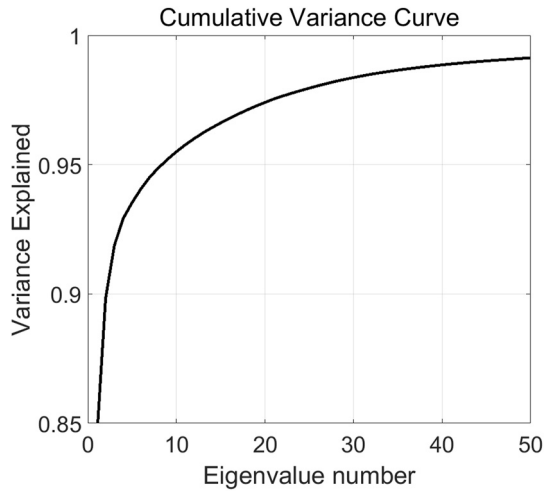
and stored in the matrices V_{val} and V_{vec}, respectively. The matrix V_{val} is a 432×432 diagonal matrix which is sorted in descending order, according to the eigenvalues, as shown in Equation 4

$$V_{val} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{d-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_d \end{bmatrix} \qquad (4)$$

where $\lambda_1 > \lambda_2 > \ldots > \lambda_d$ and $d = 432$.

The matrix V_{vec} is a 432×432 matrix which contains the PCs basis functions of EEJ variation in its columns, as shown in Equation 5

$$V_{vec} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1(d-1)} & v_{1d} \\ v_{21} & v_{22} & \cdots & v_{2(d-1)} & v_{2d} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_{(d-1)1} & v_{(d-1)2} & \cdots & v_{(d-1)(d-1)} & v_{(d-1)d} \\ v_{d1} & v_{d2} & \cdots & v_{d(d-1)} & v_{dd} \end{bmatrix} \qquad (5)$$

where PC_1 is given by column 1, PC_2 by column 2, …, PC_d by column d.

To capture the most important sources of variance within the data and therefore reduce its dimensionality, only a small number of PCs are selected to model the EEJ variation. Two methods were used to determine the number of PCs: the cumulative variance plot and the visual inspection of the PCs. The cumulative variance of the ith eigenvalue is defined as

$$CV_i = \frac{\sum_{j=1}^{i} \lambda_j}{\sum_{j=1}^{P} \lambda_j} \qquad (6)$$

where $P = 432$, the total number of eigenvalues of the matrix COV. Figure 4 shows the obtained cumulative variance plot. From a total of 432 eigenvalues, more than 95% of the variance within the data can be explained by the first 10 PCs. Figure 5 shows the first 10 PCs, organized in the longitude versus local time grid. The visual inspection of Figure 5 confirms that the first 10 PCs present structured spatiotemporal features (panels a–j). For instance, the most important contribution comes from PC1 (panel a), which resembles the average spatiotemporal variation of the EEJ, as seen in previous studies (Lühr & Manoj, 2013; Lühr et al., 2008).

3.2.1. Principal Components Fit

The PCEEJ was obtained by fitting the PCA basis functions to the hybrid ground-satellite data set. This was done by representing the available data set as a linear combination of the 10 PCA basis functions, as shown in Equation 7

$$\tilde{b} = \alpha_1 . PC_1 + \alpha_2 . PC_2 + \cdots + \alpha_{10} . PC_{10} \qquad (7)$$

where $\tilde{b}$ stands for PCEEJ model, α_i stands for the PCs amplitudes, and PC_i are the basis functions shown in Figure 5. Equation 7 builds up a linear system in which the amplitudes α_i are unknown parameters to be determined. To estimate α_i, we solved a minimization problem based on a least-squares solution, with the following estimator:

$$y = \left\{ A^T A + \mu L^T L \right\}^{-1} \left\{ A^T b \right\} \qquad (8)$$

Figure 4. Cumulative variance explained by the principal components, based on the associated eigenvalues. Only eigenvalues from 1 to 50 are shown for better visualization.
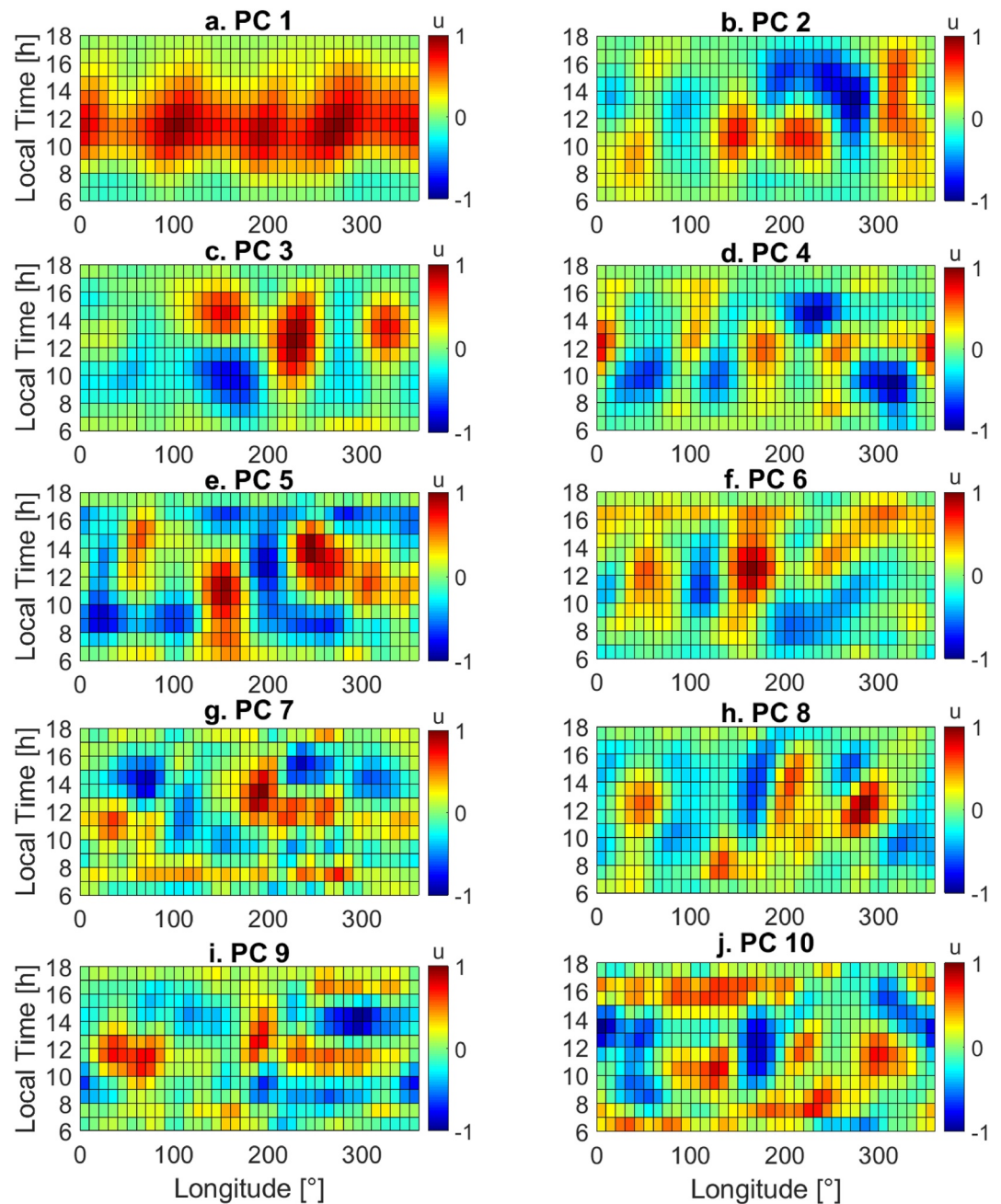
Figure 5. Principal components 1–10 (panels a–j, respectively), derived from equatorial electrojet (EEJ) satellite data. Note that the color bar shows normalized units of variation between −1 and +1.

where y represents the parameters under estimation (amplitudes of PCs), A is the linear operator matrix that contains the PCs in its columns, b is the observed data, μ is the regularization parameter, and L is the regularization matrix.

We chose a regularized scheme because the data vector b is sparse in time and space. This sparsity depends on the length of the time window considered in the analysis. Focusing on the establishment of a new technique, we decided to follow a conservative approach and use a time window of 70 days, which is approximately the period needed by the Swarm satellites to cover all local times. Smaller windows could be used, but this would affect the level of confidence of the results due to reduced data availability. This means that each solved inverse problem contains data from 70 days. A running window approach is used, so the window moves by one consecutive day.
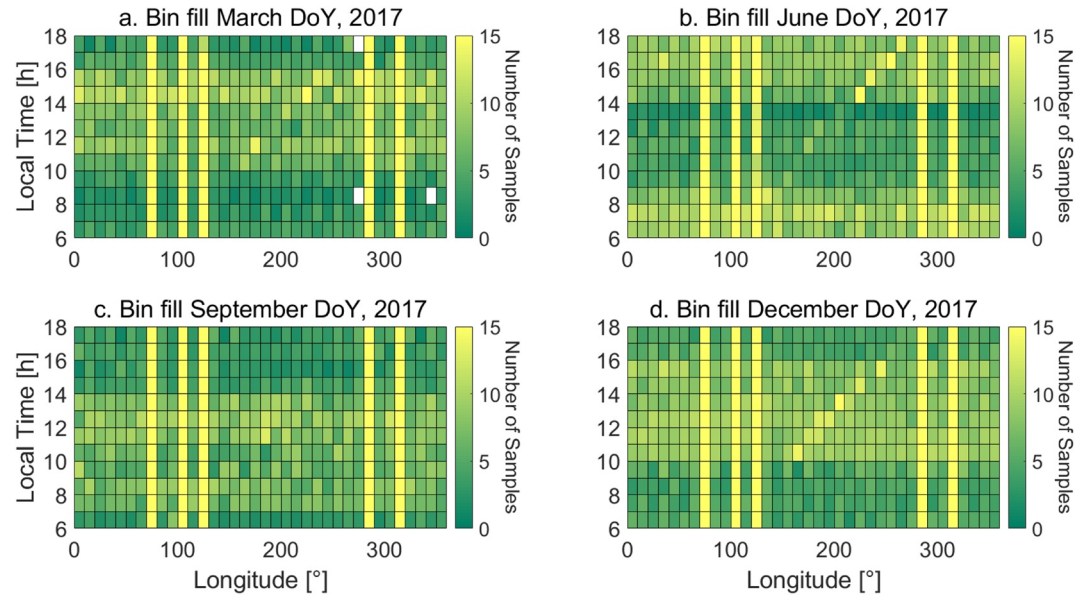
Figure 6. Longitude versus local time grid with color indicating the number of samples within each bin for different epochs of 2017: March (a), June (b), September (c), and December (d).

Even when a time window of 70 days is used, data gaps can occur and lead to an ill-posed inverse problem, depending on the available data set. To overcome possible instabilities, our regularized solution incorporates a priori information about the desired solution (Hansen, 2010). This bias is introduced in the problem by the regularization matrix L, which is a diagonal matrix containing the inverse of the PCA eigenvalues squared-root, i.e., equal to $1/\sqrt{V_{val}}$ (following Equations 17–19 of Alken et al. (2017)). In other words, we assign more weight to the PCs that explain more of the data variance when modeling the EEJ. The degree of this forcing is then controlled by the μ parameter, defined after examining the predicted data and residuals to the actual observations, an approach analogous to the L-curve logarithmic plot (Hansen, 2010). Thus, we benefit once again from the PCA technique as it provides a priori information to constrain the inverse problem well.

In addition, when performing the fit of Equation 8, we gave more weight to the data that comes from satellite than those that comes from ground. This is done to avoid possible bias arising from the data type in the inversion. On average, the ground data contributes to about 10–20% of the 432 data bins. Although it covers less than half of the total number of bins, the ground data provides a very large number of samples for the fitting, as it is always measuring at the same longitude. This unbalanced distribution between ground and satellite data can be seen in Figure 6, which shows the number of samples within each grid bin for different epochs of 2017: March (a), June (b), September (c), and December (d). In all panels, it is possible to identify five longitudinal sectors with constantly high number of samples (vertical yellow structures) that are caused by the presence of ground data. Thus, we weight the data sets according to their percentage of bin filling related to the total 432 bins.

Our final PCEEJ model is available as a data publication (Soares et al., 2022) in GFZ Data Services, where its 10 PCA basis functions and its final model values are provided.

