








Calibration of Swarm Plasma Densities Overestimation Using Neural Networks

Artem Smirnov^{1,2} , Yuri Shprits^{1,3,4} , Hermann Lühr¹ , Alessio Pignalberi⁵ , and Chao Xiong⁶ 

¹Helmholtz Centre Potsdam - GFZ German Research Centre for Geosciences, Potsdam, Germany, ²Department of Earth and Environmental Sciences, Ludwig Maximilian University of Munich, Munich, Germany, ³Institute of Physics and Astronomy, University of Potsdam, Potsdam, Germany, ⁴Department of Earth, Planetary and Space Sciences, University of California Los Angeles, Los Angeles, CA, USA, ⁵National Institute of Geophysics and Volcanology (INGV), Rome, Italy, ⁶Department of Space Physics, College of Electronic Information, Wuhan University, Wuhan, China

Key Points:

- We present a neural network-based calibration model that resolves the nighttime overestimation of ion densities by Swarm Langmuir Probes (LPs)
- This overestimation, peaking at 30°QDLat in both hemispheres, mimics the morphology of light ions diffusing downward from the plasmasphere
- The corrected LP ion densities are in excellent agreement with COSMIC radio occultation data

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. Smirnov,
artem.smirnov@gfz-potsdam.de

Citation:

Smirnov, A., Shprits, Y., Lühr, H., Pignalberi, A., & Xiong, C. (2024). Calibration of Swarm plasma densities overestimation using neural networks. *Space Weather*, 22, e2024SW003925. <https://doi.org/10.1029/2024SW003925>

Received 15 MAR 2024

Accepted 4 JUL 2024

Author Contributions:

Conceptualization: Artem Smirnov, Hermann Lühr

Data curation: Artem Smirnov

Formal analysis: Artem Smirnov

Funding acquisition: Yuri Shprits

Methodology: Artem Smirnov, Hermann Lühr, Chao Xiong

Software: Artem Smirnov

Supervision: Yuri Shprits, Hermann Lühr

Validation: Artem Smirnov, Alessio Pignalberi

Visualization: Artem Smirnov

Writing – original draft: Artem Smirnov

Writing – review & editing:

Artem Smirnov, Yuri Shprits, Hermann Lühr, Alessio Pignalberi, Chao Xiong

Abstract Recent studies have shown that the measurements of Langmuir Probes (LPs) onboard ESA's Swarm mission overestimate ion densities on the nightside by up to 50%. The overestimation is due to the assumption of oxygen-only plasma for ion density calculations, which is often violated at mid-latitudes on the nightside. In this study, we present a calibration model that resolves the nighttime overestimation by Swarm LPs. Using observations by Swarm FacePlate (FP) as a reference, we develop a neural network (NN) model that adjusts LP data to the FP measurements. The model incorporates dependence on solar and geomagnetic conditions, parameterized by the P10.7 and Hp30 indices, location, day of the year and local time. Our model reveals a distinct double-crest pattern in nighttime density overestimation by LPs, centered at ~30° quasi-dipole latitude in both hemispheres. This overestimation intensifies during low solar activity and shows strong seasonal dependence. During solstices, the crests are more pronounced in the local winter hemispheres, while during equinoxes the crests are weaker and exhibit hemispheric symmetry. This morphology aligns with the presence of light ions diffusing downward from the plasmasphere. Validating the LP data in conjunctions with Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) observations showed a much stronger agreement after applying the developed correction: for Swarm B, nighttime correlation with COSMIC increased from 0.74 to 0.93. The NN-calibrated LP data set has numerous applications in ionospheric research, and the developed model can provide useful insights into the ion composition in the topside ionosphere.

Plain Language Summary Operating since late 2013, ESA's Swarm mission has provided an extensive data set of plasma densities in the topside ionosphere, frequently used in space weather research. Recent studies have shown that the Langmuir Probes (LPs) onboard the Swarm mission overestimate ion densities during nighttime by up to 50%. This overestimation is due to the presence of light ions not accounted for in the ion density calculations. In this study, we develop a new NN-based calibration model for Swarm observations, which resolves the nighttime overestimation by the LPs for the first time. Our results reveal a distinct double-crest pattern on the nightside, centered at ~30° quasi-dipole latitude in both hemispheres, which corresponds to decreased effective ion masses and thus causes the overestimation of ion densities. This pattern closely mimics the behavior of light ions diffusing downwards from the plasmasphere. The developed NN-based calibration improves the quality of the LP observations and allows using them for improving the existing models of the topside ionosphere, among other applications. Furthermore, the developed model provides insights into the ion composition and light ion dynamics in the topside ionosphere.

1. Introduction

Earth's ionosphere is a partially ionized region of the upper atmosphere that extends from ~60 to about 1,000 km in altitude (e.g., Rishbeth & Garriott, 1969). The part of the ionosphere located above the F2-layer peak, known as the topside ionosphere, contains up to 80% of the total electron content and is particularly important for radio communications and navigation (Bilitza, 2009). Although millions of satellite plasma density observations have been collected over the past two decades, the observational coverage of the topside ionosphere remains highly non-uniform (e.g., Bilitza et al., 2022; Smirnov et al., 2021). Most ionospheric constellations operate within narrow altitude ranges, and their orbit-bound in situ measurements rarely provide three-dimensional coverage. To

© 2024. The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

ensure the consistency and reliability of ionospheric models across different altitudes, local times and solar activity levels, it is crucial to utilize data from all of the available missions. The corresponding measurements need to be adjusted to the same reference frame before their integration into the empirical models, and therefore the task of calibrating the available data sets plays a critical role in advancing our ionospheric modeling capabilities.

Operating since late 2013, the ESA Swarm mission (Friis-Christensen et al., 2006) has provided an extensive plasma density data set in the topside ionosphere. Swarm is comprised of three identical spacecraft (Alpha, Bravo and Charlie, hereafter abbreviated as A, B, and C) on near-polar low Earth orbits (LEO). During the commissioning phase, the three satellites were deployed in the pearls-on-a-string configuration, following each other at similar altitudes of around 490 km. By Spring of 2014, the spacecraft were manoeuvred into a different configuration. Swarm A and C satellites were lowered in altitude to ~460 km and put side-by-side with a 1.4° separation in longitude. In contrast, Swarm B was raised to a higher altitude of ~520 km, with its orbital plane precessing at a different rate than that of Swarm A/C (Catapano et al., 2022). The Swarm satellites slowly rotate in local time (LT), and each of the spacecraft provides a full LT coverage every 130–140 days (Pignalberi et al., 2022). The primary goal of the Swarm mission is to provide highly accurate observations of the Earth's magnetic field and its variability, including the core dynamics, geodynamo processes and shorter-scale variations (e.g., Friis-Christensen et al., 2006). The Swarm satellites are also equipped with a set of instruments allowing a comprehensive characterization of ionospheric environment, including electric fields, currents and other ionospheric plasma processes (Catapano et al., 2022; Knudsen et al., 2017; Xiong et al., 2022).

Each of the Swarm satellites carries an Electric Field Instrument (EFI) comprised of two Langmuir Probes (LPs) measuring plasma densities and temperatures and two Thermal Ion Imagers (TIIs) observing ion drifts (Knudsen et al., 2017). The initial motivation for adding LPs to the payload was to monitor the magnetic field perturbations due to the diamagnetic effect (e.g., Friis-Christensen et al., 2006; Lühr et al., 2003), but after a decade in operation the scientific output of these instruments has far surpassed the initial goal, and Swarm LPs have been used extensively for studying various ionospheric phenomena (Catapano et al., 2022; Pakhotin et al., 2022, 2023; Xiong et al., 2022). Swarm LPs are quite unconventional, as they are the first instruments ever deployed in orbit that measure plasma densities and temperatures in the so-called harmonic mode. This mode of operation requires an assumption of the effective ion mass (M_{eff}) for N_i calculations (Knudsen et al., 2017). Inaccuracies of this assumption may lead to deviations in the derived ion densities. Namely, when the assumed M_{eff} is too high compared to realistic conditions, plasma densities are underestimated, while effective ion masses that are too low lead to density overestimation (for details, see Section 2, and Pakhotin et al., 2022, 2023). LP data processing assumes all ions to be O^+ ($M_{\text{eff}} = 16$ atomic mass units (AMU)), which is generally a reasonable assumption at altitudes of 450–500 km during high solar activity levels but is frequently violated during medium and low solar activity conditions on the nightside (Smirnov et al., 2021). In particular, the upper transition height (UTH), which marks the change from oxygen- to hydrogen-dominated parts of the ionosphere, is usually located above 1,000 km in altitude, but can subside to much lower altitudes during solar minima. It has been shown that occasionally, up to 50% of ions are H^+ at altitudes of ~500 km during nighttime (Aponte et al., 2013; Heelis et al., 2009; Klenzing et al., 2011; Vaishnav et al., 2021), and therefore the assumption of oxygen-only plasma used in LP data processing can be violated.

Swarm satellites are also equipped with FacePlates (FPs), upon which the TIIs are mounted (Knudsen et al., 2017). When TIIs are not active, the FP instruments operate similarly to planar Langmuir probes, allowing derivation of plasma density without making any assumptions on plasma composition at a high resolution of 16 Hz. However, the FPs operate only for several orbits per day, and were basically switched off in October 2019 for Swarm A and B, and since July 2020 for Swarm C (see Figure 1b). Therefore, it should be emphasized that the FP measurements cannot serve as a replacement for the LP data set, which remains the main nominal source of the plasma density observations on Swarm. Consequently, it is important to pay attention to the LP calibration, as they provided one of the most widely used ionospheric data sets to date.

Several studies have been devoted to calibration and quality control of the Swarm Langmuir Probe ion densities. Lomidze et al. (2018) performed a comparison of Swarm LP and ISR observations for the overhead passes in December 2013 - June 2016 and found that Swarm LPs generally underestimated ion densities. However, due to limited operational time of the ISRs only a few dozen conjunctions could be found and the analysis of local time differences could not be performed. Later on, Smirnov et al. (2021) compared the Swarm LP plasma densities in conjunctions with Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) radio

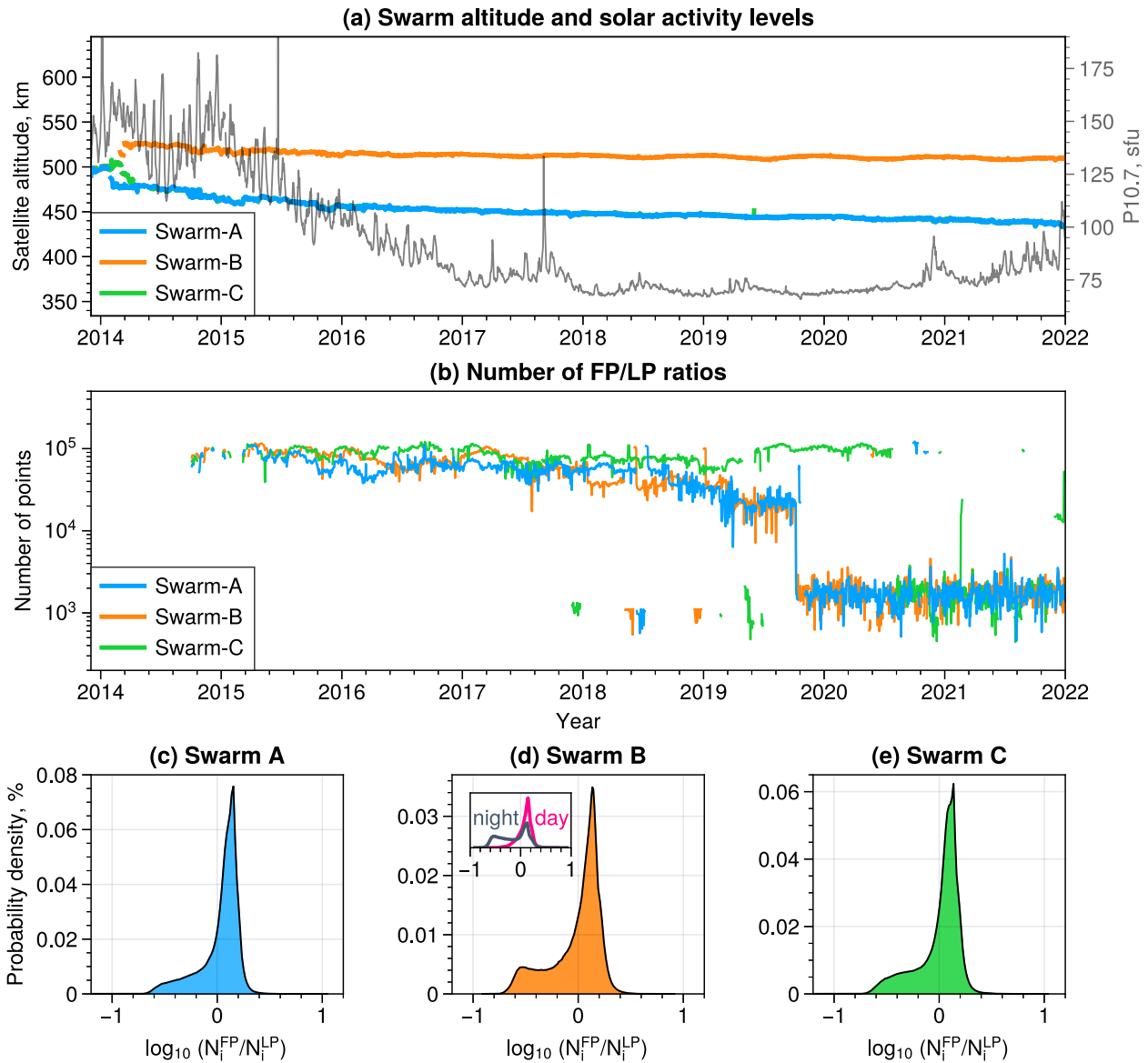


Figure 1. (a) Daily averaged altitude of the three Swarm satellites (colored lines), and the solar flux index P10.7 (gray line); (b) daily number of FP/LP ratios. The 16 Hz FP observations were interpolated on the times of LP measurements, and the resulting number of points was smoothed by 3 days. The FP instrument was operating regularly until late 2019 and since then was mostly switched off, hence the drop in the number of observations from $\sim 10^5$ to $\sim 10^3$ points per day. Panels (c, d, e) show the histograms of the FP/LP ratios. The shapes of the distributions are similar for Swarm A and C satellites, as they follow similar orbits, while the distribution of Swarm B data is strongly bimodal, due to the day-night differences, which is shown in the inset panel in subplot (d).

occultation (RO) observations and found that Swarm LPs underestimated electron densities during daytime but showed a significant overestimation during nighttime at low and middle latitudes. The authors attributed this effect to the non-negligible fraction of light ions at satellite altitudes. Furthermore, the overestimation was larger for Swarm B than for Swarm A/C satellites, due to its higher altitude and closeness to the UTH. The nighttime overestimation by the LPs could not be seen by Lomidze et al. (2018), as their analysis was performed for high solar activity periods, while this effect appeared later on toward the minimum of solar cycle 24/25. Catapano et al. (2022) performed the first comparison of the LP and FP measurements on Swarm. They reported an almost perfect agreement between the two instruments during the day, with a correlation of 0.98, but only a moderate correspondence on the nightside with a correlation of 0.47. By comparing LP data to the FP observations, Xiong et al. (2022) demonstrated that the bias of Swarm LPs was also strongly dependent on solar activity. They found

that during high solar activity periods, LPs generally underestimate ion densities, while for low solar activity densities are overestimated. Regardless of solar flux levels, the bias of Swarm LPs was dependent on local time, showing a daytime underestimation and nighttime overestimation of ion densities. Fast et al. (2023) reported a similar pattern of LT-dependent bias of Swarm LPs on a large number of conjunctions between the three Swarm satellites and Poker Flat and Resolute Bay ISRs.

