

Statistical models of the variability of plasma in the topside ionosphere: 2. Performance assessment

Luca Spogli^{1,2,*}, Yaqi Jin³, Jaroslav Urbář^{1,4}, Alan G. Wood⁵, Elizabeth E. Donegan-Lawley⁵, Lasse B.N. Clausen³, Golnaz Shahtahmassebi⁶, Lucilla Alfonsi¹, James T. Rawlings⁶, Antonio Cicone^{1,7}, Daria Kotova³, Claudio Cesaroni¹, Per Høeg³, Gareth D. Dorian⁵, Luke D. Nugent⁵, Sean Elvidge⁵, David R. Themens⁵, María José Brazal Aragón⁸, Pawel Wojtkiewicz⁸, and Wojciech J. Miloch³

¹ Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata 605, 00143 Rome, Italy

² SpaceEarth Technology, Via di Vigna Murata 605, 00143 Rome, Italy

³ Department of Physics, University of Oslo, Sem Sæland svei 24, 0371 Oslo, Norway

⁴ Institute of Atmospheric Physics of the Czech Academy of Sciences, Boční II 1401, 141 00 Prague, Czech Republic

⁵ Space Environment and Radio Engineering (SERENE) Group, University of Birmingham, B15 2TT Birmingham, UK

⁶ Department of Physics & Mathematics, Nottingham Trent University, NG7 2RD Nottingham, UK

⁷ Università Degli Studi Dell'Aquila, Via Vetoio 1, 67100 L'Aquila, Italy

⁸ GMV Innovating Solutions, ul Hrubieszowska 2, 01-209 Warsaw, Poland

Received 2 May 2023 / Accepted 10 January 2024

Abstract—Statistical models of the variability of plasma in the topside ionosphere based on the Swarm data have been developed in the “Swarm Variability of Ionospheric Plasma” (Swarm-VIP) project within the European Space Agency’s Swarm+4D-Ionosphere framework. The models can predict the electron density, its gradients for three horizontal spatial scales – 20, 50 and 100 km – along the North-South direction and the level of the density fluctuations. Despite being developed by leveraging on Swarm data, the models provide predictions that are independent of these data, having a global coverage, fed by various parameters and proxies of the helio-geophysical conditions. Those features make the Swarm-VIP models useful for various purposes, which include the possible support for already available ionospheric models and proxy of the effect of ionospheric irregularities of the medium scales that affect the signals emitted by Global Navigation Satellite Systems (GNSS). The formulation, optimisation and validation of the Swarm-VIP models are reported in Paper 1 (Wood et al. 2024, *J Space Weather Space Clim.* in press). This paper describes the performance assessment of the models, by addressing their capability to reproduce the known climatological variability of the modelled quantities, and the ionospheric weather as depicted by ground-based GNSS, as a proxy for the ionospheric effect on GNSS signals. Additionally, we demonstrate that, under certain conditions, the model can better reproduce the ionospheric variability than a physics-based model, namely the Thermosphere-Ionosphere-Electrodynamics General Circulation Model (TIE-GCM).

Keywords: Topside ionosphere / Space weather / Ionosphere atmosphere interactions / Statistical modelling

1 Introduction

The Earth’s ionosphere is a dynamical system whose complexity manifests itself in an irregular variability in time and space and an apparent randomness (or quasi-randomness) of the physical parameters characterizing its state (Materassi et al., 2019). Such complexity is due to the non-linear couplings among the different sub-systems forming the near-Earth environment. In fact, the ionosphere is characterised by its coupling

with the magnetosphere and solar wind from above, with the lower atmospheric layers and lithosphere from below, and by the changes in the neutral atmosphere in which the ionosphere lies. The need for modelling the ionospheric changes in space and time stands not only as part of our search for understanding our environment and the physical laws ruling it but also as an important input to technologies and applications for which the ionospheric medium is an issue. These include, among others, satellite-based positioning, as provided by Global Navigation Satellite Systems (GNSS), telecommunications in the HF-VHF-UHF bands, radio astronomy, Earth observation from

*Corresponding author: luca.spogli@ingv.it

space, and many others (Bilitza et al., 2022). The ionospheric models can be sorted into: (i) physics-based models, including the Thermosphere-Ionosphere-Electrodynamics General Circulation Model (TIE-GCM) (Qian et al., 2014; Maute, 2017; and references therein); (ii) empirical and semi-empirical models, including the International Reference Ionosphere Model (Bilitza et al., 2022; and references therein) and the NeQuick (Nava et al., 2005, 2008), (iii) models of the impact of the effect of plasma density irregularities on radio wave propagation, including scintillation (Priyadarshi, 2015, and references therein) and (iv) data assimilation models where a background empirical or physical model is modified based on observations (e.g. Elvidge & Angling, 2019).

The use of the in-situ and topside ionospheric information provided by data collected by Low-Earth Orbit (LEO) satellite missions has been demonstrated to be of paramount importance to support the development of ionospheric models (see., e.g. Wernik et al., 2007; Alfonsi et al., 2017; Kamal et al., 2021). Among LEO missions, Swarm is the European Space Agency's (ESA) first constellation mission for Earth Observation (Friis-Christensen et al., 2008). It was initially constituted by three identical satellites, called Alpha (A), Bravo (B), and Charlie (C), which follow quasi-Sun-synchronous near-polar orbits (87.3° inclination) at different altitudes: Swarm A and C fly at around 440 km (in 2021) above the Earth's surface and Swarm B at a higher altitude (around 510 km). These satellites were launched in November 2013. Conceived mainly as a magnetic mission, Swarm has also provided a plethora of interesting results about ionospheric variability, as reported by the thorough review by Wood et al. (2022). It has also been successfully used in the recent past to improve ionospheric modelling (Pezzopane & Pignalberi, 2019). In the Swarm Variability of Ionospheric Plasma (Swarm-VIP) project, funded by the European Space Agency ("Swarm+4D-Ionosphere" framework), the Swarm data have been exploited to develop fully empirical, statistical models of the variability of plasma in the topside ionosphere. In this regard, Swarm-VIP models are part of the model indicated by (ii) in the aforementioned list.

The Swarm-VIP models predict some of the quantities available in the IPIR (Ionospheric Plasma Irregularities) Swarm data product (Jin et al., 2022). These are the electron density, its gradients at three spatial scales – 20, 50 and 100 km – and an index which indicates the strength of the fluctuations, called IPIR Index ($IPIR_{ix}$) (Jin et al., 2022). Those spatial scales are the quantities that are significant by Urbar et al. (2023) to highlight the ionospheric response to the geospace forcing.

The formulation of the models, their optimization and validation are provided in detail within the companion paper (Wood et al., 2024; hereafter referred to as Paper 1). This paper aims to assess and investigate the capability of the models in reproducing known features of ionospheric climatology; to underline, for selected applications and case events, their usability to support GNSS-based ionospheric information, and to report a comparison against the physical modelling through TIE-GCM. Indeed, the Swarm-VIP models are conceived to contribute to the knowledge of ionospheric behaviour and to support the development of possible operational tools which can provide ionospheric information in regions scarcely covered by ground-based instrumentation. The latter follows also from the hints about the usability of IPIR parameters to support the

identification of plasma irregularities, their level of structuring and their effect on L-band signals provided by Kotova et al. (2023).

The paper is structured as follows. Section 2 reports the basic principles of the Swarm-VIP model formulation, detailed in Paper 1, but here recalled, in brief, to provide the context for this paper. Section 3 details the methods adopted and the considered datasets, sorted into two categories: (i) Swarm for the climatological assessment and (ii) GNSS-based datasets. Additionally, Section 3 recalls the TIE-GCM and how the performance assessment against it is realised. Section 4 provides selected results of the performance assessment, while Section 5 is dedicated to the discussion and conclusions.

2 Recalling of modelling basic principles

The Swarm-VIP models have been developed using a technique called Generalised Linear Modelling (GLM) that is used to predict the following quantities at Swarm A&C altitudes (~440 km): (i) the absolute value of the spatial gradients of the electron density over 20, 50, 100 km along Swarm orbit ($|Grad_{Ne@XXkm}|$, where $XX=20, 50, 100$), (ii) the electron density (N_e), and (iii) the $IPIR_{ix}$, from which categorisation of fluctuations in the ionospheric plasma density (0–3 low, 4–5 medium, and >6 high level) can be derived, as detailed in Jin et al. (2022). Those quantities are also reported in Paper 1 and are all based on the measurements performed by the Langmuir Probe onboard Swarm. These data are available at 1 Hz in the Swarm level 2 data product termed $IPDxIRR_{2F}$, available at: <ftp://swarm-diss.esa.int>.

The predicted quantities are expressed as a function of a subset of explanatory variables, among all trialled ones, which can be categorised into six families (see Sect. 2.2 of Paper 1 for the details and referring text):

1. *Solar activity*: Solar radio flux at 10.7 cm wavelength (F10.7) and the sunspot number R ;
2. *Solar wind*: Bulk speed, density, pressure, Interplanetary Magnetic Field (IMF) and Interplanetary Electric Field (IEF) time-shifted to the Earth's bow shock;
3. *Geomagnetic activity*: The aa , AE , am , AL , Ap , $ASY-D$, $ASY-H$, AU , Dst , Kp , Polar cap (north) index (PCN), $SYM-D$ and $SYM-H$ indices;
4. *Location*: Geographic latitude, magnetic latitude, local solar time and magnetic local time;
5. *Miscellaneous*: Solar Zenith Angle (SZA) and a sine function to represent the seasonal variation, going from -1 at northern midwinter to $+1$ at northern midsummer.

The full list of the explanatory variables can be found in Table S1 of Paper 1.

Each model predicts the investigated quantities in one of the four considered different regions (polar, auroral, mid-latitude, and equatorial), with the regions identified using a combination of the ionosphere region flag in the IPIR data product and of the modulus of the magnetic latitude. This translates into a total number of 5 quantities ($|Grad_{Ne@XXkm}|$ with $XX = 20, 50, 100, N_e$, and $IPIR_{ix}$) times 4 regions, equalling 20 different models. The modelling method assumes that the explanatory

variables selected are independent of one another. Ensuring that this condition was fulfilled was a substantial challenge and the relevant work is described in Section 2.6 of [Paper 1](#).

The basic form of the models is as follows:

$$|Grad_Ne@XXkm| = (\exp(\beta_0 + \beta_1 \cdot x_1 + \dots + \beta_n \cdot x_n))^m, \quad (1)$$

in which, $x_1 \dots x_n$ are the explanatory variables and $\beta_0 \dots \beta_n$ are empirically determined constants known as the parameter estimates and the power m changes model per model. The $\beta_0 \dots \beta_n$ are reported in Tables S2 and S3 of [Paper 1](#). An example of a model is provided in equation (2), which reports the formulation of the polar cap model for the $|Grad_Ne@100km|$:

$$|Grad_Ne@100km| = (\exp(\beta_0 + \beta_1 \cdot F107_81 + \beta_2 \cdot SZA + \beta_3 \cdot DOY_fn + \beta_4 \cdot Kp + \beta_5 \cdot |MLAT|))^3. \quad (2)$$

In the example above, $F107_81$ is the average value of F10.7 solar flux over 81 days (see, e.g. [Xiong et al., 2022](#)), SZA is the solar zenith angle, $|MLAT|$ is the absolute value of the magnetic latitude, Kp is the planetary K -index and DOY_fn is a sine function based on the day of year (DOY), going from -1 at northern midwinter to $+1$ at northern midsummer.

Two versions of the models were produced. These are shown in Tables S2 and S4 of [Paper 1](#). In version 2 (Table S4 of [Paper 1](#)), all of the explanatory variables (Table S1 of [Paper 1](#)) were trialled. This model was intended to increase scientific understanding of the system with explanatory variables chosen to act as proxies for a wide variety of driving processes. Version 1 of the models (Table S2 of [Paper 1](#)) uses a subset of the explanatory variables, only including those which are routinely available. The choice of explanatory variables is discussed in detail in [Paper 1](#) but, in essence, the complementary observations from Swarm are excluded. This allows a model to be created which is more generally applicable, which can predict ionospheric parameters at locations other than that where the Swarm satellite is currently observing. It is the assessment of the performance of these models (version 1, Table S2 of [Paper 1](#)) which are considered in the remainder of this paper.

The models presented here are slightly different from those reported in [Paper 1](#). In [Paper 1](#), a transformation is applied to the dependent variable, so the m th root of the dependent variable is modelled. Therefore, the m th root is shown on the right-hand side of these equations. In the present paper, the dependent variable is predicted. Therefore, the left-hand side of the equation is expressed as the m th power.

3 Datasets and methods

The performance assessment addresses three different aspects of the modelling, namely: (i) the capability of the model to reproduce the climatological variability of the considered quantities through the direct comparison with Swarm data; (ii) the capability to reproduce ionospheric weather as depicted by ground-based GNSS and to serve as proxy for the ionospheric effect on GNSS signals; and (iii) the comparison with the physics-based modelling provided by the TIE-GCM.

The Swarm-VIP models are not a single model, rather they are 20 equations which predict the variability of ionospheric

plasma at different scales and in different latitudinal regions. There are multiple versions of the models and multiple comparisons drawn within this paper (comparisons to climatology, GNSS data and a physics-based model, namely TIE-GCM). A complete assessment would involve 180 comparisons, each with its own set of plots and statistics. Summary statistics of such comparisons would average out areas of particularly good, or bad, model performance. Therefore, a selection of comparisons is presented in this paper. These are designed to show the strengths and limitations of the Swarm-VIP models. We have made a deliberate choice to include and discuss areas where model performance is poor in order to make recommendations for model improvement. It is hoped that these recommendations will be relevant to other groups engaged in statistical, climatological or machine-learning modelling activities. We further highlight here that [Paper 1](#) presents also a primary model validation against Swarm measurements evaluating all the necessary metrics (as suggested by [Liemohn et al., 2021](#)). In this paper, we concentrate on performance assessment, specifically in its application to approximate other parameters, such as those indicating the impact of medium-scale ionospheric irregularities on GNSS signals. Consequently, our focus is on a qualitative evaluation rather than a quantitative one.

3.1 Swarm

The capability of the model to reproduce the climatological variability of the considered quantities is assessed through the direct comparison with Swarm data. We use only the Swarm A datasets since the climatology obtained from the three Swarm satellites is quite similar ([Jin et al., 2020](#)). Regional models are investigated separately by comparing the model outputs with the Swarm climatology. We use the plasma density gradients available within the IPIR dataset ([Jin et al., 2022](#)) for the year 2015 as a basis for the proposed comparison.

3.2 Global Navigation Satellite Systems

The performance assessment against ground-based GNSS data is based on the comparison with ionospheric scintillation indices and Rate of Total Electron Content (TEC) changes (ROT) provided by Ionospheric Scintillation Monitor Receivers (ISMRs) owned by Istituto Nazionale di Geofisica e Vulcanologia (INGV), and with TEC N-S gradients at 20 km, 50 km, and 100 km evaluated by using TEC maps over Italy and Brazil, as described below.

ISMRs provide amplitude (S_4) and phase scintillation (σ_ϕ) indices every 1 minute and ROT every 15 s for every satellite-receiver pair ([Bougard et al., 2011](#)). Amplitude and phase scintillation indices are defined as ([Fremouw et al., 1978](#)):

$$S_4 = \sqrt{\frac{\langle SI^2 \rangle - \langle SI \rangle^2}{\langle SI \rangle^2}}, \quad (3)$$

$$\sigma_\phi = \sqrt{\langle \phi^2 \rangle - \langle \phi \rangle^2}, \quad (4)$$

in which SI is the signal intensity normalized with its low-band pass filtered and ϕ is the detrended signal carrier phase, while $\langle \dots \rangle$ indicates the ensemble average over the considered time window, i.e. 1 minute in our case ([Van Dierendonck](#)

et al., 1993). The detrending is a very delicate aspect of the σ_Φ retrieval at high latitudes (see, e.g. McCaffrey & Jayachandran, 2019; Ghobadi et al., 2020; and references therein) and it would require a dedicated fine-tuning of the adopted scheme and cut-off frequency to retrieve values of the index with the minimum contamination of the refractive effects on GNSS signals (Spogli et al., 2021a). We use here the σ_Φ evaluation based on the application with a 6th-order Butterworth filtering with a 0.1 Hz cut-off frequency, which makes the index particularly sensitive to a medium-scale ionospheric irregularities (above hundreds of meters up to a few tens of kilometres), which includes the kilometre scale, to which Swarm plasma density data from Langmuir Probes onboard Swarm are sensitive. As increases of $S4$ values are sensitive to ionospheric irregularities below the Fresnel's scale for L-band signals and GNSS observational geometry (order of a few hundreds metres), we are not considering this effect since Swarm Langmuir Probe data are not able to monitor these scales.

ROT is defined as:

$$ROT = \frac{sTEC_i - sTEC_{i-1}}{t_i - t_{i-1}}, \quad (5)$$

that is the gradient of the TEC evaluated along the slant path connecting the receiver-satellite pair ($sTEC$) between two different epochs. ROT mixes the spatial and the temporal gradients along the arc and it is here considered for comparison with the plasma density gradients at the various scales provided by the model.

The performance assessment focuses on investigating the capability of the models to follow the development of geomagnetic storms and/or the formation of ionospheric irregularities as detected by increases/changes in the parameter values provided by selected ISMR covering different model regions. In doing such a comparison, a time series of ISMR parameters are considered and models are evaluated by considering the geographic location of an Ionospheric Piercing Point (IPP) located at 350 km altitude, which is widely adopted for the thin shell approximation. We are not using Swarm altitudes for the IPP evaluation as we aim to assess the capability of the Swarm-VIP model to proxy what is experienced by GNSS receivers in their standard use, i.e. on the ground. An elevation angle mask of 30° has been applied to reduce the impact of multipath, which may mimic ionospheric scintillation as well as the impact of large ROT values for low elevation observations. The list of the considered ISMRs, including their geographic coordinates and corresponding ionospheric region considered for selection of the appropriate model region is reported in Table 1. Data from the selected ISMRs are part of the scintillation data collections available at the electronic Space Weather upper atmosphere (eSWua, <http://www.eswua.ingv.it/>) data portal managed by INGV (Upper Atmosphere and Radio Propagation Working Group, 2020).