3.3. Inverting EEJ for Tides Amplitudes

We fitted Equation 2 to the PCEEJ model in order to determine the time series of solar tides amplitudes that explain the EEJ variation. The tidal modes considered for the fit are those with zonal wavenumber s ranging from -6 to $+6$ and period n of 24, 12, 8, and 6 hr (see Equation 2). The fit is done in a least-squares sense and the estimator of the tidal amplitudes is given by Equation 9

$$p = \left\{ M^T M + \mu_{RR} I \right\}^{-1} \left\{ M^T \tilde{b} \right\} \tag{9}$$

where p is the model parameter containing the tidal amplitude information, M is the linear operator containing the trigonometric functions that describe the tidal components propagation (Equation 2), $\tilde{b}$ is the EEJ modeled

by PCEEJ, $\mu_{RR}I$ is a ridge-regression regularization term added to the problem formulation to yield more stable parameters inversion. The ridge-regression regularization term adds a priori information that forces the solution to tend to zero, which in practice just attenuates abnormally large unstable values of solution.

3.4. EEJM-2 Model Settings

The EEJM-2 model data were obtained by running the model for each DoY within the years from 2003 to 2018, with a longitude spacing of 1°, local time varying with 1 hr, lunar time calculated based on the solar time and the EUVAC parameter calculated based on daily F10.7 index data (Alken & Maus, 2007). To allow a direct comparison with the observed and the PCEEJ data, the EEJM-2 data were averaged according to the 70-day time window.

3.5. Comparison With SABER Temperature Data

Equation 2 was used to fit the SABER temperature data and retrieve solar tides amplitudes, analogously as was done for the geomagnetic EEJ data. Prior to inversion, SABER data preprocessing included the selection of data within the altitude interval from 100 to 110 km and the geographic latitude interval from −45° to +45°. Data points with abnormal values were also discarded. As SABER provides very good spatiotemporal coverage compared to the EEJ data, PCA modeling was not necessary. The parameter estimation from Equation 8 was applied for SABER data inversion. Like in the EEJ data analysis, time windows of 70 days were used when performing the fit.

While comparing inversion results obtained from EEJ and temperature data, an emphasis is given to nonmigrating tides. This is because migrating tides in the EEJ are mostly due to direct solar radiation effects on the ionospheric conductivity, rather than by migrating tides in the neutral atmosphere. Even if there is no tidal forcing from the neutral atmosphere, the ionospheric conductivity is high during day and low during night, which would lead to migrating tidal components in the EEJ. On the other hand, nonmigrating tidal components of the EEJ are strongly affected by nonmigrating tides in the neutral atmosphere (Lühr & Manoj, 2013).

4. Results and Discussion

4.1. Observed EEJ, PCEEJ, EEJM-2, and Reconstruction With Tides

Figure 7 shows a comparison between the observed EEJ, PCEEJ model, EEJM-2 model, and EEJ reconstructed with tides data for the 70-day time window centered at the 15th day from March, June, September, and December months from the year of 2017.

The observed EEJ data shown in Figures 7a–7d were obtained by calculating the median of the samples found within each bin of the grid. The EEJ spatiotemporal variation and its seasonal variation shown for the observed data are in agreement to the expected average pattern, as the occurrence of the prominent wave-4 longitudinal structure around September (Lühr & Manoj, 2013). For the 2017 period, a very high percentage of the total number of data grid bins is filled. For instance, in Figure 7, the occurrence of data gaps can only be identified around March (panel a), which shows 99% of bin filling with only three white-colored bins related to data gaps. However, for other years, this percentage can go under 70%. This occurs due to the different data sets available for data analysis but also due to differences in satellite mission sampling. In the case of the Swarm era, the constellation is sampling at very similar local times during the beginning of the mission. But then, in 2017/2018, the Swarm satellites start to sample 6 hr apart, maximizing the spatiotemporal coverage, increasing the available information for PC fit and reducing drastically the occurrence of data gaps. Thus, we take advantage of this convenient property of the Swarm mission and use the period from 2017/2018 as a benchmark since it provides the best data coverage from our data set.

The PCEEJ (panels e–h) was obtained by the approach described in Section 3.2.1. The model results are in agreement with those from the observed data in terms of amplitude and in terms of spatiotemporal behavior for the DoYs shown as examples. As expected, the PCEEJ model provides dense EEJ intensities that smoothly change with longitude and local time without the noise-like fluctuations seen in the observations.
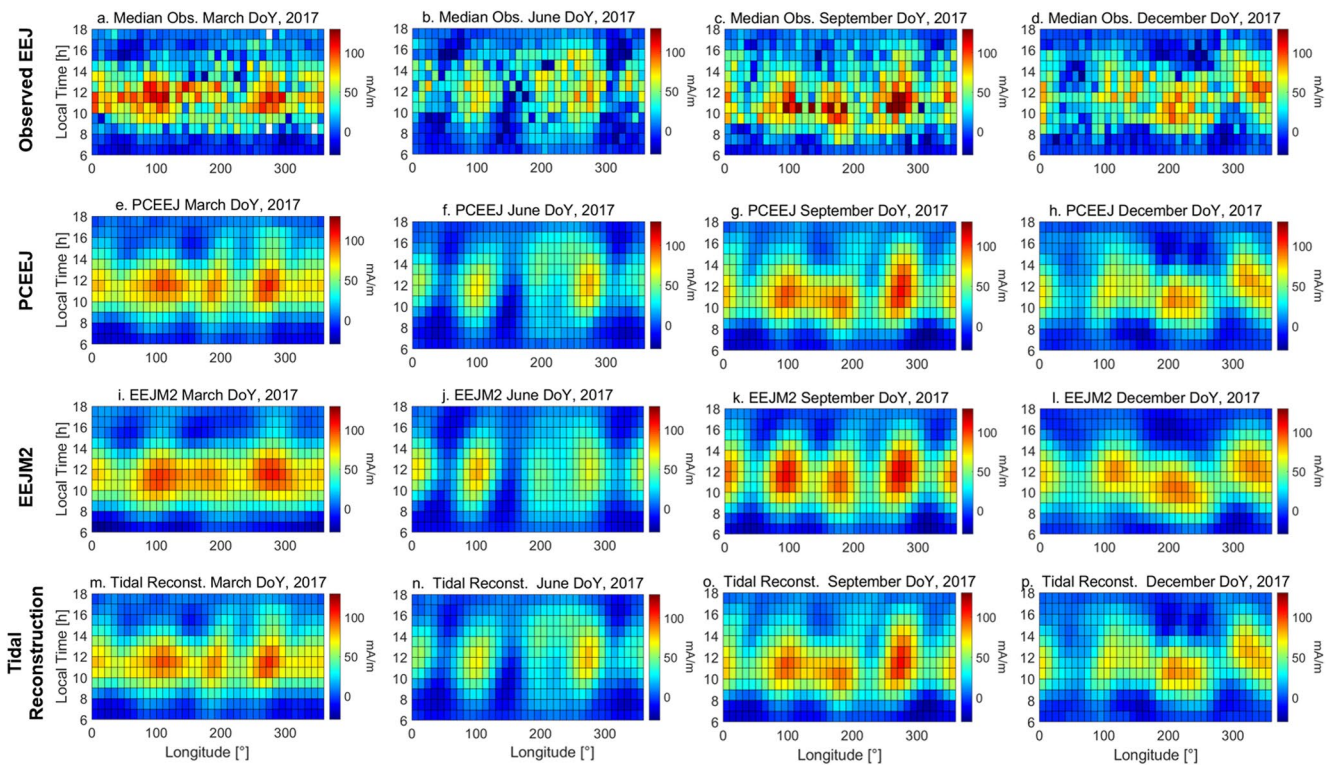
Figure 7. Observed equatorial electrojet (EEJ) data shown as median value per bin for 70 days centered on March (a), June (b), September (c), and December (d) 2017. Displays from the same selected days of 2017 are shown for the PCEEJ model (e–h), EEJM-2 (i–l), and the EEJ data reconstructed as the sum of tidal modes with |s| ≤ 6 and n = 1–6 (panels m–p).

By comparing the PCEEJ with the EEJM-2 results (panels i–l), we see that the overall long-wavelength features are in agreement. For the examples shown in Figure 7, there are minor differences between PCEEJ and EEJM-2 that can be related on how well each model can explain the observations in time and space.

The EEJ reconstructed as the sum of the tidal components with s from −6 to 6 and n from 1 to 6 are shown in Figure 7 (panels m–p). These results were derived through the inversion process described in Section 3.3. If the tidal amplitudes are properly captured and the inversion scheme is robust, the reconstructed EEJ should reproduce to a good extent the input from the PCEEJ model. Indeed, this is noted when comparing the panels from the second and fourth rows of Figure 7.

For completeness, Figures 8 and 9 show the comparison between the observed EEJ, PCEEJ model, EEJM-2 model, and EEJ reconstructed with tides data for the years 2018 and 2016, respectively. A consistent EEJ climatology can be observed from year-to-year, for all data sets, when comparing Figures 7–9. The results obtained for PCEEJ, EEJM-2, and EEJ reconstruction data for 2017 are comparable to those obtained for 2018 and 2016. Again, there is an overall agreement between PCEEJ and EEJM-2 long-wavelength features and some differences in the shorter wavelength scale.

It is interesting to note that, like 2017, the 2018 observed data indicate a very good data coverage with few data gaps. However, 2016 observed data show an increase in the data gaps due to the more redundant Swarm satellites coverage. The PCEEJ model is robust enough to deal with such increase in data gaps, as it preserves the well-known EEJ features without creating any artifact or imprint from the data gaps.