Even though the LT-dependent bias of Swarm LPs has now been well documented in literature, it has not yet been accounted for in the LP data processing (Pignalberi et al., 2022). Recently, Pakhotin et al. (2022, 2023) developed a Swarm Langmuir probe Ion drift, Density and Effective Mass (SLIDEM) data product that uses the FP current data and the orbital-motion limited (OML) theory in order to relax the O^+ -only assumption in LP ion density derivations (see also Burchill & Lomidze, 2024). However, their calculations require the FP bias to be strongly negative, and therefore the SLIDEM data can only be derived for sparse time intervals when TIIs are inactive. In the present study, we develop a climatological correction that can be used to correct the entire LP data set. We use the FacePlate observations as reference, due to their good agreement with ISR measurements across all levels of solar activity (Xiong et al., 2022), and model ratios of FP to LP ion densities. These ratios vary substantially based on geophysical conditions, and thus the LP observations cannot be calibrated using a simple linear adjustment. Instead, the correction factors should take into account the solar and geomagnetic activity, altitude, location, local time and season. FP/LP ratios can only be derived when both instruments operate simultaneously, usually for several orbits per day, which makes the availability of training points quite sparse. Therefore, it is crucial to select a technique that can learn effectively from very sparse data points. Artificial Neural Networks (NNs) have been shown to excel at this task (e.g., Chen et al., 2023; Pan et al., 2024; Smirnov et al., 2023; Zhelavskaya et al., 2017), and for this reason, our correction model for Swarm LPs is based on neural networks. It should be emphasized that the developed correction is not bound to the availability of FP data, as it provides a climatological model of the FP/LP ratios based on geophysical conditions. Therefore, it can be used to correct the entire LP data set, even when the FP data are not available.

The paper is organized as follows. In Section 2, we describe the data set and modeling methodology. Section 3 presents the results, including synthetic model runs analyzing the FP/LP differences under different levels of solar and geomagnetic activity. In Section 4, we compare the NN-calibrated LP data to independent radio occultation observations by the COSMIC mission. Section 5 presents the discussion of the results and outlines opportunities for future investigations, and the final section draws the conclusions.

2. Data Set and Machine Learning Methodology

2.1. Data Set

2.1.1. General Principles of Langmuir Probe Operations

Langmuir probes are among the most frequently used instruments to provide in situ observations of plasma densities (Lebreton et al., 2006). These relatively simple and inexpensive devices represent electrodes of planar, cylindrical or spherical shapes immersed into the plasma (Hargreaves, 1992). LPs are named after Irving Langmuir, whose group pioneered their usage for measuring plasma parameters (Mott-Smith & Langmuir, 1926; Tonks & Langmuir, 1929). LPs measure the current (I) between the probe and the spacecraft, resulting from incremental adjustments of the probe's bias voltage (V). This technique allows to sweep across a complete range of values relevant for the given conditions (Knudsen et al., 2017). As the bias varies from negative to positive, ions and electrons from the surrounding plasma are attracted toward the probe. The resulting current-voltage ($I - V$) characteristics are then used to retrieve plasma density and temperature (Brace, 1998; Lebreton et al., 2006). When the probe bias is negative, the current results predominantly from positive ions. Conversely, when the bias is strongly positive, the current is carried by electrons. These two regions of the $I - V$ characteristics are known as the ion and electron saturation regions, respectively. In between these two regimes, the exponential change in current relative to the probe's bias is proportional to the electron temperature. It is of note that the $I - V$ characteristics always rely on certain assumptions, and factors such as geometry, material properties, size of probes and posts play a crucial role in estimating the density and temperature of the plasma (Resendiz Lira & Marchand, 2021; Xiong et al., 2022).

2.1.2. Swarm LP and FP Instruments

Swarm LPs are spheres of 4 mm radius that are mounted on ~8 cm posts below the ram panels of the three spacecraft (Catapano et al., 2022; Lomidze et al., 2018). Each satellite carries two LPs, one with a nitrated titanium (TiN) surface and the other one with a gold (Au) surface (Knudsen et al., 2017). The two LPs usually operate with fixed but different gains, called the high and low gains, which differ by a factor of ~50 (Catapano et al., 2022; Knudsen et al., 2017). The gain of each LP is commanded from the ground, and usually one of the LPs is in high gain and the other in low gain mode (Lomidze et al., 2018). The two probes are separated by a distance of ~30 cm. Swarm LPs use the harmonic mode of operation to measure the $I - V$ characteristics, where the probe bias is sinusoidally modulated at a nominal frequency of 128 Hz (Knudsen et al., 2017). The harmonic cycles typically last for half a second, during which time the currents and admittances are measured. As noted by Catapano et al. (2022), Swarm is the first mission to use this method in orbit.

In the processing algorithm of Swarm LPs, the derived ion densities depend on several parameters, including the ion ram velocity (v_i) and probe radius (r_p). A crucial factor is the so-called ion admittance d_{ion} , which represents the response of current to harmonic modulations in voltage ($\partial I / \partial V_b$):

$$N_i^{LP} \propto d_{ion} = \frac{\partial I}{\partial V_b} = \frac{1}{M_{eff}} \cdot \frac{2N_i e^2 \pi r_p^2}{v_i}, \quad (1)$$

where I is the probe current, V_b is the probe bias, e is the elementary charge, and M_{eff} is the effective mass. The latter parameter can have a particularly strong influence on the ion density calculations. As described by Pakhotin et al. (2022), the contribution of individual ion species to M_{eff} is inversely proportional to their masses:

$$\frac{1}{M_{eff}} = \frac{1}{N_i} \sum_{s=1}^k N_s \frac{1}{m_s}, \quad (2)$$

where m_s is the mass and N_s is the number density of the s -th ion species. The effective ion mass is therefore very sensitive to even small concentrations of light ions in the ambient plasma. Pakhotin et al. (2022) estimated that adding just a 10% fraction of H^+ ions to an oxygen-only plasma would reduce the effective ion mass from 16 to 6.4 AMU. This change would translate into a dramatic overestimation of plasma density by the LPs.

The scientific products from Swarm LPs are provided at the Level 1B (L1B). Their processing is described in detail in Catapano et al. (2022). The ion density data utilized in this study (version “0502” of the “EFIx_LP_1B” product) are measured at a 2 Hz sampling rate, which corresponds to the spatial resolution of ~3.8 km (Smirnov et al., 2021). LP observations are provided with quality flags that indicate the instrument performance, settings and sources of possible errors (DTU, 2019). For outlier exclusion, we employed the following quality flags: $Flags_{LP} = 1$, $Flag_{Ne}$ and $Flag_{Te} = 10$ or 20, and importantly, $Flag_{Vs} = 10$ or 20. The impact of these flags on data filtering is further explored in Supporting Information S1 (Figures S1.1 and S1.2).

Plasma densities can also be derived using the FPs, which are rectangular plates each with an area of $351 \times 229 \text{ mm}^2$ and a thickness of 3.175 mm (Resendiz Lira & Marchand, 2021; Xiong et al., 2022). On each of the Swarm satellites, the FP is located on the ram panel and is used to mount the TIIs (see Figure 1 in Catapano et al. (2022)). The primary goal of the TII suite is to provide measurements of plasma drifts (Knudsen et al., 2017). When the TIIs are active, the Faceplate maintains a slightly negative bias of -1 V which allows ions to penetrate into the TIIs (Pignalberi et al., 2022). When TIIs are not active however, FP operates as a planar Langmuir probe and measures the current collected at a fixed-bias voltage (typically $\sim -3.5 \text{ V}$) for several orbits a day. Assuming that the plasma flow is supersonic in the reference frame of the spacecraft and that the FPs are in the ion saturation region (e.g., Xiong et al., 2022), the ion densities can be obtained using the following formula:

$$N_i^{FP} = \frac{I_{FP}}{eA u_i}, \quad (3)$$

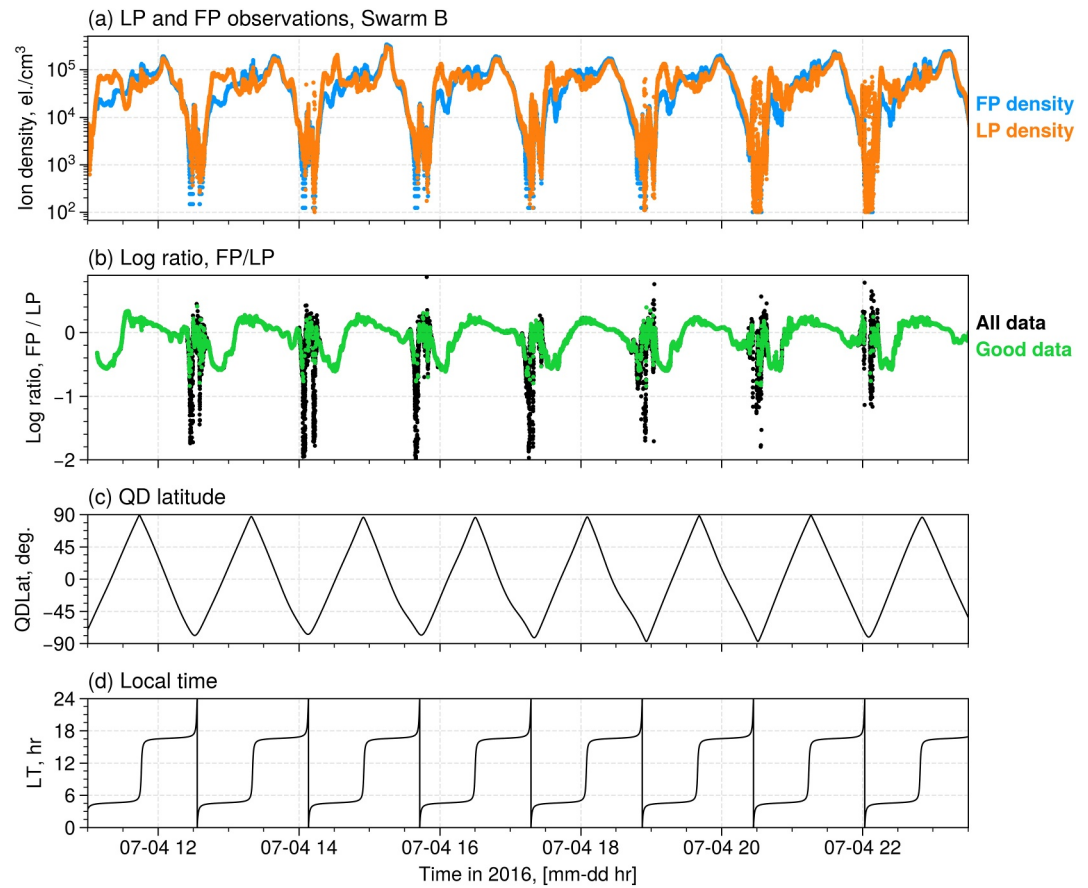


Figure 2. (a) Swarm FP and LP ion density observations, shown in blue and orange, respectively. Panel (b) shows logarithmic ratios between the two instruments. The ratios obtained from all available observations are shown as black dots, while the ratios taken after applying the quality flags to the LP observations are shown in green. The employed flags greatly reduce the number of outliers. Panel (c) demonstrates the quasi-dipole latitude, and panel (d) shows the local time along the satellite orbit.

where u_i is the ion velocity normal to the FP surface, and A is the FP area. The FP N_i data are provided with a sampling frequency of 16 Hz and resolution of 100 cm^{-3} (e.g., Wang et al., 2022). The derived densities do not depend on the composition assumptions, and as was shown by Xiong et al. (2022), do not exhibit a solar flux-dependent bias. In this study, we used the FP observations that are part of the Advanced plasma data products (“SW_EXTD_EFIA_LP_FP”). Technical explanations for the FP data set are available at (IRF, 2017).

