



OPEN

On the link between the topside ionospheric effective scale height and the plasma ambipolar diffusion, theory and preliminary results

Alessio Pignalberi¹, Michael Pezzopane¹, Bruno Nava² & Pierdavide Coisson³

Over the years, an amount of models relying on effective parameters were implemented in the challenging issue of the topside ionosphere description. These models are based on different analytical functions, but all of them depend on a parameter called effective scale height, that is deduced from topside electron density measurements. As their names state, they are effective in reproducing the topside electron density profile only when applied to the analytical function used to derive them. Then, in principle, they do not have any physical meaning. It is the goal of this paper to mathematically link the effective scale height modeled through the Epstein layer to the vertical scale height theoretically deduced from the plasma ambipolar diffusion theory. Firstly, effective and theoretical scale heights are linked through a mathematical relation by showing that they tend to each other in the topside ionosphere. Secondly, their connection is preliminarily demonstrated by calculating effective scale height values from the entire COSMIC/FORMOSAT-3 radio occultation dataset. Thirdly, a possible connection between the vertical gradient of the topside scale height (as obtained by COSMIC/FORMOSAT-3 satellites) and the electron temperature (as obtained by ESA Swarm B satellite) is studied by highlighting corresponding similarities in the diurnal, seasonal, solar activity, and latitudinal variability.

In the context of the Space Weather, the ionosphere plays a fundamental role as the medium coupling the external forcing given by solar radiation, solar wind, and magnetosphere, to the inner Earth's atmosphere. Specifically, the topside part of the ionosphere, extending from the F-layer peak (corresponding to the ionospheric electron density maximum) to the plasmasphere, is of utmost importance because it contains the largest fraction of the ionospheric total electron content^{1,2}.

In the topside, the electron density decreases monotonically as the ion population smoothly transitions from heavy O^+ ions, dominating the lower part of the F region, to lighter H^+ and He^+ ions above. Moreover, the increase of the plasma temperature in the topside and the presence of the dynamical forcing imposed by magnetic and electric fields and by collisions with neutrals, makes the picture quite complex. The rate of electron density decrease in the topside is linked to several physical and chemical phenomena explained by the plasma ambipolar diffusion theory through the introduction of a theoretical vertical scale height¹⁻⁴ (*VSH*). The *VSH* generalizes the concept of plasma scale height (H_p) by including all the relevant physical and chemical concepts in its definition^{3,4}. Specifically, *VSH* is equal to H_p only under diffusive equilibrium conditions, i.e., when the collisions with neutrals and the vertical gradient of plasma temperature can be neglected.

Despite their exact conceptual definition, both *VSH* and H_p are of difficult application for operational purposes. This is because they require a knowledge of the physical state of the plasma in terms of temperature, chemical state, and mean ion mass for the whole topside profile, that is currently unaffordable. This is why, in the past, a more direct and practical approach was developed based on the use of electron density measurements

¹Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata 605, 00143 Rome, Italy. ²The Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, 34151 Trieste, Italy. ³Université de Paris, Institut de physique du globe de Paris, CNRS, F-75005 Paris, France. ✉email: alessio.pignalberi@ingv.it

collected in the topside ionosphere by different techniques. Topside electron density measurements are limited both in space and time by the difficulties in probing an ionospheric region hidden to the widely spread ground-based ionosondes. More sophisticated and expensive techniques and instruments like topside sounders, radio occultation (RO) satellites, and incoherent scatter radars (ISR), were employed to this task. Over the years, several modeling techniques were developed to obtain the most reliable picture of the topside vertical electron density profile through an increasingly sophisticated description of the topside effective vertical scale height^{5–19} (H). These vertical scale heights gained the adjective “effective” because they are effective in reproducing the topside electron density profile only when applied to the analytical function used to derive them from either measured topside profiles or measured electron density values at low Earth orbit (LEO) satellite altitude. Different functions were applied by several authors, the most popular being exponential, Chapman, and Epstein families of functions^{12,20–24}. The first studies were based on constant H values, but the growing bulk of topside measurements suggested the need to allow for an increasing trend of H with altitude^{13,14,25–27}. This is why the most recent topside models are based on effective scale heights varying with altitude.

Despite the increasing complexity and reliability of modern topside models based on effective topside scale heights, in principle, modeled H values have no clear connection with theoretical VSH . This happens because these modeled H values do not rely on the physical quantities describing the topside electron density distribution (for example, ions distribution, plasma temperature, dynamics, and collisions), but only on the electron density profile itself which is indeed the final result of different physical mechanisms operating at the same time. This makes it difficult to associate the observed H variability to physical parameters, in order to perform a prediction of the topside ionosphere variability based on theoretical expectations^{3,4,28}. The goal of this work is to shed some light on the connection between modeled H and theoretical VSH (as derived by plasma ambipolar diffusion theory) through mathematical deductions, modeling approaches, and measured topside datasets.

Recently, Pignalberi et al.¹³ have proposed a method to derive a fully analytical expression of the effective topside scale height based on the analytical inversion of the semi-Epstein function (applied by both International Reference Ionosphere²⁹ (IRI) and NeQuick³⁰ models to represent the topside profile) as a function of the electron density and F-layer peak parameters. They applied the proposed methodology to a selected reliable dataset of COSMIC/FORMOSAT-3 (hereafter, COSMIC for brevity) RO profiles by highlighting a very distinct linear increasing trend of H for the lowest part of the topside ionosphere. By linearly fitting calculated H , they were able to obtain the H 's intercept and slope values for each COSMIC topside profile. These coefficients allow to model H , hence the topside electron density profile, in the topside region probed by COSMIC satellites: from the height of the F-layer peak to about 800 km.

In this study, the connection between the fully analytical scale height H of Pignalberi et al.¹³ and the theoretical VSH (from plasma ambipolar diffusion theory) is mathematically studied. This is done to link empirical modeled effective parameters to theoretical physical quantities. After having analytically derived the relation between effective and theoretical parameters, preliminary applications are shown based on effective parameters derived from the entire COSMIC RO dataset (by applying the approach of Pignalberi et al.¹³) and measured parameters from the European Space Agency (ESA) Swarm mission.

Linking the Epstein layer effective vertical scale height to the plasma ambipolar diffusion theory

The plasma ambipolar diffusion theory for the topside ionosphere. The motion of plasma in the topside ionosphere is dominated by diffusion along magnetic field lines. For our purposes, we can consider the ionospheric plasma as a minor gas mixture, formed by electrons and ions, immersed in the major neutral atmosphere gas. When speaking about ions in the topside, we mean a mixture of O^+ , H^+ , and He^+ ions, with relative percentage variable with altitude. Then, motions of ions and electrons can be described by the Navier–Stokes equation of motion for both charged plasma components. Here, to derive the vertical scale height, we follow the simplified description of plasma diffusion made by Rishbeth and Garriott¹ and Ratcliffe². Specifically, only the vertical diffusion (along the z axis) is here considered. However, in the ionosphere F region, the plasma diffuses along the geomagnetic field lines. The extension to the general case, i.e., plasma diffusion along sloped geomagnetic field lines, is straightforward by considering the field lines inclination^{1,2}.

Equations of motion for ions and electrons in the vertical component are:

$$\begin{cases} \frac{\partial}{\partial z}(N_i k_B T_i) = -N_i m_i g + N_i e E_z - N_i m_i \nu_{in}(W_i - W_n), \\ \frac{\partial}{\partial z}(N_e k_B T_e) = \underbrace{-N_e m_e g}_{\text{Gravity field}} - \underbrace{N_e e E_z}_{\text{Electric field}} - \underbrace{N_e m_e \nu_{en}(W_e - W_n)}_{\text{Collisions with neutrals}}, \end{cases} \quad (1)$$

where the following physical quantities are defined:

- N_i and N_e are the ions and electrons densities, respectively;
- k_B is the Boltzmann's constant;
- T_i and T_e are the ions and electrons temperatures, respectively;
- m_i and m_e are the ions and electrons masses, respectively;
- g is the acceleration due to the gravity field;
- e is the electric charge unit;
- E_z is the vertical component of the electric field;
- ν_{in} and ν_{en} are the ions-neutrals and electrons-neutrals effective collisions frequencies, respectively; collisions between charged particles are negligible in the topside ionosphere;

- W_i and W_e are the vertical component of ions and electrons drift velocities, respectively;
- W_n is the vertical component of neutrals velocity.

As described by Eq. (1), ions and electrons are both subject to their own pressure gradient, force of gravity, electromagnetic force imposed by electric fields, and collisions with neutrals. Ions and electrons are treated as perfect gases whose partial pressures are $P_i = N_i k_B T_i$ and $P_e = N_e k_B T_e$, respectively. However, electrical forces between ions and electrons prevent the separation of charges ensuring that they both diffuse at the same speed along magnetic field lines; hence the name “ambipolar” plasma diffusion. As a consequence, neutrality of charges is assumed $N_e \simeq N_i$. Collisions with neutral particles tend to hamper plasma diffusion at a rate dependent on collisions frequencies and difference in charged-neutrals velocities. From Dalgarno³¹ and Dalgarno et al.³², v_{in} and v_{en} effective collisions frequencies depend primarily on neutrals density N_n , secondly on neutrals mass m_n and temperature T_n , and dominate in the lower ionosphere (E region). Moreover, vertical velocity due to neutral winds is important in the lower ionosphere but it can be neglected in the topside; so, $W_n \simeq 0$ is assumed in our calculations and $W_i \simeq W_e \equiv W_D$, where W_D is defined as the plasma drift velocity, due to ambipolar diffusion.

By making the aforementioned assumptions, Eq. (1) become:

$$\begin{cases} \frac{\partial}{\partial z}(N_e k_B T_i) = -N_e m_i g + N_e e E_z - N_e m_i v_{in} W_D, \\ \frac{\partial}{\partial z}(N_e k_B T_e) = -N_e m_e g - N_e e E_z - N_e m_e v_{en} W_D. \end{cases} \quad (2)$$

In the topside, a mixture of O^+ and lighter H^+ and He^+ ions composes the positive charged particles population; because negative ions are not present there, only electrons compose the negative charged particles population. Since $m_e \ll m_i$ then $m_e N_e \ll m_i N_i$. Moreover, in the topside ionosphere, it is very well verified that $m_e v_{en} \ll m_i v_{in}$. So, summing up Eq. (2) for ions and electrons, we get:

$$\frac{\partial}{\partial z}(N_e k_B T_p) = -N_e m g - N_e m v W_D, \quad (3)$$

where $m_i + m_e \simeq m_i \equiv m$, $v_{in} \equiv v$, and the plasma temperature has been defined as $T_p \equiv T_e + T_i$.