In addition, it is possible to identify year-to-year EEJ amplitude changes in the observed and modeled EEJ due to solar activity changes. The 2018 amplitudes are smaller when compared to 2017, while the 2016 amplitudes are the largest.
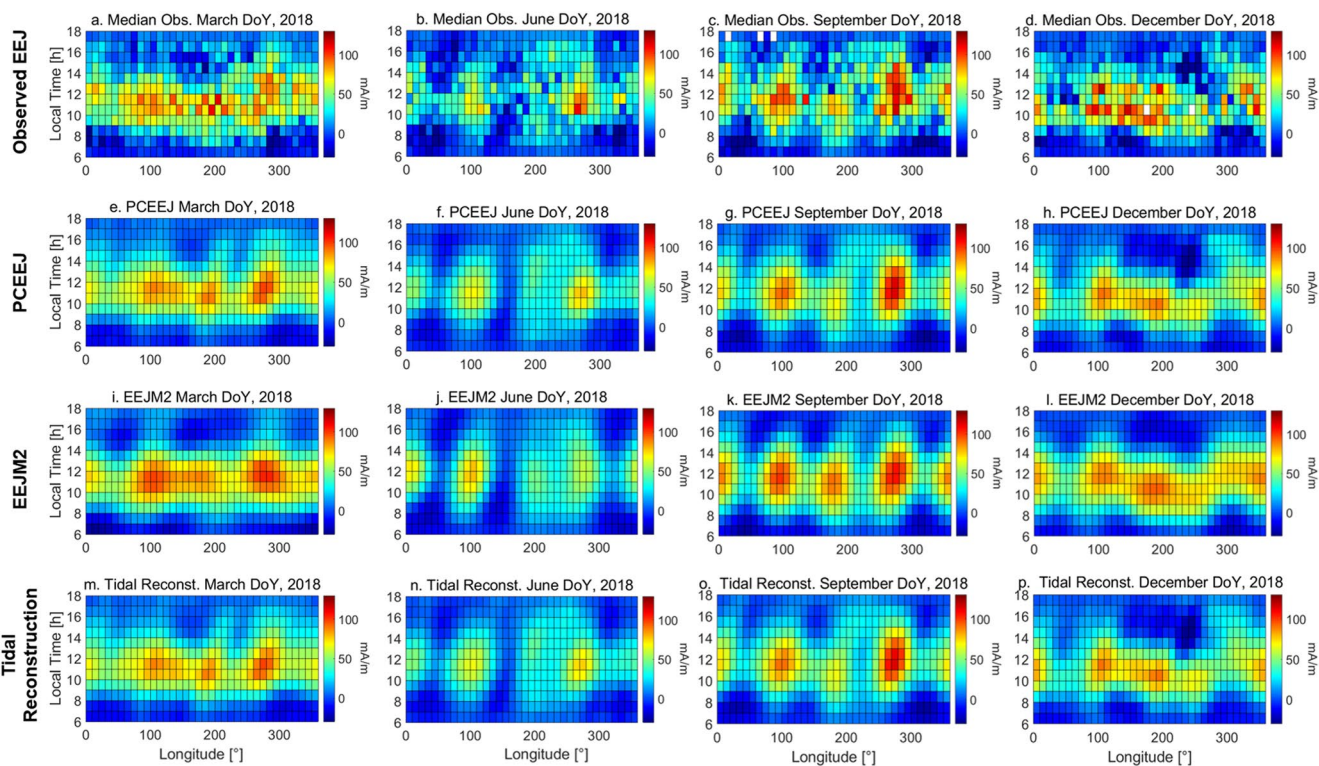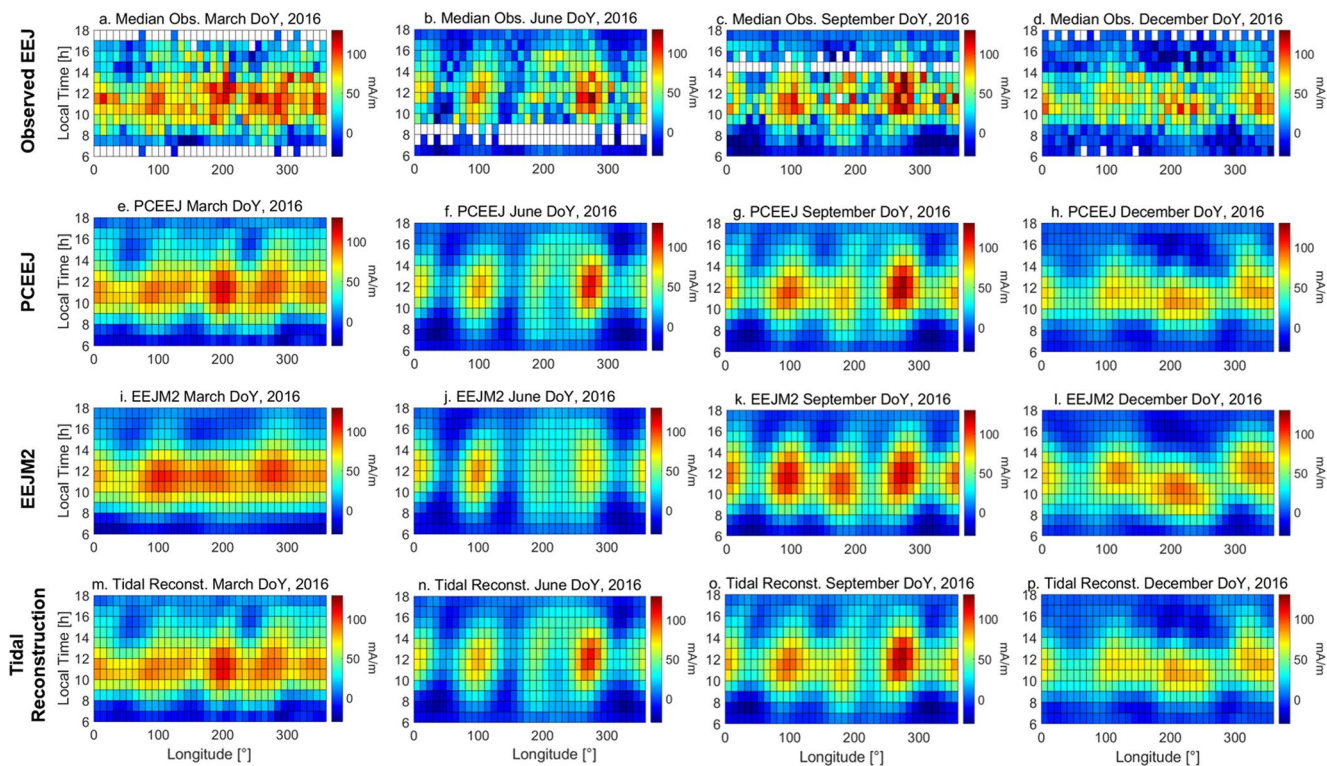
Figure 8. The same as Figure 7, but for 2018.
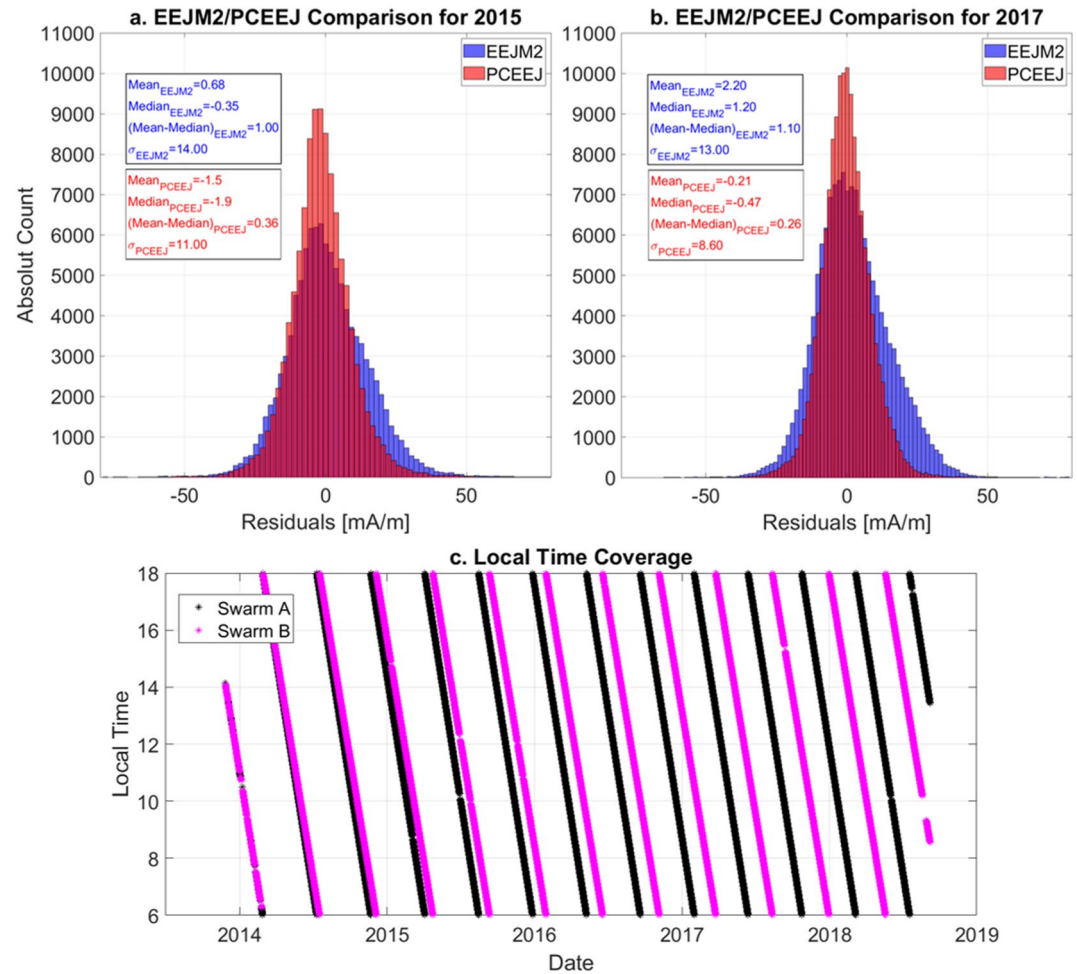

Figure 9. The same as Figure 7, but for 2016.

Figure 10. Histograms of residuals between observed data and EEJM-2 model (blue) and between observed data and PCEEJ model (red) for 2015 (a) and 2017 (b). Panel c shows the variation of the local time coverage of satellites Swarm A and B from 2013 to 2018.

4.2. PCEEJ Versus EEJM-2

4.2.1. Match to Observed Data

An important part of assessing the accuracy of the proposed PCA modeling scheme relies on the comparison with the EEJM-2 model. Figure 10a shows the distribution of the residuals of the modeled EEJ to the observed data for 2015 (PCEEJ model in red and EEJM-2 in blue). Here, all 432 grid bins from each of the 365 fits per year are considered. The low values of residuals mean and median indicate that both EEJM-2 and PCEEJ can well represent the average EEJ. The standard deviation and the difference between mean and median (as a measure of skewness of the distribution) for the PCEEJ model are smaller. Apart from the aforementioned modeling differences, these statistical differences can also be attributed to the absence of Swarm data during the construction of EEJM-2.

Figure 10b is analogous to Figure 10a, but for the year of 2017. It shows that (i) as for 2015, the PCEEJ model represents better the observed data than EEJM-2; (ii) the PC fit for 2017 is better than for 2015, as the mean, median, and standard deviation sigma values are reduced (−0.21, −0.47, and 8.60 mA/m, respectively). Meanwhile, EEJM-2 shows very similar statistical results for 2015 and 2017. The differences in the PC fit quality between 2015 and 2017 can be attributed to the improved local time coverage of the satellite data as the difference in local time sampling between the Swarm A and B that reached its maximum value of ∼6 hr around 2017 (Figure 10c).
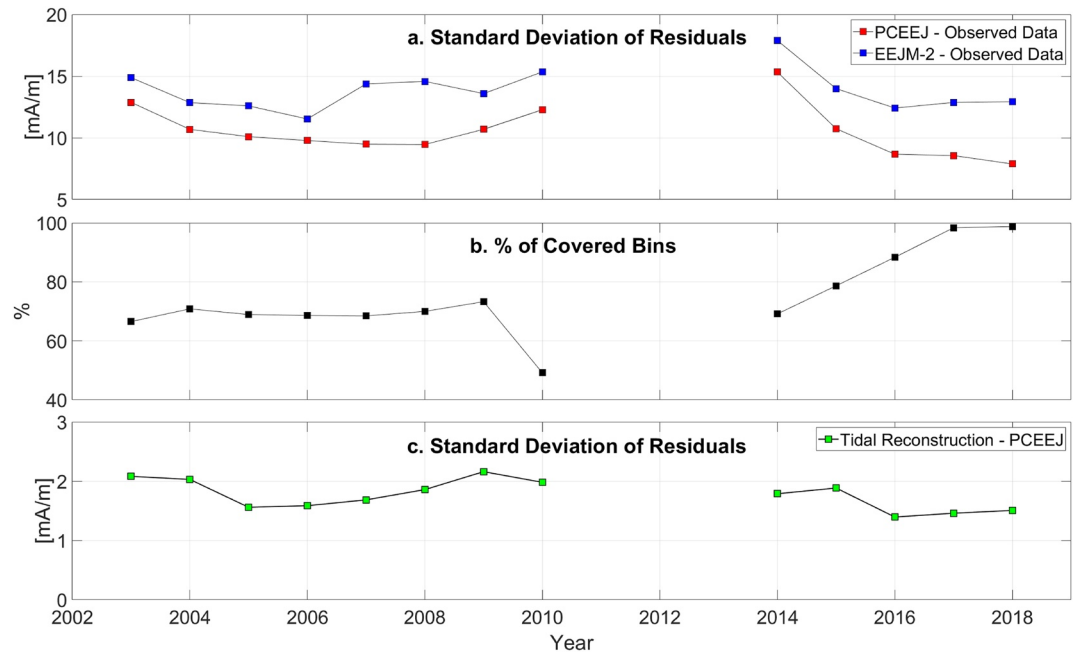
Figure 11. (a) Standard deviation of residuals between equatorial electrojet (EEJ) models (PCEEJ in red and EEJM-2 in blue) and EEJ observations, shown as time series from 2003 to 2018. (b) Percentage of data grid bins covered for each year. (c) Standard deviation of residuals between EEJ reconstructed by tidal fits and the PCEEJ model.

A yearly overview of the residuals between the models and observations is shown in Figure 11a. It shows the residuals standard deviation as time series from 2003 to 2018, where red squares are related to the PCEEJ model and blue squares are related to the EEJM-2 model. Each square represents the standard deviation of the residuals between model and observations for a complete year. The standard deviation values related to the PCEEJ model are consistently smaller than those related to EEJM-2. The PCEEJ standard deviation shows a decreasing trend in 2016, 2017, and 2018, with a minimum in 2018, which are years that benefit from the improved Swarm constellation local time coverage. However, the EEJM-2 model standard deviation shows a minimum in 2006 and an increasing trend for 2016, 2017, and 2018. These results indicate that the PCEEJ can indeed better represent the EEJ observations (for both CSØ and Swarm periods).