The ability of the model to provide a proxy for TEC spatial gradients has been investigated by retrieving maps of GNSS-based TEC gradients along the N-S direction, being the direction to which the modelled $|Grad_Ne@XXkm|$ mostly refer, due to the high inclination of the Swarm satellites. We concentrate on two regions: the Italian mid-latitudes region and the region around the expected position of the southern crest of

the Equatorial Ionization Anomaly (EIA) over Brazil. The Italian region was chosen as the European sector was one of the regions on which the ESA-funded Swarm-VIP project focussed. The Brazilian region was chosen as it was expected to be a particularly challenging location for the models and also for the presence of the South Atlantic Anomaly.

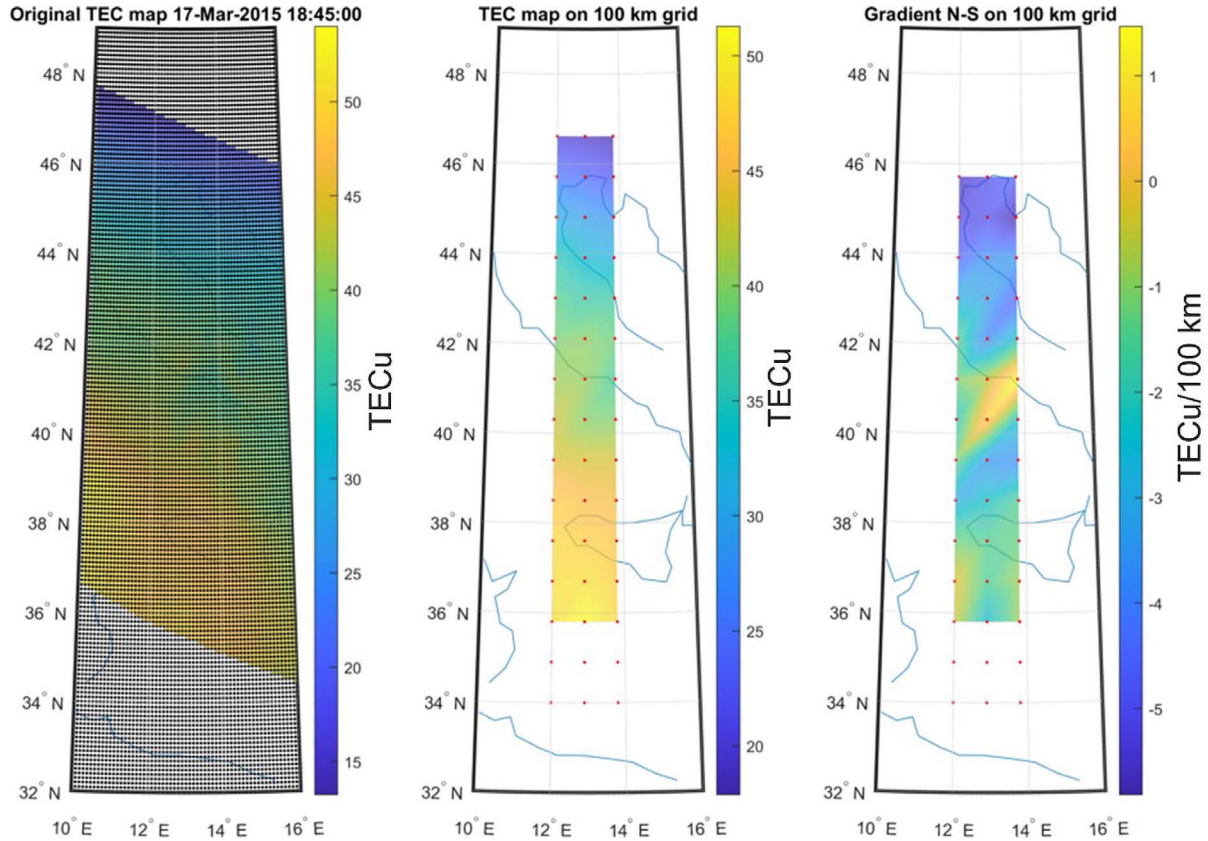
For Italy, we use the Receiver INdependent EXchange format (RINEX) data (at a 30-s sampling rate) freely provided by the Rete Integrata Nazionale GNSS (RING) network of INGV (INGV RING Working Group, 2016) by considering selected cases of ionospheric disturbances. For Brazil, RINEX data (30-s sampling rate) freely provided by the Rede Brasileira de Monitoramento Contínuo dos Sistemas GNSS (RBMC), managed by the Instituto Brasileiro de Geografia e Estatística (IBGE), are used for the period 1 September 2013 to 30 March 2014. This period is selected because it covers high solar flux conditions, and it refers to equinoctial and summer periods. This translates into the maximum probability of formation of post-sunset Equatorial Plasma Bubbles (EPBs) in the Brazilian region. The values of TEC for each receiver-satellite couple are obtained by following the method detailed in Ciraolo et al. (2007), Cesaroni et al. (2021), and Tornatore et al. (2021). The techniques to retrieve the mapping of TEC and TEC N-S gradients are detailed in Cesaroni et al. (2015). To retrieve TEC gradients along the N-S direction at 20 km, 50 km, and 100 km, TEC maps were first evaluated every 20 minutes for Brazil and every 15 minutes for Italy and obtained by considering the last 5 minutes of GNSS observations. Maps over Italy are originally on a 0.1° latitude $\times$ 0.1° longitude grid, while Brazilian maps have a 0.5° latitude $\times$ 0.5° longitude resolution. This segmentation of the maps has been already proven to be effective in depicting the TEC variability in the two considered regions (Cesaroni et al., 2021; Spogli et al., 2021b). TEC values are then re-evaluated on the map with a new grid which has a XX km ($XX = 20, 50, 100$) spacing in the N-S direction and then the N-S gradient is calculated according to formula (4) of Cesaroni et al. (2015). Hereafter we refer to the absolute value of such a TEC gradient simply as $\nabla TEC@XXkm$. An example of the retrieval of the TEC N-S gradient at 100 km over Italy is provided in Figure 1. The model performance is then evaluated in terms of the comparison between medians in the selected geographical areas of $\langle \nabla TEC@XXkm \rangle \pm \sigma(\nabla TEC@XXkm)$ and $\langle |Grad_Ne@XXkm| \rangle \pm \sigma(|Grad_Ne@XXkm|)$, in which σ indicates the standard deviation and $\langle \dots \rangle$ indicates the average.

3.3 TIE-GCM

TIE-GCM is a comprehensive, three-dimensional, first-principles model of the coupled thermosphere and ionosphere system (Dickinson et al., 1981; Richmond et al., 1992; Qian et al., 2014). It solves the three-dimensional momentum, energy, and continuity equations for neutral and ion species at each model time step. Version 2.0 of the model, used here, runs in two resolution modes: 5° and 2.5° in longitude and latitude. The higher resolution version of the model, with a spatial resolution of 2.5° in both longitude and latitude, was used in the present study. The model height spacing is at one-half scale height at 5° resolution and one-quarter scale height at 2.5° . Dang et al. (2021) have described a high-resolution “Version 2.1” of the model, which runs to a resolution of 0.625° across an altitude

Table 1. List of ISMRs, including geographic coordinates and corresponding ionospheric region considered to select the model.

Receiver ID	Location	Latitude (°N)	Longitude (°E)	Region
DMC0P	Concordia Station (Antarctica)	−75.10	123.40	Polar/Auroral
MZS0P	Mario Zucchelli Station (Antarctica)	−74.70	164.11	Auroral
SAN0P	SANAE-IV Station (Antarctica)	−71.70	−2.80	Auroral
NYA0P	Ny-Ålesund (Svalbard)	78.90	11.90	Auroral
TUC0P	San Miguel de Tucumán (Argentina)	−23.73	−65.23	Equatorial

**Figure 1.** Example of *TEC* N-S gradient calculation over Italy. Original *TEC* map using 5-minute data interval (left panel) with the 0.1° latitude $\times$ 0.1° longitude resolution (as indicated by the black grid); *TEC* map evaluated on a regular grid (red dots) having a 100 km spacing over the N-S direction (middle panel); corresponding *TEC* N-S gradient map (right panel).

range of ~ 97 km to 500–700 km (depending on solar activity levels).

TIE-GCM makes a number of assumptions, namely hydrostatic equilibrium, constant gravity, steady-state ion and electron energy equations and incompressibility on a constant pressure surface. Several empirical models are used within the TIE-GCM to specify photoelectron heating, the production of secondary electrons, and the upper boundary conditions for electron heat transfer and electron number flux. TIE-GCM needs a number of magnetospheric inputs, including particle precipitation and the ionospheric electric fields at high latitudes (which are driven from above). Atmospheric tides must be specified at the lower boundary, and by default, the Global Scale Wave Model (GSWM; Hagan et al., 1995) is used.

A more detailed description of the model is provided, and online model runs can be obtained, at NASA's Community

Coordinated Modelling Centre (CCMC) at https://ccmc.gsfc.nasa.gov/requests/IT/TIE-GCM/tiegcm_user_registration.php.

The Swarm data products predicted by the model, i.e. $|\text{Grad_Ne@100km}|$, $|\text{Grad_Ne@50km}|$, $|\text{Grad_Ne@20km}|$ and the *IPIR_{ix}* have a higher resolution than the spatial resolution of TIE-GCM, so this comparison will be restricted to predictions and observations of the electron density.

Predictions and observations compared with four week-long case studies are selected: 4–10 September 2017, 26 April–2 May 2017, 13–20 August 2017, and 23–29 October 2017. These dates are chosen as they include a range of local times (LTs) and geomagnetic activities. On each day, the average longitude is calculated for each half-orbit (from one pole to the other). The half orbit with the smallest average longitude is selected. This meant that all the passes shown below crossed the European sector. The longitudes are between 0° and 15°E .

meridian. In TIE-GCM the longitude which was closest to the average longitude of Swarm was found. The average altitudes of the pressure levels at this longitude are calculated. The pressure levels immediately above and below the altitude of the Swarm are selected. The electron density values at these altitudes and the associated scale height are used to estimate the electron density at the altitude of Swarm.

4 Results

4.1 Climatological assessment of the model performance

In this subsection, the model is tested in its ability to reproduce the climatological features of the Ne and Ne gradients. For this purpose, the model output will be evaluated by direct comparison with Swarm data.

4.1.1 Climatology in the polar caps

In this section, we compare the density gradient at 100 km calculated from the polar cap model with the gradients at 100 km from Swarm data. The equation of the model is recalled below:

$$|Grad_Ne@100km| = \left(\exp \left(-1.69 + 5.74 \cdot 10^{-3} F107_81 - 4.30 \cdot 10^{-3} SZA - 6.29 \cdot 10^{-3} DOY_fn + 3.53 \cdot 10^{-3} Kp + 6.68 \cdot 10^{-3} MLAT \right) \right)^3. \quad (6)$$

The physical meaning of each parameter is explained in the supplementary material of Paper 1 (Table S1). We use the year 2015 as a basis for the comparison. Figures 2a and 2b show the seasonal and universal time (UT) variations of the modelled density gradient (polar model) in 2015 above the geomagnetic pole in the Northern Hemisphere (83.0011° N; 83.7272° W – Fig. 2a) and geomagnetic pole in the Southern Hemisphere (74.8207° S; 124.5879° E – Fig. 2b). Figures 2c and 2d show the seasonal and UT variation of the measured average plasma density gradient at 100 km scale in the polar cap ionosphere (poleward of $\pm 75^\circ$ $MLAT$) binned by 1 h in UT and 5 days in DOY , for Northern (Fig. 2c) and Southern (Fig. 2d) hemispheres. The values of the SZA near the altitude of Swarm A above geomagnetic poles in the northern and southern hemispheres are shown in contours. The horizontal black lines show solstices and equinoxes. Due to the orbital effect, there are a few data gaps in the Southern Hemisphere at certain UT and DOY . However, Figures 2c and 2d, similar to Jin & Xiong (2020), demonstrates that large-scale density gradients have clear seasonal variations, i.e., they show maximum from equinox to winter in the Arctic and they show peak from equinox to summer in the Antarctic. In addition, they also show UT variations which are due to the offset between geographic and geomagnetic poles. This effect can be represented by the SZA as presented by black and white contours. During deep winter with the highest solar zenith angle (lowest solar elevation angle), the density gradients are the lowest.

By comparing Figures 2a and 2b with Figures 2c and 2d, differences between Swarm-VIP models and data in the Northern Hemisphere are present. However, the seasonal and UT variations of density gradients in the southern hemisphere

are well represented, i.e. the low values of density gradients are located near SZA values around 120° , while the high values are located near SZA values of 60° – 90° . The differences are likely due to the strong control of the model by $F107_81$. The high solar activity during the first half of 2015 resulted in strong density gradients during that period, and this kind of variation dominates the other model drivers (not shown). On the contrary, the SZA above the geomagnetic south pole varies a lot, and this results in a strong regulation by the SZA in that sector.

It is also worth noting that the seasonal variations of density gradients in the North and South Polar Caps are different. The density gradient in the North Polar Cap is enhanced during local winter, while the South Polar Cap shows enhancement during local summer. The model does not distinguish between hemispheres, since it has been trained using data from both sectors. Therefore, the Southern Hemisphere could dominate the model as it is usually associated with larger plasma structures (Jin & Xiong, 2020). There is a discontinuity at the year boundary in the model because, while some terms in the model such as DOY_fn are continuous across the year boundary, others are not. For example, $F107_81$ differs at the start and end of 2015 and so contributes to the discontinuity in the model at the year boundary.

4.1.2 Climatology at the auroral latitudes

In this section, we compare the density gradient calculated from the auroral model for the gradients at 100 km against the Swarm data. The equation of the model is recalled below:

$$|Grad_Ne@100km| = \left(\exp \left(-1.65 + 4.6 \cdot 10^{-3} F107_81 - 7.92 \cdot 10^{-2} DOY_fn + 7.15 \cdot 10^{-3} MLAT + 2.08 \cdot 10^{-2} SW_Press - 2.66 \cdot 10^{-3} \cdot SZA + 1.69 \cdot 10^{-2} \cdot IEF - 1.20 \cdot 10^{-4} SW_Vel \right) \right)^3. \quad (7)$$

The physical meaning of each parameter is explained in Section 2.2 of Table S1 of Paper 1.

Figures 3a–3d show the parameters that are used to feed $|Grad_Ne@100km|$ model in the auroral latitudes in the northern and southern hemispheres, respectively. As the equation at auroral latitudes does not depend on LT , the Swarm orbits are not divided by ascending/descending orbits. We use the Swarm data that are flagged as being located in the auroral latitudes from the IPIR dataset. The mean magnetic latitudes are shown in Figure 3b. Due to the dynamical feature of the auroral oval in response to the solar wind-magnetosphere-ionosphere coupling, the magnetic latitudes of the auroral region vary with solar wind dynamic wind pressure and speed. The solar wind dynamic pressure and solar wind speed, and IEF are presented (Figs. 3c–3d). Figures 3 shows the modelled (Fig. 3e) and daily averaged (Fig. 3f) density gradients in 2015. The most obvious feature of the modelled results is that they show annual variations that peak around December–January. There is also one peak in the middle of 2015 that is likely due to the high value of solar wind dynamic pressure during the storm on June 22, 2015 (Piersanti et al., 2017). Figure 3f shows that the daily averaged density gradients from Swarm also show seasonal variations that are like the modelled result. However, the averaged data are more fluctuating and with larger amplitudes than the modelled one. This result is consistent with Paper 1, where