By solving the vertical derivative in the left-hand side of Eq. (3) and by grouping all terms in N_e , the following expression for the logarithmic derivative of the electron density is obtained:

$$-\frac{1}{N_e} \frac{\partial N_e}{\partial z} = \frac{m g}{k_B T_p} + \frac{m v}{k_B T_p} W_D + \frac{1}{T_p} \frac{\partial T_p}{\partial z}. \quad (4)$$

Equation (4) describes the electron density variation with height in the plasma ambipolar diffusion state, and defines two important plasma parameters:

- The plasma scale height $H_p \equiv \frac{k_B T_p}{m g}$, that is the scale height describing the plasma distribution with height if only pressure gradient and gravity force were present;
- The plasma diffusion coefficient $D \equiv \frac{k_B T_p}{m v}$, which describes the effects of collisions for a minor plasma diffusing in the major neutrals constituents.

Correspondingly, Eq. (4) can be written as:

$$-\frac{1}{N_e} \frac{\partial N_e}{\partial z} = \frac{1}{H_p} + \frac{W_D}{D} + \frac{1}{T_p} \frac{\partial T_p}{\partial z}. \quad (5)$$

The vertical scale height VSH is defined as:

$$VSH \equiv \left(-\frac{1}{N_e} \frac{\partial N_e}{\partial z} \right)^{-1}, \quad (6)$$

and gives the rate of change of the electron density with height.

By equating Eqs. (5) and (6), and solving for VSH , we obtain:

$$VSH = \frac{k_B T_p}{m g + m v W_D + k_B \frac{\partial T_p}{\partial z}}. \quad (7)$$

Equation (7) is the vertical scale height of the topside electron density as deduced from plasma ambipolar diffusion theory. VSH represents the generalization of H_p considering also the collisions with neutrals and the vertical gradient of plasma temperature; so, including also the effects of the thermal structure and dynamics of the ionosphere. VSH is equal to H_p only under diffusive equilibrium conditions, i.e., when $W_D = 0$, and when the vertical gradient of the plasma temperature can be neglected. Titheridge³³ studying the behavior of the topside ionosphere under diffusive equilibrium (then without considering drifts and collisions), came to the conclusion that three physical parameters are the most important: (1) ion composition, (2) plasma temperature, (3) vertical gradient of the plasma temperature. Equation (7) states also the importance of considering the collisions and plasma drift velocity for a complete description. Relationships between VSH and H_p for different conditions were studied by Liu et al.^{3,4} through Millstone Hill (288.5°E, 42.6°N) and Arecibo (18.3°N, 293.2°E) ISR measurements, and were linked to changes in the shape of the F2-layer region by Luan et al.²⁸. They found that the ratio between

VSH and H_p shows distinct diurnal, seasonal, solar activity, and geographic latitude variations; thus, making the inclusion of the thermal structure and dynamics of the ionosphere, in the estimation of the scale height, of utmost importance for a better description of the topside profile. Anyway, it has to be kept in mind that the diffusive equilibrium condition is an idealized picture of the real topside ionosphere describing its asymptotic behavior in a rather simple analytical way³⁴. For a complete and exact description of the topside ionosphere, the full time-dependent continuity equations for electrons and ions should be considered.

Deriving the effective vertical scale height from topside electron density profiles. To calculate topside effective scale height values, the semi-Epstein analytical formulation is used^{13,21,23,30}. The semi-Epstein analytical formulation describes the topside electron density N_e as a function of the reduced height $z = h - hmF2$, starting from the F2-layer peak density $NmF2$ value at $z = 0$ ($h = hmF2$):

$$N_e(z) = 4NmF2 \frac{\exp\left(\frac{z}{H_{\text{Epstein}}(z)}\right)}{\left[1 + \exp\left(\frac{z}{H_{\text{Epstein}}(z)}\right)\right]^2}. \quad (8)$$

The rate of the electron density decrease in the topside is driven by the topside effective scale height $H_{\text{Epstein}}(z)$ as described in Pignalberi et al.¹³:

$$H_{\text{Epstein}}(z) = \frac{z}{\ln\left\{\frac{1}{N_e(z)} \left[(2NmF2 - N_e(z)) + 2\sqrt{NmF2^2 - N_e(z) \cdot NmF2} \right]\right\}}. \quad (9)$$

By applying Eq. (9), it was possible to calculate the topside effective scale height $H_{\text{Epstein}}(z)$ for each height z by using COSMIC retrieved $N_e(z)$, $NmF2$, and $hmF2$ values¹³. Some examples of this procedure are shown in Figs. 1, 2, and 3 for three selected COSMIC retrieved RO profiles. Blue points in the lower left panels of Figs. 1, 2, and 3 are $H_{\text{Epstein}}(z)$ values calculated by applying Eq. (9) to COSMIC retrieved electron density values, shown as blue points in the upper left panels of Figs. 1, 2, and 3. For a detailed description of the methodology refer to Pignalberi et al.¹³.

The logarithmic derivative of the electron density for a semi-Epstein layer can be analytically deduced by differentiating Eq. (8) with respect to the reduced height:

$$\frac{\partial N_e(z)}{\partial z} = \frac{4NmF2}{H_{\text{Epstein}}(z)} \frac{\exp\left(\frac{z}{H_{\text{Epstein}}(z)}\right)}{\left[1 + \exp\left(\frac{z}{H_{\text{Epstein}}(z)}\right)\right]^2} \left[\frac{1 - \exp\left(\frac{z}{H_{\text{Epstein}}(z)}\right)}{1 + \exp\left(\frac{z}{H_{\text{Epstein}}(z)}\right)} \right]. \quad (10)$$

The right-hand side of Eq. (10) can be simplified by considering Eq. (8) and by using the definition and properties of the hyperbolic tangent:

$$\frac{\partial N_e(z)}{\partial z} = -\frac{N_e(z)}{H_{\text{Epstein}}(z)} \tanh\left(\frac{z}{2H_{\text{Epstein}}(z)}\right). \quad (11)$$

Rearranging Eq. (11) to explicit the logarithmic derivative of the electron density, as done for Eq. (5), we obtain:

$$-\frac{1}{N_e(z)} \frac{\partial N_e(z)}{\partial z} = \frac{\tanh\left(\frac{z}{2H_{\text{Epstein}}(z)}\right)}{H_{\text{Epstein}}(z)}. \quad (12)$$

The reciprocal of Eq. (12)

$$VSH = \frac{H_{\text{Epstein}}(z)}{\tanh\left(\frac{z}{2H_{\text{Epstein}}(z)}\right)}. \quad (13)$$

describes the vertical scale height (Eq. (6)) for a topside electron density profile modeled through a semi-Epstein function (Eq. (8)) with a height-dependent effective scale height (Eq. (9)).

Linking the theory to retrieved effective parameters. Equation (13) links the effective scale height parameter $H_{\text{Epstein}}(z)$ to the theoretical VSH , which depends on physical variables as described by Eq. (7). It is worth noting that, by its nature, VSH is always dependent on z , even where not explicitly written. Retrieving information on the effective scale height $H_{\text{Epstein}}(z)$ is much simpler than getting information on plasma temperature, ions distribution, and other chemical and dynamical quantities needed to calculate VSH , because only the knowledge of the topside electron density profile is required. Then, through the calculation of $H_{\text{Epstein}}(z)$ and Eq. (13), it is possible to infer information about VSH and then on the physical state of the topside ionosphere, under the simplified hypotheses imposed by the plasma ambipolar diffusion theory together with the imposition of Epstein layers.

In the section “A preliminary application based on COSMIC/FORMOSAT-3 Radio Occultation and ESA Swarm satellites data”, a preliminary application of Eq. (13) based on COSMIC RO topside profiles will be given;

ionPrf_C002.2014.090.12.46.G29_2013.3520_nc
 2014/03/31, 12:45:29 UT, 12.7581 UT, 12.2445 LT
 lat=-75.86 lon=-7.7, foF2=11.06 MHz hmF2=283 km