Figure 2 shows a comparison between the LP and FP ion density observations for several orbits of Swarm B. It can be seen that although latitudinal N_i variations observed by the two instruments are in good agreement, there are some systematic differences between the two curves. Namely, the LP values are higher than the FP ones on the nightside at low and mid-latitudes. It is noteworthy that the resolution of FP data is around 100 cm^{-3} and for low densities the FP readings are discretized accordingly. Panel (b) shows the logarithmic ratios between the two instruments. Here, the LP densities are in the denominator, and thus positive log-ratios correspond to periods of underestimation by LPs, while negative log-ratios signify times when LPs overestimate ion densities. One can see that applying the LP quality flags removes most of the outliers and reduces the overall range of value from ± 2 orders of magnitudes to the range from about -0.85 to ~ 0.65 . However, even after this filtering some outliers are still present. To mitigate this, both data sets are smoothed using a 30-s sliding window, and windows containing NaN values are excluded from the final data set. This facilitates additional filtering, ensuring that the data for the climatological model are smooth and free of jumps

(see Figure S1.3 in Supporting Information S1). As shown in Figure 1 (panels c, d, e), the FP/LP ratios cover a wide range of values spanning across ~ 1.5 orders of magnitude. To standardize the data and reduce the skewness of the distributions, we take logarithms of FP/LP ratios during the preprocessing stage and subsequently perform the modeling in log-scale.

2.2. Machine Learning Methodology

2.2.1. Input Parameters

The ratios between FP and LP instruments, which represent the correction factors for LP ion densities, are known to vary with geophysical conditions (e.g., Xiong et al., 2022). We parametrize the strong dependence of FP/LP ratios on solar activity by the P10.7 solar flux index, which is a smoothed version of the 10.7 cm solar radio flux (F10.7) index (for details, see Bilitza & Xiong, 2021; Bilitza et al., 2022; Xiong et al., 2022). Furthermore, we include the dependence on geomagnetic activity by using the Hp30 index, which is derived using a methodology similar to that of the Kp index except with a higher cadence of 30 min (Yamazaki et al., 2022). While Kp has an upper limit of 9, the Hp30 index is open-ended and has been shown to exceed $H_p = 12$ for extreme storms. Hp30 values are discretized at ~ 0.33 increments. If used directly, these discrete values would introduce steps into the model output, which is not acceptable for the data calibration purposes. Thus, the Hp30 values are linearly interpolated onto the times of measurements to avoid such discontinuities.

Other inputs to the model include the geographic and quasi-dipole latitude and longitude (Laundal & Richmond, 2017), satellite altitude, LT and day of the year (DOY). We transform the latitude, longitude, local time and DOY using Fourier features technique, introduced by Tancik et al. (2020). This method replaces the raw values of the cyclic features with their sine and cosine transformations up to a given degree. It was recently applied to ionospheric modeling by Smirnov et al. (2023) and allowed the developed Neural network model of Electron density in the Topside ionosphere (NET) to reproduce even fine-scale structures, such as the mid-latitude density trough. Here, we use Fourier features up to degree 2 to account for semi-diurnal and seasonal effects. At the same time, it should be noted that at the poles, where all meridians converge, there is a discontinuity in longitudes that Fourier feature mappings may not be able to completely resolve. Hence, in the immediate vicinity of the poles, the models based on the Fourier features are suitable for interpreting the results but should be used with caution for the calibration, and it is recommended to exclude data within $\pm 2.5^\circ$ of the poles from the final calibrated data set.

2.2.2. Data Splitting and Resampling

Figure 1b shows the number of available FP/LP ratios per day throughout the duration of this study (2014–2022). The LP data set is continuous in time, while the FPs were mostly turned off in late 2019 for A and B satellites and in 2020 for Swarm C, and the number of available FP observations decreased from $\sim 10^5$ to $\sim 10^3$ per day. This makes the temporal and spatial distributions of FP/LP ratios available for training rather sparse. To ensure that the data coverage is approximately similar across different seasons and locations, we perform data resampling, described in the Supporting Information S1 (Figures S1.4–S1.9). For each of the Swarm satellites, the complete data set after resampling comprises approximately 30–40 million samples.

In empirical modeling, it is important to assess the generalization ability of the developed model and evaluate the performance not only on the fitted data but also on independent observations. The full data set is typically split into 3 subsets: the training set, used for fitting model parameters, the validation set, used for choosing the best architecture and providing an unbiased performance estimate during training iterations, and the test set. The test set does not take part in model development and is only used once, after the training and validation are completed, in order to evaluate the generalization on entirely independent data. The objective is to develop a model that is effective across all three sets. If the model performs comparably well on the test set as it does on the training set, it has a good generalization ability and minimal overfitting (Hastie et al., 2009).

For time-series and spatio-temporal problems, it is vital to split the data into training, validation and test sets in the time-domain. Drawing the samples into three subsets randomly can cause data leakage, which occurs when the subsets overlap in time and the model is trained and validated on the same events. This may lead to the selection of suboptimal hyperparameters and low generalization capabilities of the model (Camporeale, 2019). In this study, we

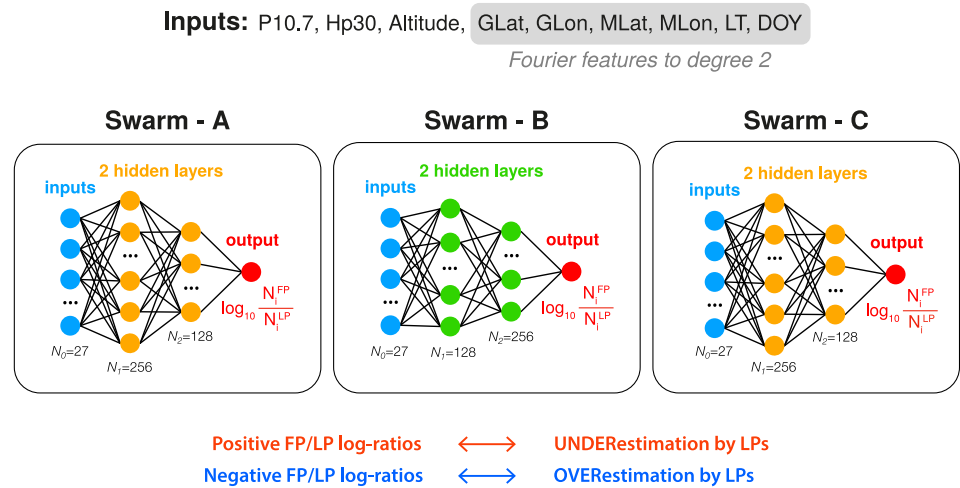


Figure 3. Schematic representation of the NN model workflow. The model inputs include the solar flux index P10.7, geomagnetic index Hp30, satellite altitude, as well as DOY, LT, and geographic and magnetic coordinates transformed to sines and cosines up to degree 2. The calibration models are developed for each of the three Swarm satellites separately and constitute 2-layer MLPs, which output the logarithmic ratios between FP and LP ion densities. The LP observations can be multiplied by these factors to perform the calibration.

first divided the data into continuous weekly intervals, and then randomly allocated those periods into the three sets. The weekly duration was selected as it is sufficiently longer than ionospheric memory but shorter than time periods over which the satellite orbit would drift significantly in local time. Such splitting allows us to train and validate the model on all solar and geomagnetic activity conditions and to utilize all LTs sampled by Swarm for training. A similar splitting technique was used, for instance, by Chu et al. (2021) and Smirnov et al. (2023). We have also tested the 14-day and 27-day intervals for the splitting (Figure S1.10 in the Supporting Information S1). It was found that increasing the duration of the time blocks did not substantially alter the model performance on independent data: for the 14-day time-intervals, the change of performance was around 1%. For the 27-day blocks, the accuracy was still high, but some non-linear features could be seen in the 2D histograms. This is likely due to the fact that over the period of 27 days, the satellites rotate in local time by around 4 hr, and the distributions for training and testing the model would differ. This phenomenon is known as the data (or, covariance) shift (e.g., Sugiyama et al., 2007), and should be mitigated. In order to ensure the coverage of all conditions seen by Swarm, after the independent evaluation on the test set (Section 3.1), the final models are retrained on a combination of the three sets.

2.2.3. Model Architecture and Pipeline

Neural networks are a broad family of mathematical models, renowned for their state-of-the-art performance in both classification and regression tasks. A key advantage of NNs is in their ability to learn from very sparse observations, fitting even the most complex non-linear relationships between input and output variables. In this study, we use multilayer perceptrons (MLPs), which are a subclass of feedforward neural networks (e.g., Goodfellow et al., 2016). MLPs consist of series of interconnected nodes, organized into layers. There are three types of layers: the input layer which receives matrices of input parameters, the output layer that generates the network's output, and one or more intermediate layers, often called “hidden” because the neurons within them are not directly observable. The process of training an MLP involves optimizing the weights and biases in the hidden layers in order to better fit the training data. This is an optimization problem, which is solved iteratively using gradient descent algorithms that incorporate backpropagation (Rumelhart et al., 1986).

In this study, we use MLPs with 2 hidden layers and develop separate calibration models for each of the satellites (Figure 3). Initial hyperparameter tuning for the Swarm A model showed that 2-layer MLPs performed similarly to 3-layer ones, and therefore we fixed the number of layers at 2 and optimized the other hyperparameters. It should also be noted that due to the similar altitude and distributions of FP/LP ratios for Swarm A and C satellites (see Figures 1c and 1e), we use the same NN architectures for these models, while the hyperparameters for the Swarm B model are optimized separately. The search domains of the hyperparameters and the optimized values can be found in the Supporting Information S1, Table S1.1. Due to the fact that MLPs are comprised of large

numbers of parameters (in our case, $\sim 10^5$ weights and biases), the models may become too expressive and overfit to training data. To combat overfitting, we use regularization techniques, namely, the Gaussian noise and Dropout, and treat the noise magnitude and percentage of dropped neurons as hyperparameters.

Each of the three models outputs the ratios of ion densities between the FP and LP instruments, modeled in logarithmic scale (Figure 3). The LP readings can be calibrated by multiplying them with the correction factors as follows:

$$LP_{\text{calibrated}} = LP_{\text{measured}} \cdot \left(\frac{FP}{LP} \right)_{\text{modeled}} \approx FP. \quad (4)$$

This correction technique does not depend on the FP data availability and relies on the FP/LP ratios obtained from the climatological NN-based model. If the correction factors are accurately reproduced by the model, the calibrated LP data should equal the FP observations. In essence, our approach can be viewed as a way to extend the FP data set to the entire duration of the Swarm mission.

3. Modeling Results

In this section, we describe the modeling results and perform global synthetic runs of the NN model in order to analyze the FP/LP ratios on a global scale under different solar activity and geomagnetic conditions.

3.1. Comparison of Observed and Predicted FP/LP Ratios

Figure 4 shows the 2D histograms comparing the observed and predicted FP/LP log-ratios for all three Swarm satellites on the training, validation and test sets. Dashed black lines in each panel represent the ideal one-to-one correspondence between observed and modeled log-ratios. One can see that most of the occurrences follow the one-to-one trend very well. In order to quantify the agreement between predictions and observations, we evaluated several metrics, including the mean biases, standard deviations and Spearman rank correlation (ρ) coefficients; their values are given in the corresponding panels. The correlations between predictions and observations are consistently above 0.9 and the model biases are close to zero (in the order of -0.002). The metrics evaluated on the training, validation and test sets show minimal differences, indicating that the model performance is almost identical on the data that were used for training and on completely independent observations. This indicates that the model has a good generalization ability and a very low degree of overfitting.

An interesting finding in Figure 4 is the difference in the distributions of FP/LP log-ratios between Swarm A/C and Swarm B satellites. In particular, in case of Swarm A and C models, one can see that most of the occurrences correspond to positive values of 0–0.2. These correspond to a general underestimation of electron density by Swarm LPs (e.g., Lomidze et al., 2018). However, the results for Swarm B in Figures 4d–4f show additional peaks of occurrences around negative values of -0.7 – -0.4 highlighted with white arrows. These values correspond to an overestimation of ion densities by Swarm LPs and originate mainly on the nightside (see also Figure 1d); their characteristics are analyzed in detail in the following subsections.

3.2. Model Response to Solar Activity

In the previous subsection, it was shown that the developed NN model can reproduce the observed FP/LP ratios very well. The trained model can now be used to conduct synthetic NN runs under different levels of solar activity. We run the Swarm A and B models (at altitudes of 460 and 515 km, respectively) for low and high solar activity conditions ($P_{10.7} = 90$ and 150 sfu, respectively). These runs are performed for the solstice and equinox conditions in order to analyze the seasonal variability.

Figure 5 presents the synthetic NN runs for the Swarm A model. The left and right columns correspond to high ($P_{10.7} = 150$ sfu) and low ($P_{10.7} = 90$ sfu) solar activity, respectively. The magenta lines demarcate the day-night terminator boundaries, and the gray lines show the quasi-dipole (QD) coordinates (the QD equator and $\pm 30^\circ$, 60° quasi-dipole latitude (QDLat)). Under high solar activity, the FP/LP log-ratios exhibit near-uniform spatial distributions with positive values of around 0.1–0.2. They show minimal seasonal dependence and stay at almost constant levels, except for slight decreases to around 0 at mid-latitudes during nighttime and a pronounced enhancement at high latitudes in the J-season outside of the terminator boundary.