For completeness, Figure 11b shows the percentage of bins from our data grid that are covered through each year. As the data from each day is organized in a grid of 432 bins, each year will correspond to a total of 157,680 bins to be filled (432 bins times 365 days). Figure 11b shows that the percentage of bins covered during the CSØ period remains always at a similar level of around 70% (except in 2010 due to the end of the CHAMP mission). On the other hand, we observe an important increase in the percentage of covered bins during the Swarm period, going from around 70% in 2014 to nearly 100% in 2018. By comparing Figures 11a and 11b, an anticorrelation between number of covered bins and standard deviation of PCEEJ residuals is visible.

The residuals between the EEJ reconstructed by tidal fits and the PCEEJ model are also addressed in Figure 11c, again as time series. The yearly averages of these residuals (not shown here) are rather small, ranging between 0 and 2 mA/m. Here, in Figure 11c, the standard deviation of these residuals is shown, always between 1 and 3 mA/m. These results indicate that the tidal fit can reproduce the PCEEJ well throughout the years investigated in this study. Besides the small residuals, it is important to note that the reconstructed EEJ is in agreement with the expected EEJ variation, as seen in Figure 7 and related discussion.

As an additional quality check, we calculated the correlation coefficient between the modeled EEJ (PCEEJ and EEJM-2) and the observed EEJ for each bin of our 36 × 12 data grid, using data from 2017. The results are shown in Figure 12a for PCEEJ and Figure 12b for EEJM-2, respectively, with the correlation coefficient values ranging from −0.5 to +1. Figure 12a clearly shows higher values of correlation (toward yellow color) when compared to Figure 12b. Figures 12c and 12d show a more quantitative analysis by displaying the distribution of the correlation coefficient values and their associated median value, which is greater for the PCEEJ case (0.8269) than for
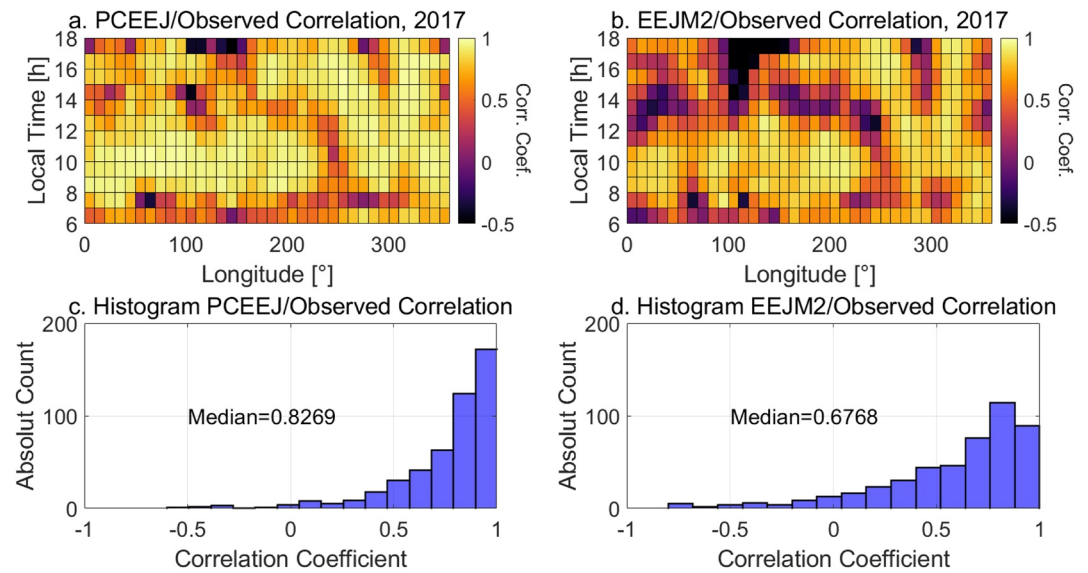
Figure 12. Correlation coefficient between observed data and PCEEJ model and between observed data and EEJM-2 model ((a) and (b) for distribution on grid and (c) and (d) for histogram, respectively).

the EEJM-2 case (0.6768). Therefore, the PCEEJ is better correlated to the observed EEJ in space and time when compared to EEJM-2, for the benchmark of 2017.

These results confirm that a better modeling of EEJ variations can be achieved with the PCA technique proposed in this study. However, this depends on the available data set and its spatiotemporal coverage.
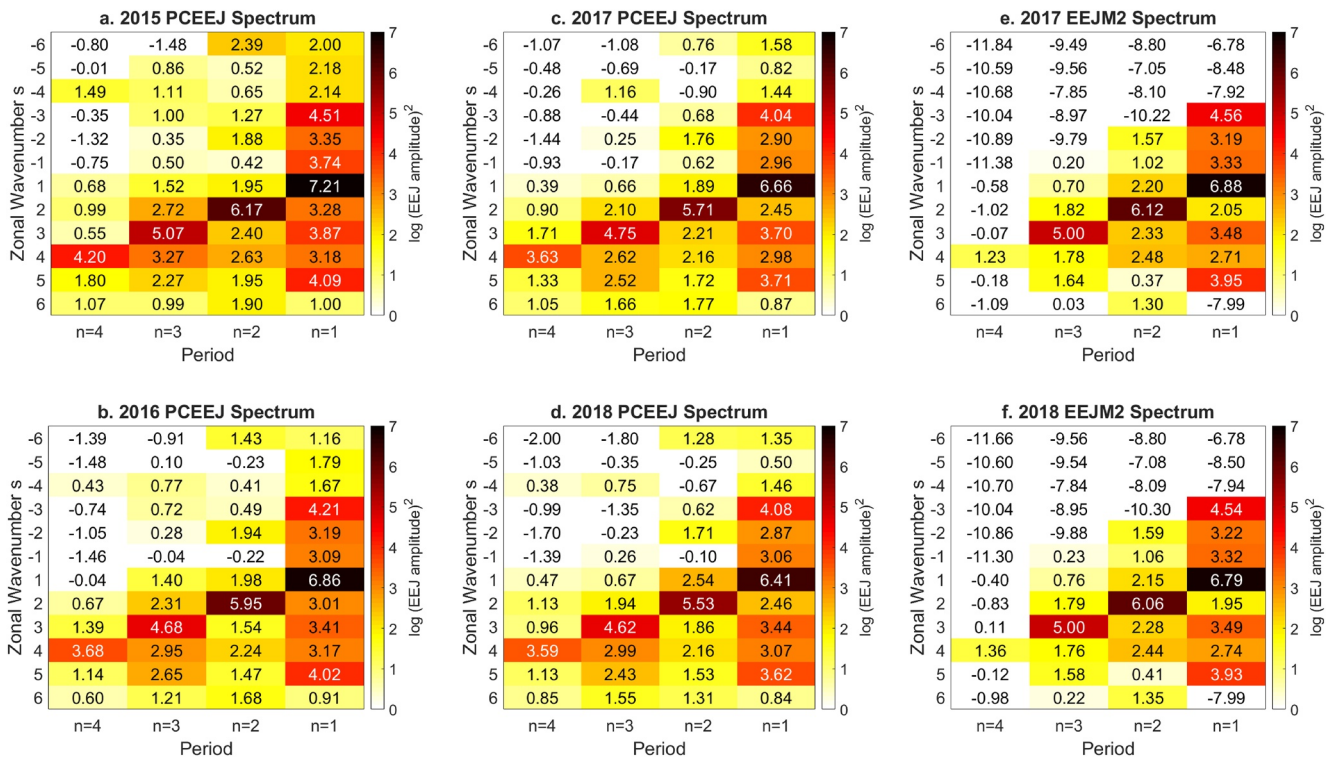
Figure 13. Average yearly equatorial electrojet (EEJ) tidal spectrum given by $\log(mA/m)^2$ for each combination of n and s, during the years of 2015 ((a) for PCEEJ), 2016 ((b) for PCEEJ), 2017 ((c) for PCEEJ and (e) for EEJM-2), and 2018 ((d) for PCEEJ and (f) for EEJM-2). The color bar indicates the values of $\log(EEJ\ amplitude)^2$ or, equivalently, $2\log(EEJ\ amplitude)$.
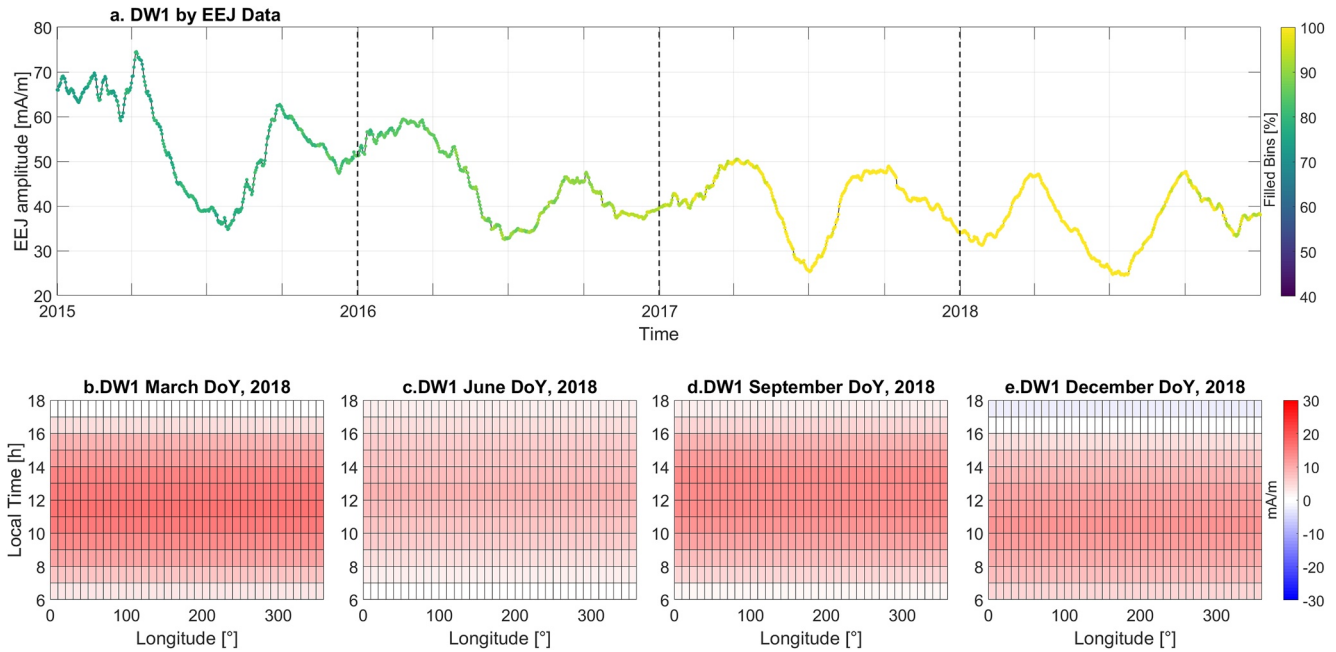
Figure 14. (a) Time series of DW1 amplitude obtained from equatorial electrojet (EEJ) data during the Swarm period (*x* axis represent time with the years as labels). Vertical dashed black lines indicate the beginning of a new year. Panels b–e indicate the contribution of DW1 to the final predicted EEJ in terms of positive and negative perturbations for selected days of year from March, June, September, and December 2018, respectively.

4.2.2. Average Tidal Composition per Year

The year-to-year variation of the average amplitude obtained for each tidal component in the EEJ is shown in Figure 13, given as log(EEJ amplitude)^2. These average amplitudes are obtained by calculating the average of the 365 values available for each year after running the inversion based on the 70-day moving window. The year-to-year variation can be observed by comparing panels a, b, c, and d, which show the PCEEJ average tidal composition for the years 2015, 2016, 2017, and 2018, respectively. For comparison purposes, the EEJM-2 average tidal composition, obtained in the same manner of the PCEEJ case, is also shown for the years 2017 and 2018 in panels e and f, respectively. Each panel in Figure 13 is formed by 48 bins and each bin represent a different tide, as indicated by the different combinations of period (*n*, *x* axis) and zonal wavenumber (*s*, *y* axis).