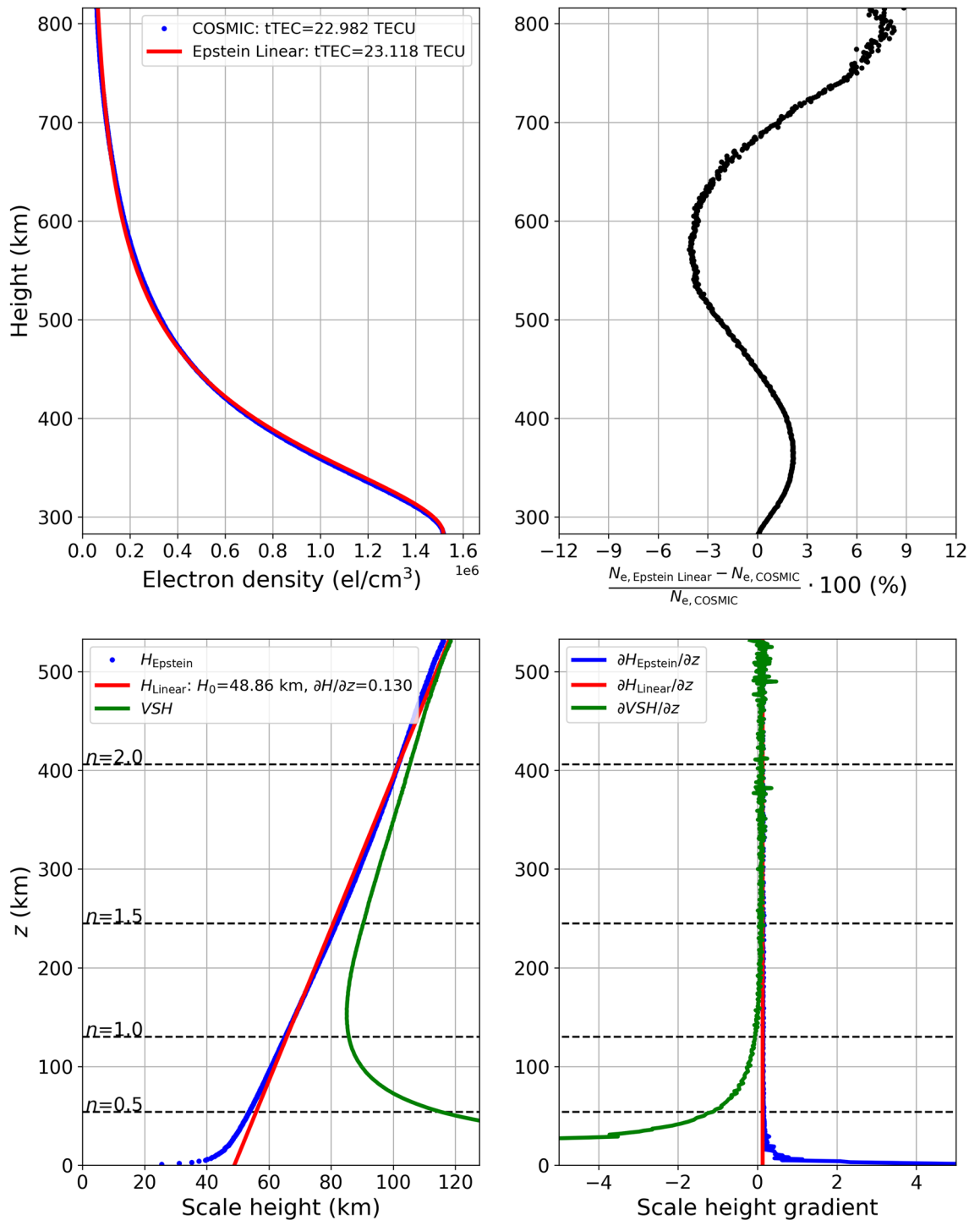


Figure 1. (Upper left panel) Topside electron density values retrieved by COSMIC (blue) and modeled by using the linear topside effective scale height $H_{Linear}(z)$ (red). (Upper right panel) Residuals percentage (black) between modeled ($N_{e, Epstein Linear}$) and measured ($N_{e, COSMIC}$) electron density values represented in the upper left panel. (Lower left panel) Topside effective scale height values ($H_{Epstein}(z)$, blue) obtained from the COSMIC measured profiles shown in the upper left panel, and corresponding modeling through a linear fit ($H_{Linear}(z)$, red). Green curves represent the theoretical VSH calculated from Eq. (13). (Lower right panel) Topside scale height gradient values for $H_{Epstein}(z)$ (blue), $H_{Linear}(z)$ (red), and VSH (green). Vertical coordinates are given as (top panels) height above Earth's surface and (bottom panels) reduced height z . Horizontal dashed lines in the bottom panels represent different values of $n = z/2H_{Epstein}(z)$. In the title, the name of the considered ionPrf file along with time, spatial, and F2-layer peak parameters are reported. Specifically, the file ionPrf_C002.2014.090.12.46.G29_2013.3520_nc represents a high-latitude topside profile.

ionPrf_C001.2015.075.11.28.G06_2014.2860_nc
 2015/03/16, 11:27:14 UT, 11.4539 UT, 12.4884 LT
 lat=42.03 lon=15.52, foF2=11.47 MHz hmF2=275 km

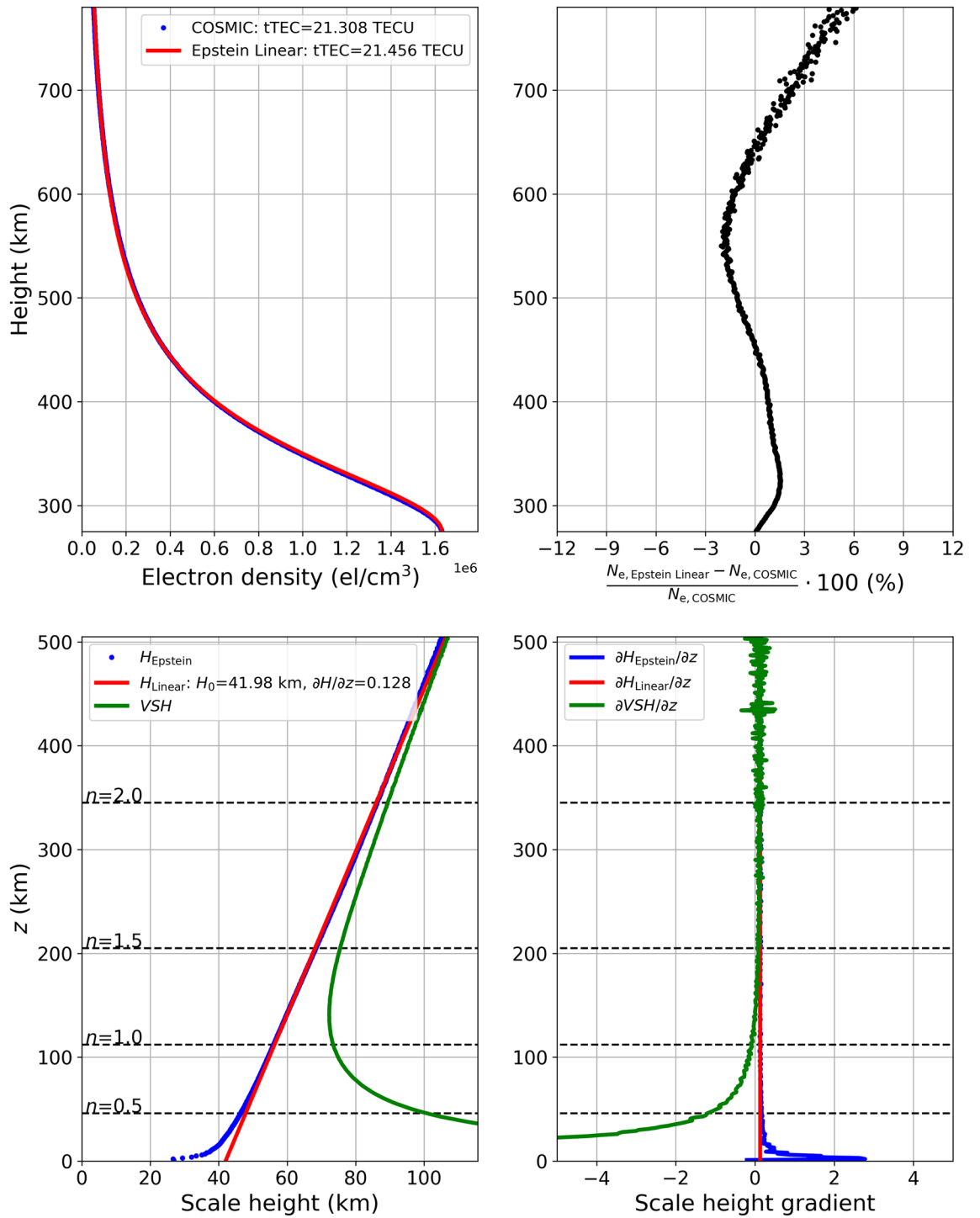


Figure 2. Same as Fig. 1 but for the file ionPrf_C001.2015.075.11.28.G06_2014.2860 representing a mid-latitude topside profile.

ionPrf_C001.2014.129.03.11.G26_2014.2860_nc
 2014/05/09, 03:12:57 UT, 3.2158 UT, 14.303 LT
 lat=4.21 lon=166.31, foF2=13.87 MHz hmF2=374 km

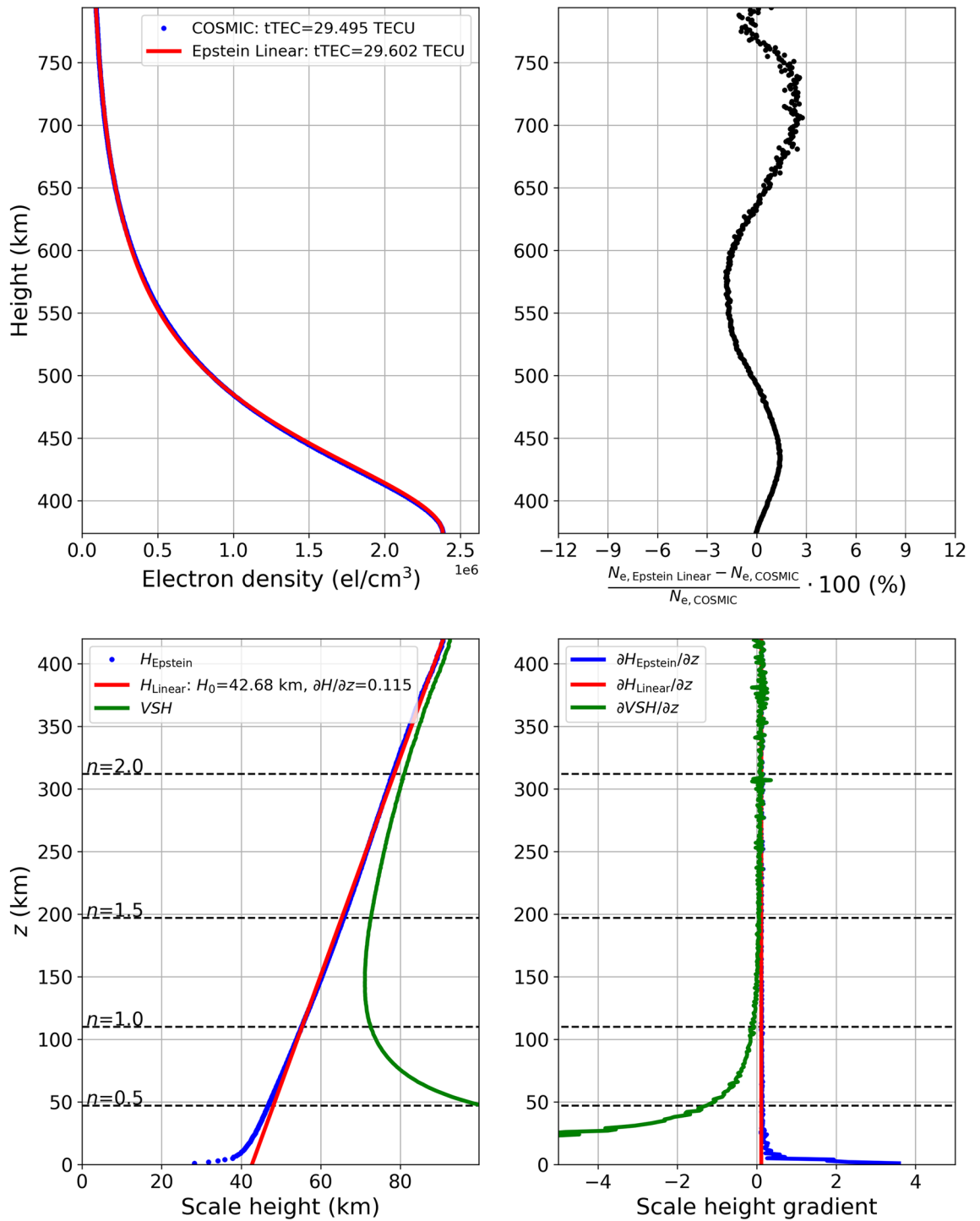


Figure 3. Same as Fig. 1 but for the file ionPrf_C001.2014.129.03.11.G26_2014.2860_nc representing a low-latitude topside profile.