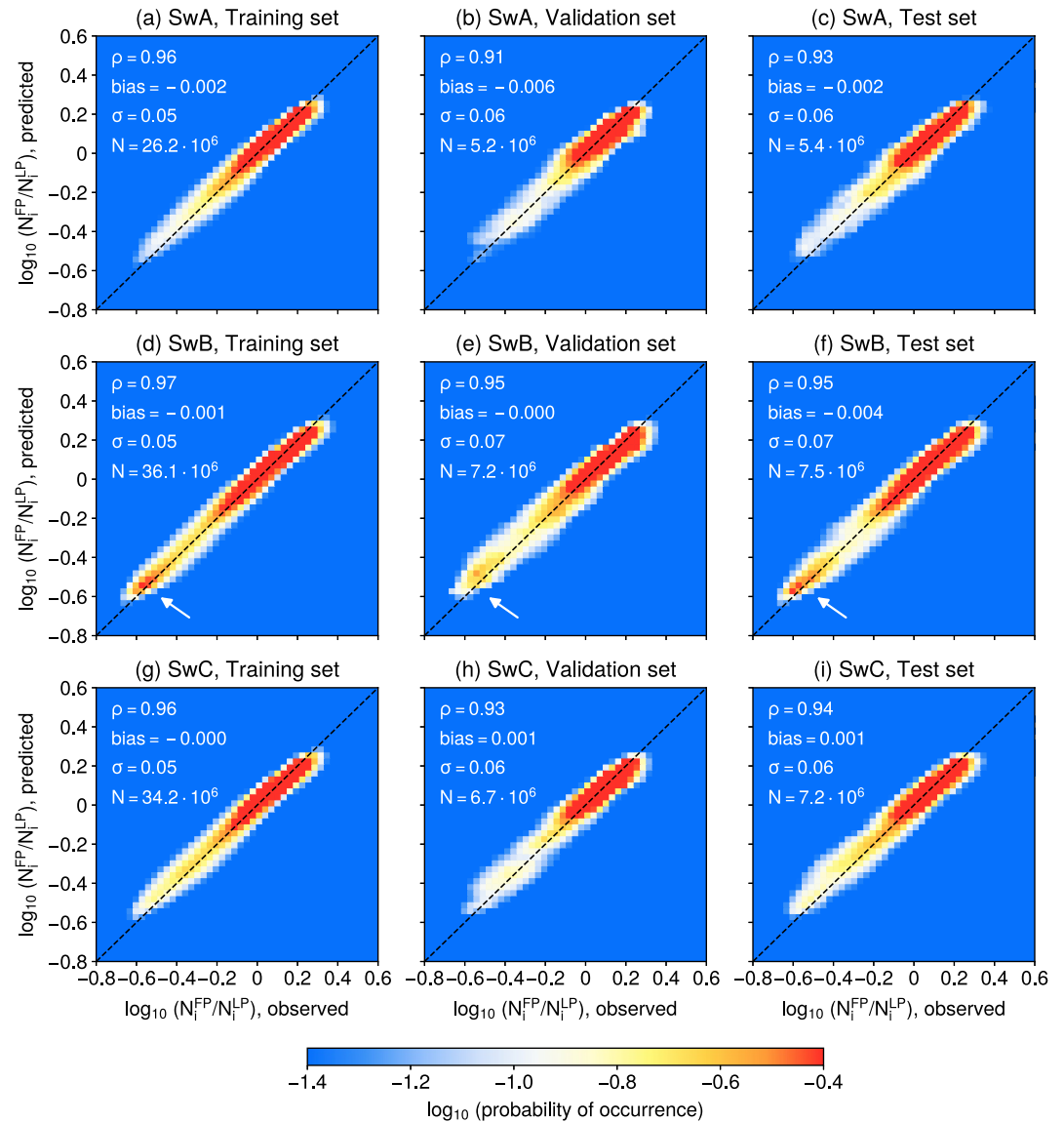


Figure 4. 2D histograms of observed versus predicted FP/LP ratios for the three models on the training (left column), validation (middle column), and test (right column) data sets. The model reproduces the observed ratios very well, with Spearman rank correlations (ρ) > 0.9 both on the training and fully independent (test) sets. Furthermore, the performance is very close on the training, validation and test sets which shows a good generalization ability of the model and low degree of overfitting. White arrows in panels (d, e, f) highlight the secondary peaks of occurrences corresponding to nocturnal light-ion crests which are strongest for Swarm B.

In contrast, the low solar activity runs reveal more complex structures (right column of Figure 5). Of particular interest is the appearance of blue crests at mid-latitudes on the nightside. They correspond to negative FP/LP log-ratios, with values ranging from -0.1 down to ~ -0.4 . These crests are centered around $\pm 30^\circ$ QDLat, and therefore should be differentiated from the equatorial ionization anomaly (EIA) crests, where the corresponding maxima lie around $\pm 15^\circ$ QDLat (Xiong et al., 2019; Zhong et al., 2019). In general, these crests can be observed between 22 and 06 LT. They show the strongest negative values after midnight and disappear abruptly before sunrise. This pattern exhibits strong seasonal dependence. During December and June solstices (D- and J-seasons, respectively), negative values are more pronounced in the local winter hemispheres. During equinoxes (E-seasons), the crests are almost identical in both hemispheres and generally weaker than during solstices.

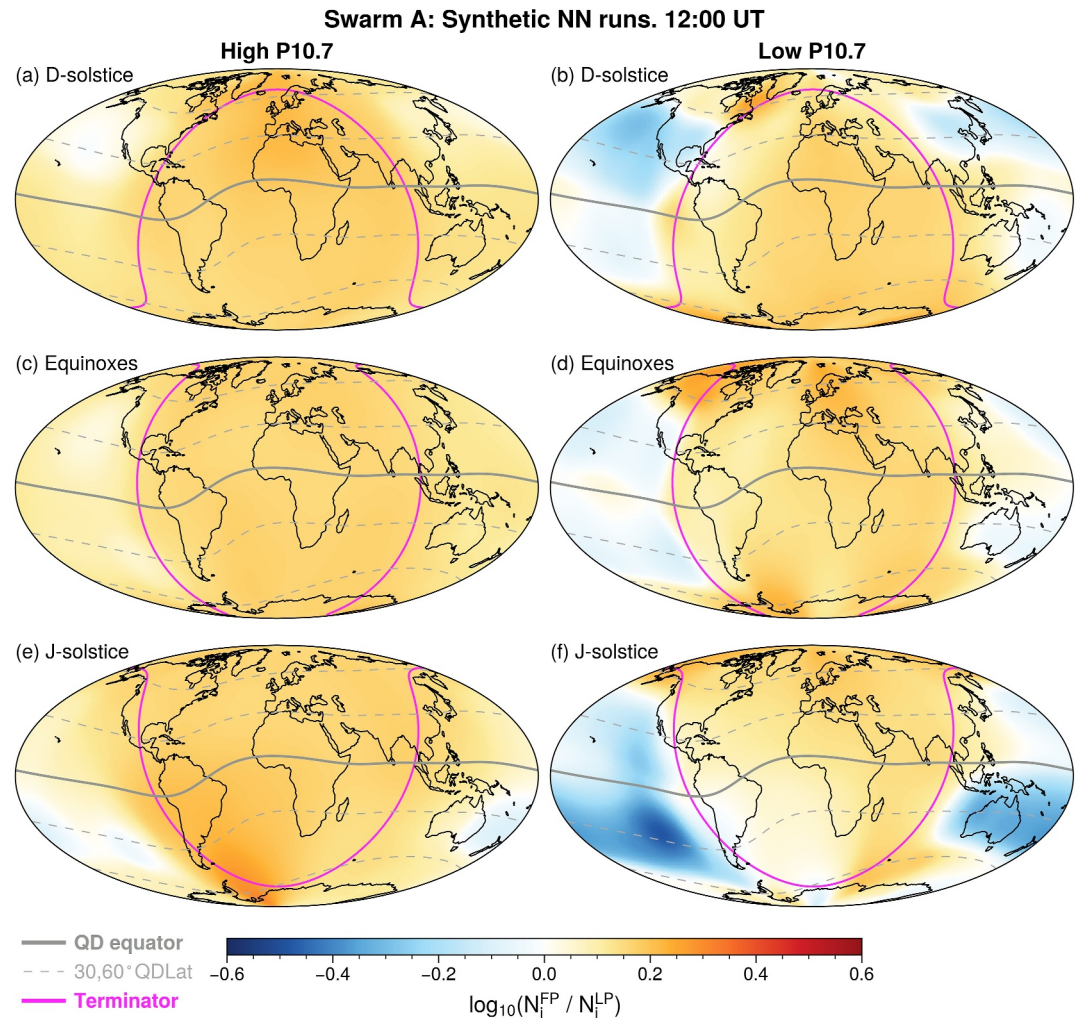


Figure 5. Synthetic model runs for Swarm A (altitude = 460 km). The left column shows runs for high solar activity ($P_{10.7} = 150$ sfu), while the right column shows the results for low solar activity ($P_{10.7} = 90$ sfu). The rows correspond to the three Lloyd seasons (results for the vernal and autumnal equinoxes were averaged for panels c and d). The FP/LP log-ratios are mostly uniform for high solar activity (~ 0.1 – 0.2). During low solar activity, a dual-crest structure on the nightside appears. It has a clear seasonal dependence: the local winter crests are more pronounced during solstices, while during the equinoxes the crests are similar in both hemispheres.

On the dayside, the log-ratios are positive (around 0.1 – 0.2) and mostly uniform, although there are regions with enhanced values around the morning terminator. In particular, in Figure 5b, slightly increased values can be seen at the geomagnetic equator during sunrise. This region corresponds to the so-called morning overshoot of electron temperature (e.g., Stolle et al., 2011). This indicates a local increase of the mean ion mass above that of atomic oxygen. In the discussion section we will offer some interpretation for this phenomenon. Additionally, enhanced positive values can be observed at high latitudes before the morning terminator.

Figure 6 displays synthetic NN runs for the Swarm B model, following the same format as Figure 5. In the case of Swarm B, which is higher in altitude by ~ 50 km, the nighttime crests are noticeable even during high solar activity with values around -0.1 . During low solar activity, the nighttime values at mid-latitudes are strongly negative, reaching ~ -0.6 in logarithmic scale. The log-ratios of -0.6 indicate overestimation of ion densities by the LPs by a factor of ~ 4 . The corresponding crests appear earlier after sunset compared to the Swarm A model and exhibit the largest magnitudes after midnight. The crests show similar seasonal behavior as for the Swarm A model, being stronger in the local winter hemispheres during solstices, and showing hemispheric symmetry during equinoxes. The large positive log-ratios at high latitudes that occur at sunrise are also stronger than those at Swarm A altitudes.

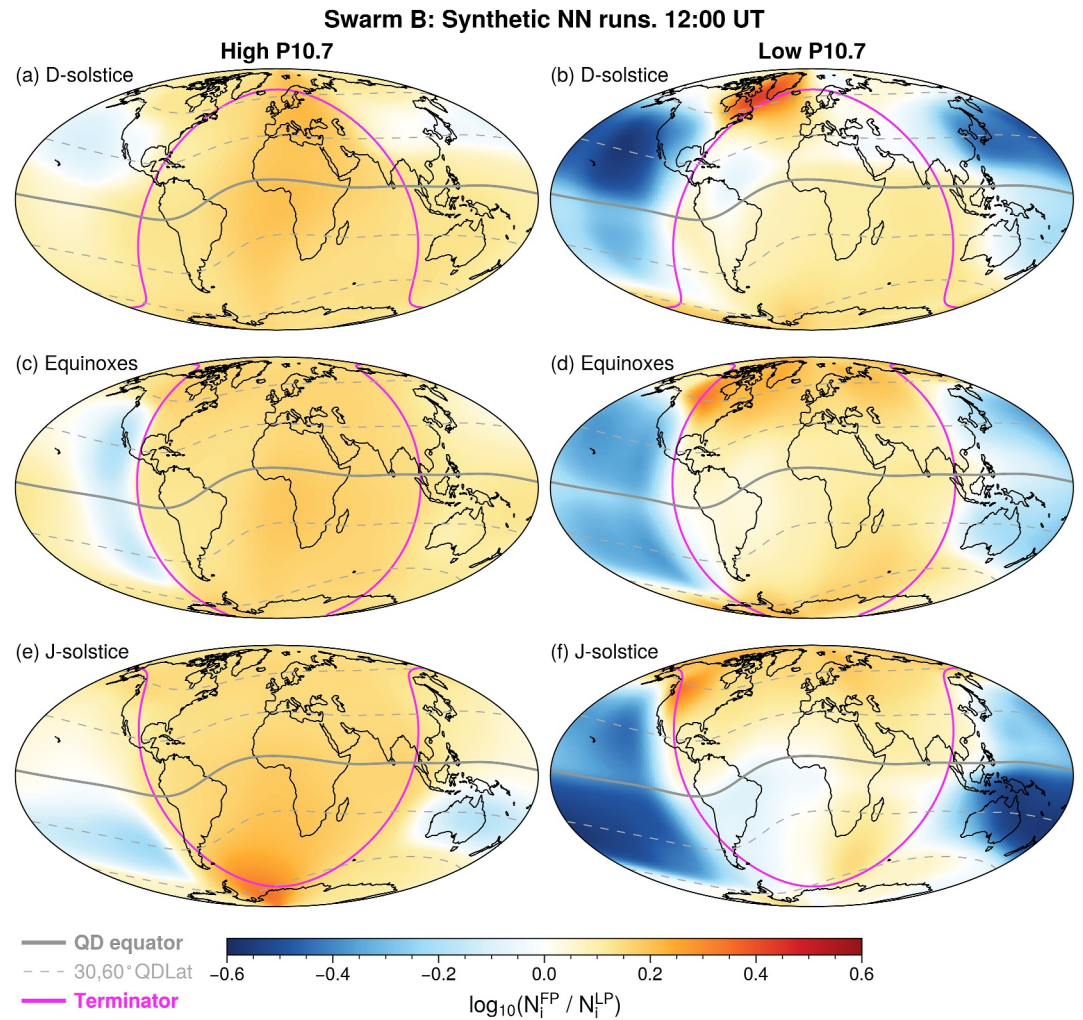


Figure 6. Same as Figure 5 but for Swarm B (altitude = 515 km). The negative log-ratios on the nightside during low solar activity are much stronger than for Swarm A (~ -0.6 compared to ~ -0.2), due to the higher altitude of Swarm B: during solar minima the UTH on the nightside subsides to lower altitudes, and a non-negligible fraction of light ions can be seen at Swarm B altitudes.

3.3. Model Response to Geomagnetic Activity

Xiong et al. (2019) reported that the nighttime mid-latitude plasma density enhancements, occurring due to the downward ambipolar diffusion of light ions, tend to disappear during geomagnetically active times. This is of particular interest for the present study, and therefore, we also investigate the dependence of the FP/LP ratios on geomagnetic activity. The parametrization of our FP/LP model includes a dependence on the Hp30 index. This allows us to conduct synthetic NN runs for low and high Hp30 values (0.67 and 6, respectively) and analyze the model response to geomagnetic activity while keeping all other parameters fixed. For these runs, we use equinoctial conditions and low solar activity with $P10.7 = 90$ sfu.