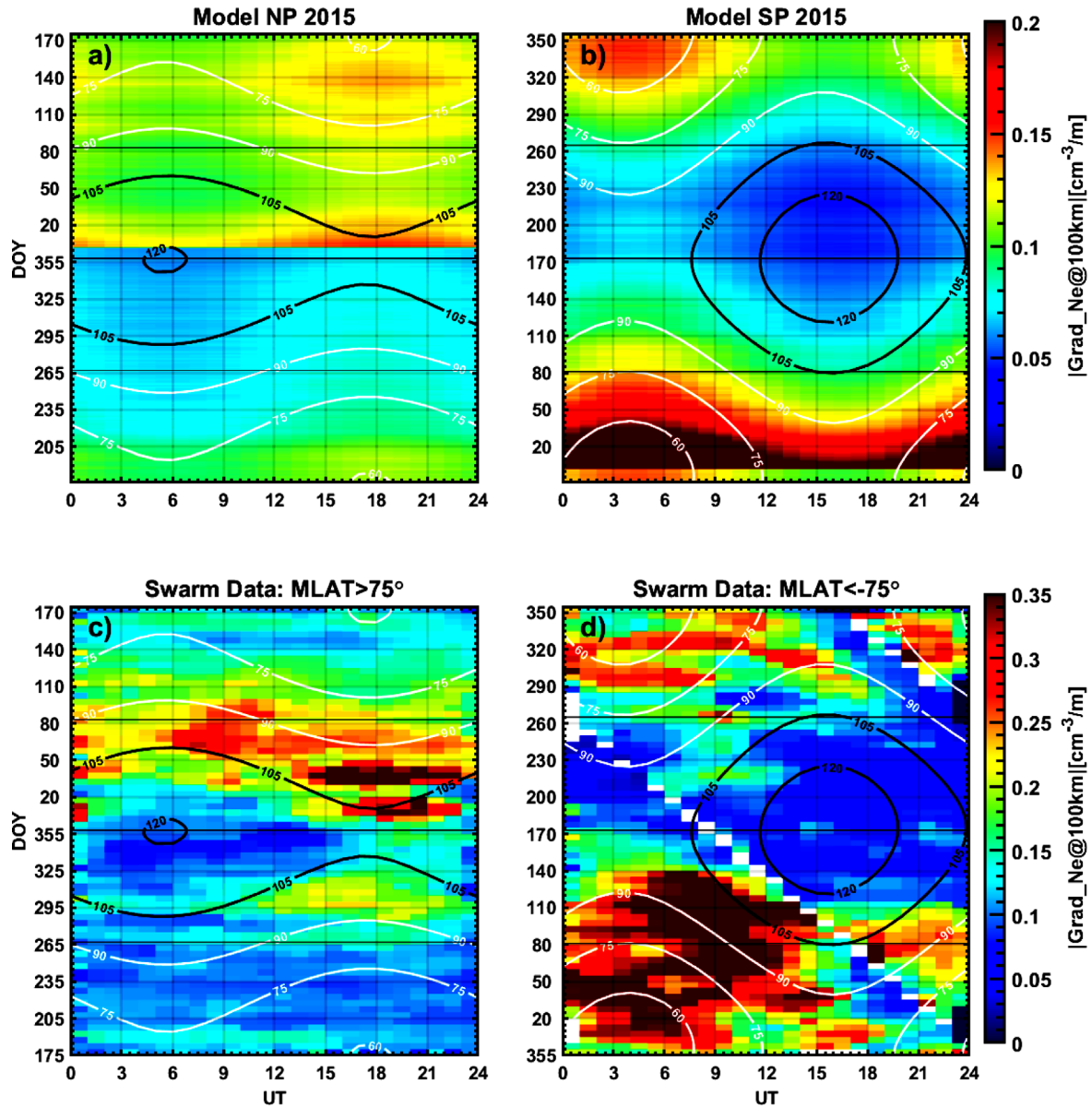


Figure 2. Seasonal and UT variations of modelled density gradient (polar model) in 2015 above the geomagnetic pole in the northern hemisphere (NP, 83.0011° N; 83.7272° W – a) and geomagnetic pole in the southern hemisphere (SP, 74.8207° S; 124.5879° E – b). Seasonal and UT variation of the averaged plasma density gradient at 100 km in the polar cap ionosphere (poleward of $\pm 75^\circ$ MLAT) binned by 1 h in UT and 5 days in the day of the year, for Northern (c) and Southern (d) hemispheres. The solar zenith angle near the altitude of Swarm A above geomagnetic poles in the northern and southern hemispheres are shown in contours. The horizontal black lines show solstices and equinoxes.

Table 3 showed that the precision of this model (defined according to Liemohn et al., 2021) was markedly less than 1, indicating that the model predictions were less variable than the observations. Interestingly, the daily averaged density gradients do show a peak on June 22, 2015.

Figure 4 presents scatter plots of modelled and daily averaged density gradients from Swarm A at auroral latitudes against $F107_{81}$ and DOY_{fn} for the Northern (black) and Southern (red) Hemispheres to show their dependence on these two different parameters. Clearly, both the model and the data show dependence on a proxy for the solar activity (in this case $F107_{81}$). When the solar activity increases, the density

gradients increase as well. The Swarm data have a wider range of values than the model. This indicates that other parameters are also playing a role in the generation of plasma structures. When it comes to the dependence of DOY_{fn} , the model shows a clear correlation with DOY_{fn} , and the correlation is more obvious in the Southern Hemisphere.

4.1.3 Climatology at mid-latitudes

In this section, we compare the density gradient calculated from the mid-latitude model for the Ne gradients at 100 km against the Swarm data. The equation of the model is recalled below:

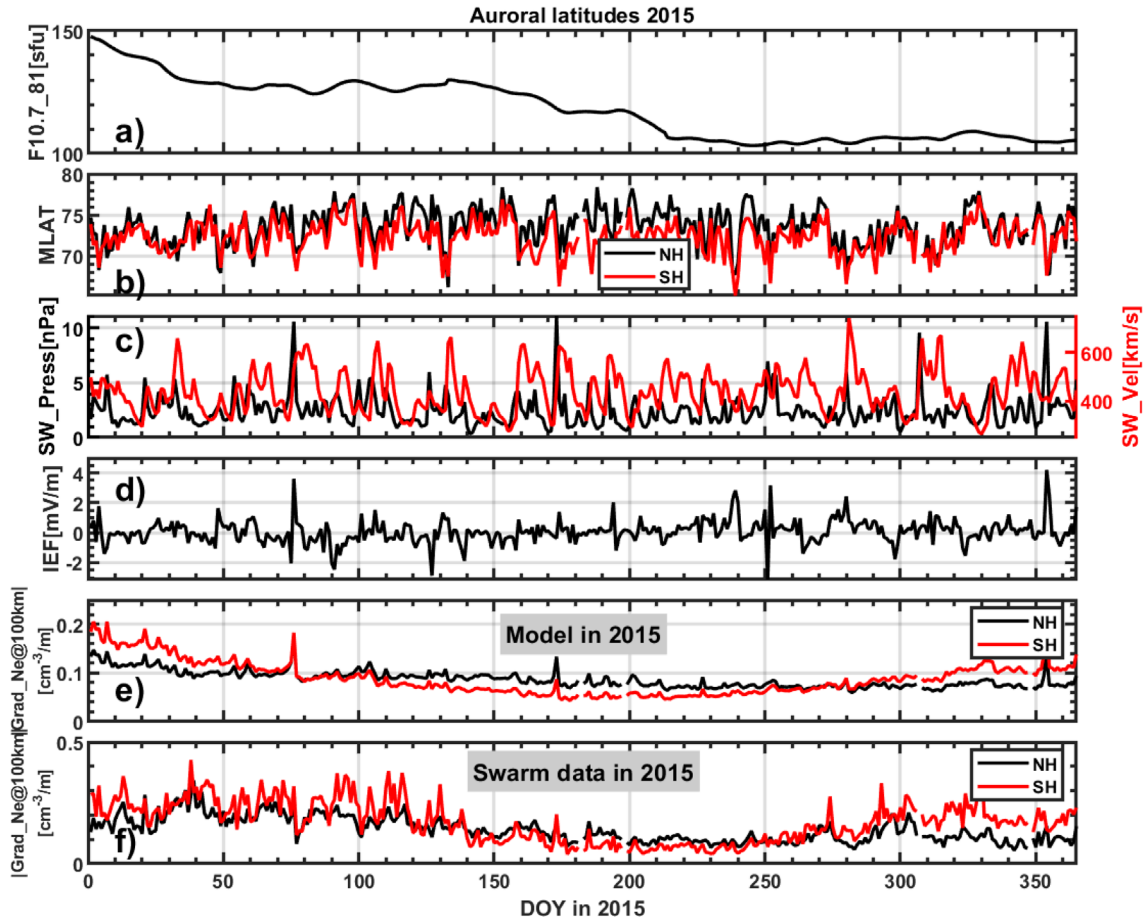


Figure 3. The input for the auroral model is the daily averaged values of Swarm A (a–d). The magnetic latitudes are daily averaged of the ionosphere region 2 (aurora) from the IPIR dataset. The modelled (auroral model) density gradients at 100 km (e) and the daily averaged density gradients at 100 km (f) at auroral latitudes of the Northern (NH, black) and Southern (SH, red) Hemispheres using data from Swarm A.

$$|Grad_Ne@100km| = \left(\exp \left(-5.97 \cdot 10^{-1} + 9.93 \cdot 10^{-4} F107_81 - 8.67 \cdot 10^{-3} ST + 3.45 \cdot 10^{-6} Newell - 2.02 \cdot 10^{-2} DOY_fn - 5.15 \cdot 10^{-4} LAT + 2.03 \cdot 10^{-3} IMF_Abs_By \right) \right)^7. \quad (8)$$

The physical meaning of each parameter is explained in [Paper 1](#) (Sect. 2.2 and Table S1).

[Figures 5a–5d](#) show the parameters for calculating density gradients at middle latitudes. As the model is dependent on solar time (ST), we only show the results corresponding to the ascending (northward) orbits. This is because Swarm satellites cross different LT sectors during ascending and descending orbits. The ST of the ascending orbits in the Northern and Southern Hemispheres are shown in [Figure 5b](#). To make it symmetric about local solar noon, the value of the ST is used if $ST < 12$, while the value used is $24 - ST$ for $ST > 12$. The modelled and daily averaged density gradients are presented in [Figures 5e](#) and [5f](#). The ratio of $|Grad_Ne@100km|$ calculated by the model and Swarm data is about 0.1–0.9, with a mean value of 0.27. For the mid-latitude region, the seasonal variations of Swarm data are obviously different from the model. The model uses a sine function with a period of one year (mainly driven by its anti-correlation with ST). The purpose of this function was

to represent the annual anomaly, whereas the Swarm data clearly shows evidence of variations in other periods with two clear peaks. This is the major discrepancy between the model and the data. In a previous study the technique of Generalised Linear Modelling was applied to the high-latitude ionosphere ([Dorrian et al., 2019](#)), a function to represent the semi-annual anomaly was trialled but was not then included in any of the models. Such a function was not included as an explanatory variable in the present study for two reasons. Firstly, it was weakly correlated with the SZA and so, based on the modelling method used, a choice had to be made between using this function and the SZA . The SZA has a more obvious physical interpretation and was selected. Secondly, it was anticipated that any such variations would be captured by the observations of the thermospheric density made by Swarm (models shown in Table S4 of [Paper 1](#)). As discussed in [Paper 1](#), a thermospheric data product with a higher temporal resolution is needed and this will be explored in a future study.

[Figure 6](#) shows the scatter plots of modelled (mid-latitude model) and daily averaged density gradients versus $F107_81$ and DOY_fn . Similarly to auroral latitudes, both the model and data are dependent on $F107_81$. The bottom panels show the dependence on the day of the year. For the model, a clear dependence on DOY_fn is seen. However, two groups

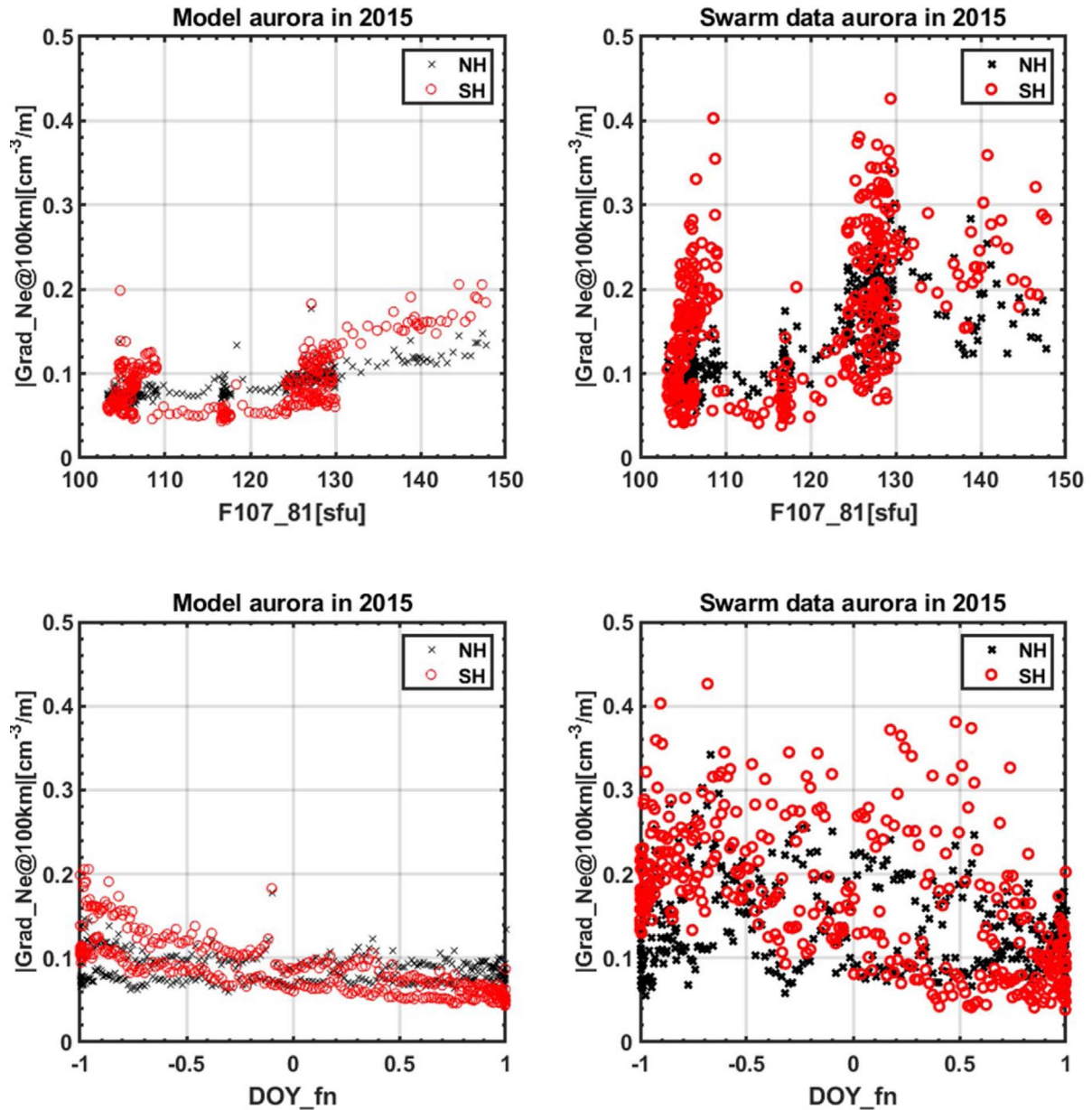


Figure 4. The scatter plots of modelled (auroral model) and daily averaged density gradients from Swarm A at auroral latitudes against *F107_81* and *DOY_fn* for Northern (black) and Southern (red) Hemispheres.

of data exist. This can explain the differences in the seasonal variations, i.e. a sine function with a period of one year is not enough to explain the seasonal variations of mid-latitude density gradients.

4.1.4 Climatology at the equatorial latitudes

In this section, we compare the density gradients calculated from the equatorial-latitudes model for the gradients at 100 km against the Swarm data. Figure 7 shows the climatology of electron density and density gradient at 100 km scale at low latitudes, as a function of geographic latitude and local time. The data are averaged in 2015 using Swarm A. The averaged plasma density and its gradient are enhanced from 09:00 to 22:00 LT, due to EIA (MacDougall, 1969). After 18:00 LT, the plasma density

gradients are intensified and cover a wider area, which includes the equator. This is the signature of the plasma density gradients embedded in the post-sunset EPBs that can be detected at Swarm altitudes (see Wood et al., 2022; and references therein).

To overcome the limitation of the scarce capability of the models to reproduce the signatures of the post-sunset equatorial plasma bubbles, three additional categories of model in the equatorial region, one to represent daytime, one to represent night-time and one to represent the evening, when EPBs were more likely to occur (Li et al., 2021), are developed. The three different considered LT sectors separately could be set as 01:00–08:00 LT (night), 08:00–18:00 LT (day) and 18:00–01:00 LT (bubbles). The expressions for the $|Grad_Ne@100km|$ models for the three LT regions are for the dayside sector (08:00–18:00 LT):

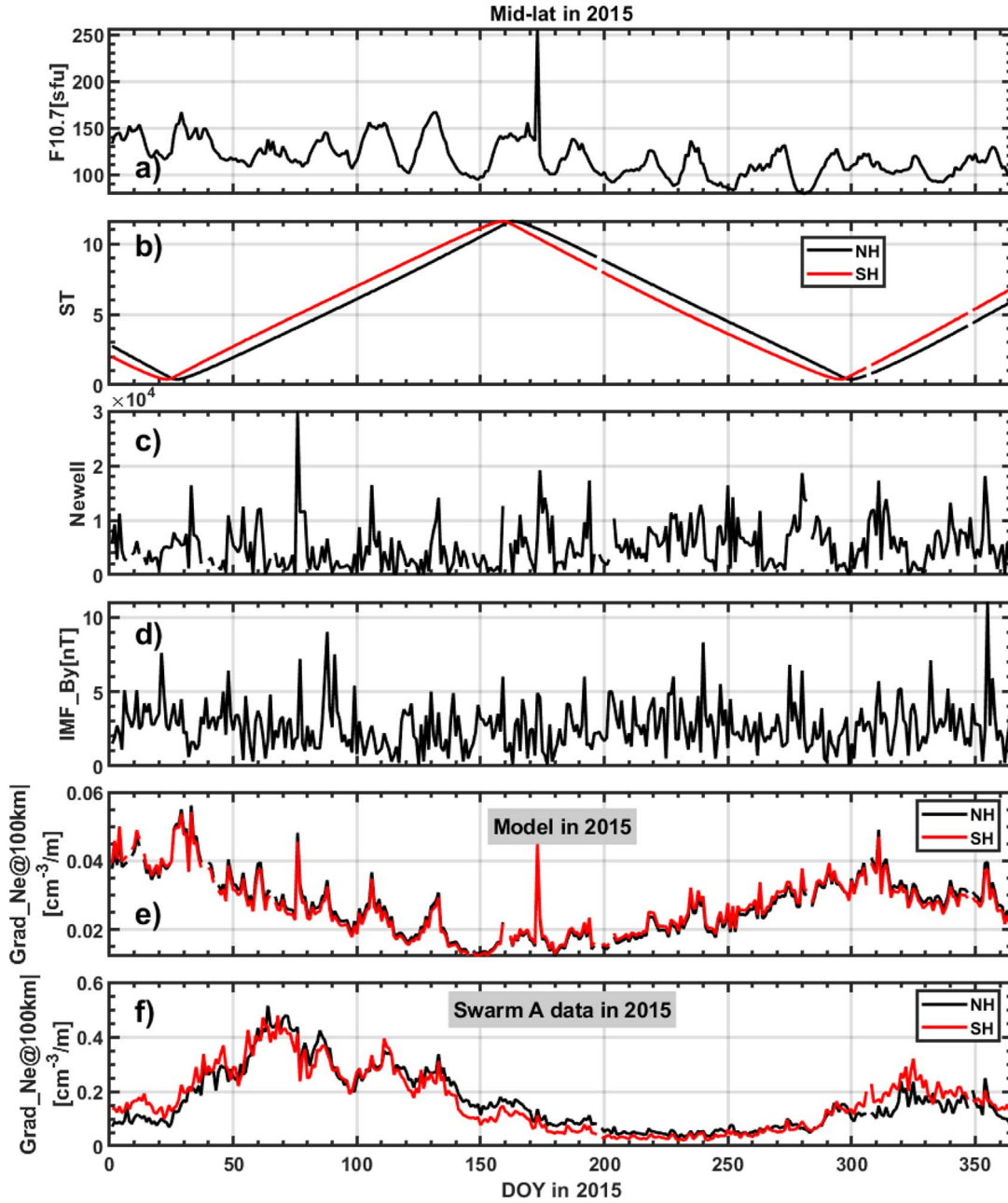


Figure 5. The input for the mid-latitude model as daily averaged values of Swarm A (a–d). The modelled (mid-latitude model) density gradients at 100 km (e) and the daily averaged density gradients at 100 km (f) at mid-latitudes of the Northern (black) and Southern (red) Hemispheres using data from Swarm A.