n	$\tanh(n)$	$VSH _{z=n \cdot 2H_{\text{Epstein}}(z)}$	$\frac{n \tanh(n) + \sinh^2(n)}{\tanh(n) \sinh^2(n)}$	$\frac{1}{2 \sinh^2(n)}$	$\frac{\partial VSH}{\partial z} _{z=n \cdot 2H_{\text{Epstein}}(z)}$
0	0	∞	∞	∞	∞
1	0.762	$1.313 \cdot H_{\text{Epstein}}(z)$	2.037	0.362	$2.037 \cdot \frac{\partial H_{\text{Epstein}}(z)}{\partial z} - 0.362$
2	0.964	$1.037 \cdot H_{\text{Epstein}}(z)$	1.189	0.038	$1.189 \cdot \frac{\partial H_{\text{Epstein}}(z)}{\partial z} - 0.038$
3	0.995	$1.005 \cdot H_{\text{Epstein}}(z)$	1.035	0.005	$1.035 \cdot \frac{\partial H_{\text{Epstein}}(z)}{\partial z} - 0.005$
4	0.999	$1.001 \cdot H_{\text{Epstein}}(z)$	1.006	0.001	$1.006 \cdot \frac{\partial H_{\text{Epstein}}(z)}{\partial z} - 0.001$

Table 1. Numerical values of VSH and $\partial VSH/\partial z$ as a function of $H_{\text{Epstein}}(z)$ and $\partial H_{\text{Epstein}}(z)/\partial z$, respectively, and related numerical coefficients defined in Eq. (17), for different values of $n = z/2H_{\text{Epstein}}(z)$.

however, it is here interesting to deduce some general mathematical properties of VSH and $H_{\text{Epstein}}(z)$. First, the behavior of the ratio between $H_{\text{Epstein}}(z)$ and VSH at infinity is studied:

$$\lim_{z \rightarrow \infty} \frac{H_{\text{Epstein}}(z)}{VSH} = \lim_{z \rightarrow \infty} \tanh\left(\frac{z}{2H_{\text{Epstein}}(z)}\right) = 1 \Rightarrow H_{\text{Epstein}}(z) \xrightarrow{z \rightarrow \infty} VSH. \tag{14}$$

Equation (14) shows that $H_{\text{Epstein}}(z)$ tends to VSH at infinity but, it is interesting to highlight how rapidly this happens; then, a relation between the vertical derivative of VSH and $H_{\text{Epstein}}(z)$ has to be found. This is why, Eq. (13) is differentiated with respect to the reduced height:

$$\frac{\partial VSH}{\partial z} = \frac{\frac{\partial H_{\text{Epstein}}(z)}{\partial z}}{\tanh\left(\frac{z}{2H_{\text{Epstein}}(z)}\right)} - \frac{H_{\text{Epstein}}(z) - z \frac{\partial H_{\text{Epstein}}(z)}{\partial z}}{2H_{\text{Epstein}}(z) \sinh^2\left(\frac{z}{2H_{\text{Epstein}}(z)}\right)}. \tag{15}$$

Also, in this case the corresponding behavior at infinity is studied:

$$\lim_{z \rightarrow \infty} \frac{\partial VSH}{\partial z} = \lim_{z \rightarrow \infty} \frac{\frac{\partial H_{\text{Epstein}}(z)}{\partial z}}{\tanh\left(\frac{z}{2H_{\text{Epstein}}(z)}\right)} - \frac{H_{\text{Epstein}}(z) - z \frac{\partial H_{\text{Epstein}}(z)}{\partial z}}{2H_{\text{Epstein}}(z) \sinh^2\left(\frac{z}{2H_{\text{Epstein}}(z)}\right)} = \frac{\partial H_{\text{Epstein}}(z)}{\partial z}. \tag{16}$$

Equation (16) shows that $\frac{\partial VSH}{\partial z}$ tends to $\frac{\partial H_{\text{Epstein}}(z)}{\partial z}$ at infinity, then both the values and the derivatives with height of $H_{\text{Epstein}}(z)$ and VSH tend to be equal very distant from the F2-layer peak (virtually at infinity).

To be valuable for operational purposes, it is desirable that relations (14) and (16) are valid (within few percent of error) also in the lower topside ionosphere, i.e., in the region between the F2-layer peak (highly variable between around 200 and 400 km) and the upper transition height (the height that separates the ionosphere from the plasmasphere, highly variable in a wide range of altitudes around 1000 km). In order to verify this desired behavior, we further developed the calculation, with the change of variable $z = n \cdot 2H_{\text{Epstein}}(z)$, $n \in \mathbb{R}_0^+$, and evaluated VSH and corresponding derivatives through Eqs. (13) and (15):

$$\begin{cases} VSH|_{z=n \cdot 2H_{\text{Epstein}}(z)} = \frac{H_{\text{Epstein}}(z)}{\tanh(n)}, \\ \frac{\partial VSH}{\partial z}|_{z=n \cdot 2H_{\text{Epstein}}(z)} = \frac{\partial H_{\text{Epstein}}(z)}{\partial z} \left[\frac{n \tanh(n) + \sinh^2(n)}{\tanh(n) \sinh^2(n)} \right] - \frac{1}{2 \sinh^2(n)}. \end{cases} \tag{17}$$

Numerical values of VSH and $\frac{\partial VSH}{\partial z}$ for increasing n values (then for increasing heights, where now the height is given as multiples of $H_{\text{Epstein}}(z)$), with corresponding numerical values of coefficients of Eq. (17), are given in Table 1. Table 1 clearly shows how VSH and $\frac{\partial VSH}{\partial z}$ tend very quickly to $H_{\text{Epstein}}(z)$ and $\frac{\partial H_{\text{Epstein}}(z)}{\partial z}$, respectively. Already for $n=2$, i.e., for $z = 4 H_{\text{Epstein}}(z)$, VSH and $H_{\text{Epstein}}(z)$ are the same within 4% and are virtually the same for $n=3$. For $n=0$, i.e., at the F2-layer peak, VSH and H_{Epstein} are not comparable because of the mathematical definition of the semi-Epstein layer and the corresponding effective scale height, which is indefinite at $z=0$. However, this is only a mathematical issue which does not cause any problem for operational applications. In fact, for the N_e calculation (Eq. (8)) what is important is the ratio between z and H_{Epstein} , and this ratio behaves well for $z \rightarrow 0$ (as showed by Pignalberi et al.¹³). A demonstration of the mathematical results shown by Table 1 is given in the section “Deriving the effective vertical scale height from Radio Occultation data” based on retrieved COSMIC RO topside electron density profiles.

A preliminary application based on COSMIC/FORMOSAT-3 radio occultation and ESA Swarm satellites data

The mathematical findings obtained in the section “Linking the Epstein layer effective vertical scale height to the plasma ambipolar diffusion theory” are here substantiated by applying them to actual measured data. For this purpose, the very reliable and wide dataset of topside electron density vertical profiles given by COSMIC RO is exploited. The Pignalberi et al.¹³ approach is first applied to the entire COSMIC dataset to retrieve effective scale height parameters on a global basis and for different helio-geophysical conditions. After that, a preliminary

comparison between these retrieved effective scale height parameters and physical quantities measured by ESA's Swarm satellites is given.

Deriving the effective vertical scale height from radio occultation data. Pignalberi et al.¹³ performed a study based on a selected dataset of 382 COSMIC/FORMOSAT-3 RO profiles matching simultaneously measured F2-layer peak parameters ($NmF2$ and $hmF2$) by co-located ionosondes. This allowed to work on a very reliable dataset of topside electron density profiles³⁵. The Pignalberi et al.¹³ technique is here applied to the entire COSMIC RO dataset, recorded from 2006 to 2018, to retrieve preliminary information on the effective vertical scale height dependence on diurnal, seasonal, solar activity, and spatial variabilities.

COSMIC/FORMOSAT-3 was a six microsatellites constellation launched on 15 April 2006 and deployed into a circular orbit (with 72° of inclination) at about 800 km of altitude (reached at the end of 2007) and a separation angle of 30° in longitude between neighboring satellites³⁶. The mission was a collaborative project between the National Space Organization in Taiwan and the University Corporation for Atmospheric Research in the United States. Each satellite carried a Global Positioning System (GPS) RO receiver, composed by four antennas, capable of measuring the phase delay of radio waves from GPS satellites as they are occulted by the Earth's atmosphere, thus providing an accurate determination of the ionospheric vertical electron density profile. COSMIC RO data (*ionPrf* files) were downloaded from the COSMIC Data Analysis and Archive Center (CDAAC, <https://cdaac-www.cosmic.ucar.edu/cdaac/products.html>).

In this work, all available COSMIC *ionPrf* files from 22 April 2006 to 31 December 2018 were used, a total of 3,626,729 COSMIC retrieved electron density profiles. For the selection of reliable COSMIC topside profiles, a specific filtering process was developed to remove profiles affected by issues related to the failure of the spherical symmetry assumption in the Abel inversion procedure³⁷. The interested reader can refer to the section "Methods" for a detailed description of the filtering process. After filtering, the COSMIC dataset used in this study reduced to 1,791,676 topside profiles (49.4% of the total).

In Pignalberi et al.¹³, $H_{\text{Epstein}}(z)$ values retrieved by COSMIC RO topside electron density values through Eq. (9) were linearly fitted to calculate corresponding slope and intercept values. The topside effective scale height thus obtained, called $H_{\text{Linear}}(z)$, is function of the reduced height z :

$$\begin{cases} H_{\text{Linear}}(z) = \frac{\partial H_{\text{Linear}}}{\partial z} z + H_{0, \text{Linear}}, \\ H_{0, \text{Linear}} = H_{\text{Linear}}(z = 0) = H_{\text{Linear}}(h = hmF2). \end{cases} \quad (18)$$

In Eq. (18), $\partial H_{\text{Linear}}/\partial z$ and $H_{0, \text{Linear}}$ are the slope and intercept values obtained after applying the linear fit procedure of Pignalberi et al.¹³. The slope represents the gradient of the modeled topside scale height $\partial H_{\text{Linear}}/\partial z$, while the intercept $H_{0, \text{Linear}}$ represents the value of $H_{\text{Linear}}(z)$ at the F2-layer peak ($h = hmF2$).