Concerning the PCEEJ results, most of the tidal amplitudes tend to decrease from 2015 to 2018, based on the color-coded label. This trend can be mainly attributed to solar cycle effects, as 2015 is closer to the solar maximum than 2018. This feature was also reported by Lühr and Manoj (2013), who compared the EEJ tides amplitudes from periods of solar maximum and minimum, based on two data subsets of 5-year averages (2000–2005 and 2005–2010). Here, we are directly comparing the results of independent years. The PCEEJ and EEJM-2 2017 and 2018 spectra show similar amplitude distribution for the main migrating and nonmigrating tides. However, the EEJM-2 has many nonmigrating tides with negligible amplitudes, as indicated by Figures 13e and 13f white-colored bins with tiny logarithmic values (considered as constantly zeroed time series). The individual case of the nonmigrating tide DE4 ($n = 1$, $s = -4$) is discussed in Section 4.4 as one example of important discrepancy between PCEEJ and EEJM-2 tidal composition.

4.3. Migrating Tides Time Series

Figures 14a and 15a show the time series obtained after EEJ inversion for the DW1 and SW2 migrating tides, respectively. Both figures show the results for the Swarm period, from 2015 to 2018, with vertical dashed black lines indicating the beginning of a new year, and the color bar in panel a showing the percentage of filled bins for each day (i.e., amount of data available for inversion). By comparing Figures 14a and 15a, DW1 shows larger amplitudes than SW2 and that both components present a similar seasonal variation with maxima at equinoxes and minima at northern hemisphere summer. These features agree very well with those reported by the average
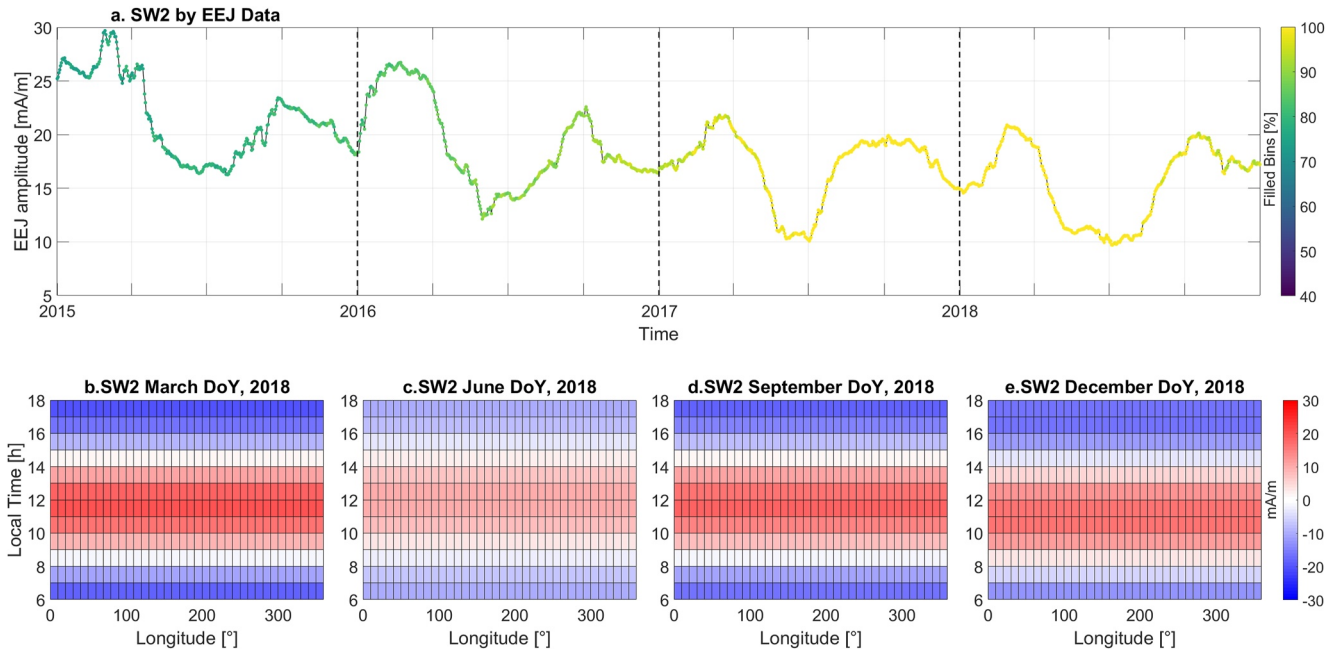
Figure 15. The same as Figure 14, but for SW2 tide.

spectra derived by Lühr and Manoj (2013). Both DW1 and SW2 time series present a decreasing trend from 2015 to 2018, related to the solar cycle effect already seen in Figure 13. The other migrating tides, as TW3 and QW4 (not shown here), present similar seasonal variations with reduced amplitudes.

The panels b, c, d, and e found in Figures 14 and 15 indicate the contribution of the tidal component to the EEJ in terms of positive and negative perturbations for selected days of year from March, June, September, and December 2018. By Figures 14b–14d and 14e, it is possible to see that DW1 acts as the positive background. It adds some negative perturbation in early morning or late afternoon periods. On the other hand, a semidiurnal pattern can be seen as negative and positive perturbations from the SW2 tide in Figures 15b–15d and 15e. No longitudinal variation is observed for both DW1 and SW2 EEJ perturbations because these are migrating tides, which by definition do not depend on longitude.

4.4. Nonmigrating Tides Time Series

In this section, the time series of selected nonmigrating tides obtained after inversion of EEJ geomagnetic data will be presented, discussed, and compared to analogous time series obtained after inversion of SABER temperature data. First, examples from the Swarm period will be shown for DE3, DE2, DE4, SW4, and TW1 tides. Then, the results from the CSØ period will also be show for DE3 and SW4 tides.

Figure 16 shows the results for the DE3 tide obtained for the Swarm period, known as one of the most important nonmigrating components. With large amplitude and being the primary cause of the so-called wave-4 longitudinal structure in the ionosphere, the DE3 signal has been investigated in a variety of different data sets and studies (England et al., 2006; Kil et al., 2007; Lühr et al., 2008; Singh et al., 2018). Figure 16a shows the DE3 time series obtained from EEJ data. Figure 16b shows the spatiotemporal variation of DE3 amplitude obtained from SABER temperature data. Vertical dashed lines indicate the beginning of a new year, in black for Figure 16a, and in white for Figure 16b. A comparison of Figures 16a and 16b indicates a remarkable correlation between EEJ and temperature data analyses for DE3. In both analyses, the DE3 signal peaks around August, a behavior that is in agreement with previous studies based on EEJ data (Lühr & Manoj, 2013) and on SABER temperature data (Forbes et al., 2008). As in the case of the migrating tides, the DE3 amplitude of the EEJ shows some dependency on solar activity, i.e., the amplitude shows a decreasing trend from the year 2015 to 2018. This trend is not seen in the DE3 amplitude of the SABER temperature. Studies found little solar-activity effect on upward-propagating tides from the lower atmosphere (e.g., Oberheide et al., 2009). Figures 16c–16e and 16f are the perturbations
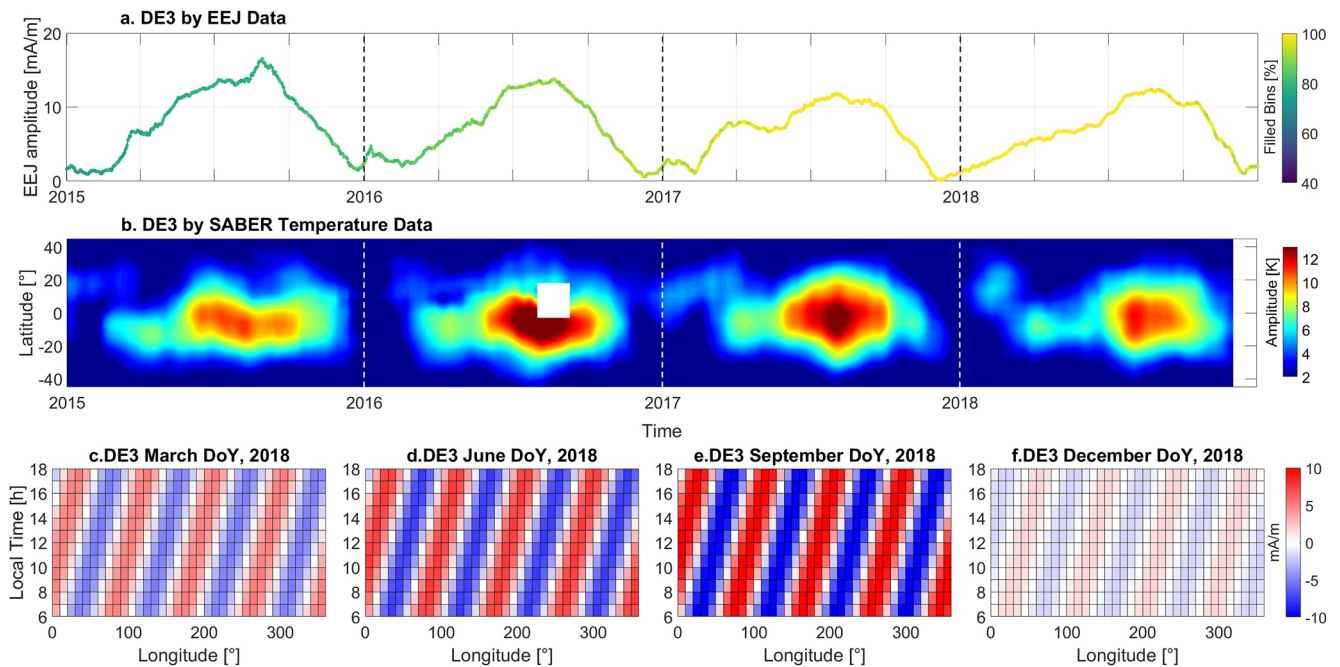
Figure 16. Time series of DE3 amplitude obtained from equatorial electrojet (EEJ) data (a) and from SABER temperature data (b) during the Swarm period (*x* axis represent time with the years as labels). SABER temperature observations range from 100 to 110 km. Vertical dashed lines indicate the beginning of a new year. Panels c–f indicate the contribution of DE3 to the final predicted EEJ in terms of positive and negative perturbations for selected days of year from March, June, September, and December 2018, respectively.

caused by the DE3 tide to the final EEJ reconstructed by tides during days from March, June, September, and December 2018, that indicate a clear wave-4 longitudinal structure.