$$|Grad_Ne@100km| = \left(\exp \left(-1.69 + 1.39 \cdot 10^{-2} Sunspot.27 - 2.53 \cdot 10^{-2} LAT + -2.81 \cdot 10^{-3} \cdot SYM_H - 5.53 \cdot 10^{-3} SZA - 9.82 \cdot 10^{-2} DOY_fn \right) \right)^4, \quad (9)$$

while for the EPB sector (18:00–01:00 LT):

$$|Grad_Ne@100km| = \left(\exp \left(-4.17 \cdot 10^{-1} + 3.86 \cdot 10^{-3} F107_81 - 1.11 \cdot 10^{-2} LAT - 3.19 \cdot 10^{-3} SZA + 3.84 \cdot 10^{-4} IMF_Clock \right) \right)^4, \quad (10)$$

and for the nightside sector (01:00–08:00 LT):

$$|Grad_Ne@100km| = \left(\exp \left(-4.95 \cdot 10^{-1} - 2.90 \cdot 10^{-2} ST + S4.84 \cdot 10^{-4} Sunspot + 6.62 \cdot 10^{-3} IMF_Bt - 3.87 \cdot 10^{-3} MLAT - 1.94 \cdot 10^{-4} SYM_H - 2.29 \cdot 10^{-3} IMF_By - 6.95 \cdot 10^{-4} SYM_D \right) \right)^6. \quad (11)$$

The known season/longitude variations of equatorial ionospheric irregularities are not included in the models. [Figure 8](#)

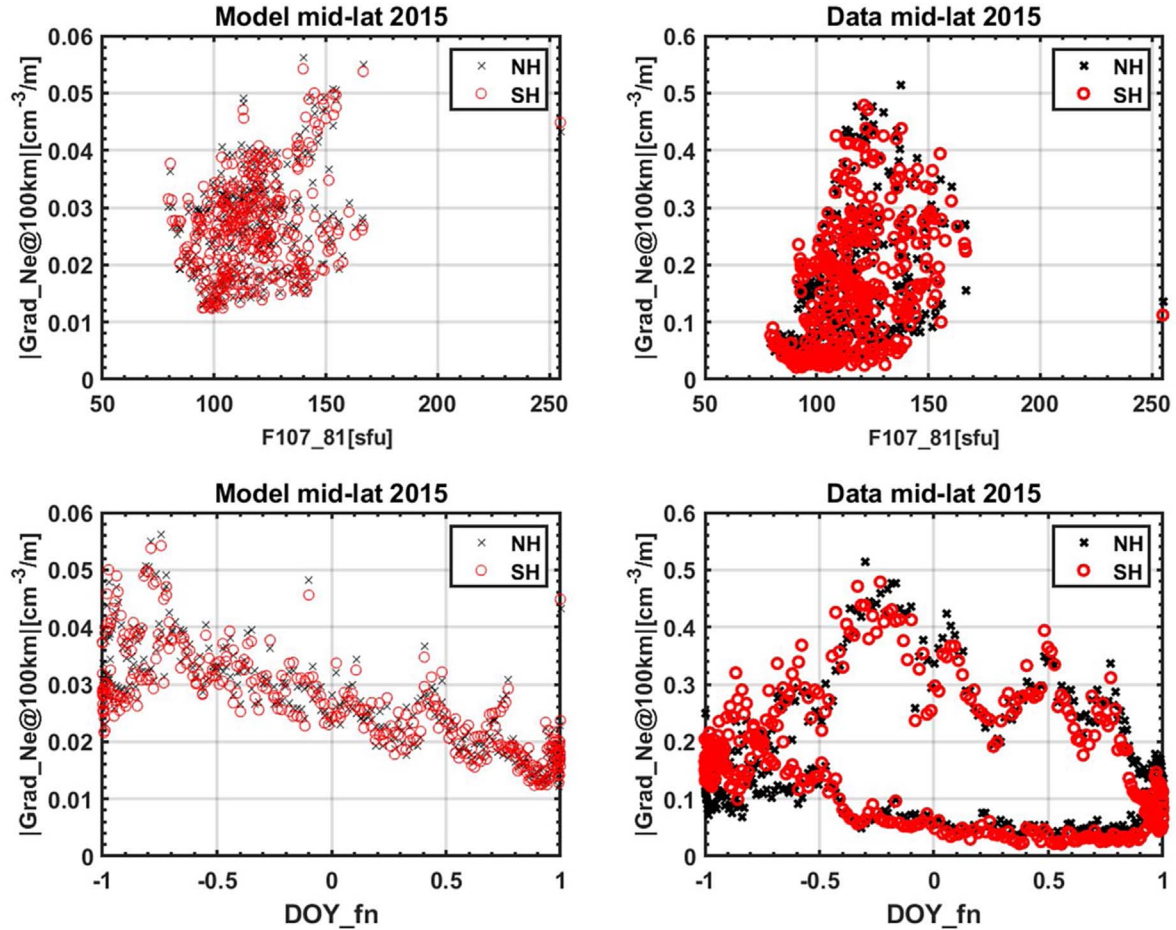


Figure 6. The scatter plots of modelled (left plots) and daily averaged (right plots) density gradients at 100 km from Swarm A for the ascending orbits as a function of $F107_81$ (top plots) and DOY_fn (bottom plots), separately for Northern (black) and Southern (red) Hemispheres.

is analogous to Figure 7 except that the equatorial models have been sorted according to the different LT s. Figure 8 is produced using the averaged helio-geophysical indices in 2015. This is to match the time for the averaged Swarm data in 2015 (Fig. 7). For the dayside sector 08–18 LT , the DOY was selected as the Spring equinox (March 22) in 2015. Note that selecting other DOY will only affect absolute values of density gradients but not the general pattern. The models reproduce the general features, like daytime enhancements of both N_e and $|Grad_Ne@100km|$ during daytime and the increase and latitudinal spreading after local post-sunset hours (18:00 LT), but the double crest signature is completely lost.

4.2 Performance assessment against ground-based GNSS observations

In this section, the model is tested to verify its capability of reproducing ionospheric weather as depicted by ground-based GNSS data, by concentrating on σ_ϕ and ROT data from ISMRs and on TEC spatial gradients evaluated in the N-S direction at the three scales of 20, 50, and 100 km.

To verify the ability of the model to follow the development of a geomagnetic storm and represent the consequent formation of the medium-scale irregularities in the high-latitude ionosphere, we focus on the September 2017 storm. The severe geomagnetic storm, caused by a series of X and M class flares and associated coronal mass ejections, had its main phase during 7–8 September 2017 and its effects are among the most studied of the last solar cycle (see., e.g. Vanlommel & Van der Linden, 2017; Linty et al., 2018; Alfonsi et al., 2021; de Paula et al., 2022; Sato et al., 2019). The formation of irregularities and their impact on GNSS has been reported at high latitudes in both the Northern and Southern Hemispheres, in terms of enhancements of the scintillation indices recorded by ground-based GNSS receivers (Linty et al., 2018; Ghobadi et al., 2020; Vilà-Valls et al., 2020; D’Angelo et al., 2021). To the scope, Figure 9 shows time series of σ_ϕ (black) for every GPS satellite in view with elevation $>30^\circ$ by MZSOP, SANOP (b), DMCOP (c), and NYAOP (d) receivers (see Table 1) and of the corresponding IPIR index, $IPIR_ix$ (red), from auroral model (Figs. 9a–9d) and polar cap model (Figs. 9e–9h), for the period 5–10 September 2017. Events that occur rarely pose a challenge for statistical models to capture. These dates were chosen as they included a

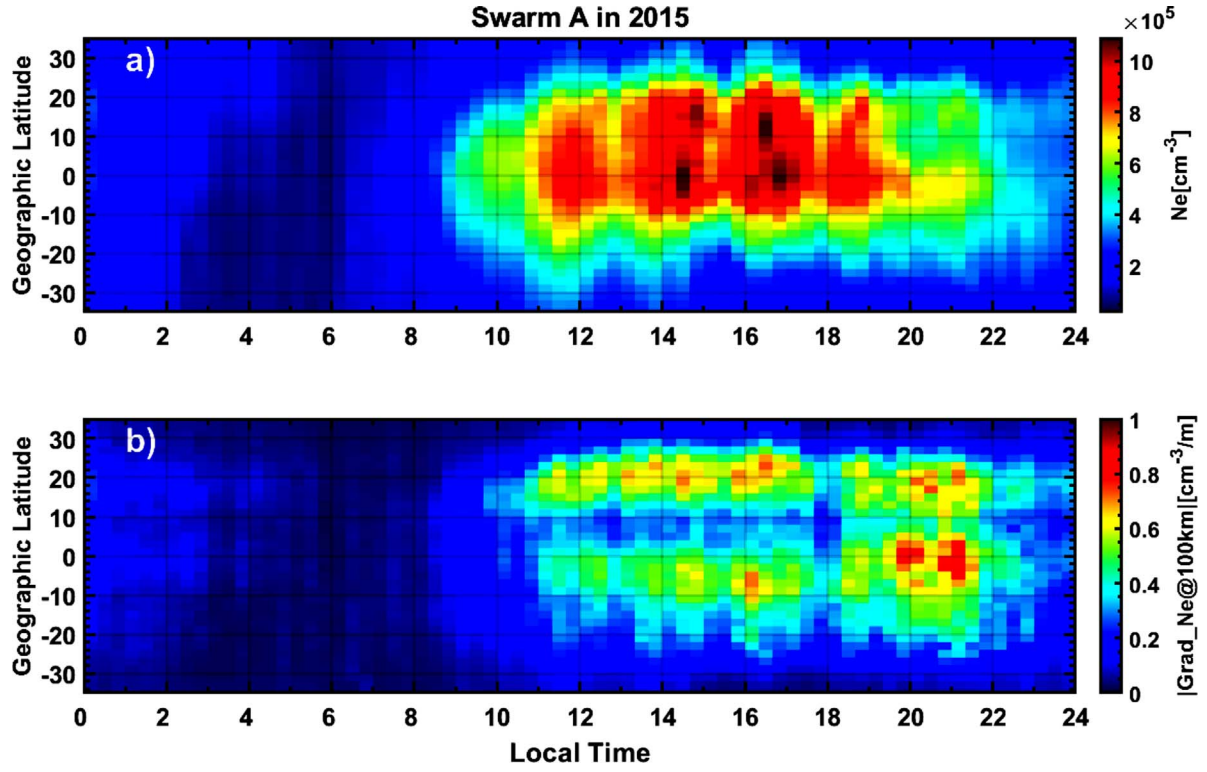


Figure 7. The average N_e (top plot) and density gradient at 100 km (bottom plot) as a function of geographic latitude and local time. The data from Swarm A in 2015 are divided into bins of 2° in latitude and 20 min in local time.

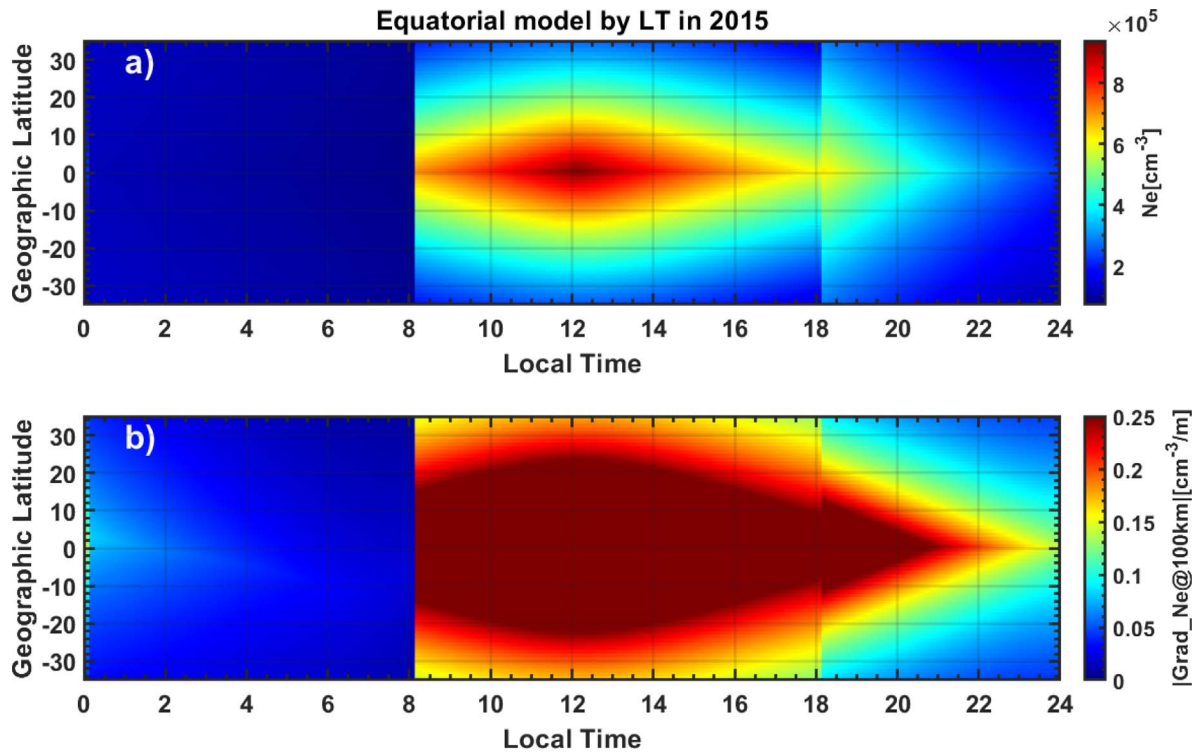


Figure 8. Same as Figure 7, but for the equatorial models being sorted according to the different LT .