Both the Swarm A and B models exhibit a strong response to the increase in geomagnetic activity, as shown in Figure 7. In both cases, the nighttime mid-latitude plasma density enhancements become much less pronounced and corresponding values change from -0.4 to ~ -0.1 . It is also noticeable that the crests become confined to lower latitudes. If for geomagnetically quiet conditions they are visible up to $\sim 60^\circ$ QDLat, during active times they are only observed until $45\text{--}50^\circ$ QDLat. For

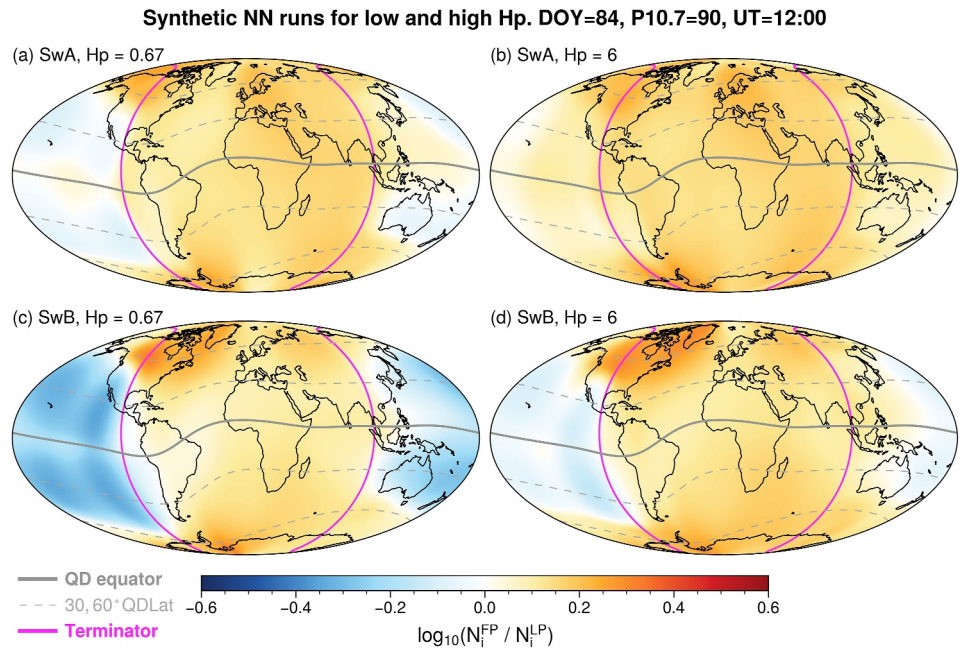


Figure 7. Synthetic model runs for geomagnetically quiet and active times ($H_p30 = 0.67$ and 6 , respectively), for equinoctial conditions. An increase in the H_p index makes the negative nighttime ratios less pronounced, likely due to an increase in $hmF2$ during active conditions.

comparison, we conducted synthetic runs of the NET ionospheric model (Smirnov et al., 2023) for the peak height of the F2-layer ($hmF2$) under the same conditions. The results demonstrate that $hmF2$ increases globally with geomagnetic activity, except for a few regions such as that of the pre-reversal enhancement (PRE) electric fields. During nighttime, $hmF2$ increases by ~ 60 km at low and middle latitudes. This indicates that the oxygen-dominated part of the ionosphere rises to higher altitudes of 350–400 km compared to 300–330 km during quiet times.

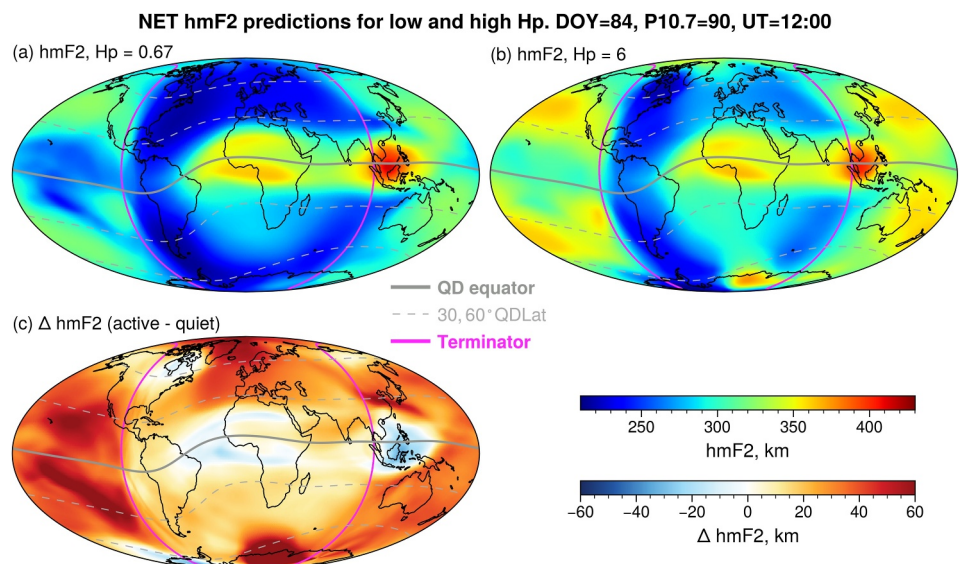


Figure 8. Synthetic runs of the NET ionospheric model for the F2-peak height ($hmF2$) for (a) geomagnetically quiet ($H_p = 0.67$, $Dst = -5$ nT) and (b) disturbed ($H_p = 6$, $Dst = -60$ nT) conditions. All other variables are fixed, and $P10.7 = 90$ sfu. Panel (c) shows the difference between the storm-time and quiet $hmF2$ levels.

4. Independent Validation by COSMIC Radio Occultation Data

In Section 3.1, it was established that the NN-based calibration model can accurately reproduce the FP/LP ratios, and therefore the LP data set can be adjusted to the FP data for the entire duration of the study (2014–2022). It is now necessary to ensure that this newly corrected data set is consistent with other global data sources in the topside ionosphere. To this end, we perform an independent validation of the NN-calibrated LP measurements against the radio occultation electron density profiles (EDPs) from the COSMIC mission. The LT-dependent bias of Swarm Langmuir probes was first identified in comparison with COSMIC data (Smirnov et al., 2021); therefore, validating the NN-corrected LP data against COSMIC would determine whether the newly developed correction model successfully removes these discrepancies.

To validate the calibrated Swarm LP ion densities by COSMIC/RO observations, we use the same conjunctions as in Smirnov et al. (2021). The authors found over 3,800 collocations for each of the Swarm satellites with COSMIC EDPs, by selecting data within $\pm 1.25^\circ \times 2.5^\circ$ GLat and GLon and ± 7.5 min universal time in 2013–2020. These dimensions ensure that the selected observations are close in location and time and not affected by strong ionospheric gradients. Observations corresponding to intervals of high geomagnetic activity ($K_p > 3$) are excluded from these comparisons (more details on this methodology can be found in Smirnov et al. (2021)). Scatter plots of the COSMIC/RO versus Swarm LP plasma densities at conjunctions are shown in Figure 9. The left column corresponds to the Swarm densities before calibration, and the right column presents the comparison after the NN-based correction. As in Smirnov et al. (2021), we plot the points coming from 06 to 18 hr LT in orange, and the nighttime points in blue.

The nighttime overestimation of ion densities by Swarm LPs manifests as clusters of points above the one-to-one correspondence lines, and this is particularly highlighted in the Swarm B comparison, where the overestimation is strongest (Figure 9e). For Swarm A and C, the nighttime correlation increases from 0.86 to 0.89 before correction up to 0.92 after NN-based calibration. The effect of the NN-calibration is most evident for Swarm B, where the correlation with COSMIC on the nightside increases from 0.74 to 0.93. The developed calibration technique significantly improves the LP ion density data, and the corrected data are in much better agreement with COSMIC for all three Swarm satellites.

Figures 9b, 9d, and 9f shows that the calibrated LP observations are slightly higher than the COSMIC measurements, likely because the Swarm FP data, used as a baseline, potentially requires its own calibration. Nevertheless, the calibrated Swarm LP data show linear agreement with COSMIC in all three cases, and this bias can be eliminated by a simple linear adjustment. Xiong et al. (2022) previously noted a slight positive bias of Swarm FPs in comparison with the ISR data, however, this bias was roughly constant over time and did not depend on solar activity levels. To remove the difference between the NN-calibrated LP and COSMIC densities, we shift NN-corrected LP ion densities to match the COSMIC observations. The NN-corrected LP data essentially represent the extended FP data set, and therefore the derived factors can be viewed as calibration coefficients for the FP instruments. The adjustment is made by the following equation:

$$\log_{10}(y_{\text{calibrated}}) = c \cdot \log_{10}(y) + d, \quad (5)$$

and the slope c and intercept d of the calibration trends can be found in Table 1. The final results after the adjustment are shown in Figure 10. The correlation coefficients are identical for all 3 satellites on the dayside (equal to 0.97) and very similar on the nightside (0.92 for Swarm A/C and 0.93 for Swarm B). Additional metrics evaluated for Swarm-COSMIC conjunctions before and after the developed correction are presented in Table 2. In all cases, the developed LP correction technique improves the agreement with COSMIC. For instance, in case of Swarm B, 83.3% of the points fell within a factor of 2 from collocated COSMIC observations, while after the correction, 93.4% met this criterion. In order to quantify the improvement achieved by this calibration, we evaluate the skill score, given by the formula:

$$SS = 1 - \frac{\sum_{i=1}^N (LP_i^{\text{cor}} - \text{COSMIC}_i)^2}{\sum_{i=1}^N (LP_i^{\text{uncor}} - \text{COSMIC}_i)^2}, \quad (6)$$

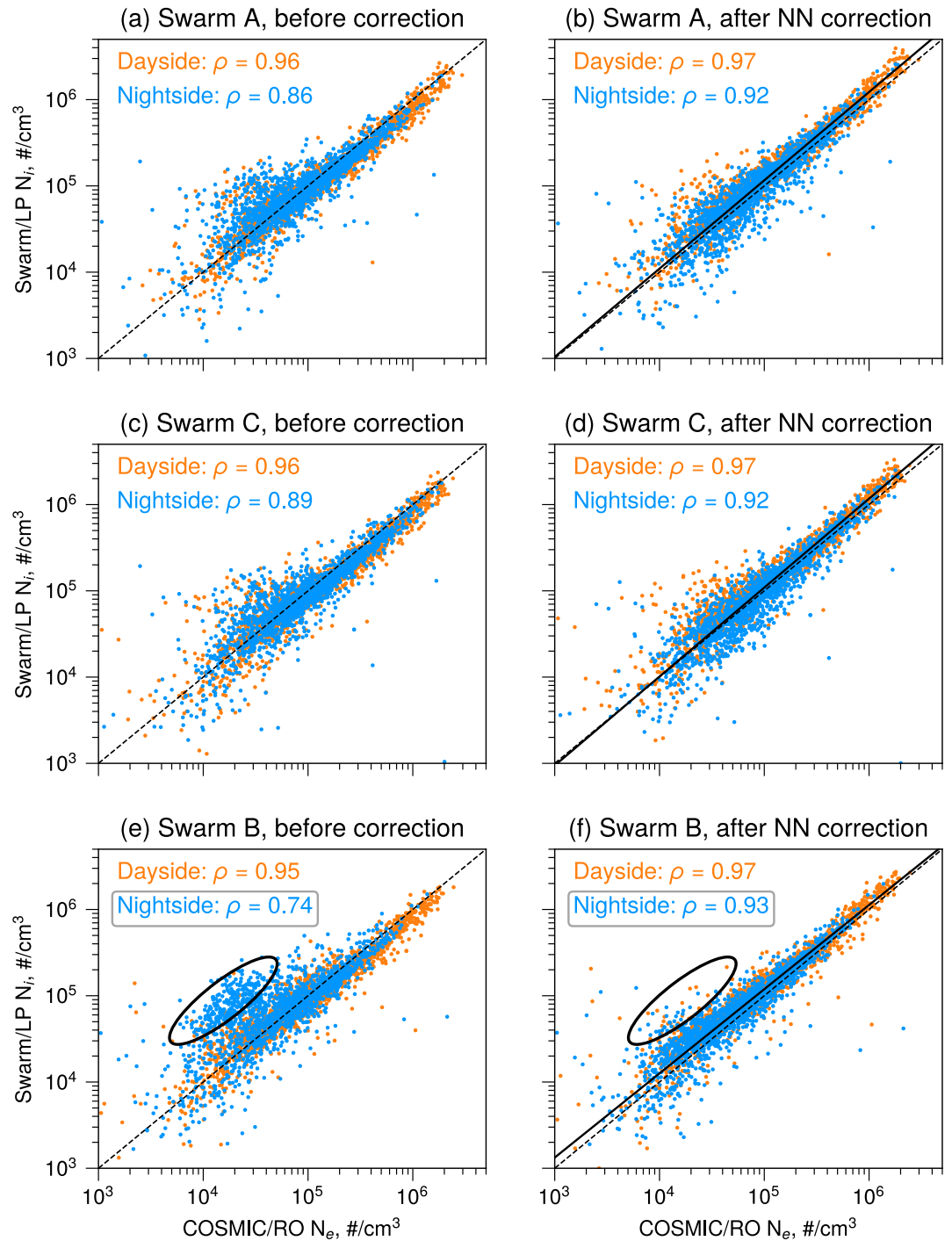


Figure 9. Independent validation of the NN- corrected LP data in conjunctions with COSMIC. The left column shows the scatter plots of COSMIC versus Swarm LP observations before the correction, and the right column shows the results after applying the NN-based correction developed in this study. The developed correction successfully removes the nighttime overestimation by Swarm LPs. For Swarm B, the correlation with COSMIC data on the night side increases from 0.74 to 0.93 (panels e, f).

Table 1
Calibration Coefficients Used to Adjust the NN-Corrected Swarm-LP Ion Densities to COSMIC Observations Using Equation 5

Satellite	c	d
Swarm A	0.977	0.052
Swarm B	1.025	-0.201
Swarm C	0.970	0.115

where LP^{cor} denotes the corrected Swarm LP values, and LP^{uncor} represents the uncorrected LP measurements. In our case, the skill score values (Table 2) show that the developed technique improves the agreement with COSMIC by around 25%.