This procedure is here applied to each of the 1,791,676 selected COSMIC topside profiles, thus allowing for compiling a very large dataset of $\partial H_{\text{Linear}}/\partial z$ and $H_{0, \text{Linear}}$. Some examples of application of Eq. (18) are shown in Figs. 1, 2, and 3 for three COSMIC RO profiles recorded at high (Fig. 1), mid (Fig. 2), and low (Fig. 3) latitudes. In both examples, $H_{\text{Linear}}(z)$ values are obtained applying the linear fit of Eq. (18) to $H_{\text{Epstein}}(z)$ values obtained through Eq. (9). Modeled topside electron density values are thus calculated replacing $H_{\text{Epstein}}(z)$ with H_{Linear} in Eq. (8), and are represented by the red curve in the upper left panels of Figs. 1, 2, and 3.

Figures 1, 2, and 3 show how $H_{\text{Epstein}}(z)$ values retrieved from COSMIC N_e measurements exhibit a clear linear trend with height; while only a slight departure from this behavior is visible for the few first tens of kilometers above the F2-layer peak. As a consequence, the linear fitting procedure represented by Eq. (18) allows to reliably describe the topside effective scale height and to properly reproduce the topside electron density profile. Green curves in the lower left panels represent VSH values calculated through Eq. (13). As it was already mathematically demonstrated in the section "Linking the theory to retrieved effective parameters" and illustrated by Table 1 values, VSH markedly departs from $H_{\text{Epstein}}(z)$ just above the F2-layer peak (for $z \rightarrow 0$ or $n \rightarrow 0$) but rapidly approaches it at higher altitudes by assuming quite similar values above $n > 2$. In the lower right panels, the corresponding behavior of vertical gradients for $H_{\text{Epstein}}(z)$, $H_{\text{Linear}}(z)$, and VSH , are shown. Also in this case, theoretical VSH values depart from effective scale height values just above the F2-layer peak, but then rapidly tend to the effective values above.

Figures 1, 2, and 3 testify that it is possible to relate in a very accurate way effective scale height values to theoretical VSH ones for a wide range of altitudes from some hundreds of kilometers above the F2-layer peak to the plasmasphere domain.

The comparison made by Pignalberi et al.¹³ between electron density values measured by COSMIC and those modeled by using $H_{\text{Linear}}(z)$ showed that it is possible to accurately reproduce the ionospheric topside electron density profile by using a semi-Epstein layer with a topside scale height linearly dependent on the height, at least from $hmF2$ to about 800 km, considering a limited dataset of only 382 RO profiles.

For each of the 1,791,676 selected COSMIC topside profiles, the topside total electron content (tTEC) is calculated by integrating the topside electron density values from $hmF2$ to the satellite height. tTEC values are calculated for measured and modeled topside profiles:

$$\begin{cases} \text{tTEC}_{\text{measured}} = \int_{hmF2}^{h_{\text{COSMIC}}} N_{e, \text{COSMIC}} dh, \\ \text{tTEC}_{\text{modeled}} = \int_{hmF2}^{h_{\text{COSMIC}}} N_{e, \text{Epstein Linear}} dh, \end{cases} \quad (19)$$

Epstein Linear model

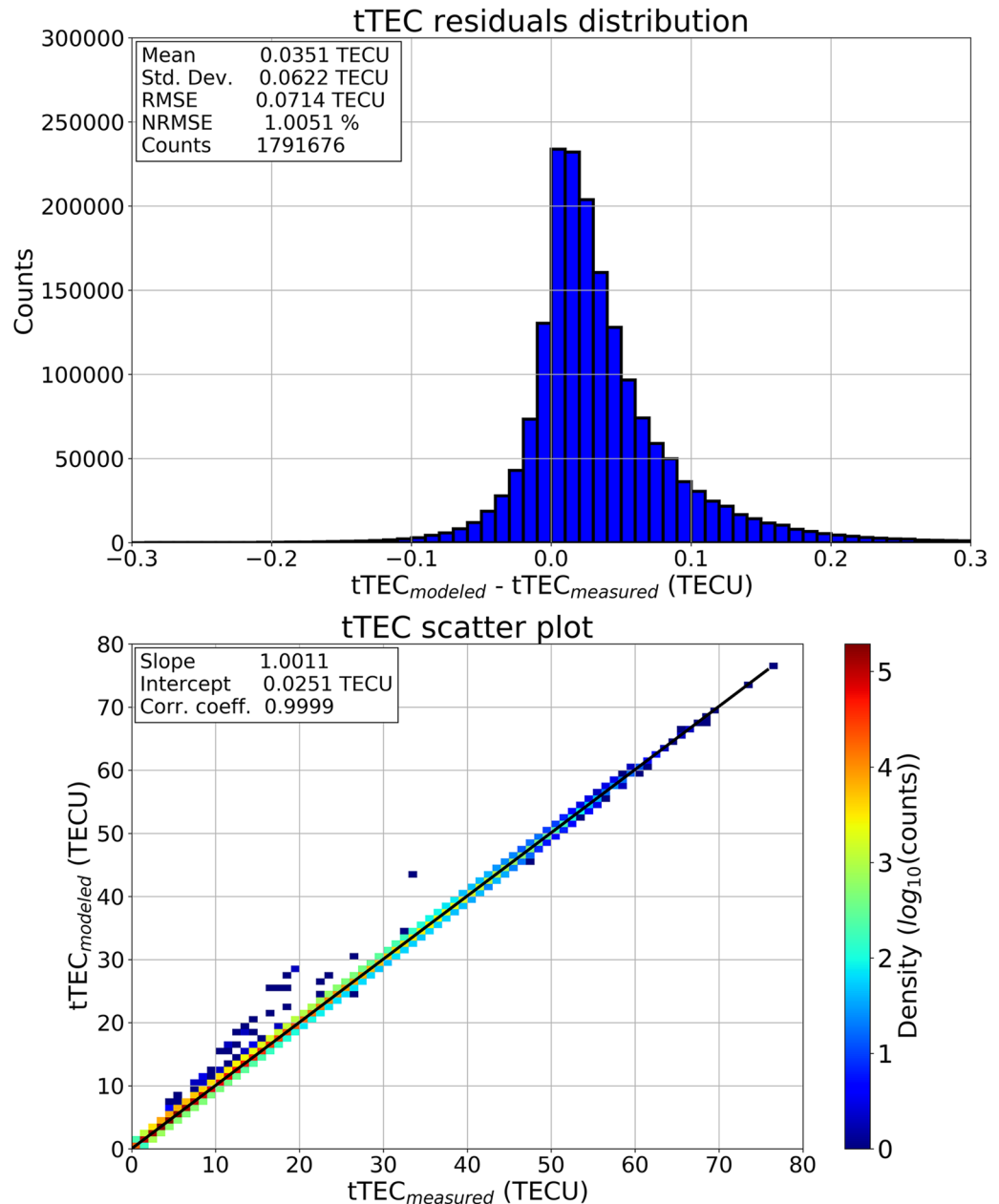


Figure 4. (Top panel) Histogram of residuals between modeled and measured tTEC values. (Bottom panel) Density plot of modeled and measured tTEC values. Black line represents the best linear fit line. The measured tTEC values are related to COSMIC RO profiles, while the modeled ones are obtained by applying the Epstein Linear formulation.

where $N_{e,COSMIC}$ are electron density values measured by COSMIC, while $N_{e,Epstein\ Linear}$ are those modeled by Eq. (8) with topside effective scale height values from Eq. (18).

To perform a statistical validation, tTEC Root Mean Square Error (RMSE) and Normalized Root Mean Square Error (NRMSE) values are calculated, expressed in TECU (1 TECU = 10^{16} el/m²) and in percentage, for the entire COSMIC dataset

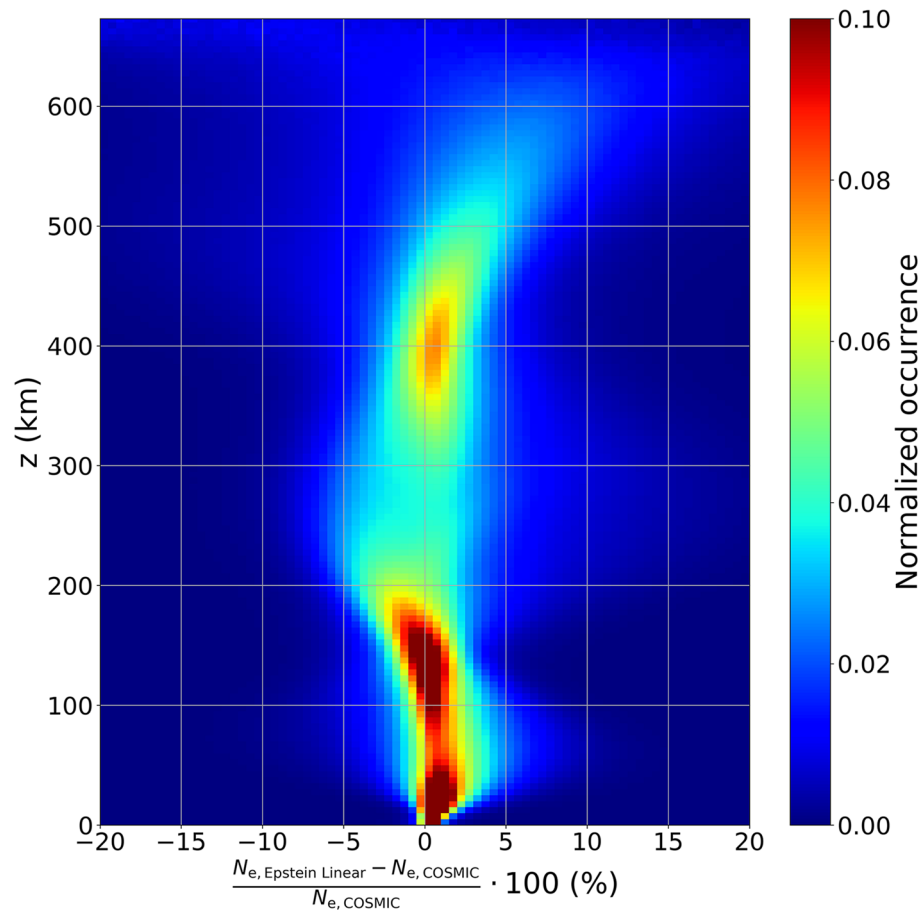


Figure 5. Density plot of residuals percentage between modeled ($N_{e, \text{Epstein Linear}}$) and measured ($N_{e, \text{COSMIC}}$) electron density values as a function of the reduced height z . The number of occurrences in each bin is normalized with respect to the number of occurrences for the same altitude.

$$\begin{cases} \text{RMSE [TECU]} = \sqrt{\frac{\sum_{i=1}^N (\text{tTEC}_{\text{modeled},i} - \text{tTEC}_{\text{measured},i})^2}{N}}, \\ \text{NRMSE [\%]} = \sqrt{\frac{\sum_{i=1}^N \left(\frac{\text{tTEC}_{\text{modeled},i} - \text{tTEC}_{\text{measured},i}}{\text{tTEC}_{\text{measured},i}} \cdot 100 \right)^2}{N}}, \end{cases} \quad (20)$$

where N is 1,791,676, the total number of selected COSMIC profiles.