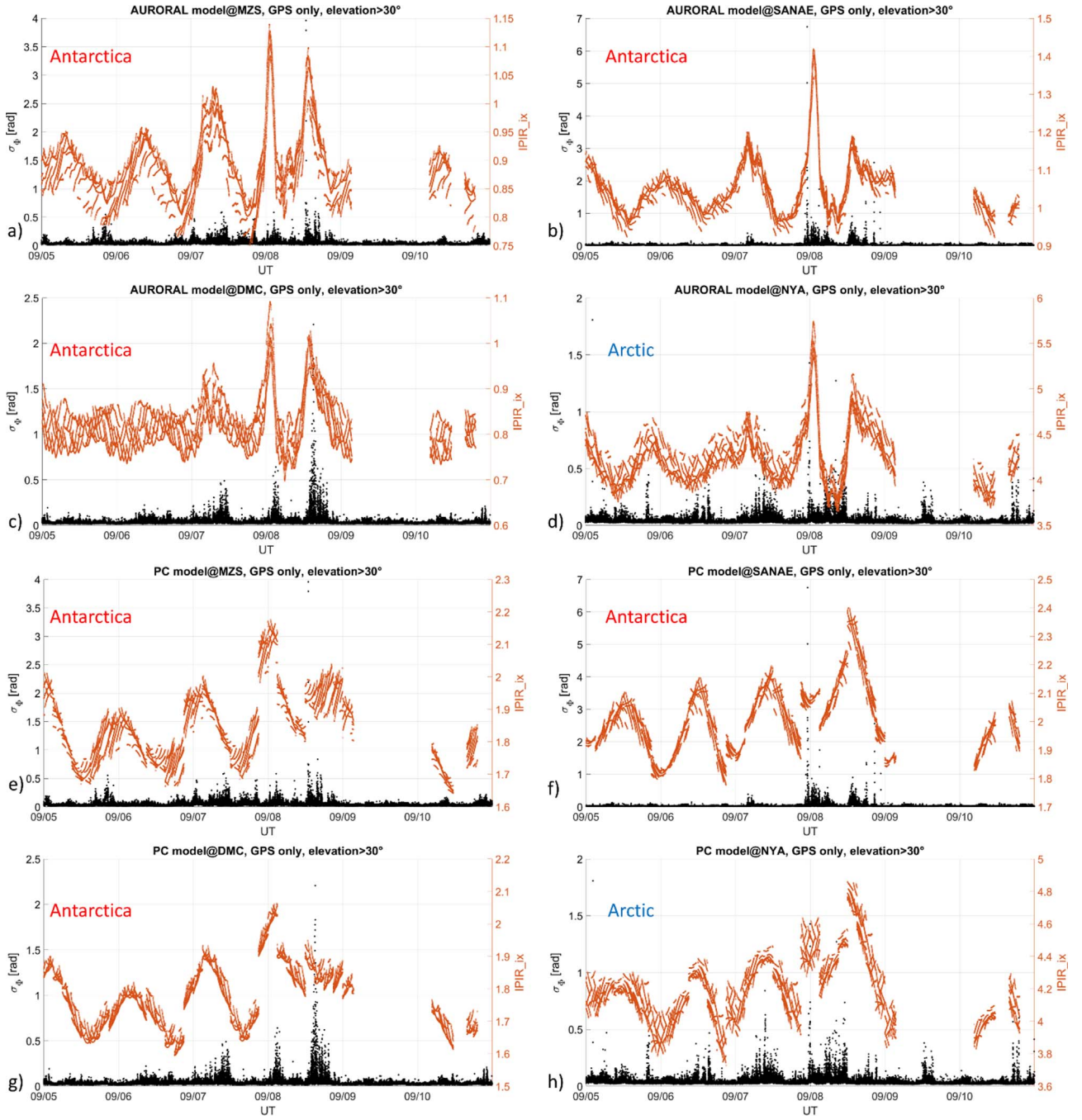


Figure 9. Time series of σ_ϕ (black) for every GPS satellite in view with elevation $>30^\circ$ as observed from the receivers at (a) MZS0P, (b) SAN0P, (c) DMC0P, (d) and NYA0P and of the corresponding IPIR index (red) from auroral model (panels a–d) and polar cap model (e–h), for the period 5–10 September 2017.

number of flares and heightened geomagnetic activity, both of which represent challenging conditions for the Swarm-VIP models. The IPIR index, $IPIR_{ix}$, was used within this comparison as Kotova et al. (2023) highlighted the usability of this index as a proxy for large-scale plasma variations, which can lead to phase fluctuations in GNSS signals, by leveraging on a statistical characterization upon 23 ground-based GNSS scintillation receivers at polar, auroral and low latitudes. The

expressions of the $IPIR_{ix}$ model for the auroral and polar regions, respectively, are the following:

$$\begin{aligned}
 IPIR_{ix} = & \exp \left(4.87 \cdot 10^{-1} + 2.83 \cdot 10^{-3} F107_{81} \right. \\
 & + 9.60 \cdot 10^{-3} MLAT - 8.02 \cdot 10^{-3} MLT + 1.38 \cdot 10^{-2} SW_{Press} \\
 & \left. - 2.64 \cdot 10^{-2} DOY_{fn} + 1.24 \cdot 10^{-2} IEF \right), \quad (12)
 \end{aligned}$$

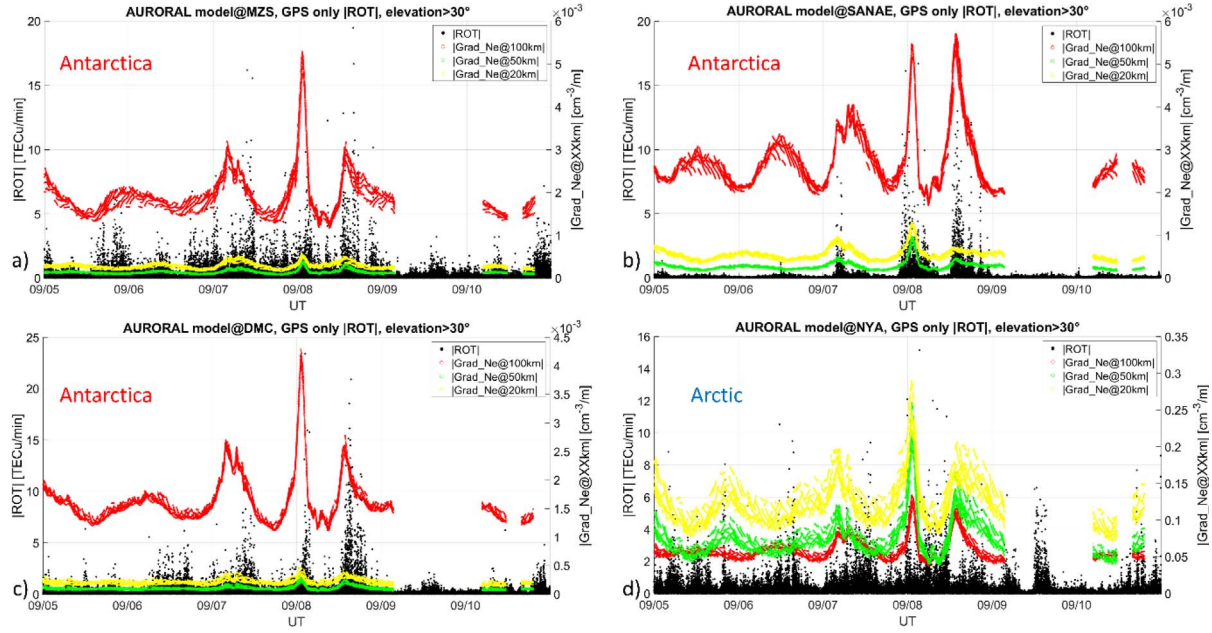


Figure 10. Time series of $|ROT|$ (black) for every GPS satellite in view with elevation $>30^\circ$ as observed from the receivers at (a) MZS0P, (b) SAN0P, (c) DMC0P and (d) NYA0P and of the corresponding $|Grad_Ne@XXkm|$ ($XX = 20, 50$, and 100 km) predicted by the auroral model, according to the colour coding reported in the legend, for the period 5–10 September 2017.

$$\begin{aligned}
 IPIR_{ix} = & \exp \left(9.67 \cdot 10^{-1} + 3.32 \cdot 10^{-3} F107.81 \right. \\
 & - 3.11 \cdot 10^{-3} SZA + 5.03 \cdot 10^{-3} MLAT + 1.93 \cdot 10^{-3} Kp \\
 & \left. - 2.01 \cdot 10^{-2} DOY_{fn} + 1.92 \cdot 10^{-3} SW_{Den} \right). \quad (13)
 \end{aligned}$$

The MZS0P, SAN0P, and NYA0P receivers are mostly auroral and cusp stations while DMC0P is mostly within the cap (see, e.g. Spogli et al., 2009; Prikryl et al., 2011; D’Angelo et al., 2018; De Franceschi et al., 2019). The time series of $IPIR_{ix}$ has a gap because all the values of the IEF and/or other solar wind parameters for the preceding 2-hour period were missing. This model requires the averaging of these parameters over the preceding 2-hour interval.

Enhancements of σ_ϕ for all receivers are found mostly on 7 and 8 September, indicating the presence of ionospheric irregularities covering the range spanned by Swarm observations. Such enhancements have good agreement with increases in the model predictions of $IPIR_{ix}$. Some scintillation is found on late 5 September for MZS0P and NYA0P, in correspondence with an increase of the polar cap model predictions (in anti-phase with the auroral model). There is a daily peak of the $IPIR_{ix}$ which does not have any translation into enhancements of the phase scintillation index.

Figure 10 presents time series of $|ROT|$ (in black) for every GPS satellite with elevation $>30^\circ$ by MZS0P (a), SAN0P (b), DMC0P (c), and NYA0P (d) receivers and of the corresponding $|Grad_Ne@XXkm|$ ($XX = 20, 50$, and 100 km) from auroral model, according to the colour coding reported in legend, and for the period 5–10 September 2017. $|ROT|$ is the measure of the TEC variation along each arc, hence mixing spatial and temporal TEC gradients, while $|Grad_Ne@XXkm|$ represents the irregularities on the N-S direction at the given spatial scale

(20, 50, and 100 km). We remind the reader that $|ROT|$ enhancements are triggered mostly by the passage of the GNSS signal through mesoscale irregularities, i.e. those having scale sizes in the few km up to few tens of km range. The exact scale causing the phase variation is also dependent on the relative geometry between the ray path at the IPP and the bulk direction of the irregularity. So the purpose of the comparison is to assess the ability of the model to represent irregularities on a spatial scale close to the one investigated by means of $|ROT|$, especially the 20 km one. Additionally, $|ROT|$ is a physical measure of a similar kind of $|Grad_Ne@XXkm|$. As already mentioned, it is not the purpose of the proposed assessment to get a one-by-one comparison of the investigated quantities, rather than check the possibility to use the model to proxy some GNSS effects induced by the presence of the irregularities on compatible scale sizes.

Similar to Figure 9, the effect of the storm in both hemispheres was the triggering of ionospheric irregularities, as shown by $|ROT|$ enhancements. Here the investigated models differ from what was shown in the previous figure. The auroral models for the three gradients of Ne are expressed as:

$$\begin{aligned}
 |Grad_Ne@100km| = & \left(\exp(-1.65 + 4.64 \cdot 10^{-3} F107.81 \right. \\
 & - 7.92 \cdot 10^{-2} DOY_{fn} + 7.15 \cdot 10^{-3} MLAT + 2.08 \cdot 10^{-3} SW_{Press} \\
 & \left. - 2.66 \cdot 10^{-3} SZA + 1.69 \cdot 10^{-2} IEF - 1.20 \cdot 10^{-4} SW_{Vel}) \right)^3, \quad (14)
 \end{aligned}$$

$$\begin{aligned}
 |Grad_Ne@50km| = & \left(\exp(-8.55 \cdot 10^{-1} + 1.76 \cdot 10^{-3} F107.81 \right. \\
 & + 5.04 \cdot 10^{-3} MLAT - 2.80 \cdot 10^{-2} DOY_{fn} - 8.39 \cdot 10^{-1} SW_{Press} \\
 & \left. - 4.86 \cdot 10^{-3} MLT - 7.53 \cdot 10^{-3} IEF) \right)^8, \quad (15)
 \end{aligned}$$

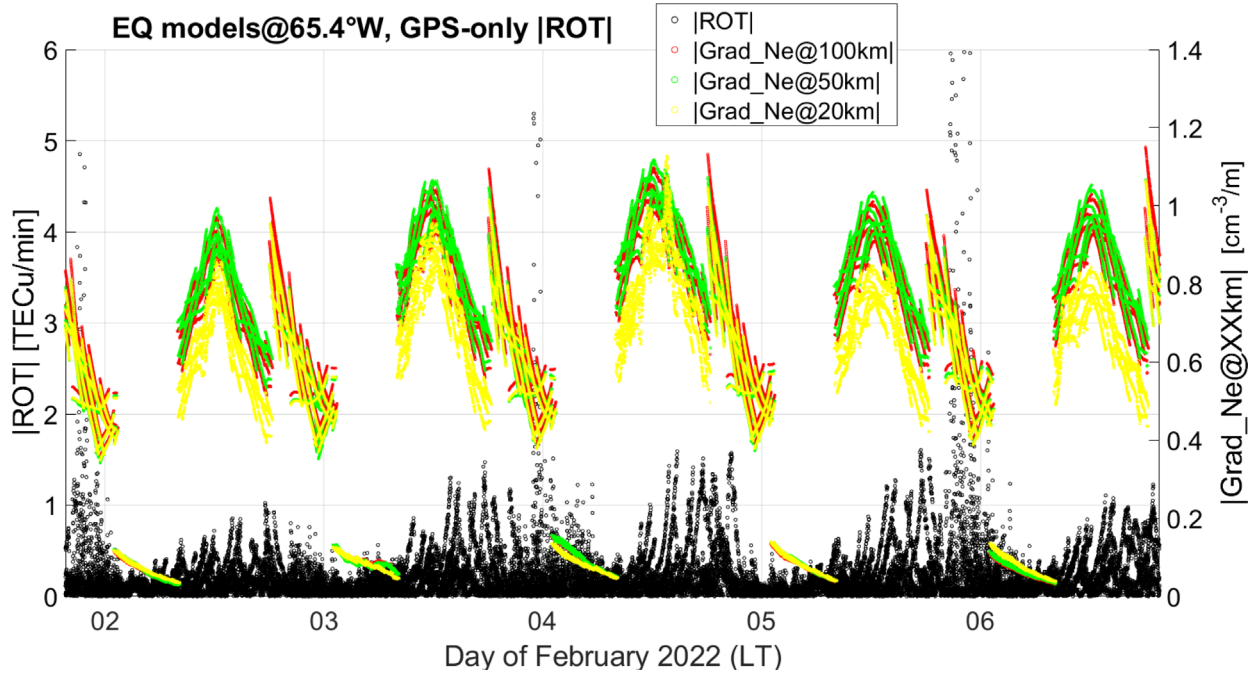


Figure 11. Time series (in LT) of $|ROT|$ (black), for GPS satellites in view with elevation $>30^\circ$ by TUCOP receiver and of the corresponding $|Grad_Ne@XXkm|$ (XX = 20, 50, and 100 km) from the LT-dependent equatorial models, according to the colour coding reported in legend, for the period 2–6 February 2022 UT.

$$|Grad_Ne@20km| = \left(\exp(-3.10 \cdot 10^{-1} + 6.98 \cdot 10^{-3} F107.81 + 1.95 \cdot 10^{-2} MLAT - 1.07 \cdot 10^{-1} DOY_{fn} + 3.21 \cdot 10^{-2} SW_Press - 1.76 \cdot 10^{-2} MLT - 1.17 \cdot 10^{-2} IMF_Bz) \right)^2. \quad (16)$$

The figure intends also to report on the inter-hemispheric asymmetry between the Northern Hemisphere station NYAOP (d) and the other Southern Hemisphere measurements by the three receivers at MZSOP (b), SANOP (b) and DMCOP (c) and of the corresponding $|Grad_Ne@XXkm|$ from auroral model, according to the colour coding reported in legend, for the period 5–10 September 2017. Also in this case, the auroral model can reproduce the presence of irregularities with scale sizes of a few km's scales and can follow the development of the geomagnetic storms in the high-latitude sector.

To assess the equatorial model, the period 2–6 February 2022 in the Southern Hemisphere has been selected, as it has the right seasonality (close to the equinox and local summer) and solar flux conditions (ascending phase of the solar cycle) to have enhanced probability of formation of ionospheric irregularities embedded in the post-sunset EPBs (see, e.g. Cesaroni et al., 2015; Muella et al., 2017; Li et al., 2021; Macho et al., 2022). This variability, coupled with the proximity to the South Atlantic Anomaly, represents a particularly challenging set of conditions for the model. Figure 11 shows the time series of $|ROT|$ (black) for every GPS satellite with elevation $>30^\circ$ by TUCOP (top) receiver, located under the crest of the EIA in the Argentinean sector, and of the corresponding $|Grad_Ne@XXkm|$ from the LT-dependent equatorial models, according to the colour coding reported in legend, for the period 2–6 February 2022.

For the sake of brevity, the 9 analytical expressions of the $|Grad_Ne@XXkm|$ LT-dependent equatorial models are not explicitly reported, and the reader is referred to Table S3 of Paper 1.

The $|ROT|$ time series for the TUCOP receiver is featured by enhancements in local post-sunset hours ($LT = UT - 3$) up to 5 TECu/min and more. This is the typical signature of the presence of the post-sunset EPBs. For what concerns the models, the discontinuities are due to the different formulations of the models in the three considered LT sectors (01:00–08:00, 08:00–18:00, and 18:00–01:00). All models present a primary peak at the local noon, having a weak correspondence with $|ROT|$ data which tends to be enhanced slightly after local noon. A secondary peak, present at later hours by the 08:00–18:00 model, fits with the increase in $|ROT|$ due to plasma bubbles, although the 08:00–18:00 model tends to indicate that the maximum gradients occur five hours before the peaks in $|ROT|$. During the night between 2 and 3 February and the night between 4 and 5 February 2022, no signature of EPBs is found in $|ROT|$ data. This is due to the day-to-day variability of the EPB occurrence (Li et al., 2021), while the model, being climatological in nature, predicts their presence every day.