5. Discussion

Several recent studies have reported discrepancies in Swarm LP ion density observations at nighttime, with overestimation at low and middle latitudes reaching 50% (Catapano et al., 2022; Fast et al., 2023; Pignalberi et al., 2022;

Smirnov et al., 2021; Xiong et al., 2022). In this study, we developed the first calibration technique for Swarm LP observations that resolves these discrepancies and can be used to correct the entire LP data set. Our calibration approach is to adjust the LP data to FP observations by developing NN-based correction models for each of the three Swarm satellites. These models are trained on sparse times when FP measurements are available and then extended to the entire duration of the LP data set.

The nighttime overestimation by LPs comes from the assumption of oxygen-only plasma in the ion density calculations (Catapano et al., 2022; Smirnov et al., 2021; Xiong et al., 2022). One of the recently developed products that relaxes the O^+ -only assumption is the SLIDEM product (Burchill & Lomidze, 2024; Pakhotin et al., 2022, 2023). It uses the OML theory and FP current measurements that allow for an improved derivation of full ion densities and effective ion masses. However, the SLIDEM methodology requires the bias to be strongly negative (around -3.5 V), which means that the availability of the SLIDEM densities is limited to times when TIIs are inactive. It should be emphasized that our technique is markedly different from the SLIDEM methodology, and instead can be viewed as a way to extend the FP data set for the entire duration of the Swarm mission. We perform climatological modeling of the logarithmic ratios between FP and LP ion densities, which mainly relate to changes in effective ion mass at Swarm altitudes. Namely, when the FP/LP log-ratios are negative, effective ion masses are lowered and LPs overestimate ion densities, while for positive log-ratios the ion masses are increased and the LP densities are underestimated (see Equation 2 and Section 2.1.2 for details). At the same time, the SLIDEM results, obtained from physical considerations, provide a useful means of comparison with our purely data-driven approach.

Developing a climatological model of the FP/LP ratios allows us to analyze their behavior on global scale. Our results demonstrate that the regions where Swarm LPs overestimate N_i form a clear double-crest structure on the nightside with peaks around 30° QDLat in both hemispheres. This pattern corresponds to decreased effective ion masses and most likely arises due to the downward ambipolar diffusion of light ions from the plasmasphere, which can be explained as follows. After sunset, the pressure of O^+ ions in the topside ionosphere rapidly decreases. In contrast, the pressure of light ions at higher altitudes remains almost constant (Heelis et al., 2022). This leads to a pressure gradient force that drives the plasmaspheric particles, mainly the H^+ and He^+ ions, downward and creates a dual-hemispheric pattern at mid-latitudes. This downward diffusion is one of the main mechanisms

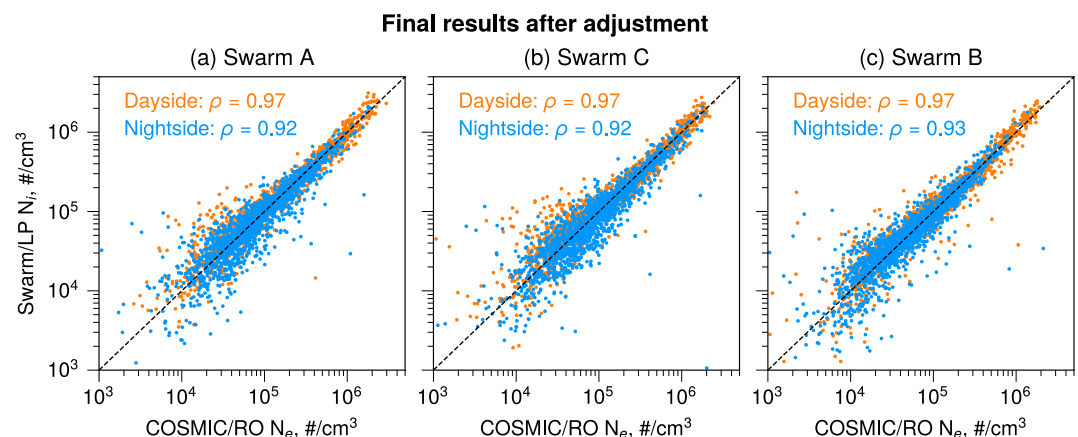


Figure 10. Comparison of the NN-calibrated LP data to COSMIC electron densities after the linear adjustment.

Table 2
Metrics Evaluated at the Conjunctions of the Three Swarm Satellites and COSMIC

Metric	Swarm A	Swarm C	Swarm B
Median bias, (cm ⁻³)	-11,910/ -733	-14,240/ -621	-4,046/ -142
Mean bias, (cm ⁻³)	-29,571/ -4,225	-32,967/ -2,997	-11,302/ 1,749
MAE, (cm ⁻³)	25,069/ 15,767	25,526/ 17,339	21,634/ 10,851
σ , (cm ⁻³)	96,106/ 88,173	93,228/ 86,175	82,762/ 72,195
Spearman ρ	0.93/ 0.95	0.93/ 0.95	0.86/ 0.96
% of values within a factor of 2	90.7/ 92.6	90.8/ 91.5	83.3/ 93.4
Skill score	0.23	0.24	0.25

Note. The metrics for the uncalibrated Swarm LP observations and those after the adjustment presented in this study are shown on the left and right, respectively. The better values for each metric are highlighted in bold. The developed Swarm/LP correction improves the agreement with COSMIC in all cases. (MAE = Median absolute error).

that allow the ionospheric F2-layer to be maintained at night (Titheridge, 1968b), and the existence of the corresponding nocturnal mid-latitude stripes has been well documented in literature since the 1960s (Balan et al., 1994; Titheridge, 1968b; Xiong et al., 2019; Zhong et al., 2019). In our study, the light-ion influence can be seen from $\sim 5^\circ$ QDLat up to about 60° QDLat. The poleward edges of the crests coincide fairly well with locations of the mid-latitude density trough (see Figure S1.11 in the Supporting Information S1). In turn, the midlatitude trough is known to map along magnetic field lines to the outer boundary of the plasmasphere (Heilig et al., 2022). This boundary, known as the plasmopause, is marked by a strong decrease of plasma density over a short distance of ~ 0.5 L-shell (Carpenter, 1963, 1966). The processes that govern the particle dynamics and transport beyond the plasmopause are vastly different from those inside the plasmasphere. Specifically, beyond the plasmopause the particles are moved by the magnetospheric convection (Nishida, 1966) and are not expected to experience the same downward transport as the plasmaspheric ones. The light-ion crests at Swarm altitudes are therefore not observed at latitudes above the approximate footprint of the plasmopause. Therefore, the spatial morphology of the nighttime ion density overestimation by Swarm LPs aligns very well with the presence of light ions transported downwards due to the ambipolar diffusion.

Notably, there appears a relatively sudden change from oxygen- to hydrogen-dominated layers in the topside ionosphere. This boundary was in the early years tentatively called “diffusive barrier,” for example, by Titheridge (1968a) and Park (1970). Nowadays, the charge exchange ($O + H^+ \rightleftharpoons O^+ + H$) is assigned as the main process for effectively separating the parts of the ionosphere. The ionization potentials of hydrogen and oxygen have quite similar values, and therefore hydrogen ions are soon neutralized when entering an oxygen-rich layer. The same is true for oxygen ions when rising into the hydrogen sphere. Presently, the term ‘upper transition height’ is more common for this boundary. Observations by the C/NOFS satellite revealed that the UTH is descending to around 500 km during solar minima (Heelis et al., 2009; Klenzing et al., 2011). The existence of this boundary is one of the explanations why the nighttime mid-latitude crests are weaker in Swarm A/C data compared to Swarm B. Swarm A and C are commonly below the UTH, thus light ions cannot easily reach their altitudes, while Swarm B is located closer to the UTH, which makes the impact of light ions more evident.

Our results demonstrate that the nocturnal crests are most pronounced when plasma densities at mid-latitudes are lowest, namely, during low solar activity and in the local winter hemispheres. For these conditions, the background O^+ density is reduced due to lower production rates. Consequently, the downward diffusion is more efficient and the light-ion fraction becomes more significant in the total composition, leading to a decrease in M_{eff} and the observed negative FP/LP log-ratios in Figures 5–7. Conversely, the light ion crests are weaker when the O^+ density is higher, such as in local summer hemispheres during solstices or in both hemispheres during equinoxes. This seasonal behavior is in very good agreement with the previous study by Zhong et al. (2019), who reported a similar seasonal pattern of the mid-latitude stripes using COSMIC data.

The light-ion crests at Swarm altitudes are strongly influenced by geomagnetic activity. Specifically, during geomagnetically active times these crests weaken and become confined to lower magnetic latitudes. There are several contributing processes that are worth discussing. One important factor is the position of the plasmopause. During geomagnetic storms the enhanced convection electric field erodes the plasmasphere, and the plasmopause

moves inward to lower L-shells (Nishida, 1966). The light ion crests investigated in this study should only be observed inside the plasmopause (L_{pp}). One of the empirical models that describes the L_{pp} position is the Carpenter and Anderson (1992) model:

$$L_{pp} = 5.6 - 0.46 \cdot Kp_{\max}, \quad (7)$$

where Kp_{\max} is the maximum Kp value over the past 36 hr. When Kp increases from 0.67 to 6, the plasmopause moves inward from $L = 5.3$ to $L = 2.8$. The dipole latitudes corresponding to the plasmopause footprints can be calculated as $\lambda_{pp} = \arccos(\sqrt{1/L})$, and have values of 64° and 53° , respectively. This means that with increasing geomagnetic activity, the poleward edges of the crests should shift by 11° toward the equator. This simple estimation agrees very well with our results in Figures 7c and 7d, where the poleward edges move from 60° to 50° QDLat. Other more recent empirical models, including O'Brien and Moldwin (2003) and Heilig and Lühr (2013), allow for the MLT-dependent plasmopause modeling, and the connection between the poleward edges of the light ion crests and the λ_{pp} positions can be explored in the future.

Another important aspect which contributes to the reduced strength of light-ion crests is the global increase in $hmF2$ during geomagnetic storms (Figure 8). As the $hmF2$ increases, the oxygen-dominated part of the ionosphere extends to higher altitudes, and therefore, the influence of a smaller light-ion fraction in the total ion composition would be seen as less pronounced. There are two candidate mechanisms that explain why $hmF2$ increases during storm-times: (a) high-latitude heating causing equatorward thermospheric winds resulting in traveling ionospheric disturbances (TIDs) that increase $hmF2$, or (b) local storm-time electric fields lifting the F-layer up due to the $E \times B$ drift. As noted by Pakhotin et al. (2023), these two processes have different latitudinal signatures: the former mechanism would lead to the $hmF2$ increase propagating from high latitudes toward the equator, while the opposite would be seen in the latter process. Pakhotin et al. (2023) performed a superposed epoch analysis of the effective ion mass for geomagnetic storms during the Swarm era and found a progressive increase of M_{eff} originating at high-latitudes and propagating toward the equator, suggestive of the thermospheric winds and TIDs as a prevailing mechanism. Our NN-based model currently uses the instantaneous Hp30 values without considering the time-history. For detailed investigations on the propagation of this $hmF2$ increase and its relation to the ion mass, it would be beneficial to include the time-derivative of the Hp and Dst indices into both the NET $hmF2$ model and the FP/LP calibration model developed in the present study. This would make detailed investigations of these phenomena possible.

Regions of positive FP/LP log-ratios present several interesting patterns. One example of such increase is visible at 450 km around geomagnetic equator at sunrise (Figure 5b). This region corresponds to a morning overshoot of the electron temperature (Stolle et al., 2011; Yang et al., 2020). The enhanced electron temperature sets up an ambipolar potential (e.g., Kervalishvili & Lühr, 2018). This vertical electric field reduces the gravitational pull on the ions. As a consequence, the effective g in the barometric formula is reduced, which causes larger scale-heights for the ion species. Therefore, heavier ions can reach the altitude of Swarm. This is in good agreement with the increase in M_{eff} around the same region seen in the recently developed SLIDEM product (Pakhotin et al., 2022). Another region of interest, corresponding to strongly positive log-ratios, is at high latitudes around the morning terminator. It is pronounced both in the model results and observations. Although it is currently unclear if the nature of this increase is geophysical or instrumental, it appears consistently across several studies that are based on different methodologies. Pakhotin et al. (2023) observed a similar increase in M_{eff} using SLIDEM data. If these enhanced FP/LP log-ratios are indeed of geophysical origin, the presence of molecular ions, which have masses higher than oxygen, could explain the increased M_{eff} . However, the question of why molecular ions are present in appreciable concentrations in those regions, at altitudes as high as 450–510 km, remains to be answered, and therefore this topic warrants further investigation.

The calibration of Swarm LP ion densities undertaken in this study has shown that the FP observations provide a very good baseline for the calibration. We compared the corrected LP ion densities to independent COSMIC radio occultation observations, and found that the LT-dependent bias of Swarm LPs was successfully resolved by applying the developed correction. The strongest improvement is achieved for Swarm B, where the nighttime correlation with COSMIC increased from 0.74 to 0.93 after using the developed calibration model.