Figure 17 shows the results for the DE2 nonmigrating tide time series during the Swarm period. EEJ and temperature data analyses are in very good agreement, indicating two peaks per year: one around June solstice and other
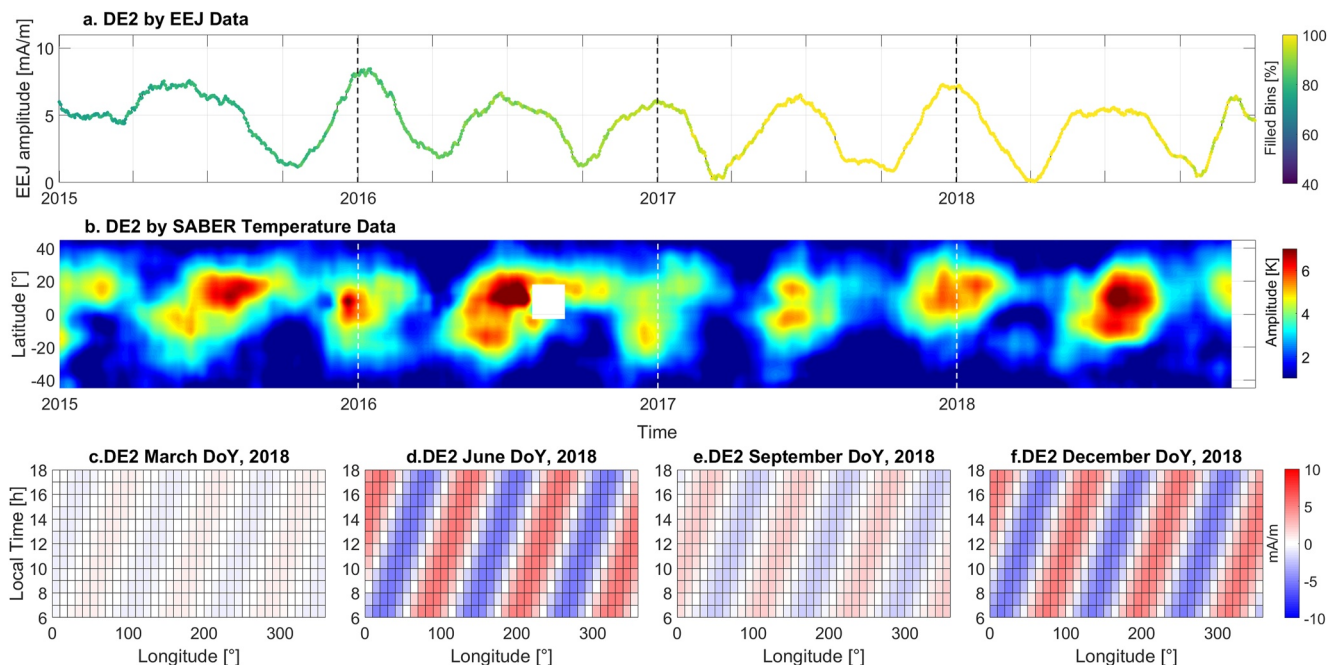
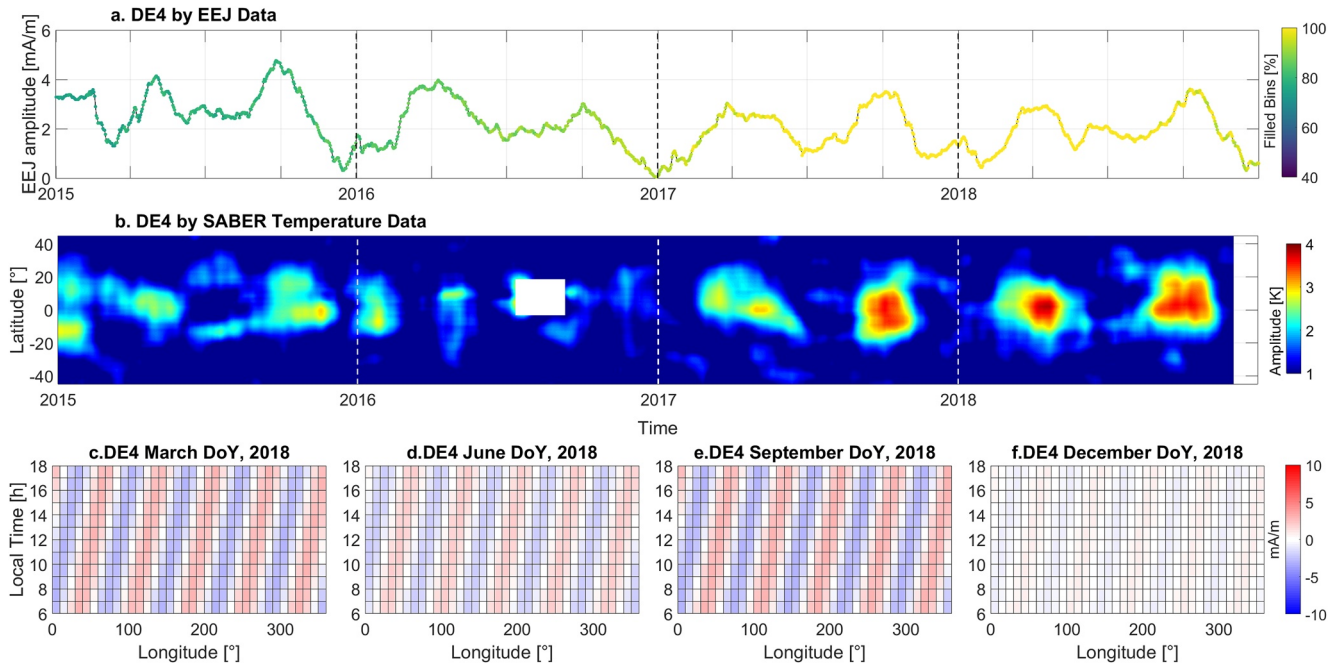**Figure 17.** The same as Figure 16, but for DE2 tide.

Figure 18. The same as Figure 16, but for DE4 tide.

around December solstice. This behavior is consistent with the average DE2 pattern by Lühr and Manoj (2013). In our analysis, further subyear temporal variations are revealed due to the use of the 70-day analysis window. The DE2 perturbation displays in panels c, d, e, and f indicate the wave-3 longitudinal structure.

Figure 18 shows the results for the DE4 nonmigrating tide time series during the Swarm period. Although the DE4 amplitudes are weaker than those from DE3 and DE2, its signatures in the EEJ and temperature data are also in good agreement, with two equinoctial peaks per year (Figures 18a and 18b). This behavior is consistent
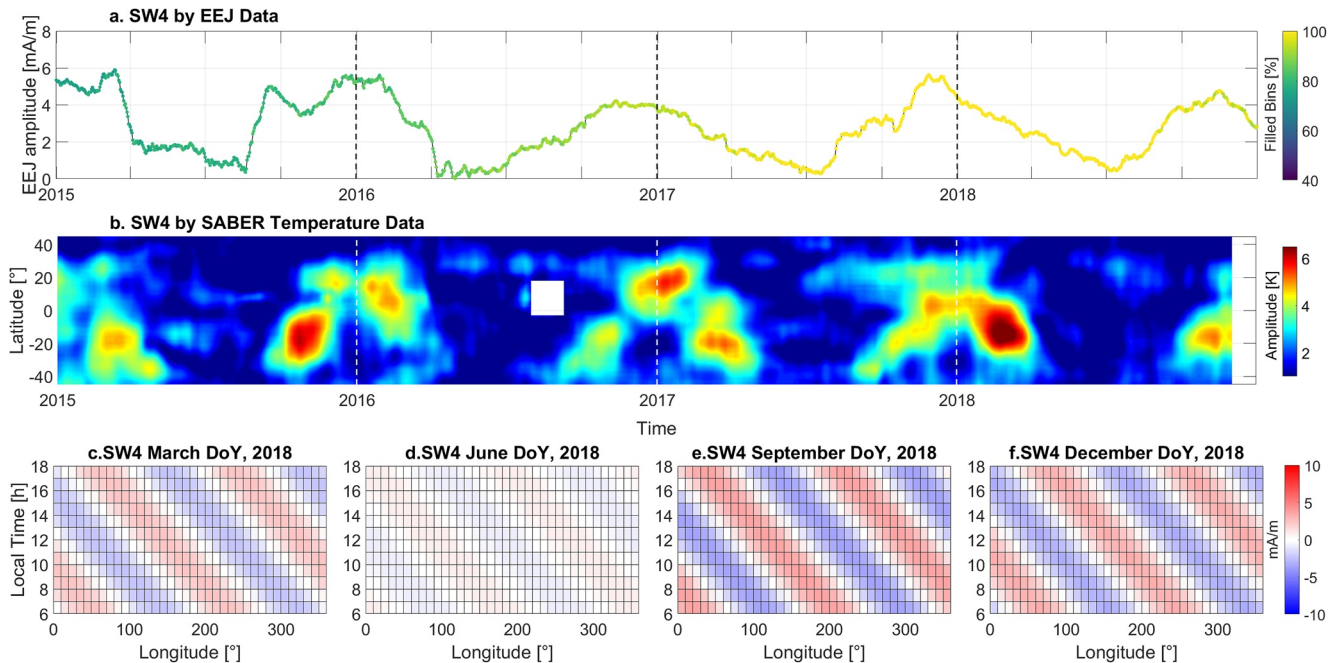
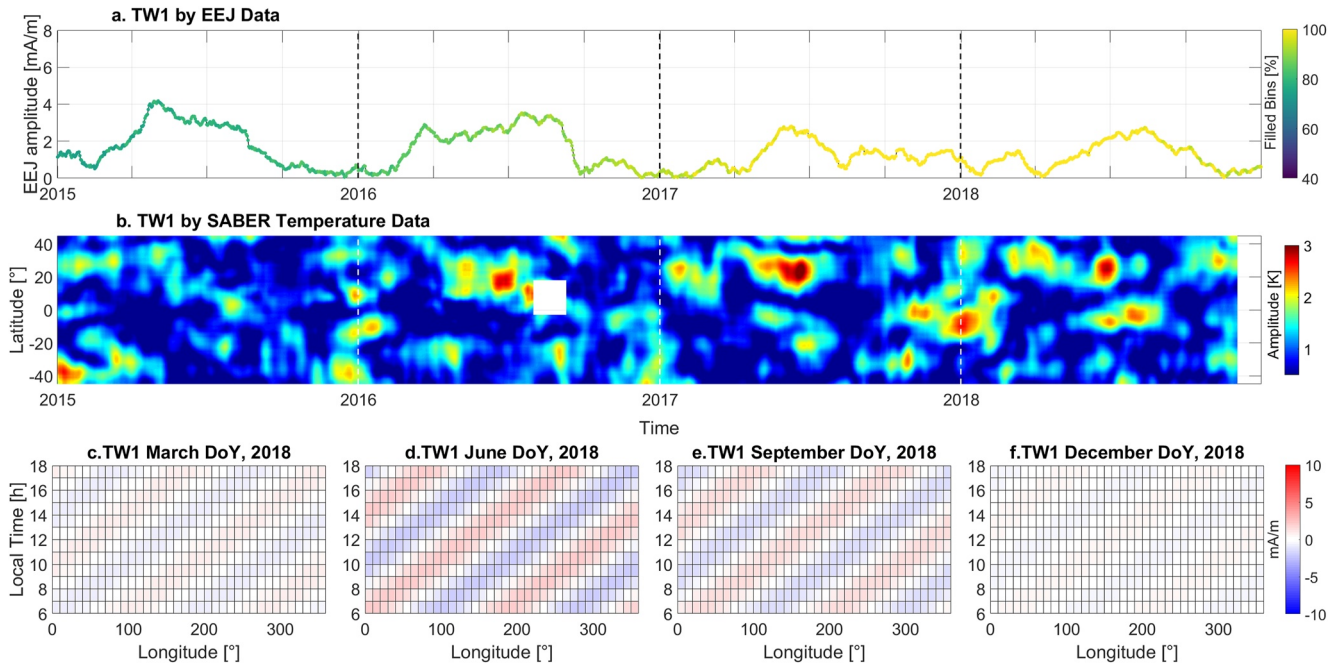**Figure 19.** The same as Figure 16, but for SW4 tide.

Figure 20. The same as Figure 16, but for TW1 tide.

with the average DE4 pattern by Lühr and Manoj (2013). The DE4 is one example of tide which has significant amplitudes in PCEEJ data and negligible tiny amplitudes in EEJM-2 data, as seen in Figure 13. The PCEEJ DE4 amplitudes and its good match to the temperature data suggests that the model is able to represent realistically even those nonmigrating tides with smaller amplitudes. The DE4 perturbation displays in panels c, d, e, and f indicate a wave-5 longitudinal structure.

Figure 19 shows another example of good correlation between EEJ and temperature tides, but now for the SW4 nonmigrating tide during the Swarm period. Both Figures 19a and 19b indicate a major peak around December, which is in agreement with the average SW4 pattern obtained by Lühr and Manoj (2013). In our analysis, some year-to-year variation in SW4 can be observed, which cannot be attributed solely to solar activity, e.g., the high values observed for December 2017 near solar minimum. It is known that the SW4 tide is generated by the nonlinear interaction between the stationary planetary wave SPW2 and the migrating semidiurnal tide SW2 (Forbes et al., 2008; Teitelbaum & Vial, 1991). This means that the year-to-year variation of SW4 may be contributed by other mechanisms than solar activity. Planetary or Rossby waves are mainly caused by airflow over large-scale topographic features (Holton, 2004). If the planetary waves amplitude is large enough in the ionospheric dynamo region, they will affect current systems such as the EEJ. The relative contribution of planetary wave activity and solar activity to the year-to-year variation of SW4 needs more investigation. The SW4 perturbation displays in panels c, d, e, and f indicates its wave-2 longitudinal structure.