Concerning the performance of the models against TEC spatial gradients in the North-South direction ($V_{TEC@XXkm}$), whose retrieval is reported in Section 3.2, Figure 12 reports the median of TEC N-S gradient at 20 km (a) and 100 km (c) and of the LT-dependent equatorial models of $|Grad_Ne@20km|$ (b) and of $|Grad_Ne@100km|$ (d) in the geographical sector ($54^\circ W$, 17° – $20^\circ S$), i.e. in the Brazilian sector, as a function of day and LT. This geographical sector is characterized to be the closest, among the available ones, to the expected position of the southern crest of the EIA. The yellow solid lines indicate

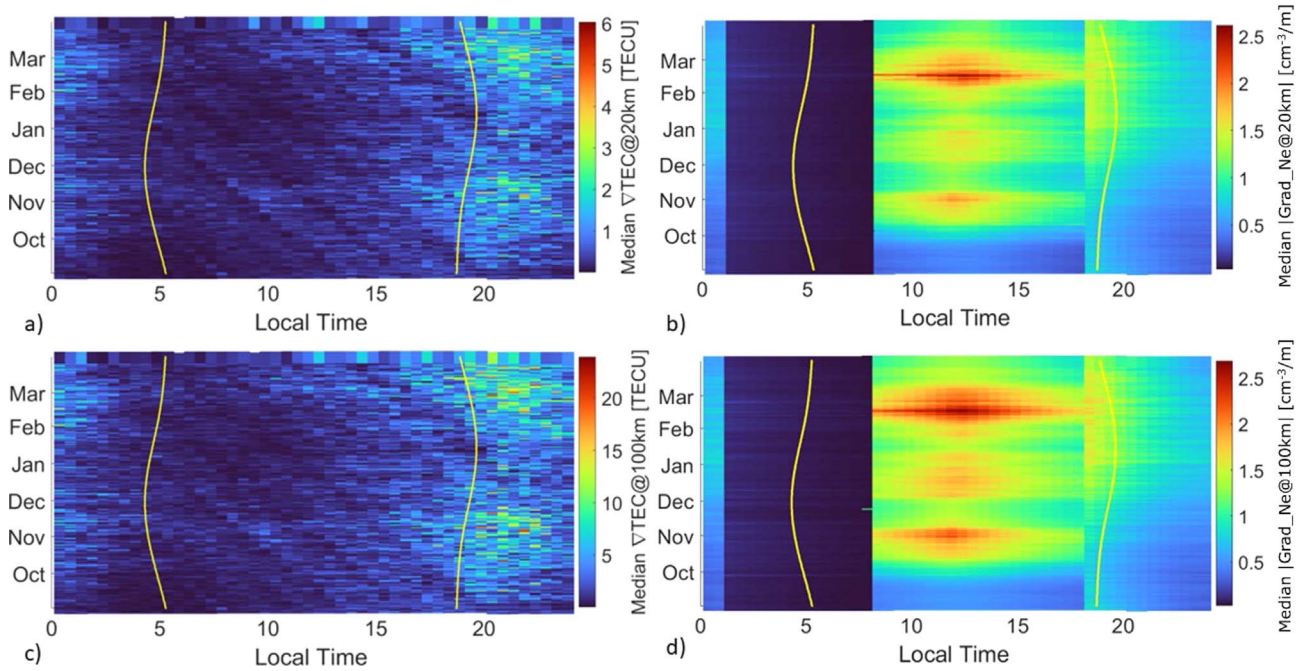


Figure 12. Median of TEC N-S gradient at 20 km (a) and 100 km (c) and of the model $|\text{Grad_Ne@20km}|$ (b) and $|\text{Grad_Ne@100km}|$ (d) from the equatorial model in the geographical sector (54°W , $17^\circ\text{--}20^\circ\text{S}$) as a function of day and LT . The yellow solid lines indicate the sunrise and sunset times evaluated at 120 km. The considered time period is from 1 September 2013 to 30 March 2014.

the sunrise and sunset times evaluated at 120 km altitude. The considered period is from 1 September 2013 to 30 March 2014. Similarly to what was stated in Figure 11, the seasonal conditions during the period from 1 September 2013 to 30 March 2014 (southern hemisphere equinoxes and summer) are such to have the maximum probability of formation of post-sunset equatorial plasma bubbles.

TEC N-S gradients at the considered spatial scales peak after the local post-sunset hours, present a day-to-day variability, and favour the equinoctial months. This suggests that those gradients are related to the presence of EPBs. The models peak at local noon after 18:00 LT (i.e. in correspondence with the post-sunset hours) and values are enhanced until local midnight. The peak at local noon in the models maximizes during local winter and equinoxes, favouring the March equinox for both 20 km and 100 km gradients. The increase after 18:00 LT in the models has a maximum in the March equinox. Also, the TEC data reports a peak in the post-sunset hours that is more pronounced in the March equinox. The peak at noon is completely absent in the data. This is likely because, around noon and in such a narrow geographical sector, TEC is mainly dominated by the EIA crest, which covers a large area. This may result in very little meridional gradients when considering an integrated quantity like TEC . Conversely, a larger variability is expected when considering that the model addresses just the topside sector of the ionosphere.

To check how this climatological behaviour translates into specific cases, Figure 13 shows the time series of the median TEC N-S gradient at 20 km (bottom), 50 km (middle) and 100 km (top) and of the corresponding Ne gradient medians from the equatorial model in the geographical sector (47°W , $17^\circ\text{--}20^\circ\text{S}$). Shaded areas indicate the ± 1 sigma spread around the median. The selected period refers to 21–27 January

2014. As in the case of the relation between the peaks in $|\text{ROT}|$ presented in Figure 11, the $|\text{Grad_Ne@XXkm}|$ models show a shift in the time of occurrence of the night-time $|\text{Grad_Ne@XXkm}|$ maximum which occurs three hours earlier than the onset of EPBs. These features are present in the climatology of evaluated Swarm Ne gradients data on which the model is based. This is, as stated, in contradiction with the scintillation/EPB occurrence measurements not only on a day-to-day basis but also climatologically (offset by a few hours). Nevertheless, whenever investigating the cross-correlation (not shown) of the complete statistics of the median values of both $|\nabla \text{TEC@XXkm}|$ and $|\text{Grad_Ne@XXkm}|$, the larger correlation is found in correspondence with the zero lag. At a 3-hour lag, it presents a 15% decrease, dropping below 50% for lags exceeding 10 h. Additionally, there was another diurnal peak at ± 24 h, similar in magnitude to the one observed at a 3-hour lag.

An additional reason for the differences between the model and the observations in this sector is likely due to the presence of the South Atlantic Anomaly (Aruliah, personal communication). In the model development, no measure of longitude (geographic or geomagnetic) was trialled as an explanatory variable due to the characteristics of the Swarm orbit. During a year, Swarm samples all LT and longitude sectors. However, it only samples a given LT sector in a given longitude sector once every 131 days, which corresponds to two or three intervals per year. It is not feasible to trial both LT and longitude using a dataset that spans 2 years, and it is not currently feasible to extend this dataset without compromising the ability of the model to consider times of higher solar activity. Therefore, the behaviour of the equatorial ionosphere was considered on average, whereas it would be advantageous to model the region around the South Atlantic Anomaly separately. As the Swarm mission

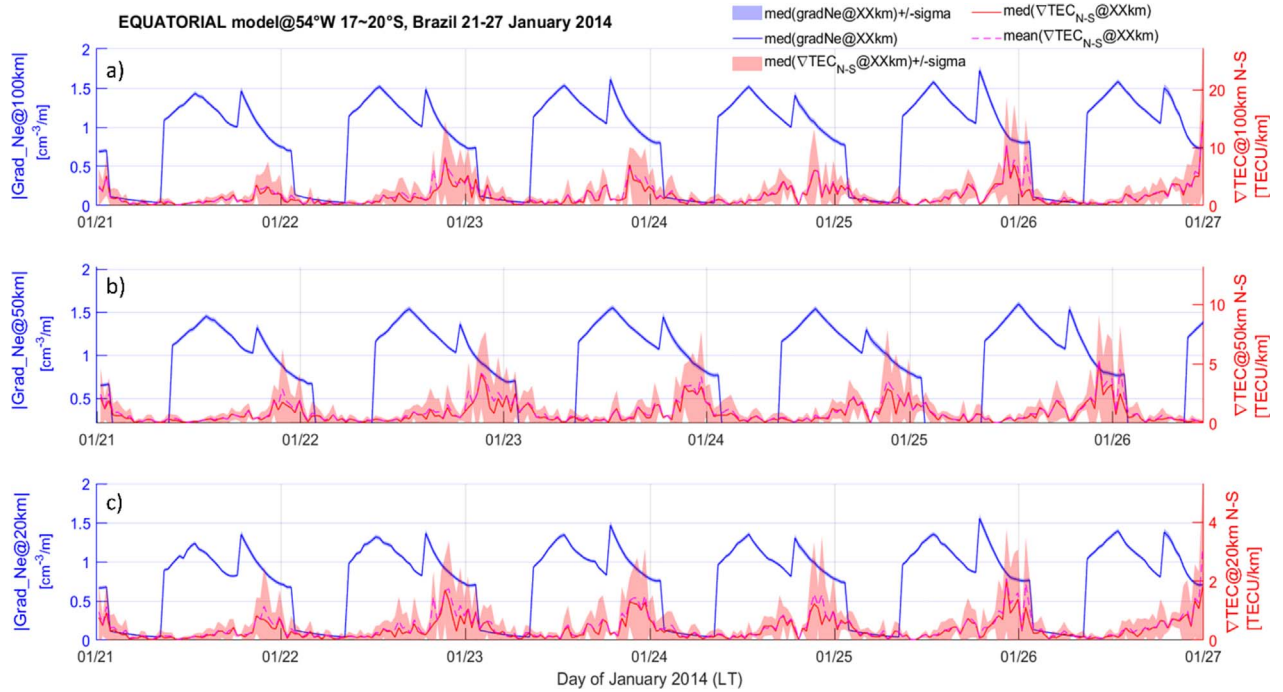


Figure 13. Time series of the median of the *TEC* N-S gradient at 20 km (bottom), 50 km (middle) and 100 km (top) and the corresponding *Ne* gradient medians from three equatorial models in the geographical sector (47°W, 17–28°S). Shaded areas indicate the ± 1 sigma spread around the median. The selected period refers to 21–27 January 2014.

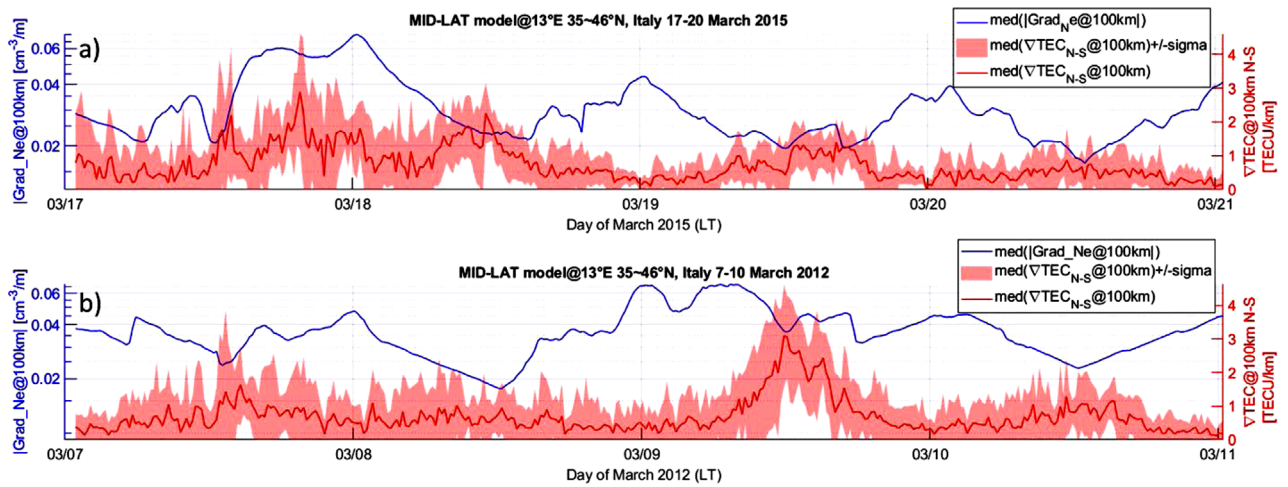


Figure 14. Time series of median of the *TEC* N-S gradient at 100 km and of the corresponding *Ne* gradient medians from mid-latitude model in the geographical sector (13°E, 35–46°N). Shaded areas indicate the ± 1 sigma spread around the median. Periods refer to 17–20 March 2015 (a) and 7–10 March 2012 (b).

continues during solar cycle 25, the additional observations will make this a feasible proposition.

To check the behaviour of the mid-latitudes models, we also evaluated the performance of the $|Grad_Ne@100km|$ against *TEC* N-S gradients over Italy (as described in Sect. 3.2). The expression of such a model is again in Table S2 of Paper 1. Figure 14 shows the time series of the median of the *TEC* N-S gradient at 100 km and the corresponding *Ne* gradient median values from the mid-latitude model in the geographical sector (13°E, 35–46°N). Shaded areas indicate the ± 1 sigma spread

around the median. We report only the model for $|Grad_Ne@100km|$ as no significant differences with the other *Ne* gradient models are observed for mid-latitudes models. The selected periods refer to 17–20 March 2015 (Fig. 14a) and 7–10 March 2012 (Fig. 14b), which are both characterized by the presence of a geomagnetic storm. The former is the widely studied 2015 St. Patrick's Day Storm which produced different types of Large-Scale Traveling Ionospheric Disturbances (LSTIDs) in the European sector (Borries et al., 2016; Zakharenkova et al., 2016); the latter has been studied by

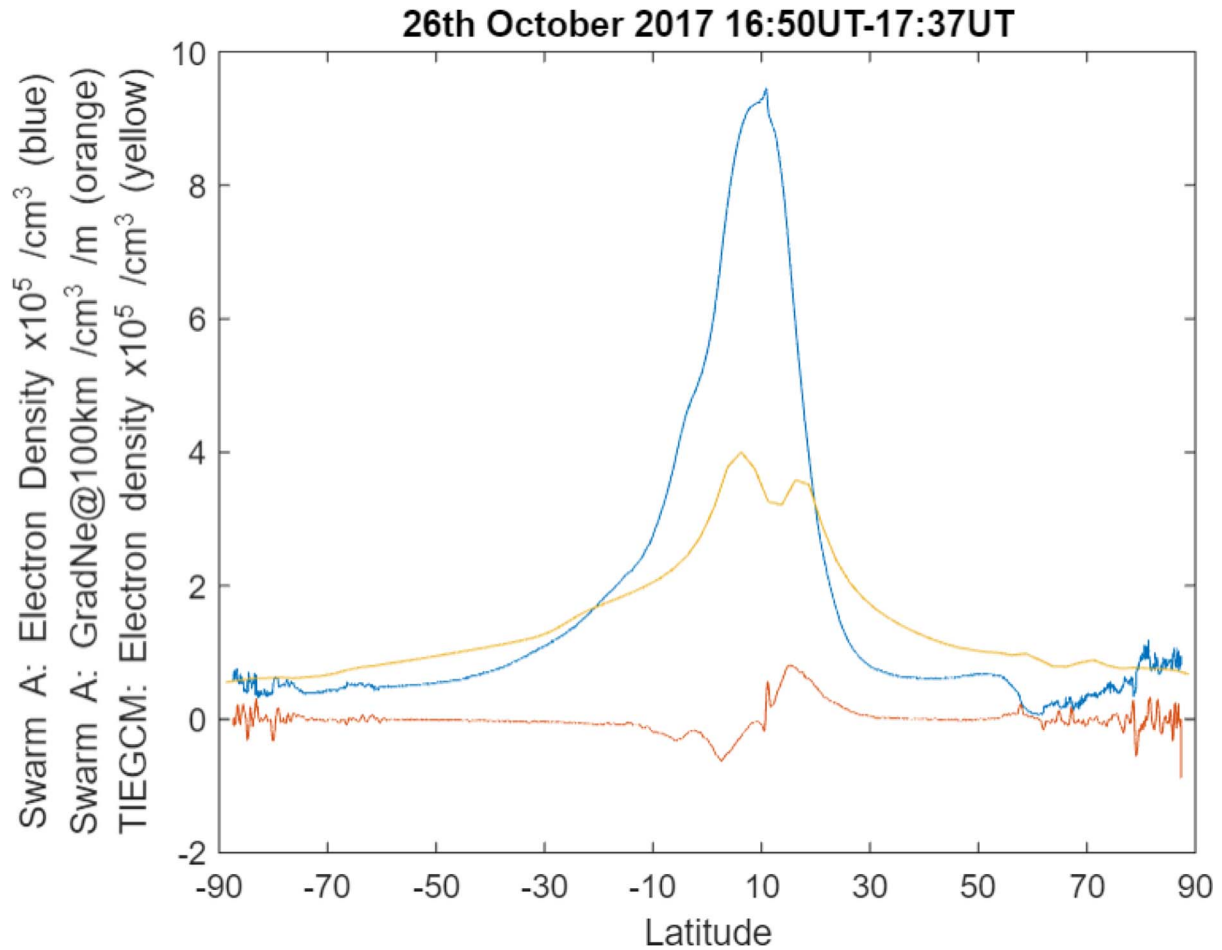


Figure 15. The electron density (N_e) and $|Grad_Ne@100km|$ observed by Swarm A with the electron density (N_e) estimated by TIE-GCM along the Swarm A orbital path for 26th October 2017 16:50 UT–17:37 UT.

Belehaki et al. (2017), reporting again LSTIDs activity over Europe linked to auroral electrojet intensification following the storm. Concerning the case of the St. Patrick's Day storm case (Fig. 14a), concurrent enhancements of the investigated quantities are found in the first phase of the storm (17 to the half of 18 March). To evaluate the ability of the mid-latitude models to account for the effect of a partial solar eclipse, the period includes 20 March 2015 between 09:00 and 10:00 UT during which such an eclipse occurred over Europe. The ionospheric effects induced by the eclipse in the European sector (including Italy) are reported in the literature (see, e.g. Pezzopane et al., 2015; Verhulst et al., 2016; Stankov et al., 2017). The effect of the solar eclipse has been modelled by setting for the respective hour the value of the F10.7 parameter to zero to feed the model. As only $F107_81$ is input to the model, no meaningful differences have been found in the overall behaviour, as expected. On the contrary, in the March 2012 case, there is a midnight-noon inversion of the peak gradient from the model and data, which limits the capability of the model to reproduce the storm effect in terms of the induced spatial gradients. We recall that under the comparison between TEC N-S grad and $|Grad_Ne@XXkm|$ lies a strong assumption, i.e. that enhancements of gradients in TEC can be reflected in gradients at the same spatial scale in the topside ionosphere, and that these

gradients occur simultaneously. Deviations from these assumptions can occur under disturbed conditions, contributing to the possible failure of the model in the March 2012 case event.