Top panel of Fig. 4 shows the histogram of residuals between modeled and measured tTEC values, with the corresponding statistical values: $\text{RMSE} = 0.0714$ TECU, $\text{NRMSE} = 1.0051\%$. Moreover, the distribution is well peaked around zero (residuals mean = 0.0351 TECU) with a very low dispersion (residuals standard deviation = 0.0622 TECU). Bottom panel of Fig. 4 shows the scatter plot of modeled versus measured tTEC values. Scattered tTEC values are binned on a 1 TECU \times 1 TECU grid, and the number of points in each bin is shown in color-coded logarithmic scale. The best linear fit is drawn as a solid black line. The scatter plot exhibits a clear one-to-one dependence between measured and modeled tTEC values (slope = 1.0011, intercept = 0.0251 TECU, Pearson correlation coefficient = 0.9999). These results testify that the application of Pignalberi et al.¹³ methodology on COSMIC RO topside profiles is very effective in modeling the topside effective vertical scale height through a linear fit procedure. Moreover, this is a further verification that the vertical scale height exhibits a very clear linear trend, at least for the altitudes probed by COSMIC satellites, from $hmF2$ to about 800 km.

Looking more closely at the residuals of Fig. 4, we can see that modeled values show a slight tendency to overestimate measured values. This is mainly due to the fact that the linear fit of calculated $H_{\text{Epstein}}(z)$ values is made starting from $hmF2 + 50$ km upwards¹³, thus producing a slight overestimation of modeled scale height values just above $hmF2$, which in turns leads to an overestimation of measured N_e values. On the contrary, tTEC values outside the linear fit in the scatter plot of Fig. 4 (the dark blue bins just above the fit line) cannot be explained as above. These profiles do not exhibit a linear trend of the scale height with altitude. They can be considered as outliers and they are not visible in the histogram of Fig. 4 because outside the displayed axis range. Anyhow, these outliers affect very little the full statistics because they are relatively few (tens of topside profiles).

Figure 5 shows the normalized occurrence of residuals percentage between modeled and measured electron density values, for different reduced heights, for the entire COSMIC selected dataset. Notably, it is expected that

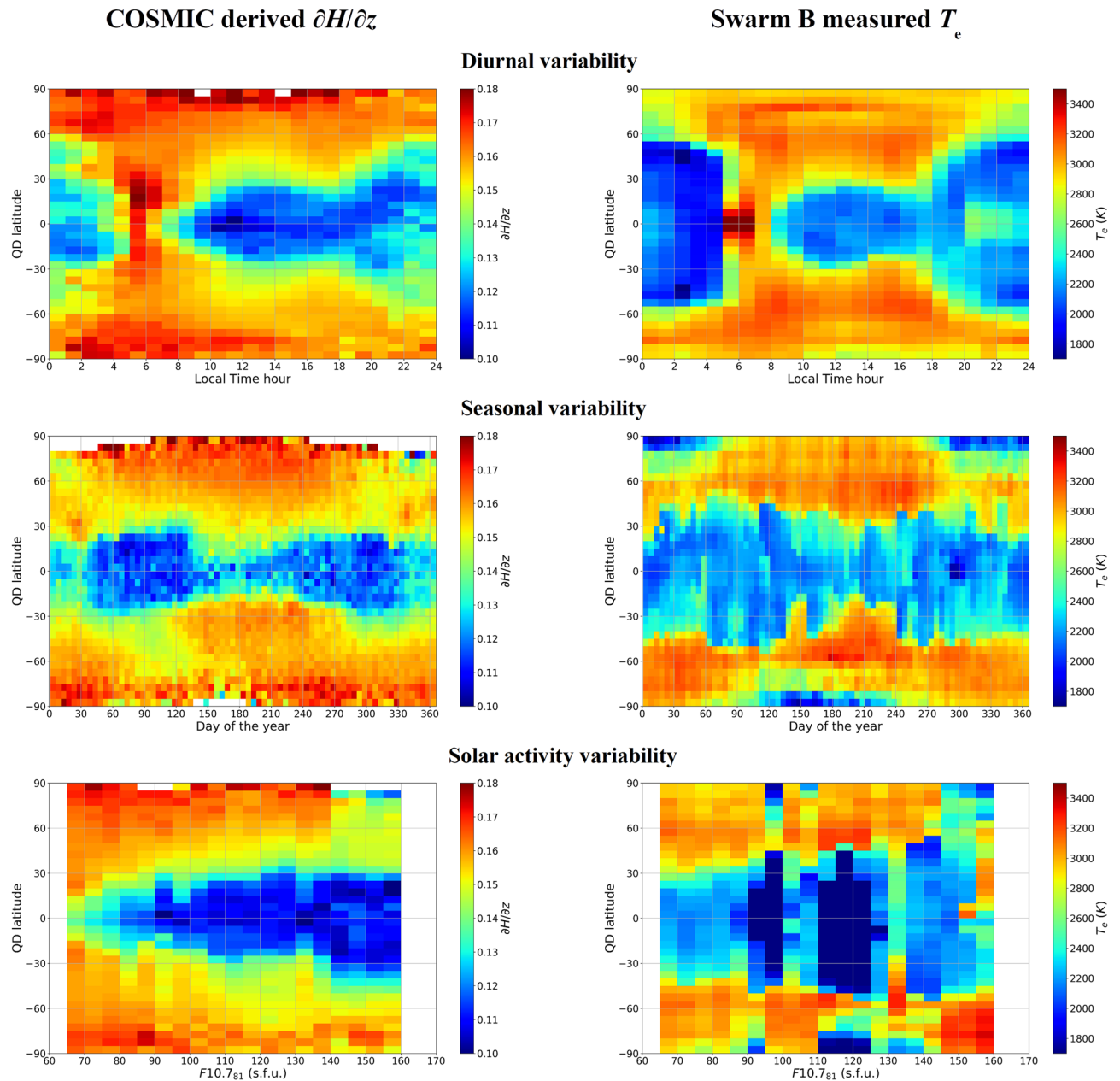


Figure 6. (Left panels) Topside vertical scale height gradient values as deduced from COSMIC RO dataset (2006–2018). Binned median scale height gradient values are shown as a function of QD latitude (y -axis) and of (x -axis) Local Time hour (top panels), day of the year (middle panels), and F10.7₈₁ (bottom panels). (Right panels) Swarm B measured T_e dataset (2014–2019) binned like the topside vertical scale height gradient values.

the spreading of data increases with altitude, because all profiles are constrained at the peak ($z=0$). Nevertheless, up to $z=500$ km the profiles are remarkably well reproduced. Specifically, Fig. 5 highlights that most of the occurrences lie in a narrow range between $\pm 5\%$. As a consequence, topside electron density values modeled by the $H_{\text{Linear}}(z)$ approximation (Eq. (18)) allow for a reliable description of the topside profile in the altitude range probed by COSMIC satellites.

Results of Figs. 4 and 5 confirm what preliminarily found by Pignalberi et al.¹³, i.e., the scale height retrieved from COSMIC RO profiles exhibits a clear linear trend for the lower topside region, the one from the F2-layer peak to about 800 km of height. The COSMIC RO dataset used for this work includes different diurnal, seasonal, and solar activity conditions spanning from equatorial to auroral latitudes. Then, the linear trend of the topside scale height is a very well defined topside ionospheric feature, regardless of geophysical conditions. It is worth noting that these results are valid until about 800 km of height (the maximum height covered by COSMIC satellites); for higher altitudes it has been demonstrated that a departure from the linearity takes place¹⁴.

Comparison with ESA Swarm data. From Eq. (7), it is clear that several physical quantities account for the VSH variability. Specifically, plasma temperatures and corresponding vertical gradient values, ions distribu-

tion, ions-neutrals collision frequency, and plasma vertical drift velocity. Moreover, the variation with height of these quantities, and also of the gravity acceleration, should be taken into account. As a preliminary analysis, we tried to find possible connections between *VSH* (and corresponding vertical gradient) and some physical quantities measured by ESA's Swarm satellites. Specifically, here we present a preliminary comparison between the electron temperature recorded by ESA's Swarm satellites and the vertical gradient of *VSH*. Electron temperature can be considered as a proxy of plasma temperature. We focus on the vertical gradient of *VSH* because the analysis of the section "Deriving the effective vertical scale height from Radio Occultation data" has demonstrated that the topside ionosphere exhibits a very clear linear trend of the scale height and then a constant scale height gradient. Because the scale height gradient is constant for most of the altitudes sounded by COSMIC satellites, it is fair to use the comparison with Swarm satellites measurements collected between about 450 and 520 km of altitude.

Swarm³⁸ is a three-satellites constellation launched at the end of 2013 by ESA in a LEO circular near-polar orbit. Two of them, Swarm A and C, are orbiting the Earth side by side at the same altitude of about 460 km, with an inclination of 87.4° and an east–west separation of 1–1.5° in longitude. Swarm B is flying about 60 km higher, with an inclination of 88° on a different orbit. They are all equipped with identical instruments consisting of high-resolution sensors for measurements of both geomagnetic and electric fields, as well as plasma density and temperature.

Here, we consider Level 1b electron temperature T_e measurements at 2 Hz rate recorded by the Swarm's Langmuir probes³⁹ from the beginning of 2014 to the end of 2019. Swarm's data are freely accessible at <ftp://swarm-diss.eo.esa.int>. Detailed information on Swarm's Langmuir Probes data are provided in Knudsen et al.⁴⁰ and Lomidze et al.⁴¹.

Right panels of Fig. 6 show median Swarm B T_e values binned as a function of Quasi-Dipole⁴² magnetic latitude on y-axis (QD), and of Local Time, day of the year, and 81-days running mean of the solar index F10.7⁴³ (F10.7₈₁). Left panels of Fig. 6 show instead median values of the calculated topside effective scale height gradient $\partial H_{\text{Linear}}/\partial z$ derived from COSMIC dataset, as described in the section "Deriving the effective vertical scale height from Radio Occultation data", and binned as Swarm data. The COSMIC dataset provides a good and quite uniform coverage of different diurnal and seasonal conditions for the QD latitude range $\pm 70^\circ$. In terms of solar activity the dataset is instead slightly biased toward low solar activity values because most of RO profiles were recorded at the beginning of the mission, i.e., for years of low solar activity (2006–2010).