So far, we have presented examples with very good agreement between the results obtained from EEJ and temperature data analyses. One fact in common between DE3, DE2, DE4, and SW4 is that these are tides with relatively strong signal and/or a simple spatiotemporal distribution. Other components, however, present small amplitudes and/or more complex temporal variation without a clear seasonal pattern, making it difficult to compare EEJ and SABER temperature results. The TW1 nonmigrating tide in Figure 20 is one of these examples. The comparison between the time series obtained with EEJ and temperature data sets does not indicate a very clear correlation as seen for DE3, DE2, DE4, and SW4. In this case, the spatiotemporal variation of the TW1 amplitude in the SABER analysis is more heterogeneous, and the agreement with the temporal variation of the TW1 amplitude in the EEJ is limited. Although there is no striking correlation between EEJ and SABER temperature results for TW1, it is important to note that the seasonal variation of the TW1 amplitude in the EEJ shown in Figure 20a is in very good agreement with the seasonal variation presented in Lühr and Manoj (2013), with a peak around June solstice. The TW1 perturbation displays in panels c, d, e, and f indicate the wave-2 longitudinal structure.
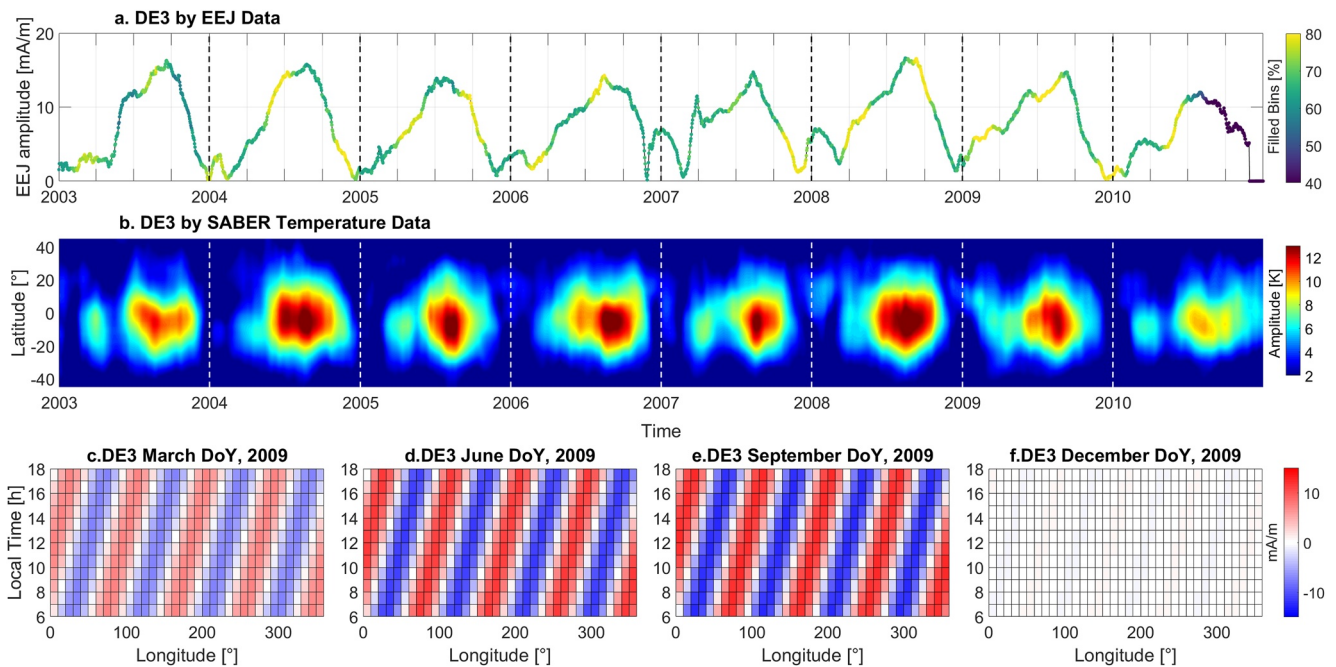
Figure 21. Time series of DE3 amplitude obtained from equatorial electrojet (EEJ) data (a) and from SABER temperature data (b) during the CSØ period (x axis represent time with the years as labels). SABER temperature observations range from 100 to 110 km. Vertical dashed lines indicate the beginning of a new year. Panels c–f indicate the contribution of DE3 to the final predicted EEJ in terms of positive and negative perturbations for selected days of year from March, June, September, and December 2009, respectively.

Figures 21 and 22 show results for the nonmigrating tides DE3 and SW4 for the CSØ period, respectively. The amplitudes are comparable with those obtained for the Swarm period (Figures 16 and 19). Again, there is a good match between the time series of the amplitudes derived from EEJ and SABER temperature data. For example, the amplitude of the DE3 in both EEJ and SABER temperature shows an annual variation with the maximum
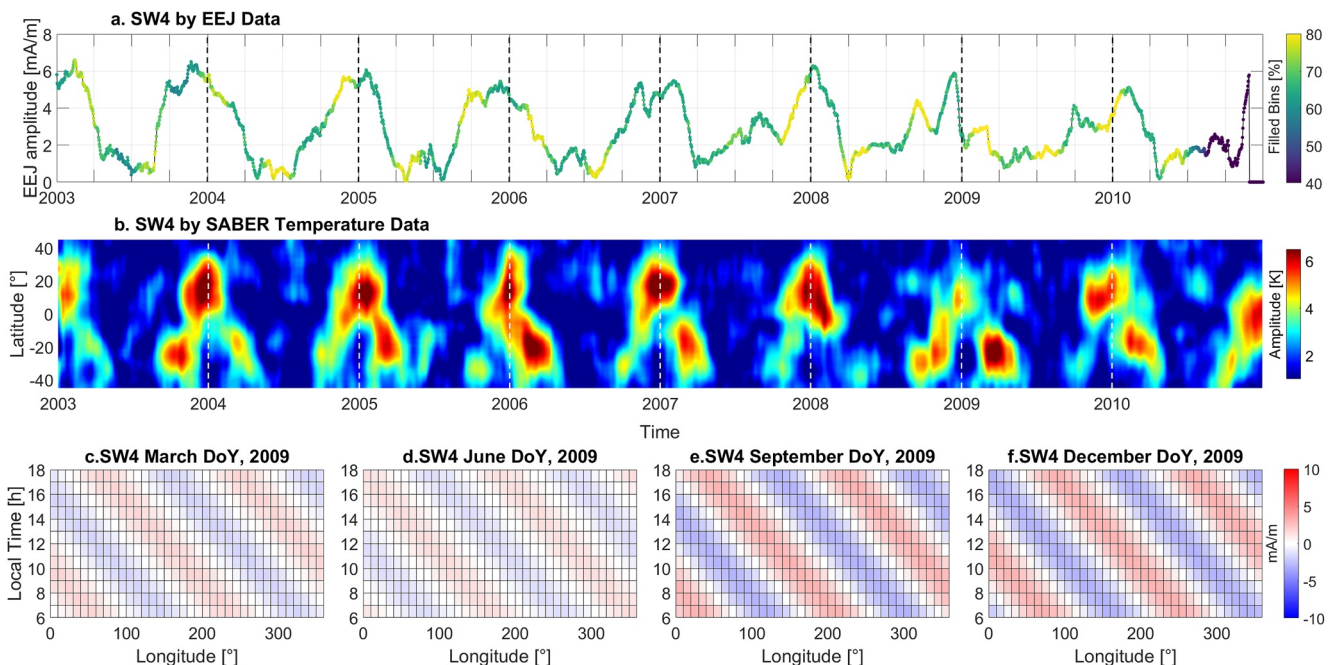

Figure 22. The same as Figure 21, but for SW4 tide.

in August-September. Both EEJ and SABER temperature DE3 amplitudes show relatively short-term variations with a local maximum in December, especially in the years 2007 and 2008. The significant decrease in the amount of filled bins in 2010, represented by the panel a color bar, is due to the end of the CHAMP satellite mission. As shown and discussed for the Swarm period, the amplitudes of DE3 and SW4 do not show strong dependence on solar cycle activity.

The tides signatures obtained from SABER and EEJ analyses show an obvious correlation, but they result from different processes. In SABER analysis, the obtained tides signatures demonstrate the tides themselves in the atmosphere. In EEJ analysis, the obtained tides signatures represent a secondary effect, namely the influence of the tides in the ionosphere.

5. Summary and Conclusions

A novel technique to model the intensity of the EEJ is proposed. The method involves PCA of EEJ intensities observed by Swarm, CHAMP, and Ørsted satellites from 2000 to 2019, and fitting of the obtained PCs to hybrid ground-satellite data over a given period of 70 days. Our statistical analysis shows that the new EEJ model can reproduce observations better than the climatological model EEJM-2. This is because the EEJM-2 was designed to represent primarily the EEJ climatology, while our PCEEJ model was designed to reproduce the EEJ for a specific period of individual years. In contrast, the performance of the PCEEJ depends on the local time and longitudinal coverage of the EEJ data from satellites and ground stations.

The PCEEJ is used to examine tidal signatures in the EEJ, and their relations to tides in atmospheric temperature as observed by TIMED/SABER. On average, the tidal composition of the EEJ derived from our new model is consistent with the 5-year climatology presented by Lühr and Manoj (2013). Our method can provide the amplitude time series of various tidal modes in the EEJ for individual years. It is found that seasonal variations of major nonmigrating tidal modes such as DE3, DE2, and SW4 are consistent with those in SABER temperature. Due to its construction aspects, the PCEEJ can provide realistic EEJ tidal composition and corresponding temporal variation, even for those components with reduced amplitude, as seen for the DE4 and TW1 examples. Thus, our model can be used to monitor the ionospheric effect of nonmigrating tides that propagate from the middle atmosphere.

The present study used a 70-day time window for PCA modeling of the EEJ. The application of a shorter time window, down to time scales as short as days, would be possible if more EEJ data with a suitable spatiotemporal distribution is available.

Data Availability Statement

HUA, KOU, ABG, TTB, TAM, PHU, and GUA data used in this paper are available at INTERMAGNET (www.intermagnet.org). YAP, MUT, DAV, and BCL data were provided by the MAGDAS network (http://data.icswse.kyushu-u.ac.jp/). PKT data were provided by the SEALION network (https://aer-nc-web.nict.go.jp/sealion/index.html). PIU data were provided by the LISN network (http://lisn.igp.gob.pe/). SAM data were provided by the WAMNET network (http://www.bcmt.fr/wamnetnetwork.html). TIR data were provided by the WDC-Mumbai network (http://wdciig.res.in/WebUI/Home.aspx). TND data were provided by the WDC-Edinburgh (http://www.wdc.bgs.ac.uk/). KTB data were provided by the 210° meridian chain network (https://stdb2.isee.nagoya-u.ac.jp/mm210/). TTB data are partly available at GFZ Data Services (http://dataservices.gfz-potsdam.de/portal/). The CHAMP satellite data are accessible online http://isdc.gfz-potsdam.de/. The Swarm satellite mission data are available online at http://www.earth.esa.int/swarm. The Ørsted and SAC-C data are available at https://ftp.spacecenter.dk/data/magnetic-satellites/. The SABER temperature observations are available online at http://saber.gats-inc.com/. F10.7 solar flux data are available at http://www.spaceweather.ca/solarflux/sx-en.php. Kp index data are available at https://www.gfz-potsdam.de/en/kp-index/.

References

Alken, P., & Maus, S. (2007). Spatio-temporal characterization of the equatorial electrojet from CHAMP, Ørsted, and SAC-C satellite magnetic measurements. *Journal of Geophysical Research*, 112, A09305. https://doi.org/10.1029/2007JA012524

Alken, P., Maus, S., Vigneron, P., Sirol, O., & Hulot, G. (2013). Swarm SCARF equatorial electric field inversion chain. *Earth Planets and Space*, 65(11), 1309–1317. https://doi.org/10.5047/eps.2013.09.008

Alken, P., Maute, A., Richmond, A. D., Vanhamäki, H., & Egbert, G. D. (2017). An application of principal component analysis to the interpretation of ionospheric current systems. *Journal of Geophysical Research: Space Physics*, *122*, 5687–5708. https://doi.org/10.1002/2017JA024051

Anandarao, B. G., & Raghavarao, R. (1987). Structural changes in the currents and fields of the equatorial electrojet due to zonal and meridional winds. *Journal of Geophysical Research*, *92*(A3), 2514–2526. https://doi.org/10.1029/ja092ia03p02514

Chapman, S. (1951). The equatorial electrojet as detected from the abnormal electric current distribution about Huancayo, Peru and elsewhere. *Archiv fur Meteorologie, Geophysik, und Bioklimatologie Serie A Meteorologie und Geophysik*, *4*(1), 368–390. https://doi.org/10.1007/BF02246814

Chapman, S., & Raja Rao, K. S. (1965). The H and Z variations along and near the equatorial electrojet in India, Africa and the Pacific. *Journal of Atmospheric and Terrestrial Physics*, *27*(4), 559–581. https://doi.org/10.1016/0021-9169(65)90020-6

Cnossen, I., & Richmond, A. D. (2013). Changes in the Earth's magnetic field over the past century: Effects on the ionosphere-thermosphere system and solar quiet (Sq) magnetic variation. *Journal of Geophysical Research: Space Physics*, *118*, 849–858. https://doi.org/10.1029/2012JA018447