4.3 Performance assessment against TIE-GCM

In this section, the modelled Ne is compared against the TIE-GCM output. No other modelled parameters are compared since they are below the spatial resolution of TIE-GCM. An example of a comparison between Swarm A observations and TIE-GCM output is shown in Figure 15. This illustrates a few differences which were repeatedly seen in the wider analysis. In general, when TIE-GCM is compared with Swarm data, it can depict enhanced electron densities at equatorial latitudes close to local noon, persisting into the afternoon and evening, as expected from the diurnal variation, however, there can be significant differences between the values predicted by TIE-GCM and observed by Swarm. There are times when TIE-GCM represents the Swarm observations well, such as when the solar zenith angle is low. However, when the ionosphere is variable, TIE-GCM does not always accurately capture that variability, for example at the high latitudes in Figure 15. Even when conditions are quiet, TIE-GCM does not always

Table 2. Goodness of fit statistics compared between the Swarm-VIP models and TIE-GCM for a series of case studies in 2017. The calculation of these goodness-of-fit statistics is shown in [Paper 1](#).

Model	Region	Goodness-of-fit statistics				
		rRMSE	MSA	rME	Precision	Correlation
Swarm-VIP models	Polar	0.16	110	0.02	0.55	0.65
	Auroral	0.15	110	0.01	0.70	0.62
	Mid	0.16	111	0.00	0.76	0.48
	Equatorial	0.14	110	−0.01	0.72	0.43
TIE-GCM	Polar	0.23	112	−0.10	0.57	0.36
	Auroral	0.22	113	−0.12	0.61	0.36
	Mid	0.18	112	0.01	0.72	0.32
	Equatorial	0.16	112	0.11	1.03	0.67

capture the large-scale structure of the ionosphere as shown in the additional figures in the [Supplementary material](#).

The TIE-GCM and Swarm-VIP models are compared against Swarm observations for four weeks long case studies in a longitude sector that includes the European region (0°E–15°E longitude): 4–10 September 2017, 26 April – 2 May 2017, 13–20 August 2017, and 23–29 October 2017, with these intervals chosen to cover a range of *LT* sectors and geomagnetic conditions. The 13–20 August 2017 interval is chosen as these orbits sampled the noon and midnight *LT* sectors (approximately aligned on the 00:00–12:00 *LT* axis). The 23–29 is chosen as these orbits sampled the dawn and dusk *LT* sectors (approximately aligned on the 06:00–18:00 *LT* axis). There was a period of enhanced geomagnetic activity between 4 and 10 September 2017, and so these dates were selected. During this interval, Swarm A sampled *LT* approximately aligned with the 10:00–22:00 *LT* axis. Observations from 26 April–2 May 2017 are also selected, as the same *LT* sectors are sampled, but under quiet geomagnetic conditions. To enable a like-for-like comparison between the Swarm-VIP models and TIE-GCM, the goodness of fit statistics discussed against Swarm plasma density observations is recalculated for the Swarm-VIP models for this interval.

Ideally, dates from the evaluation dataset (detailed in [Paper 1](#)) would be used for this comparison. However, the optimization and evaluation dataset is split by DOY, with odd DOYs used for optimization and even DOYs used for evaluation. This is necessary to ensure that a suitable range of heliogeophysical conditions are present in both the optimisation and evaluation datasets for model development and testing. Unfortunately, it also meant that, if dates from the evaluation database were used for the comparison with TIE-GCM then, if 7 days of data were required, this would be drawn from a period of 13 days. In this interval, the *LT* of the orbit would vary by 2.4 h. To minimise these variations in *LT* while using 7 days of data, a continuous interval of observations is required.

As described in [Paper 1](#), the models are originally fitted using data from the training dataset. Then the model parameters are re-estimated using data from the optimization dataset. Therefore, the training dataset is used to determine what parameters are chosen, but not their estimates. To enable the comparison with TIE-GCM, dates from the training dataset are selected for this statistical comparison. This ensures that the comparison is based on four intervals covering a range of geomagnetic activities and *LT*s but where the variation with *LT* within each interval is minimised. However, it does mean that the

goodness-of-fit statistics for the Swarm-VIP models are slightly better than they would be if the models were evaluated against an independent dataset. For this reason, the comparison is intended to be illustrative, focusing on 4 weeks of data at a specific longitude sector and a range of *LT*s.

The databases used for the Swarm-VIP models include a total of 17,339 independent data points which are used for these tests. In order to enable a like-for-like comparison between the Swarm-VIP models and TIE-GCM, the four key measures of the goodness-of-fit of the model predictions to the data already used in [Paper 1](#) are used in the present study. These are the relative Root Mean Square Error (rRMSE) and Median Symmetric Accuracy (MSA) for the accuracy; the Mean Error (ME) for the bias; the precision and the Pearson Linear Correlation Coefficient for the association. All those goodness of fit statistics are introduced in [Liemohn et al. \(2021\)](#) and are reported in equations 15–19 of [Paper 1](#). These parameters are calculated for the Swarm-VIP models for this interval. These re-calculated values are shown in [Table 2](#) and are broadly similar to those shown in [Table 3](#) of [Paper 1](#). The similarity of these goodness-of-fit statistics to those shown in [Paper 1](#) gives confidence that the improvement to the goodness-of-fit statistics which results from evaluating the Swarm-VIP models against the training dataset is minor.

The model predictions from TIE-GCM at the closest grid point on a latitude-longitude-universal time grid to the location of the observation made by Swarm are selected and the same goodness-of-fit statistics are calculated. The modelled electron density is re-scaled to the mean altitude of Swarm A in this interval. Within each latitudinal region, goodness-of-fit statistics from the Swarm-VIP models and TIE-GCM can be compared. The most striking difference in the goodness-of-fit statistics between the models is seen in the correlations. The Swarm-VIP models show a moderate improvement over TIE-GCM in the polar, auroral and mid-latitude sectors, while TIE-GCM shows a moderate improvement over the Swarm-VIP model in the equatorial sector. The higher correlations indicate that more of the trend in the observations is captured by the model. The model with the higher correlation in each latitudinal region also has a slightly better accuracy, as measured by the rRMSE and the MSA. The precision is closer to 1 for the Swarm-VIP models than TIE-GCM in the auroral and mid-latitude regions, indicating that the variability in the model and observations is more similar for the Swarm-VIP models than for TIE-GCM. In the polar and equatorial regions, the precision is better (closer to 1) for TIE-GCM. Substantial biases are seen for TIE-GCM in

the auroral and equatorial regions, but not for any other model in any other region. Bias in physical models is relatively common (i.e. Elvidge et al., 2023), raising the possibility that the Swarm-VIP models could be used to calibrate the physical models. This possibility will be explored in a subsequent paper which will also conduct a more detailed comparison of TIE-GCM and the Swarm-VIP models.

A possible explanation for the differences in the goodness-of-fit statistics between the Swarm-VIP models and TIE-GCM is that a statistical model can respond more quickly to changes in the driving conditions, and the Swarm-VIP models are more heavily driven by such terms away from the equatorial region. In addition, the Swarm-VIP models can model not just N_e , but proxies for structures below the TIE-GCM grid size. However, TIE-GCM does have another significant advantage over the Swarm-VIP models. Since it is a physics-based model, it has the potential to better contribute to our physical understanding of the system.

5 Discussion and conclusions

In the frame of the Swarm-VIP project, 20 models have been developed to predict electron density, its gradients at three spatial scales – 20, 50, and 100 km – and the strength of the density of the fluctuations (through $IPIR_{ix}$) in the polar cap, auroral, mid-latitude and equatorial regions. The paper leverages on the formulation, optimisation and validation of the models provided in the companion paper (Paper 1) for identifying performance strengths of the models. This translates into addressing three different aspects: (i) the ability of the model to reproduce the climatological variability of the considered quantities through the direct comparison with Swarm data; (ii) the ability to reproduce ionospheric weather as depicted by ground-based GNSS and serve as a proxy for the ionospheric effect on GNSS signals; and (iii) the comparison with the physics-based modelling provided by the TIE-GCM. Selected results are discussed and detailed in this paper and Paper 1.

Concerning (i), the overall climatological features are generally well reproduced by the models. For the polar models, the main discrepancies are found to be related to the strong control of the models by $F107_{81}$ and the lack of inter-hemispheric asymmetry in the models. This impact can be mitigated by training models using separate datasets from each hemisphere. The most significant discrepancy between the mid-latitude data and the model is that the model significantly underestimates $|Grad_Ne@100km|$ by a factor that ranges from 0.1 to 0.9. Other discrepancies can be attributed to the fact that the model uses a sine function with a period of one year, while the Swarm data show seasonal variations with two peaks. The equatorial models have two formulations: the first one uses the GLM approach independently on the LT , while the second applies it in three LT sectors to better reproduce the effect of the post-sunset equatorial plasma bubbles. The three LT -dependent equatorial models differ significantly from the Swarm and GNSS measurements in terms of the time difference between the daily peaks of the $|Grad_Ne@100km|$ models and the corresponding gradients derived from Swarm N_e measurements and GNSS TEC observations.

Concerning (ii), in general, the high-latitude models can represent enhancements of the $IPIR_{ix}$ and $|Grad_Ne@XXkm|$

quantities that sometimes match enhancements of σ_ϕ and $|ROT|$ during the various phases of the September 2017 storm. This is more evident for the auroral model, which seems to be the one able to represent the considered ISMR data, which covers both the auroral and polar cap sectors. We remind the reader that the Swarm-VIP models can reproduce medium-scale irregularities, i.e. no matching with measured amplitude scintillation events is expected. This is somehow expected as those irregularities are most likely to form within the auroral oval, at the edges of auroral forms (see, e.g. Moen et al., 2013; Enengl et al., 2023; and references therein). For the equatorial LT -dependent models of the $|Grad_Ne@XXkm|$, signatures of the post-sunset equatorial plasma bubbles are represented by the models, but at times up to three hours earlier than the occurrence of the EPBs as observed in the σ_ϕ and $|ROT|$ measurements. However, due to the climatological nature of the models, the day-to-day variability of the bubble formation remains mostly unmodelled (Li et al., 2021). Additionally, a peak at noon of the model (favouring local winter and equinoxes) is reported and found not to be related to the TEC meridional gradients in the considered narrow area under the southern crest of the equatorial ionisation anomaly. This can be ascribed to a larger variability of the topside ionosphere than the one obtained while considering an integrated quantity like TEC over a narrow ionospheric sector.

Concerning (iii), the Swarm-VIP models of electron density show an improvement over TIE-GCM in the polar, auroral and mid-latitude sectors, but not in the equatorial sector. One possible explanation for this could be the greater dependency of polar, auroral, and mid-latitude models on driving conditions in contrast to equatorial models. Consequently, they may display more effective reactions to geospace forcing than the equatorial model, which reacts less to such forcing, ultimately leading to outperforming the TIE-GCM. It may also be that the influence of the thermosphere on the ionosphere is stronger at these latitudes and this will also be explored. Another possibility is likely to be the presence of the South Atlantic Anomaly. Ideally, a measure of longitude should be trialled as an explanatory variable, or a separate model should be developed for this region. However, the characteristics of the Swarm orbit (sampling a given LT sector in each longitude sector only once every 131 days) means that it is not currently feasible to do this without compromising the ability of the model to consider times of higher solar activity. As the Swarm mission continues during solar cycle 25, the additional observations will make this a feasible proposition. Substantial biases are seen for TIE-GCM in the auroral and equatorial regions. Biases in physical models are relatively common (i.e. Elvidge et al., 2023), raising the possibility that the Swarm-VIP models could be used to calibrate the physical models. This possibility will be explored in a subsequent paper.

A perfect agreement between the models and the data is neither expected nor observed. These models of the plasma structures are deterministic. However, there are also random variations in the ionospheric structures which cannot be captured by these models. The explanatory variables are proxies for the driving processes. They approximate these processes, rather than exactly replicating them, and this will cause a discrepancy. Nevertheless, the Swarm-VIP models successfully represent a wide variety of ionospheric features. The comparisons to other

datasets and models in this paper can be used to identify potentially fruitful areas for further model development. The most important of these is to trial thermospheric density products at temporal/spatial resolution that are similar to the ionospheric data products. Applying different lags to the solar wind data products and between the *DOY_{fn}* and the *|Grad_{Ne}@XX_{km}|* models may also be a fruitful avenue for future research. We plan to explore these avenues in a future publication.

In summary, our effort shows that the developed climatological models for the ionospheric plasma density reflect the instantaneous values reasonably well. Modelling the plasma density gradients, however, is more of a challenge. The basic features of the topside electron density are captured by our models, but, since the models are primarily climatological, most of the instantaneous variability is not captured. Future developments of the Swarm-VIP models are expected to improve their performance. Specifically, the use of the thermospheric density and the ionospheric current systems has been proven in Paper 1 to be effective in improving the models. However, the Swarm data products directly providing this information cannot be used for models aimed at being independent of the Swarm tracks that generated them. Exogenous data sources would be needed, specifically, thermospheric density at high spatial/temporal resolution, as well as temperature and/or velocity data or even indications of ionospheric currents would be of interest.

As a concluding remark, in general, physics-based models, like TIE-GCM, have the advantage of contributing to our physical understanding of the system with respect to empirical, statistical models like the presented ones. However, the Swarm-VIP models can model not only the electron density, which is the physical observable given by physics-based models, but also provide proxies for ionospheric irregularities impacting the technological systems, for example, global navigation satellite systems.

Acknowledgements

This work is within the framework of the Swarm Variability of Ionospheric Plasma (Swarm-VIP) project, funded by ESA in the “Swarm+4D-Ionosphere” framework (ESA Contract No. 4000130562/20/I-DT). YJ, DK, and WJM acknowledge funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (ERC Consolidator Grant agreement No. 866357, POLAR-4DSpace). We thank the Italian National Research Council (CNR) for hosting the GNSS receivers in “Dirigibile Italia” station at Ny-Ålesund, the Kartverket for hosting the GNSS receiver in Ny-Ålesund, the University of Tromsø for hosting the GNSS receiver in Longyearbyen and SANSA Space Science for hosting the GNSS receiver at SANAE-IV station. We also thank PNRA (Programma Nazionale di Ricerche in Antartide) for supporting the GNSS monitoring at Mario Zucchelli, Concordia and SANAE IV stations and FACET-UNT/CONICET for supporting the GNSS monitoring at San Miguel de Tucumán. We are grateful to GRAPE (GNSS Research and Application for Polar Environment), the Expert Group of the Scientific Committee on Antarctic Research (SCAR) for supporting this study. The editor thanks two anonymous reviewers for their assistance in evaluating this paper.

Data availability statement

GNSS data from the Italian “Rete Integrata GNSS” are available at the INGV RING website: <http://ring.gm.ingv.it>. Data from GNSS Ionospheric Scintillation Monitor Receivers are available at the electronic Space Weather Upper Atmosphere (eSWUA) data portal (<http://www.eswua.ingv.it/>), which operated by the Upper Atmosphere Physics and Radio propagation group of INGV. GNSS data from the Rede Brasileira de Monitoramento Contínuo dos Sistemas GNSS (RBMC), managed by the Instituto Brasileiro de Geografia e

Estatística (IBGE), are available at https://geofit.ibge.gov.br/informacoes_sobre_posicionamento_geodesico/rbmc/. The Swarm data can be accessed through: <ftp://swarm-diss.esa.int>. Explanatory variables used to input the model are provided by: Space Physics Data Facility (SPDF, <https://spdf.gsfc.nasa.gov/pub/data/omni/>); Laboratory for Atmospheric and Space Physics at the University of Colorado (<https://lasp.colorado.edu/home/>); International Service of Geomagnetic Indices (<http://isgi.unistra.fr/>). TIE-GCM online runs can be obtained, at NASA’s Community Coordinated Modelling Centre (CCMC) at https://ccmc.gsfc.nasa.gov/requests/IT/TIE-GCM/tiegcm_user_registration.php.

Supplementary material

The supplementary material of this article is available at <https://www.swsc-journal.org/10.1051/swsc/2024003/olm>.

Figure S1: The electron density and *GradNe@100km* observed by Swarm A with the electron density estimated by TIE-GCM along the Swarm A orbital path for 29th April 2017 09:18 UT–10:05 UT.

Figure S2: As Figure S1 but for 29th April 2017 20:59 UT–21:46 UT.

Figure S3: As Figure S1 but for 16th August 2017 00:16 UT–01:03 UT.

Figure S4: As Figure S1 but for 16th August 2017 11:57 UT–12:44 UT.

Figure S5: As Figure S1 but for 26th October 2017 05:09 UT–05:56 UT.

Figure S6: As Figure S1 but for 26th October 2017 16:50 UT–17:37 UT.

Figure S7: As Figure S1 but for 7th September 2017 10:17 UT–11:03 UT.

Figure S8: As Figure S1 but for 7th September 2017 21:57 UT–22:44 UT.