As demonstrated in the sections "Linking the theory to retrieved effective parameters" and "Deriving the effective vertical scale height from Radio Occultation data", the equalities $\partial H_{\text{Epstein}}/\partial z \equiv \partial H_{\text{Linear}}/\partial z \equiv \partial VSH/\partial z$ can be assumed valid already after a few scale heights above the F2-layer peak. Then, for the sake of brevity, in the following we will refer to $\partial H/\partial z$ as the topside vertical scale height gradient. At the altitudes of Swarm B satellite (initial altitude of 520 km) we are usually in a region where a perfect match between effective and theoretical parameters has not yet been achieved (for example, look at lower panels of Figs. 1, 2, and 3 where n is between 1.2 and 1.7 at Swarm B altitude). This is why, here we limit ourselves only to the identification of similar climatological patterns between $\partial H/\partial z$ and T_e .

Figure 6 shows a remarkable similarity between $\partial H/\partial z$ values derived by COSMIC and T_e measured by Swarm B, when considering the diurnal, seasonal, and solar activity variability for different QD magnetic latitudes. $\partial H/\partial z$ and T_e values exhibit an identical diurnal pattern from equatorial to auroral latitudes. Undoubtedly, the electron temperature (the same holds for plasma temperature) plays a crucial role in driving the diurnal trend of the vertical scale height gradient in the topside ionosphere. Very similar considerations can be made for the seasonal variability. However, in this case T_e values show a different latitudinal extension of the band of lower values at low latitudes, compared to that of $\partial H/\partial z$. Such differences might be attributed to the fact that Swarm B is a near-polar orbit quasi-solar synchronous satellite, which means that the sampling is not evenly distributed. The solar activity plots are the ones showing the worst similarity. This is partly due to the fact that COSMIC and Swarm B datasets cover different years; then, different solar activity levels are unevenly sampled. Both COSMIC and Swarm B datasets are biased toward low solar activity levels, but COSMIC dataset allows for a better coverage of medium and high solar activity levels than Swarm B. Anyway, similar features are shown also in this case with lower values at low latitudes and higher approaching high latitudes, with a tendency to enlarge the latitudes range of low values for high solar activity.

Results of Fig. 6 are only a first attempt to apply what has been discussed about the linking between effective and theoretical parameters and the corresponding comprehension of physical properties of the topside ionosphere. More in depth studies and comparisons with other measured data are needed.

Conclusions

This paper has shown that it is possible to mathematically link the theoretical vertical scale height *VSH*, that is connected to several physical parameters through the plasma ambipolar diffusion theory, to the effective scale height *H*, as retrieved from the analysis of COSMIC topside electron density profiles through the Epstein formulation. Specifically, the effective scale height $H_{\text{Epstein}}(z)$ derived from the semi-Epstein function has been used as the effective parameter because of its fully analytical description. A mathematical relation between *VSH* and $H_{\text{Epstein}}(z)$ has then been obtained and corresponding mathematical properties have been studied. It has been demonstrated how *VSH* and $H_{\text{Epstein}}(z)$ (and corresponding vertical gradients) tend to each other in the topside ionosphere, from some hundreds of kilometers above the F2-layer peak to the plasmasphere. This means that from the study of $H_{\text{Epstein}}(z)$, information can be deduced for *VSH* and then for the physical quantities involved in its variability. Following the approach proposed by Pignalberi et al.¹³, effective scale height values and gradients have been calculated for the entire COSMIC RO dataset. Finally, topside scale height gradient values from the COSMIC dataset have been compared with electron temperature values measured by ESA's Swarm B satellite. It was found that both exhibit very similar diurnal, seasonal, and solar activity trends as a function of QD magnetic

latitude. This study represents only a first attempt to unveil the physical properties of the topside ionosphere starting from calculated effective values. Conversely, the exploitation of different measurement techniques and satellites missions capable of retrieving physical parameters in the topside ionosphere, can help in improving the topside description made by ionospheric models like NeQuick and IRI through the proposed methodology. More in depth studies are needed to fully characterize the effective scale height variability on different heliogeophysical parameters and to link it to the physical quantities involved in the plasma ambipolar diffusion theory.

Methods

On the selection and filtering of COSMIC/FORMOSAT-3 RO topside profiles. The initial dataset of 3,626,729 COSMIC ionPrf files, recorded from 22 April 2006 to 31 December 2018, underwent a selection and filtering process constituted by several steps.

Specifically, the initial selection consisted in discarding profiles for which one (or more) of the following conditions is met:

1. $h_{\text{COSMIC}} < hmF2 + 150$ km; i.e., we require a topside profile at least 150 km wide to make a reliable fit of the topside scale height;
2. $N_e < 0$ for at least one point at $h > hmF2$;
3. It was not possible to apply the linear fitting procedure of the topside scale height (see Eq. (18)) due to missing or corrupted data in the topside profile;
4. The ionPrf file is corrupted.

After this initial filtering procedure, the original COSMIC dataset reduced to 3,069,418 electron density profiles. Then, 557,311 profiles (about 15.4% of the analyzed dataset) were discarded.

Afterwards, the dataset of 3,069,418 profiles underwent a second filtering procedure where COSMIC profiles were discarded if at least one of the following conditions is met:

- (a) $foF2 < 0.1$ MHz or $foF2 > 22$ MHz;
- (b) $hmF2 < 150$ km or $hmF2 > 650$ km;
- (c) $\frac{\partial H_{\text{linear}}}{\partial z} < 0$. This condition avoids profiles for which the topside N_e tends to increase for most of the topside profile, which is not physically acceptable;
- (d) $|Lat_{hmF2} - Lat_{hmF2+150km}| \geq 5^\circ$ or $|Lon_{hmF2} - Lon_{hmF2+150km}| \geq 10^\circ$, where Lat_{hmF2} and $Lat_{hmF2+150km}$ are the geographic latitudes, and Lon_{hmF2} and $Lon_{hmF2+150km}$ the geographic longitudes, of electron density values recorded at $hmF2$ and 150 km above $hmF2$. In this way, too slanted topside profiles were discarded.

After this second filtering procedure, 26,618 (0.87% of 3,069,418) profiles were discarded.

Afterwards, we developed a specific filter to estimate the *noise*, at different spatial scales, of the COSMIC topside electron density profiles. We measured the noise level using the standard deviation of the relative differences between the measured electron density profile and its corresponding smoothed one. The smoothed topside electron density profile is calculated as the running mean of the topside profile with running windows of different length: *small* = 10 km, *medium* = 75 km, and *large* = 150 km.

The filtering algorithm is composed by the following steps:

1. The COSMIC topside profile is preliminary vertically interpolated to obtain an even height resolution of 1 km. In this way, we have $N = h_{\text{COSMIC}} - hmF2$ topside electron density measurements, indexed through the index k running on the whole topside profile;
2. The smoothed electron density values $\overline{N}_{e,k}$ are then calculated:

$$\overline{N}_{e,k} = \frac{1}{2j+1} \sum_{i=-j}^j N_{e,k+i}, \quad (21)$$

where $N_{e,k+i}$ are the values falling inside the window of width $2j+1$ centered on the index k . The number of N_e values falling inside the window is 11 (small window), 76 (medium window), and 151 (large window). For the sake of simplicity, the profile smoothed with the small window is called $\overline{N}_{e,\text{small}}$, the profile smoothed with the medium window $\overline{N}_{e,\text{medium}}$, and the profile smoothed with the large window $\overline{N}_{e,\text{large}}$;

3. The relative differences (relative residuals) between measured and smoothed electron density values are calculated:

$$N_{e,\text{residuals_small}} = \frac{N_e - \overline{N}_{e,\text{small}}}{\overline{N}_{e,\text{small}}} \cdot 100, \quad (22a)$$

$$N_{e,\text{residuals_medium}} = \frac{N_e - \overline{N}_{e,\text{medium}}}{\overline{N}_{e,\text{medium}}} \cdot 100, \quad (22b)$$

$$N_{e,\text{residuals_large}} = \frac{N_e - \overline{N_{e,\text{large}}}}{\overline{N_{e,\text{large}}}} \cdot 100. \quad (22c)$$

4. The standard deviation of residuals of Eqs. (22a–c) is calculated, that is what is considered the *noise* in the topside profile:

$$\text{Noise_small} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (N_{e,\text{residuals_small},i} - \overline{N_{e,\text{residuals_small}}})^2}, \quad (23a)$$

$$\text{Noise_medium} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (N_{e,\text{residuals_medium},i} - \overline{N_{e,\text{residuals_medium}}})^2}, \quad (23b)$$

$$\text{Noise_large} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (N_{e,\text{residuals_large},i} - \overline{N_{e,\text{residuals_large}}})^2}, \quad (23c)$$

where $\overline{N_{e,\text{residuals_small}}}$ is the mean value of $N_{e,\text{residuals_small}}$ calculated over the whole topside profile (the same holds for the medium and large windows);

5. A profile is considered too noisy, and then discarded, when at least one of the three parameters calculated through Eqs. (23a–c) exceeds definite noise threshold values T_{window}

After a preliminary testing phase on the Pignalberi et al.¹³ dataset, in which every topside profile was visually checked, we found that a good compromise is to choose the following noise threshold values:

$$\begin{cases} T_{\text{small}} = 2 \\ T_{\text{medium}} = 3 \\ T_{\text{large}} = 4 \end{cases}. \quad (24)$$

If at least one of these three thresholds is exceeded, the corresponding COSMIC topside profile is discarded.

After applying the above described filtering technique, 1,251,124 (40.76% of 3,069,418) profiles were discarded.

Finally, the COSMIC dataset considered in this work is constituted by 1,791,676 (58.37% of 3,069,418) topside profiles.