Colomb, F. R., Alonso, C., Hofmann, C., & Nollmann, I. (2004). SAC-C mission: An example of international cooperation. *Advances in Space Research*, *34*(10), 2194–2199. https://doi.org/10.1016/j.asr.2003.10.039

Doumouya, V., Cohen, Y., Arora, B. R., & Yumoto, K. (2003). Local time and longitude dependence of the equatorial electrojet magnetic effects. *Journal of Atmospheric and Solar-Terrestrial Physics*, *65*(14–15), 1265–1282. https://doi.org/10.1016/j.jastp.2003.08.014

Doumouya, V., Vassal, J., Cohen, Y., Fambitakoye, O., & Menvielle, M. (1998). Equatorial electrojet at African longitudes: First results from magnetic measurements. *Annales Geophysicae*, *16*, 658–666. https://doi.org/10.1007/s00585-998-0658-9

England, S. L., Maus, S., Immel, T. J., & Mende, S. B. (2006). Longitudinal variation of the E region electric fields caused by atmospheric tides. *Geophysical Research Letters*, *33*, L21105. https://doi.org/10.1029/2006GL027465

Fambitakoye, O., & Mayaud, P. N. (1976). Equatorial electrojet and regular daily variation SR—I. A determination of the equatorial electrojet parameters. *Journal of Atmospheric and Terrestrial Physics*, *38*(1), 1–17. https://doi.org/10.1016/0021-9169(76)90188-4

Fang, T.-W., Akmaev, R., Fuller-Rowell, T., Wu, F., Maruyama, N., & Millward, G. (2013). Longitudinal and day-to-day variability in the iono-sphere from lower atmosphere tidal forcing. *Geophysical Research Letters*, *40*, 2523–2528. https://doi.org/10.1002/grl.50550

Finlay, C. C., Olsen, N., Kotsiaros, S., Gillet, N., & Tøffner-Clausen, L. (2016). Recent geomagnetic secular variation from Swarm and ground observatories as estimated in the CHAOS-6 geomagnetic field model. *Earth Planets and Space*, *68*(1), 112. https://doi.org/10.1186/s40623-016-0486-1

Forbes, J. M. (1981). The equatorial electrojet. *Reviews of Geophysics*, *19*(3), 469–504. https://doi.org/10.1029/RG019i003p00469

Forbes, J. M., Zhang, X., Palo, S., Russell, J., Mertens, C. J., & Mlynczak, M. (2008). Tidal variability in the ionospheric dynamo region. *Journal of Geophysical Research*, *113*, A02310. https://doi.org/10.1029/2007JA012737

Friis-Christensen, E., Luhr, H., & Hulot, G. (2006). Swarm: A constellation to study the Earth's magnetic field. *Earth Planets and Space*, *58*(4), 351–358. https://doi.org/10.1186/BF03351933

Friis-Christensen, E., Luhr, H., Knudsen, D., & Haagmans, R. (2008). Swarm—An Earth observation mission investigating geospace. *Advances in Space Research*, *41*(1), 210–216. https://doi.org/10.1016/j.asr.2006.10.008

Hamid, N. S. A., Liu, H., Uozumi, T., & Yoshikawa, A. (2015). Empirical model of equatorial electrojet based on ground-based magnetometer data during solar minimum in fall. *Earth Planets and Space*, *67*, 205. https://doi.org/10.1186/s40623-015-0373-1

Hansen, P. C. (2010). *Discrete inverse problems*. Philadelphia: SIAM Society for Industrial and Applied Mathematics.

Holton, J. R. (2004). *An introduction to dynamic meteorology* (4th ed.), (Vol. 88). Elsevier Academic Press.

Kil, H., Oh, S.-J., Kelley, M. C., Paxton, L. J., England, S. L., Talaat, E., et al. (2007). Longitudinal structure of the vertical E × B drift and ion density seen from ROCSAT-1. *Geophysical Research Letters*, *34*, L14110. https://doi.org/10.1029/2007GL030018

Kusnierkiewicz, D. Y. (2003). *An overview of the TIMED spacecraft* (Vol. 24, No. 2). JHU/APL Technical Digest.

Liu, H.-L. (2016). Variability and predictability of the space environment as related to lower atmosphere forcing. *Space Weather*, *14*, 634–658. https://doi.org/10.1002/2016SW001450

Lühr, H., & Manoj, C. (2013). The complete spectrum of the equatorial electrojet related to solar tides: CHAMP observations. *Annales Geophys-icae*, *31*(8), 1315–1331. https://doi.org/10.5194/angeo-31-1315-2013

Lühr, H., Maus, S., & Rother, M. (2004). Noon-time equatorial electrojet: Its spatial features as determined by the CHAMP satellite. *Journal of Geophysical Research*, *109*, A01306. https://doi.org/10.1029/2002JA009656

Lühr, H., Rother, M., Hausler, K., Alken, P., & Maus, S. (2008). The influence of nonmigrating tides on the longitudinal variation of the equatorial electrojet. *Journal of Geophysical Research*, *113*, A08313. https://doi.org/10.1029/2008JA013064

Manoj, C., Lühr, H., Maus, S., & Nagarajan, N. (2006). Evidence for short spatial correlation lengths of the noontime equatorial electrojet inferred from a comparison of satellite and ground magnetic data. *Journal of Geophysical Research*, *111*, A11312. https://doi.org/10.1029/2006JA011855

Matzka, J., Siddiqui, T. A., Lilienkamp, H., Stolle, C., & Veliz, O. (2017). Quantifying solar flux and geomagnetic main field influence on the equatorial ionospheric current system at the geomagnetic observatory Huancayo. *Journal of Atmospheric and Solar-Terrestrial Physics*, *163*, 120–125. https://doi.org/10.1016/j.jastp.2017.04.014

Matzka, J., Stolle, C., Yamazaki, Y., Bronkalla, O., & Morschhauser, A. (2021). The geomagnetic Kp index and derived indices of geomagnetic activity. *Space Weather*, *19*, e2020SW002641. https://doi.org/10.1029/2020SW002641

Maus, S., & Lühr, H. (2005). Signature of the quiet-time magnetospheric magnetic field and its electromagnetic induction in the rotating Earth. *Geophysical Journal International*, *162*(3), 755–763. https://doi.org/10.1111/j.1365-246X.2005.02691.x

Maus, S., Yin, F., Lühr, H., Manoj, C., Rother, M., Rauberg, J., et al. (2008). Resolution of direction of oceanic magnetic lineations by the sixth-generation lithospheric magnetic field model from CHAMP satellite magnetic measurements. *Geochemistry, Geophysics, Geosystems*, *9*, Q07021. https://doi.org/10.1029/2008GC001949

Miyahara, S., & Ooishi, M. (1997). Variation of Sq induced by atmospheric tides simulated by a middle atmosphere general circulation model. *Journal of Geomagnetism and Geoelectricity*, *49*, 77–87. https://doi.org/10.5636/jgg.49.77

Miyoshi, Y., Pancheva, D., Mukhtarov, P., Jin, H., Fujiwara, H., & Shinagawa, H. (2017). Excitation mechanism of non-migrating tides. *Journal of Atmospheric and Solar-Terrestrial Physics*, *156*, 24–36. https://doi.org/10.1016/j.jastp.2017.02.012

Neubert, T., Mandea, M., Hulot, G., von Frese, R., Primdahl, F., Jørgensen, J. L., et al. (2001). Ørsted satellite captures high-precision geomag-netic field data. *Eos Transactions American Geophysical Union*, *82*(7), 81–88. https://doi.org/10.1029/01EO00043

Oberheide, J., Forbes, J. M., Häusler, K., Wu, Q., & Bruinsma, S. L. (2009). Tropospheric tides from 80 to 400 km: Propagation, interannual variability, and solar cycle effects. *Journal of Geophysical Research*, *114*, D00I05. https://doi.org/10.1029/2009JD012388

Oberheide, J., Forbes, J. M., Zhang, X., & Bruinsma, S. L. (2011). Climatology of upward propagating diurnal and semidiurnal tides in the ther-mosphere. *Journal of Geophysical Research*, *116*, A11306. https://doi.org/10.1029/2011JA016784

Onwumechili, C. A. (1997). *The equatorial electrojet*. Newark, NJ: Gordon and Breach.

Reigber, C., Lühr, H., & Schwintzer, P. (2002). CHAMP mission status. *Advances in Space Research*, *30*(2), 129–134. https://doi.org/10.1016/S0273-1177(02)00276-4

Richmond, A. D. (1973). Equatorial electrojet—I. Development of a model including winds and instabilities. *Journal of Atmospheric and Terrestrial Physics*, *35*(6), 1083–1103. https://doi.org/10.1016/0021-9169(73)90007-x

Rigoti, A., Chamalaun, F. H., Trivedi, N. B., & Padilha, A. L. (1999). Characteristics of the equatorial electrojet determined from an array of magnetometers in N-NE Brazil. *Earth Planets and Space*, *51*(2), 115–128. https://doi.org/10.1186/bf03352216

Russell, J. M., III., Mlynczak, M. G., Gordley, L. L., Joseph, J., Tansock, Jr., & Esplin, R. W. (1999). Overview of the SABER experiment and preliminary calibration results. *Proceedings of SPIE, 3756*. https://doi.org/10.1117/12.366382

Singh, D., Gurubaran, S., & He, M. (2018). Evidence for the influence of DE3 tide on the occurrence of equatorial counterelectrojet. *Geophysical Research Letters*, *45*, 2145–2150. https://doi.org/10.1002/2018GL077076

Soares, G., Yamazaki, Y., Cnossen, I., Matzka, J., Pinheiro, K. J., Morschhauser, A., et al. (2020). Evolution of the geomagnetic daily variation at Tatuoca, Brazil, from 1957 to 2019: A transition from Sq to EEJ. *Journal of Geophysical Research: Space Physics*, *125*, e2020JA028109. https://doi.org/10.1029/2020JA028109

Soares, G., Yamazaki, Y., Matzka, J., Pinheiro, K., Morschhauser, A., Stolle, C., & Alken, P. (2018). Equatorial counter electrojet longitudinal and seasonal variability in the American sector. *Journal of Geophysical Research: Space Physics*, *123*, 9906–9920. https://doi.org/10.1029/2018JA025968

Soares, G., Yamazaki, Y., Morschhauser, A., Matzka, J., & Pinheiro, K. (2022). *PCEEJ equatorial electrojet model data for years 2003–2010 and 2014–2018*. GFZ Data Services. https://doi.org/10.5880/GFZ.2.3.2022.003

Sugiura, M., & Poros, D. J. (1969). An improved model equatorial electrojet with a meridional current system. *Journal of Geophysical Research*, *74*(16), 4025–4034. https://doi.org/10.1029/ja074i016p04025

Teitelbaum, H., & Vial, F. (1991). On tidal variability induced by nonlinear interaction with planetary waves. *Journal of Geophysical Research*, *96*(A8), 14169–14178. https://doi.org/10.1029/91JA01019

Yamazaki, Y., & Maute, A. (2017). Sq and EEJ—A review on the daily variation of the geomagnetic field caused by ionospheric dynamo currents. *Space Science Reviews*, *206*(1–4), 299–405. https://doi.org/10.1007/s11214-016-0282-z

Yamazaki, Y., Richmond, A. D., Maute, A., Liu, H.-L., Pedatella, N., & Sassi, F. (2014a). On the day-to-day variation of the equatorial electrojet during quiet periods. *Journal of Geophysical Research: Space Physics*, *119*, 6966–6980. https://doi.org/10.1002/2014JA020243

Yamazaki, Y., Richmond, A. D., Maute, A., Wu, Q., Ortland, D. A., Yoshikawa, A., et al. (2014b). Ground magnetic effects of the equatorial electrojet simulated by the TIE-GCM driven by TIMED satellite data. *Journal of Geophysical Research: Space Physics*, *119*, 3150–3161. https://doi.org/10.1002/2013JA019487

Yamazaki, Y., Stolle, C., Matzka, J., Liu, H., & Tao, C. (2018). Interannual variability of the daytime equatorial ionospheric electric field. *Journal of Geophysical Research: Space Physics*, *123*, 4241–4256. https://doi.org/10.1029/2017JA025165

Zhang, X., Forbes, J. M., Hagan, M. E., Russell, J. M., III, Palo, S. E., Mertens, C. J., & Mlynczak, M. G. (2006). Monthly tidal temperatures 20–120 km from TIMED/SABER. *Journal of Geophysical Research*, *111*, A10S08. https://doi.org/10.1029/2005JA011504