References

- Alfonsi L, Cesaroni C, Spogli L, Regi M, Paul A, et al. 2021. Ionospheric disturbances over the Indian sector during 8 September 2017 geomagnetic storm: Plasma structuring and propagation. *Space Weather* **19**(3): e2020SW002607. <https://doi.org/10.1029/2020SW002607>.
- Alfonsi L, Wernik AW, Materassi M, Spogli L. 2017. Modelling ionospheric scintillation under the crest of the equatorial anomaly. *Adv Space Res* **60**(8): 1698–1707. <https://doi.org/10.1016/j.asr.2017.05.021>.
- Belehaki A, Kutiev I, Marinov P, Tsagouri I, Koutroumbas K, Elias P. 2017. Ionospheric electron density perturbations during the 7–10 March 2012 geomagnetic storm period. *Adv Space Res* **59**(4): 1041–1056. <https://doi.org/10.1016/j.asr.2016.11.031>.
- Bilitza D, Pezzopane M, Truhlik V, Altadill D, Reinisch BW, Pignalberi A. 2022. The International Reference Ionosphere model: A review and description of an ionospheric benchmark. *Rev Geophys* **60**(4): e2022RG000792. <https://doi.org/10.1029/2022RG000792>.
- Borries C, Mahrous AM, Ellahouny NM, Badeke R. 2016. Multiple ionospheric perturbations during the Saint Patrick’s Day storm 2015 in the European-African sector. *J Geophys Res Space Phys* **121**(11): 11–333. <https://doi.org/10.1002/2016JA023178>.
- Bougard B, Sleewaegen J-M, Spogli L, Veetil SV, Monico JF. 2011. CIGALA: challenging the solar maximum in Brazil with PolaRxS. In: Proceedings of the ION GNSS 2011, Institute of Navigation, Portland, OR, September 20–23, pp. 2572–2579. Available at <http://hdl.handle.net/11449/73031>.
- Cesaroni C, Spogli L, De Franceschi G. 2021. IONORING: Real-time monitoring of the total electron content over Italy. *Remote Sens* **13**(16): 3290. <https://doi.org/10.3390/rs13163290>.
- Cesaroni C, Spogli L, Alfonsi L, De Franceschi G, Ciraolo L, et al. 2015. L-band scintillations and calibrated total electron content gradients over Brazil during the last solar maximum. *J Space Weather Space Clim* **5**: A36. <https://doi.org/10.1051/swsc/2015038>.
- Ciraolo L, Azpilicueta F, Brunini C, Meza A, Radicella SM. 2007. Calibration errors on experimental slant total electron content (TEC) determined with GPS. *J Geod* **81**(2): 111–120. <https://doi.org/10.1007/s00190-006-0093-1>.
- D’Angelo G, Piersanti M, Pignalberi A, Coco I, De Michelis P, et al. 2021. Investigation of the physical processes involved in GNSS amplitude scintillations at high latitude: a case study. *Remote Sens* **13**(13): 2493. <https://doi.org/10.3390/rs13132493>.
- D’Angelo G, Piersanti M, Alfonsi L, Spogli L, Clausen LBN, et al. 2018. The response of high latitude ionosphere to the 2015 St. Patrick’s day storm from in situ and ground based observations. *Adv Space Res* **62**(3): 638–650. <https://doi.org/10.1016/j.asr.2018.05.005>.

- Dang T, Zhang B, Lei J, Wang W, Burns A, et al. 2021. Azimuthal averaging–reconstruction filtering techniques for finite-difference general circulation models in spherical geometry. *Geosci Model Dev* **14**(2): 859–873. <https://doi.org/10.5194/gmd-14-859-2021>.
- De Franceschi G, Spogli L, Alfonsi L, Romano V, Cesaroni C, Hunstad I. 2019. The ionospheric irregularities climatology over Svalbard from solar cycle 23. *Sci Rep* **9**(1): 1–14. <https://doi.org/10.1038/s41598-019-44829-5>.
- de Paula ER, Martinon AR, Carrano C, Moraes AO, Neri JA, et al. 2022. Solar flare and radio burst effects on GNSS signals and the ionosphere during September 2017. *Radio Sci* **57**(10): 1–15. <https://doi.org/10.1029/2021RS007418>.
- Dorrian GD, Wood AG, Ronskley A, Aruliah A, Shahtahmassebi G. 2019. Statistical modeling of the coupled F-region ionosphere-thermosphere at high latitude during polar darkness. *J Geophys Res Space Phys* **124**(2): 1389–1409. <https://doi.org/10.1029/2018JA026171>.
- Dickinson RE, Ridley EC, Roble RG. 1981. A three-dimensional general circulation model of the thermosphere. *J Geophys Res Space Phys* **86**(A3): 1499–1512. <https://doi.org/10.1029/JA086iA03p01499>.
- Elvidge S, Granados SR, Angling MJ, Brown MK, Themens DR, Wood AG. 2023. Multi-model ensembles for upper atmosphere models. *Space Weather* **21**(3): e2022SW003356. <https://doi.org/10.1029/2022SW003356>.
- Elvidge S, Angling MJ. 2019. Using the local ensemble transform Kalman filter for upper atmospheric modelling. *J Space Weather Space Clim* **9**: A30. <https://doi.org/10.1051/swsc/2019018>.
- Enengl F, Kotova D, Jin Y, Clausen LB, Miloch WJ. 2023. Ionospheric plasma structuring in relation to auroral particle precipitation. *J Space Weather Space Clim* **13**: 1. <https://doi.org/10.1051/swsc/2022038>.
- Fremouw EJ, Leadabrand RL, Livingston RC, Cousins MD, Rino CL, Fair BC, Long RA. 1978. Early results from the DNA Wideband satellite experiment – complex-signal scintillation. *Radio Sci* **13**(1): 167–187. <https://doi.org/10.1029/RS013i001p00167>.
- Friis-Christensen E, Lühr H, Knudsen D, Haagmans R. 2008. Swarm – an Earth observation mission investigating geospace. *Adv Space Res* **41**(1): 210–216. <https://doi.org/10.1016/j.asr.2006.10.008>.
- Ghobadi H, Spogli L, Alfonsi L, Cesaroni C, Cicone A, Linty N, Romano V, Cafaro M. 2020. Disentangling ionospheric refraction and diffraction effects in GNSS raw phase through fast iterative filtering technique. *GPS Solut* **24**: 85. <https://doi.org/10.1007/s10291-020-01001-1>.
- Hagan ME, Forbes JM, Vial F. 1995. On modeling migrating solar tides. *Geophys Res Lett* **22**(8): 893–896. <https://doi.org/10.1029/95GL00783>.
- INGV RING Working Group. 2016. *Rete Integrata Nazionale GPS (RING) [Data set]*. Istituto Nazionale di Geofisica e Vulcanologia (INGV). <https://doi.org/10.13127/ring>.
- Jin Y, Xiong C. 2020. Interhemispheric asymmetry of large-scale electron density gradients in the polar cap ionosphere: UT and seasonal variations. *J Geophys Res Space Phys* **125**: e2019JA027601. <https://doi.org/10.1029/2019JA027601>.
- Jin Y, Xiong C, Clausen L, Spicher A, Kotova D, et al. 2020. Ionospheric plasma irregularities based on in situ measurements from the Swarm satellites. *J Geophys Res Space Phys* **125**(7): e2020JA028103. <https://doi.org/10.1029/2020JA028103>.
- Jin Y, Kotova D, Xiong C, Brask SM, Clausen LB, et al. 2022. Ionospheric plasma irregularities – IPIR – data product based on data from the Swarm satellites. *J Geophys Res Space Phys* **127**(4): e2021JA030183. <https://doi.org/10.1029/2021JA030183>.
- Kamal S, Jakowski N, Hoque MM, Wickert J. 2021. A high latitude model for the E Layer dominated ionosphere. *Remote Sens* **13**(18): 3769. <https://doi.org/10.3390/rs13183769>.
- Kotova D, Jin Y, Spogli L, Wood AG, Urbar J, et al. 2023. Electron density fluctuations from Swarm as a proxy for ground-based scintillation data: a statistical perspective. *Adv Space Res* **72**: 5399–5415. <https://doi.org/10.1016/j.asr.2022.11.042>.
- Li G, Ning B, Otsuka Y, Abdu MA, Abadi P, et al. 2021. Challenges to equatorial plasma bubble and ionospheric scintillation short-term forecasting and future aspects in east and southeast Asia. *Surv Geophys* **42**(1): 201–238. <https://doi.org/10.1007/s10712-020-09613-5>.
- Liemohn MW, Shane AD, Azari AR, Petersen AK, Swiger BM, Mukhopadhyay A. 2021. RMSE is not enough: Guidelines to robust data-model comparisons for magnetospheric physics. *J Atmos Sol Terr Phys* **218**: 105624. <https://doi.org/10.1016/j.jastp.2021.105624>.
- Linty N, Minetto A, Dosis F, Spogli L. 2018. Effects of phase scintillation on the GNSS positioning error during the September 2017 storm at Svalbard. *Space Weather* **16**(9): 1317–1329. <https://doi.org/10.1029/2018SW001940>.
- MacDougall JW. 1969. The equatorial ionospheric anomaly and the equatorial electrojet. *Radio Sci* **4**(9): 805–810. <https://doi.org/10.1029/RS004i009p00805>.
- Macho EP, Correia E, Spogli L, Muella MTAH. 2022. Climatology of ionospheric amplitude scintillation on GNSS signals at south American sector during solar cycle 24. *J Atmos Sol Terr Phys* **231**: 105872. <https://doi.org/10.1016/j.jastp.2022.105872>.
- Materassi M, Forte B, Coster AJ, Skone S, (Eds.). 2019. *The dynamical ionosphere: a systems approach to ionospheric irregularity*. Elsevier, Amsterdam, The Netherlands. <https://doi.org/10.1016/C2017-0-01069-8>.
- Maute A. 2017. Thermosphere-ionosphere-electrodynamics general circulation model for the ionospheric connection explorer: TIEGCM-ICON. *Space Sci Rev* **212**(1): 523–551. <https://doi.org/10.1007/s11214-017-0330-3>.
- McCaffrey AM, Jayachandran PT. 2019. Determination of the refractive contribution to GPS phase “scintillation”. *J Geophys Res Space Phys* **124**(2): 1454–1469. <https://doi.org/10.1029/2018JA025759>.
- Moen J, Oksavik K, Alfonsi L, Daabakk Y, Romano V, Spogli L. 2013. Space weather challenges of the polar cap ionosphere. *J Space Weather Space Clim* **3**: A02. <https://doi.org/10.1051/swsc/2013025>.
- Muella MTAH, Duarte-Silva MH, Moraes AO, de Paula ER, de Rezende LFC, Alfonsi L, Affonso BJ. 2017. Climatology and modeling of ionospheric scintillations and irregularity zonal drifts at the equatorial anomaly crest region. *Ann Geophys* **35**: 1201–1218. <https://doi.org/10.5194/angeo-35-1201-2017>.
- Nava B, Coisson P, Amarante GM, Azpilicueta F, Radicella SM. 2005. A model assisted ionospheric electron density reconstruction method based on vertical TEC data ingestion. *Ann Geophys* **48**(2): 313–320. <http://hdl.handle.net/2122/905>.
- Nava B, Coisson P, Radicella SM. 2008. A new version of the NeQuick ionosphere electron density model. *J Atmos Sol Terr Phys* **70**(15): 1856–1862. <https://doi.org/10.1016/j.jastp.2008.01.015>.
- Pezzopane M, Pignalberi A. 2019. The ESA Swarm mission to help ionospheric modeling: a new NeQuick topside formulation for mid-latitude regions. *Sci Rep* **9**: 12253. <https://doi.org/10.1038/s41598-019-48440-6>.
- Pezzopane M, Pietrella M, Pignalberi A, Tozzi R. 2015. 20 March 2015 solar eclipse influence on sporadic E layer. *Adv Space Res* **56**(10): 2064–2072. <https://doi.org/10.1016/j.asr.2015.08.001>.
- Piersanti M, Alberti T, Bemporad A, Berrilli F, Bruno R, et al. 2017. Comprehensive analysis of the geoeffective solar event of 21 June 2015: Effects on the magnetosphere, plasmasphere, and ionosphere systems. *Solar Phys* **292**(11): 1–56. <https://doi.org/10.1007/s11207-017-1186-0>.
- Prikryl P, Spogli L, Jayachandran PT, Kinrade J, Mitchell CN, et al. 2011. Interhemispheric comparison of GPS phase scintillation at high latitudes during the magnetic-cloud-induced geomagnetic storm of 5–7 April 2010. *Ann Geophys* **29**: 2287–2304. <https://doi.org/10.5194/angeo-29-2287-2011>.
- Priyadarshi S. 2015. A review of ionospheric scintillation models. *Surv Geophys* **36**(2): 295–324. <https://doi.org/10.1007/s10712-015-9319-1>.
- Qian L, Burns AG, Emery BA, Foster B, Lu G, et al. 2014. The NCAR TIE-GCM: A community model of the coupled thermosphere/ionosphere system. In: *Modeling the ionosphere – thermosphere system*. Huba J, Schunk R, Khazanov G, (Eds.), John Wiley & Sons, Washington. pp. 73–83. <https://doi.org/10.1002/9781118704417.ch7>.
- Richmond AD, Ridley EC, Roble RG. 1992. A thermosphere/ionosphere general circulation model with coupled electrodynamics. *Geophys Res Lett* **19**(6): 601–604. <https://doi.org/10.1029/92GL00401>.
- Sato H, Jakowski N, Berdermann J, Jiricka K, Heßelbarth A, Banyś D, Wilken V. 2019. Solar radio burst events on 6 September 2017 and its impact on GNSS signal frequencies. *Space Weather* **17**(6): 816–826. <https://doi.org/10.1029/2019SW002198>.
- Spogli L, Alfonsi L, De Franceschi G, Romano V, Aquino MHO, Dodson A. 2009. Climatology of GPS ionospheric scintillations over high and mid-latitude European regions. *Ann Geophys* **27**: 3429–3437. <https://doi.org/10.5194/angeo-27-3429-2009>.
- Spogli L, Ghobadi H, Cicone A, Alfonsi L, Cesaroni C, et al. 2021a. Adaptive phase detrending for gnss scintillation detection: a case study over Antarctica. *IEEE Geosci Remote Sens Lett* **19**: 1–5. <https://doi.org/10.1109/LGRS.2021.3067727>.

- Spogli L, Sabbagh D, Regi M, Cesaroni C, Perrone L, Alfonsi L, et al. 2021b. Ionospheric response over Brazil to the August 2018 geomagnetic storm as probed by CSES-01 and Swarm satellites and by local ground-based observations. *J Geophys Res Space Phys* **126**: e2020JA028368. <https://doi.org/10.1029/2020JA028368>.
- Stankov SM, Bergeot N, Berghmans D, Bolsée D, Bruyninx C, et al. 2017. Multi-instrument observations of the solar eclipse on 20 March 2015 and its effects on the ionosphere over Belgium and Europe. *J Space Weather Space Clim* **7**: A19. <https://doi.org/10.1051/swsc/2017017>.
- Tornatore V, Cesaroni C, Pezzopane M, Alizadeh MM, Schuh H. 2021. Performance evaluation of VTEC GIMs for regional applications during different s. *Remote Sens* **13**(8): 1470. <https://doi.org/10.3390/rs13081470>.
- Upper Atmosphere Physics and Radio Propagation Working Group, Cesaroni C, De Franceschi G, Marocci C, Pica E, Romano V, Spogli L. 2020. *Electronic Space Weather upper atmosphere database (eSWua) – GNSS scintillation data, version 1.0*. Istituto Nazionale di Geofisica e Vulcanologia (INGV). Available at <https://doi.org/10.13127/eswua/gnss>.
- Urbar J, Spogli L, Cicone A, Clausen LB, Jin Y, et al. 2023. Multi-scale response of the high-latitude topside ionosphere to geospace forcing. *Adv Space Res* **72**: 5490–5502. <https://doi.org/10.1016/j.asr.2022.06.045>.
- Van Dierendonck AJ, Klobuchar J, Hua Q. 1993. Ionospheric scintillation monitoring using commercial single frequency C/A code receivers. In: Proceedings of the 6th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1993), Salt Lake City, UT, 22–24 September, pp. 1333–1342.
- Vanlommel P, Van der Linden R. 2017. *STCE newsletter 4 Sep 2017–10 Sep 2017*, Solar-Terrestrial Centre of Excellence (STCE). Available at <http://www.stce.be/newsletter/> (15 September 2017).
- Verhulst TG, Sapundjiev D, Stankov SM. 2016. High-resolution ionospheric observations and modeling over Belgium during the solar eclipse of 20 March 2015 including first results of ionospheric tilt and plasma drift measurements. *Adv Space Res* **57**(11): 2407–2419. <https://doi.org/10.1016/j.asr.2016.03.009>.
- Vilà-Valls J, Linty N, Closas P, Dosis F, Curran JT. 2020. Survey on signal processing for GNSS under ionospheric scintillation: Detection, monitoring, and mitigation. *Navig J Inst Navi* **67**(3): 511–535. <https://doi.org/10.1002/navi.379>.
- Wernik AW, Alfonsi L, Materassi M. 2007. Scintillation modeling using in situ data. *Radio Sci* **42**(01): 1–21. <https://doi.org/10.1029/2006RS003512>.
- Wood AG, Alfonsi L, Clausen LB, Jin Y, Spogli L, et al. 2022. Variability of ionospheric plasma: results from the ESA Swarm mission. *Space Sci Rev* **218**(6): 1–44. <https://doi.org/10.1007/s11214-022-00916-0>.
- Wood AG, Donegan-Lawley EE, Clausen LBN, Spogli L, Urbář J, et al. 2024. Statistical models of the variability of plasma in the topside ionosphere: Paper 1: formulation and optimisation. *J. Space Weather Space Clim.* <https://doi.org/10.1051/swsc/2024002>.
- Xiong C, Jiang H, Yan R, Lühr H, Stolle C, Yin F, et al. 2022. Solar flux influence on the in-situ plasma density at topside ionosphere measured by swarm satellites. *J Geophys Res Space Phys* **127**(5): e2022JA030275. <https://doi.org/10.1029/2022JA030275>.
- Zakharenkova I, Astafyeva E, Cherniak I. 2016. GPS and GLONASS observations of large-scale traveling ionospheric disturbances during the 2015 St. Patrick's Day storm. *J Geophys Res Space Phys* **121**(12): 12138–12156. <https://doi.org/10.1002/2016JA023332>.

Cite this article as: Spogli L, Jin Y, Urbář J, Wood AG, Donegan-Lawley EE, et al. 2024. Statistical models of the variability of plasma in the topside ionosphere: 2. Performance assessment. *J. Space Weather Space Clim.* **14**, 4. <https://doi.org/10.1051/swsc/2024003>.