Received: 29 July 2020; Accepted: 21 September 2020

Published online: 16 October 2020

References

- Rishbeth, H. & Garriott, O. *Introduction to Ionospheric Physics. International Geophysics Series* Vol. 14 (Academic Press, New York, 1969).
- Ratcliffe, J. A. *An Introduction to the Ionosphere and Magnetosphere* (Cambridge University Press, Cambridge, 1972).
- Liu, L. et al. Variations of topside ionospheric scale height over Millstone Hill during the 30-day incoherent scatter radar experiment. *Ann. Geophys.* **25**, 2019–2027. <https://doi.org/10.5194/angeo-25-2019-2007> (2007).
- Liu, L. et al. An analysis of the scale heights in the lower topside ionosphere based on the Arecibo incoherent scatter radar measurements. *J. Geophys. Res. Space Phys.* **112**, A06307. <https://doi.org/10.1029/2007JA012250> (2007).
- Coisson, P., Radicella, S. M., Leitinger, R. & Nava, B. Topside electron density in IRI and NeQuick: features and limitations. *Adv. Space Res.* **37**, 937–942. <https://doi.org/10.1016/j.asr.2005.09.015> (2006).
- Hernández-Pajares, M. et al. Electron density extrapolation above F2 peak by the linear Vary-Chap model supporting new Global Navigation Satellite Systems-LEO occultation missions. *J. Geophys. Res. Space Phys.* **122**, 9003–9014. <https://doi.org/10.1002/2017JA023876> (2017).
- Hu, A. et al. Modeling of topside ionospheric vertical scale height based on ionospheric radio occultation measurements. *J. Geophys. Res. Space Phys.* **124**, 4926–4942. <https://doi.org/10.1029/2018JA026280> (2019).
- Li, Q. et al. α -Chapman scale height: Longitudinal variation and global modeling. *J. Geophys. Res. Space Phys.* **124**, 2083–2098. <https://doi.org/10.1029/2018JA026286> (2019).
- Nsumei, P., Reinisch, B. W., Huang, X. & Bilitza, D. New Vary-Chap profile of the topside ionosphere electron density distribution for use with the IRI model and the GIRO real time data. *Radio Sci.* **47**, RS0L16. <https://doi.org/10.1029/2012RS004989> (2012).
- Olivares-Pulido, G., Hernandez-Pajares, M., Aragon-Angel, A. & Garcia-Rigo, A. A linear scale height Chapman model supported by GNSS occultation measurements. *J. Geophys. Res. Space Phys.* **121**, 7932–7940. <https://doi.org/10.1002/2016JA022337> (2016).
- Pezzopane, M. & Pignalberi, A. The ESA Swarm mission to help ionospheric modeling: a new NeQuick topside formulation for mid-latitude regions. *Sci. Rep.* **9**, 12253. <https://doi.org/10.1038/s41598-019-48440-6> (2019).
- Pignalberi, A., Pezzopane, M. & Rizzi, R. Modeling the lower part of the topside ionospheric vertical electron density profile over the European region by means of Swarm satellites data and IRI UP method. *Space Weather* **16**, 304–320. <https://doi.org/10.1002/2017SW001790> (2018).
- Pignalberi, A. et al. On the analytical description of the topside ionosphere by NeQuick: modeling the scale height through COSMIC/FORMOSAT-3 selected data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **13**, 1867–1878. <https://doi.org/10.1109/JSTAR.2020.2986683> (2020).
- dos Santos Prol, F. et al. Linear Vary-Chap topside electron density model with topside sounder and radio-occultation data. *Surv. Geophys.* <https://doi.org/10.1007/s10712-019-09521-3> (2019).

15. Reinisch, B. W. *et al.* Modeling the IRI topside profile using scale heights from ground-based ionosonde measurements. *Adv. Space Res.* **34**(9), 2026–2031. <https://doi.org/10.1016/j.asr.2004.06.012> (2004).
16. Stankov, S. M. & Jakowski, N. Topside ionospheric scale height analysis and modelling based on radio occultation measurements. *J. Atmos. Sol. Terr. Phys.* **68**(2), 134–162. <https://doi.org/10.1016/j.jastp.2005.10.003> (2006).
17. Wang, S., Huang, S. & Fang, H. Topside ionospheric VARY-chap scale height retrieved from the COSMIC/FORMOSAT-3 data at midlatitudes. *Adv. Space Res.* **56**(5), 893–899. <https://doi.org/10.1016/j.asr.2015.04.021> (2015).
18. Wang, S., Huang, S. & Fang, H. New method for deriving the topside ionospheric Vary-Chap scale height. *Radio Sci.* **50**, 866–875. <https://doi.org/10.1002/2015RS005724> (2015).
19. Wu, M. J. *et al.* New Vary-Chap scale height profile retrieved from COSMIC radio occultation data. *J. Geophys. Res. Space Phys.* **125**, e2019JA027637. <https://doi.org/10.1029/2019JA027637> (2020).
20. Chapman, S. The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth. *Proc. Phys. Soc.* **43**(1), 26–45. <https://doi.org/10.1088/0959-5309/43/1/305> (1931).
21. Epstein, P. S. Reflection of waves in an inhomogeneous absorbing medium. *Proc. Natl. Acad. Sci.* **16**(10), 627–637. <https://doi.org/10.1073/pnas.16.10.627> (1930).
22. Fonda, C., Coisson, P., Nava, B. & Radicella, S. M. Comparison of analytical functions used to describe topside electron density profiles with satellite data. *Ann. Geophys.* **48**(3), 491–495. <https://doi.org/10.4401/ag-3213> (2005).
23. Rawer, K. Synthesis of ionospheric electron density profiles with Epstein functions. *Adv. Space Res.* **8**(4), 191–199. [https://doi.org/10.1016/0273-1177\(88\)90239-6](https://doi.org/10.1016/0273-1177(88)90239-6) (1988).
24. Stankov, S. M. *et al.* A new method for reconstruction of the vertical electron density distribution in the upper ionosphere and plasmasphere. *J. Geophys. Res. Space Phys.* **108**(A5), 1164. <https://doi.org/10.1029/2002JA009570> (2003).
25. Depuev, V. & Pulinets, S. A global empirical model of the ionospheric topside electron density. *Adv. Space Res.* **34**(9), 2016–2020. <https://doi.org/10.1016/j.asr.2004.05.006> (2004).
26. Leitinger, R., Radicella, S. M., Hochegger, G. & Nava, B. Diffusive equilibrium models for the height region above the F2 peak. *Adv. Space Res.* **29**(6), 809–814. [https://doi.org/10.1016/S0273-1177\(02\)00036-4](https://doi.org/10.1016/S0273-1177(02)00036-4) (2002).
27. Nava, B., Radicella, S. M., Pulinets, S. & Depuev, V. Modelling bottom and topside electron density and TEC with profile data from topside ionograms. *Adv. Space Res.* **27**(1), 31–34. [https://doi.org/10.1016/S0273-1177\(00\)00137-X](https://doi.org/10.1016/S0273-1177(00)00137-X) (2001).
28. Luan, X. *et al.* A study of the shape of topside electron density profile derived from incoherent scatter radar measurements over Arecibo and Millstone Hill. *Radio Sci.* **41**, RS4006. <https://doi.org/10.1029/2005RS003367> (2006).
29. Bilitza, D. *et al.* International reference ionosphere 2016: from ionospheric climate to real-time weather predictions. *Space Weather* **15**, 418–429. <https://doi.org/10.1002/2016SW001593> (2017).
30. Nava, B., Coisson, P. & Radicella, S. M. A new version of the NeQuick ionosphere electron density model. *J. Atmos. Sol. Terr. Phys.* **70**, 1856–1862. <https://doi.org/10.1016/j.jastp.2008.01.015> (2008).
31. Dalgarno, A. Ambipolar diffusion in the F-region. *J. Atmos. Sol. Terr. Phys.* **26**, 939 (1964).
32. Dalgarno, A., McElroy, M. B. & Walker, J. C. The diurnal variation of ionospheric temperatures. *Planet. Space Sci.* **15**, 331–345 (1967).
33. Titheridge, J. Ion transition heights from topside electron density profiles. *Planet. Space Sci.* **24**(3), 229–245. [https://doi.org/10.1016/0032-0633\(76\)90020-9](https://doi.org/10.1016/0032-0633(76)90020-9) (1976).
34. Bauer, S. Diffusive equilibrium in the topside ionosphere. *Proc. IEEE* **57**(6), 1114–1118. <https://doi.org/10.1109/PROC.1969.7163> (1969).
35. Shaikh, M. M., Nava, B. & Haralambous, H. On the use of topside RO-derived electron density for model validation. *J. Geophys. Res. Space Phys.* **123**, A025132. <https://doi.org/10.1029/2017JA025132> (2018).
36. Anthes, R. A. *et al.* The COSMIC/FORMOSAT-3 mission: early results. *Bull. Am. Meteorol. Soc.* **89**, 313–333. <https://doi.org/10.1175/BAMS-89-3-313> (2008).
37. Shaikh, M., Notarpietro, R. & Nava, B. The impact of spherical symmetry assumption on radio occultation data inversion in the ionosphere: an assessment study. *Adv. Space Res.* **53**(4), 599–608. <https://doi.org/10.1016/j.asr.2013.10.025> (2014).
38. Friis-Christensen, E., Lühr, H. & Hulot, G. Swarm: a constellation to study the Earth's magnetic field. *Earth Planets Space* **58**, 351–358. <https://doi.org/10.1186/BF03351933> (2006).
39. Swarm L1b Product Definition. https://earth.esa.int/documents/10174/1514862/Swarm_L1b_Product_Definition (2018).
40. Knudsen, D. J. *et al.* Thermal ion imagers and Langmuir probes in the Swarm electric field instruments. *J. Geophys. Res. Space Phys.* **122**, 2655–2673. <https://doi.org/10.1002/2016JA022571> (2017).
41. Lomidze, L., Knudsen, D. J., Burchill, J., Kouznetsov, A. & Buchert, S. C. Calibration and validation of Swarm plasma densities and electron temperatures using ground-based radars and satellite radio occultation measurements. *Radio Sci.* **53**, 15–36. <https://doi.org/10.1002/2017RS006415> (2018).
42. Laundal, K. M. & Richmond, A. D. Magnetic coordinate systems. *Space Sci. Rev.* **206**, 27. <https://doi.org/10.1007/s11214-016-0275-y> (2017).
43. Tapping, K. F. The 10.7 cm solar radio flux ($F_{10.7}$). *Space Weather* **11**, 394–406. <https://doi.org/10.1002/swe.20064> (2013).

Acknowledgements

This research is partially supported by the Italian MIUR-PRIN grant 2017APKP7T on Circumterrestrial Environment: Impact of Sun-Earth Interaction. The authors thank the COSMIC/FORMOSAT-3 team for making freely available Radio Occultation data by means of COSMIC Data Analysis and Archive Center (CDAAC, <https://cdaac-www.cosmic.ucar.edu/cdaac/products.html>). Thanks to the European Space Agency for making Swarm data publicly available via <ftp://swarmdiss.eo.esa.int>, and for the considerable efforts made for the Langmuir probes data calibration. $F_{10.7}$ values used in this study were downloaded from OMNIWeb Data Explorer website (<https://omniweb.gsfc.nasa.gov/form/dx1.html>) maintained by the Space Physics Data Facility of the Goddard Space Flight Center (NASA).

Author contributions

A.P.: conceptualization, methodology, software, validation, formal analysis, investigation, writing—original draft. M.P.: supervision, writing—review & editing, funding acquisition, investigation, validation. B.N.: writing—review & editing, investigation, validation. P.C.: writing—review & editing, investigation, validation.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020