To briefly recapitulate the results, Figure 11 shows a summary plot for the Swarm B LP calibration. We plot the comparison of averaged ion densities within $\pm 60^\circ$ QDLat for the ascending and descending half-orbits to compare

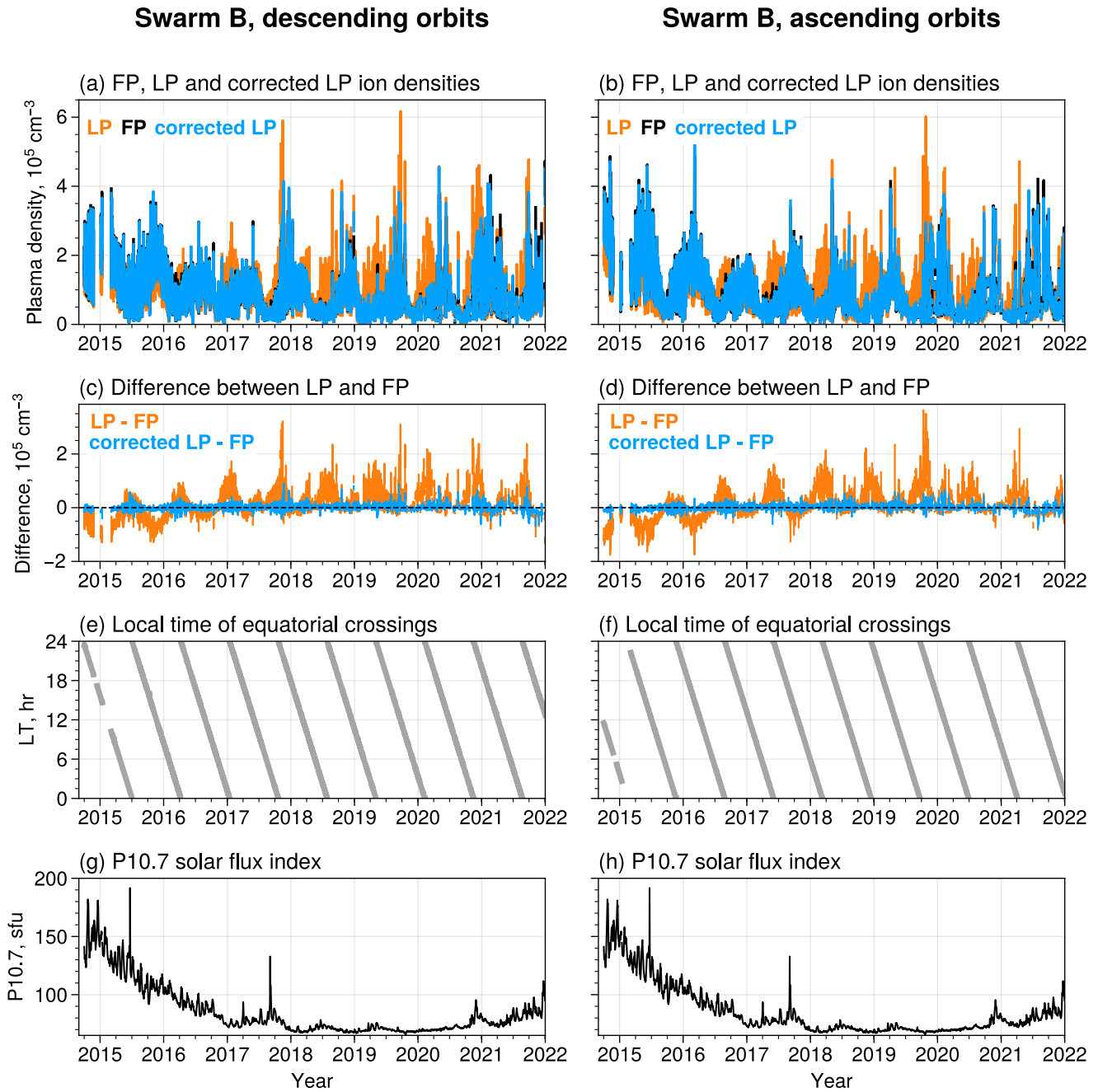


Figure 11. Summary comparison of uncorrected and NN-calibrated LP ion densities with FP measurements for Swarm B. Panels (a) and (b) show ion densities averaged within $\pm 60^\circ$ QDLat for the descending and ascending half-orbits. The LP data are shown in orange, the reference FP data are plotted in black and the NN-corrected LP densities are given in blue. Panels (c) and (d) demonstrate the difference between LP and FP observations for the uncorrected LP (orange) and NN-corrected LP data (blue). The dashed black line indicates zero difference. The two bottom rows show the local time of the equatorial crossings, and the P10.7 solar flux index.

the FP data with the raw and corrected LP observations. Here we are most interested in the effects of oxygen-hydrogen transition at low and middle latitudes and therefore do not consider observations at high latitudes. Panels (c) and (d) show differences between the two instruments before and after applying the developed NN-based calibration. One can see that the uncorrected LP data exhibit deviations from FP measurements that are strongly dependent on solar activity levels and local time. In particular, prior to 2017, LPs mostly underestimated electron densities, except at nighttime where there was a slight overestimation. These periods corresponded to high solar activity. The preliminary calibration results by Lomidze et al. (2018) against the ISR data were based

on these periods. Therefore, their calibration coefficients were uniform in local time, as the LT-effects could not be resolved from a limited number of conjunctions with the ISRs. Later on, however, the solar activity decreased and by the minimum of solar cycle 24/25, the morphology of the LP-FP difference changed dramatically. Starting from 2018, one could see a good agreement between LP and FP on the dayside (Figures 11d and 11e) but a strong positive bias of LPs when the Swarm satellites were on the nightside. The value of this bias averaged for the entire half-orbits is of the order of $2 \cdot 10^5 - 3 \cdot 10^5 \text{ cm}^{-3}$, with local differences being even more pronounced. However, all of these signatures are resolved by using the NN-based calibration developed in this study. One can see that both for periods of high and low solar activity, the difference of the corrected FP and LP values is around zero. Therefore, the developed NN-based correction improves the LP ion density in comparison to both the FP observations and independent COSMIC data and presents a very powerful tool for data calibration.

5.1. Future Work

The work on calibrating LP ion densities using neural networks, started in this paper, offers many opportunities for future investigations. In particular, the present study is based on data from the start of FP operations in October 2014 until January 2022. The Swarm mission has undergone several orbital manoeuvres in 2023, and therefore once the orbital changes are finalized, the models may be retrained based on the full data set. Additionally, the model training set mainly comes from the relatively quiet solar cycle 24. The current solar cycle 25 has, to date, been stronger than the previous one (Lugaz et al., 2023), and therefore the calibration model should be retrained as the cycle progresses to take higher solar activity conditions into account. Furthermore, as shown in Figure 1b, between 2020 and 2021 the number of FP observations, and hence the points for training our calibration model, decreased significantly, from $\sim 10^5$ to $\sim 10^3$ points per day. This may bias NNs to favor periods with more data over those with fewer points. Thus, it may be beneficial to make a special adjustment to the loss function to compensate for the reduced data coverage. One possible technique to achieve this is to introduce weights inversely proportional to the number of observations per day, and use these weights inside the loss function. This will help to make the model more balanced and will give equal emphasis to all periods, even those with fewer data points.

The NN architecture can be updated in the future in order to produce a unified calibration model for all Swarm satellites. It is possible to combine these models and evaluate, whether systematic differences exist between LPs, by experimenting with two such models - one driven purely by geophysical inputs (including altitude), and another one incorporating the satellite ID as input (e.g., using the one-hot-encoding (James et al., 2013)). Similar performance would indicate that the instruments are consistent, whereas the differences in results may show additional instrumental offsets. In the present study, we have shown that Swarm A and C satellites are very consistent and therefore inter-satellite differences are not expected.

Another important step is to perform additional validation of the corrected LP ion densities by the independent data. In this study, we already undertook the independent validation by the COSMIC observations. It would also be useful to validate the NN-corrected data by the incoherent scatter radar observations, which represent the “gold standard” of ionospheric density measurements in the topside ionosphere. Recently, Fast et al. (2023) compared the LP data with the high-latitude ISRs (PFISR and RISR) using several thousand conjunctions. These comparisons can be extended to other ISRs, and all Swarm ion density products can undergo this independent validation. It is worth mentioning that for some radars, the number of conjunctions may be limited and therefore statistical climatological averages can also provide means of comparisons. Moreover, the ISR observations can be used to calculate the percentages of light ions in the total ion composition. These observations can also be utilized to validate the FP/LP ratios modeled in this study in terms of effective ion masses.

In this study, it was shown that the FP/LP models exhibit a strong response to geomagnetic activity. The developed models can be retrained in the future to account for the progression of geomagnetic storms in time. In this study, we used the Hp30 index, due to its potential in representing the magnetospheric convection which is of direct relevance for plasmaspheric dynamics (Thomsen, 2004), but additional indices can be considered in the future, including the SYM-H index and its time-derivative. These inputs are frequently employed in magnetospheric modeling (e.g., Stephens et al., 2019) and can also be included into the NN-based ionospheric simulations. After adding these variables to both the Swarm NN-based calibration model and the NET topside model, a comparative analysis could be performed that could answer the question regarding the behavior of the midlatitude light-ion bands and how they disappear during disturbed conditions. In particular, one would be able to see if

during periods of high Kp the increase in *hmF2* occurs at high or low latitudes first, and therefore one could assess whether thermospheric winds or local electric fields provide the main mechanism of *hmF2* increase.

6. Conclusions

In this study, we developed a neural network model to calibrate ion densities measured by Swarm Langmuir Probes. Our findings can be summarized as follows:

- Swarm LPs overestimate ion densities on the nightside due to the assumption of O⁺-only plasma in the data processing. As a reference for the calibration, we use the FP observations that are free of composition assumptions. The ratios between the two instruments relate to effective ion masses in the topside ionosphere.
- The nighttime overestimation by the LPs presents a double-crest structure with peaks around ~30° QDLat in both hemispheres. This pattern mimics the morphology of light ions diffusing downward from the plasmasphere. The crests map along the magnetic field lines to plasmaspheric L-shells, and are only observed at latitudes below the plasmopause footprints.
- The nighttime overestimation is stronger for Swarm B, owing to its higher altitude and closeness to the UTH. Due to the charge exchange with neutral oxygen, the H⁺ ions are less prevalent at the altitudes of Swarm A/C that are below the UTH, and therefore the crests are weaker for Swarm A and C satellites.
- The light-ion crests intensify when the background O⁺ densities at mid-latitudes are low. They also show a strong seasonal dependence. During solstices, the crests are stronger in the local winter hemispheres, while during equinoxes, the crests show hemispheric symmetry and are weaker due to generally higher O⁺ densities.
- The light-ion crests diminish with increasing geomagnetic activity due to a global increase in *hmF2*, and become confined to lower latitudes due to the erosion of the plasmasphere and the plasmopause moving inward.
- The NN-based calibration developed in this study resolves the nighttime overestimation by the LPs. The NN-corrected ion densities are in excellent agreement with both the FP observations and independent COSMIC RO data.

The correction model developed in this study significantly improves the quality of LP observations, which constitute one of the most widely used ionospheric data sets. Swarm LP data are frequently used for modeling of the topside ionosphere (e.g., Bilitza & Xiong, 2021; Pezzopane et al., 2024), and the calibrated LP observations may contribute to the improvement of the widely used ionospheric models, such as the IRI (Bilitza et al., 2022). Furthermore, the combined interpretation of the developed calibration model in combination with other models or observations can give useful hints about the ion composition and the dynamics of the topside ionosphere. In that regard, our study shows the importance of both the FP and LP instruments on Swarm, and therefore, it would be beneficial to increase the operational times of Swarm FPs. Last but not least, our study highlights the need for more ion composition data in the topside ionosphere, and thus it is recommended for the new and planned ionospheric missions to include instruments that allow probing different ion species.

Data Availability Statement

The model and the source code are available in open access (Smirnov, 2024). The data used here are publicly available. COSMIC data can be downloaded via UCAR <https://cdaac-www.cosmic.ucar.edu/>. Swarm data were obtained from ESA <https://earth.esa.int/web/guest/swarm/data-access>. The F10.7 index was downloaded from the OMNIWeb database omniweb.gsfc.nasa.gov. The Hp30 index is provided by GFZ Potsdam <https://kp.gfz-potsdam.de/hp30-hp60>.

References

- Aponte, N., Brum, C. G., Sulzer, M. P., & González, S. A. (2013). Measurements of the O⁺ to H⁺ transition height and ion temperatures in the lower topside ionosphere over Arecibo for equinox conditions during the 2008–2009 extreme solar minimum. *Journal of Geophysical Research: Space Physics*, 118(7), 4465–4470. <https://doi.org/10.1002/jgra.50416>
- Balan, N., Bailey, G., Nair, R. B., & Titheridge, J. (1994). Nighttime enhancements in ionospheric electron content in the northern and southern hemispheres. *Journal of Atmospheric and Terrestrial Physics*, 56(1), 67–79. [https://doi.org/10.1016/0021-9169\(94\)90177-5](https://doi.org/10.1016/0021-9169(94)90177-5)
- Bilitza, D. (2009). Evaluation of the IRI-2007 model options for the topside electron density. *Advances in Space Research*, 44(6), 701–706. <https://doi.org/10.1016/j.asr.2009.04.036>
- Bilitza, D., Pezzopane, M., Truhlik, V., Altadill, D., Reinisch, B. W., & Pignalberi, A. (2022). The international reference ionosphere model: A review and description of an ionospheric benchmark. *Reviews of Geophysics*, 60(4), e2022RG000792. <https://doi.org/10.1029/2022rg000792>

Acknowledgments

Artem Smirnov is supported by Deutsche Forschungsgemeinschaft (DFG), project number 520916080. Open Access funding enabled and organized by Projekt DEAL.

- Bilitza, D., & Xiong, C. (2021). A solar activity correction term for the IRI topside electron density model. *Advances in Space Research*, 68(5), 2124–2137. <https://doi.org/10.1016/j.asr.2020.11.012>
- Brace, L. H. (1998). Langmuir probe measurements in the ionosphere. *Measurement Techniques in Space Plasmas: Particles*, 102, 23–35. <https://doi.org/10.1029/gm102p0023>
- Burchill, J. K., & Lomidze, L. (2024). Calibration of Swarm ion density, drift, and effective mass measurements. *Earth and Space Science*, 11(3), e2023EA003463. <https://doi.org/10.1029/2023ea003463>
- Camporeale, E. (2019). The challenge of machine learning in space weather: Nowcasting and forecasting. *Space Weather*, 17(8), 1166–1207. <https://doi.org/10.1029/2018sw002061>
- Carpenter, D., & Anderson, R. (1992). An ISEE/whistler model of equatorial electron density in the magnetosphere. *Journal of Geophysical Research*, 97(A2), 1097–1108. <https://doi.org/10.1029/91ja01548>
- Carpenter, D. L. (1963). Whistler evidence of a ‘knee’ in the magnetospheric ionization density profile. *Journal of Geophysical Research*, 68(6), 1675–1682. <https://doi.org/10.1029/jz068i006p01675>
- Carpenter, D. L. (1966). Whistler studies of the plasmapause in the magnetosphere: I. Temporal variations in the position of the knee and some evidence on plasma motions near the knee. *Journal of Geophysical Research*, 71(3), 693–709. <https://doi.org/10.1029/jz071i003p00693>
- Catapano, F., Buchert, S., Qamili, E., Nilsson, T., Bouffard, J., Siemes, C., et al. (2022). Swarm Langmuir probes’ data quality validation and future improvements. *Geoscientific Instrumentation, Methods and Data Systems*, 11(1), 149–162. <https://doi.org/10.5194/gi-11-149-2022>
- Chen, Z., Zhou, K., Li, H., Wang, J.-S., Ouyang, Z., & Deng, X. (2023). Global TEC map fusion through a hybrid deep learning model: RFGAN. *Space Weather*, 21(1), e2022SW003341. <https://doi.org/10.1029/2022sw003341>
- Chu, X., Ma, D., Bortnik, J., Tobiska, W. K., Cruz, A., Bouwer, S. D., et al. (2021). Relativistic electron model in the outer radiation belt using a neural network approach. *Space Weather*, 19(12), e2021SW002808. <https://doi.org/10.1029/2021sw002808>
- DTU. (2019). *Swarm 11b product definition*. National Space Institute, Technical University of Denmark (DTU). Retrieved from <https://earth.esa.int/eogateway/documents/20142/37627/swarm-level-1b-product-definition-specification.pdf/12995649-fbcb-6ae2-5302-2269fecf5a08>
- Fast, H., Koustov, A., & Gillies, R. (2023). Validation of swarm Langmuir probes by incoherent scatter radars at high latitudes. *Remote Sensing*, 15(7), 1846. <https://doi.org/10.3390/rs15071846>
- Friis-Christensen, E., Lühr, H., & Hulot, G. (2006). Swarm: A constellation to study the Earth’s magnetic field. *Earth Planets and Space*, 58(4), 351–358. <https://doi.org/10.1186/bf03351933>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Hargreaves, J. K. (1992). *The solar-terrestrial environment: An introduction to geospace—the science of the terrestrial upper atmosphere, ionosphere, and magnetosphere*. Cambridge university press.
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (Vol. 2). Springer.
- Heelis, R., Coley, W., Burrell, A., Hairston, M., Earle, G., Perdue, M., et al. (2009). Behavior of the O⁺/H⁺ transition height during the extreme solar minimum of 2008. *Geophysical Research Letters*, 36(18), L00C03. <https://doi.org/10.1029/2009gl038652>
- Heelis, R., Depew, M., Chen, Y.-J., & Perdue, M. (2022). Ionospheric connections (ICON) ion velocity meter (IVM) observations of the equatorial ionosphere at solar minimum. *Space Science Reviews*, 218(8), 68. <https://doi.org/10.1007/s11214-022-00936-w>
- Heilig, B., & Lühr, H. (2013). New plasmapause model derived from CHAMP field-aligned current signatures. *Annales Geophysicae*, 31(3), 529–539. <https://doi.org/10.5194/angeo-31-529-2013>
- Heilig, B., Stolle, C., Kervalishvili, G., Rauberg, J., Miyoshi, Y., Tsuchiya, F., et al. (2022). Relation of the plasmapause to the midlatitude ionospheric trough, the sub-auroral temperature enhancement and the distribution of small-scale field aligned currents as observed in the magnetosphere by THEMIS, RBSP, and ARASE, and in the topside ionosphere by swarm. *Journal of Geophysical Research: Space Physics*, 127(3), e2021JA029646. <https://doi.org/10.1029/2021ja029646>
- IRF. (2017). Faceplate plasma density. Retrieved from https://swarm-diss.eo.esa.int/#swarm%2FAdvanced%2FPlasma_Data%2F16_Hz_Faceplate_plasma_density
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.
- Kervalishvili, G. N., & Lühr, H. (2018). Climatology of air upwelling and vertical plasma flow in the terrestrial cusp region: Seasonal and imp-dependent processes. Magnetic fields in the solar system: Planets, moons and solar wind interactions (pp. 293–329).
- Klenzing, J., Simões, F., Ivanov, S., Heelis, R., Bilitza, D., Pfaff, R., & Rowland, D. (2011). Topside equatorial ionospheric density and composition during and after extreme solar minimum. *Journal of Geophysical Research*, 116(A12), A12330. <https://doi.org/10.1029/2011ja017213>
- Knudsen, D., Burchill, J., Buchert, S., Eriksson, A., Gill, R., Wahlund, J.-E., et al. (2017). Thermal ion imagers and Langmuir probes in the Swarm electric field instruments. *Journal of Geophysical Research: Space Physics*, 122(2), 2655–2673. <https://doi.org/10.1002/2016ja022571>
- Laundal, K. M., & Richmond, A. D. (2017). Magnetic coordinate systems. *Space Science Reviews*, 206(1), 27–59. <https://doi.org/10.1007/s11214-016-0275-y>
- Lebreton, J.-P., Stverak, S., Travnicek, P., Maksimovic, M., Klinge, D., Merikallio, S., et al. (2006). The ISL Langmuir probe experiment processing onboard DEMETER: Scientific objectives, description and first results. *Planetary and Space Science*, 54(5), 472–486. <https://doi.org/10.1016/j.pss.2005.10.017>
- Lomidze, L., Knudsen, D. J., Burchill, J., Kouznetsov, A., & Buchert, S. C. (2018). Calibration and validation of Swarm plasma densities and electron temperatures using ground-based radars and satellite radio occultation measurements. *Radio Science*, 53(1), 15–36. <https://doi.org/10.1002/2017rs006415>
- Lugaz, N., Liu, H., Carter, B. A., Gannon, J., Zou, S., & Morley, S. K. (2023). New space companies meet a “normal” solar maximum. *Space Weather*, 21(9), e2023SW003702. <https://doi.org/10.1029/2023sw003702>
- Lühr, H., Rother, M., Maus, S., Mai, W., & Cooke, D. (2003). The diamagnetic effect of the equatorial Appleton anomaly: Its characteristics and impact on geomagnetic field modeling. *Geophysical Research Letters*, 30(17), 1906. <https://doi.org/10.1029/2003gl017407>
- Mott-Smith, H. M., & Langmuir, I. (1926). The theory of collectors in gaseous discharges. *Physical Review*, 28(4), 727–763. <https://doi.org/10.1103/physrev.28.727>
- Nishida, A. (1966). Formation of plasmapause, or magnetospheric plasma knee, by the combined action of magnetospheric convection and plasma escape from the tail. *Journal of Geophysical Research*, 71(23), 5669–5679. <https://doi.org/10.1029/jz071i023p05669>
- O’Brien, T., & Moldwin, M. (2003). Empirical plasmapause models from magnetic indices. *Geophysical Research Letters*, 30(4), 1152. <https://doi.org/10.1029/2002gl016007>
- Pakhotin, I., Burchill, J., Förster, M., & Lomidze, L. (2022). The swarm Langmuir probe ion drift, density and effective mass (SLIDEM) product. *Earth Planets and Space*, 74(1), 109. <https://doi.org/10.1186/s40623-022-01668-5>

- Pakhotin, I., Burchill, J., Förster, M., & Lomidze, L. (2023). Light ion dynamics in the topside ionosphere and plasmasphere during geomagnetic storms. *Earth Planets and Space*, 75(1), 62. <https://doi.org/10.1186/s40623-023-01818-3>
- Pan, Q., Xiong, C., Lühr, H., Smirnov, A., Huang, Y., Xu, C., et al. (2024). Machine learning based modeling of thermospheric mass density. *Space Weather*, 22(5), e2023SW003844. <https://doi.org/10.1029/2023sw003844>
- Park, C. (1970). Whistler observations of the interchange of ionization between the ionosphere and the protonosphere. *Journal of Geophysical Research*, 75(22), 4249–4260. <https://doi.org/10.1029/ja075i022p04249>
- Pezzopane, M., Pignalberi, A., Pietrella, M., Haralambous, H., Prol, F., Nava, B., et al. (2024). An update of the nequick-corr topside ionosphere modeling based on new datasets. *Atmosphere*, 15(4), 498. <https://doi.org/10.3390/atmos15040498>
- Pignalberi, A., Pezzopane, M., Coco, I., Piersanti, M., Giannattasio, F., De Michelis, P., et al. (2022). Inter-calibration and statistical validation of topside ionosphere electron density observations made by CSES-01 mission. *Remote Sensing*, 14(18), 4679. <https://doi.org/10.3390/rs14184679>
- Resendiz Lira, P. A., & Marchand, R. (2021). Simulation inference of plasma parameters from Langmuir probe measurements. *Earth and Space Science*, 8(3), e2020EA001344. <https://doi.org/10.1029/2020ea001344>
- Rishbeth, H., & Garriott, O. K. (1969). *Introduction to ionospheric physics*. Introduction to ionospheric physics.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- Smirnov, A. (2024). *Model files and codes for "Calibration of Swarm plasma densities overestimation using neural networks" (2.0)*. Zenodo. <https://doi.org/10.5281/zenodo.13259795>
- Smirnov, A., Shprits, Y., Prol, F., Lühr, H., Berrendorf, M., Zhelavskaya, I., & Xiong, C. (2023). A novel neural network model of Earth's topside ionosphere. *Scientific Reports*, 13(1), 1303. <https://doi.org/10.1038/s41598-023-28034-z>
- Smirnov, A., Shprits, Y., Zhelavskaya, I., Lühr, H., Xiong, C., Goss, A., et al. (2021). Inter-calibration of the plasma density measurements in Earth's topside ionosphere. *Journal of Geophysical Research: Space Physics*, 126(10), e2021JA029334. <https://doi.org/10.1029/2021ja029334>
- Stephens, G., Sitnov, M., Korth, H., Tsyganenko, N., Ohtani, S., Gkioulidou, M., & Ukhorskiy, A. (2019). Global empirical picture of magnetospheric substorms inferred from multimission magnetometer data. *Journal of Geophysical Research: Space Physics*, 124(2), 1085–1110. <https://doi.org/10.1029/2018ja025843>
- Stolle, C., Liu, H., Truhlik, V., Lühr, H., & Richards, P. (2011). Solar flux variation of the electron temperature morning overshoot in the equatorial F region. *Journal of Geophysical Research*, 116(A4), A04308. <https://doi.org/10.1029/2010ja016235>
- Sugiyama, M., Krauledat, M., & Müller, K.-R. (2007). Covariate shift adaptation by importance weighted cross validation. *Journal of Machine Learning Research*, 8(5).
- Tancik, M., Srinivasan, P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., et al. (2020). Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems*, 33, 7537–7547.
- Thomsen, M. (2004). Why kp is such a good measure of magnetospheric convection. *Space Weather*, 2(11). <https://doi.org/10.1029/2004sw000089>
- Titheridge, J. (1968a). Calculations of diurnal changes in the exosphere. *Journal of Atmospheric and Terrestrial Physics*, 30(11), 1843–1856. [https://doi.org/10.1016/0021-9169\(68\)90027-5](https://doi.org/10.1016/0021-9169(68)90027-5)
- Titheridge, J. (1968b). The maintenance of the night ionosphere. *Journal of Atmospheric and Terrestrial Physics*, 30(11), 1857–1875. [https://doi.org/10.1016/0021-9169\(68\)90028-7](https://doi.org/10.1016/0021-9169(68)90028-7)
- Tonks, L., & Langmuir, I. (1929). A general theory of the plasma of an arc. *Physical Review*, 34(6), 876–922. <https://doi.org/10.1103/physrev.34.876>
- Vaishnav, R., Jin, Y., Mostafa, M. G., Aziz, S. R., Zhang, S.-R., & Jacobi, C. (2021). Study of the upper transition height using ISR observations and IRI predictions over Arecibo. *Advances in Space Research*, 68(5), 2177–2185. <https://doi.org/10.1016/j.asr.2020.10.010>
- Wang, H., He, Y., & Lühr, H. (2022). Analysis of ionospheric compressional waves and electron density oscillation during storm periods using Swarm observations. *Journal of Geophysical Research: Space Physics*, 127(8), e2022JA030409. <https://doi.org/10.1029/2022ja030409>
- 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. *Journal of Geophysical Research: Space Physics*, 127(5), e2022JA030275. <https://doi.org/10.1029/2022ja030275>
- Xiong, C., Lühr, H., Sun, L., Luo, W., Park, J., & Hong, Y. (2019). Long-lasting latitudinal four-peak structure in the nighttime ionosphere observed by the Swarm constellation. *Journal of Geophysical Research: Space Physics*, 124(11), 9335–9347. <https://doi.org/10.1029/2019ja027096>
- Yamazaki, Y., Matzka, J., Stolle, C., Kervalishvili, G., Rauberg, J., Bronkalla, O., et al. (2022). Geomagnetic activity index HPO. *Geophysical Research Letters*, 49(10), e2022GL098860. <https://doi.org/10.1029/2022gl098860>
- Yang, T.-Y., Park, J., Kwak, Y.-S., Oyama, K.-I., Minow, J. I., & Lee, J. (2020). Morning overshoot of electron temperature as observed by the Swarm constellation and the international space station. *Journal of Geophysical Research: Space Physics*, 125(2), e2019JA027299. <https://doi.org/10.1029/2019ja027299>
- Zhelavskaya, I. S., Shprits, Y. Y., & Spasojević, M. (2017). Empirical modeling of the plasmasphere dynamics using neural networks. *Journal of Geophysical Research: Space Physics*, 122(11), 11–227. <https://doi.org/10.1002/2017ja024406>
- Zhong, J., Lei, J., Yue, X., Luan, X., & Dou, X. (2019). Middle-latitude band structure observed in the nighttime ionosphere. *Journal of Geophysical Research: Space Physics*, 124(7), 5857–5873. <https://doi.org/10.1029/2018ja026059>