One-Dimensional Variational Ionospheric Retrieval Using Radio Occultation Bending Angles

Culverwell, I. D.; Healy, S. B.; Elvidge, S.

One-Dimensional Variational Ionospheric Retrieval Using Radio Occultation Bending Angles: 1. Theory

I. D. Culverwell[1] , S. B. Healy[2], and S. Elvidge[3]

[1]Met Office, Exeter, UK, [2]European Centre for Medium-Range Weather Forecasts, Reading, UK, [3]Space Environment and Radio Engineering (SERENE), University of Birmingham, Birmingham, UK

Abstract A new one-dimensional variational (1D-Var) retrieval method for ionospheric GNSS radio occultation (GNSS-RO) measurements is described. The forward model implicit in the retrieval calculates the bending angles produced by a one-dimensional ionospheric electron density profile, modeled with multiple "Vary-Chap" layers. It is demonstrated that gradient based minimization techniques can be applied to this retrieval problem. The use of ionospheric bending angles is discussed. This approach circumvents the need for Differential Code Bias (DCB) estimates when using the measurements. This new, general retrieval method is applicable to both standard GNSS-RO retrieval problems, and the truncated geometry of EUMETSAT's Metop Second Generation (Metop-SG), which will provide GNSS-RO measurements up to about 600 km above the surface. The climatological a priori information used in the 1D-Var is effectively a starting point for the 1D-Var minimization, rather than a strong constraint on the final solution. In this paper the approach has been tested with 143 COSMIC-1 measurements. We find that the method converges in 135 of the cases, but around 25 of those have high "cost at convergence" values. In the companion paper (Elvidge et al., 2023), a full statistical analysis of the method, using over 10,000 COSMIC-2 measurements, has been made.

Plain Language Summary This paper presents a new way of estimating the density of electrons in the ionosphere—the part of the Earth's atmosphere in which atoms are ionized by radiation from the Sun. Radio signals sent between GNSS navigational satellites and receivers in a low orbit around the Earth are delayed, and their paths are bent, by the presence of electrons in the ionosphere. Previous attempts to use these observations to estimate the electron density have been based on the delay of the signals. The new approach outlined in this paper uses the bending incurred by the radio waves instead. Such "bending angles" have been used to infer some properties of the lower atmosphere for many years, and are widely available. We demonstrate that by extending these measurements to greater heights, they can provide useful information about the ionosphere as well.

1. Introduction

The radio occultation instrument on EUMETSAT's Metop Second Generation (Metop-SG) satellites will measure up to 600 km above the Earth's surface, and provide measurements for space weather applications. The retrieval of ionospheric profile information from GNSS radio occultation (GNSS-RO) is well established (e.g., Hajj & Romans, 1998; Schreiner et al., 1999), using the Abel transform, which, for a circularly symmetric electron densities N_e and (therefore) refractive indices n, and bending angles α that are available at all impact parameters a, is given by

$$n(x) = \exp\left\{ \frac{1}{\pi} \int_x^\infty \frac{\alpha(a)}{\sqrt{a^2 - x^2}}\, da \right\}. \tag{1}$$

However, the 600 km upper limit for the Metop-SG observations complicates the application of the Abel transform, and alternative ionospheric retrieval techniques are needed for this particular truncated geometry.

Lyu et al. (2019) have presented two new approaches to this problem. The first, AVHIRO (Abel-VaryChap Hybrid density profile from topside Incomplete RO data), is a least-squares minimization which generates electron density profiles and the "carrier phase ambiguity" by comparing predicted and observed carrier phase differences between the L1 and L2 signals. Electron densities above 500 km are modeled by a VaryChap electron density layer (see Section 4). AVHIRO is currently too slow for operational applications, although this aspect is being developed (Dr Hernández-Pajares, pers. comm.). The second method, SEEIRO (Simple Estimation of Electron

density profile from topside Incomplete RO data), which iterates toward an electron density profile by modeling the densities above 500 km by a (simplified) VaryChap layer, whose parameters are estimated from Abel-inverted lower altitude electron densities at the previous iteration, is fast enough for operational implementation but probably not yet accurate enough.

In this work, we present a new method based on a variational (or "optimal estimation") approach (Rodgers, 2000). The one dimensional variational (1D-Var) retrieval approach is used extensively in neutral atmosphere GNSS-RO applications (e.g., Healy & Eyre, 2000; Palmer et al., 2000). Variational retrieval is more flexible than the Abel transform inversion, as it does not rely on an idealized measurement geometry. Specifically, the truncated Metop-SG geometry can be accounted for in a straightforward manner as part of the forward problem, $\mathbf{y} = \mathcal{H}(\mathbf{x})$, which maps the electron density profile information, $\mathbf{x}$, to an observation $\mathbf{y}$—bending angles here. This is in contrast to most GNSS-RO ionospheric retrievals, which, with the exception of Hajj and Romans (1998), are based on slant total electron contents (STECs), S. This is defined as the electron density N_e integrated along the path P between the GNSS and LEO satellites, thus:

$$S = \int_P N_e(s)ds. \tag{2}$$

However, STEC values are *relative* quantities. The ionospheric delay experienced by a radio occultation signal depends on its frequency. This would appear to allow the ionospheric effect to be isolated, and the STEC to be calculated, by taking the difference between the delays at two separate frequencies. This is not directly possible, however, because of a remaining a phase ambiguity, which arises because the phase measurements are known only modulo the wavelength (e.g., Dyrud et al. (2008)), and because of differential code biases (DCBs), introduced by both satellite and receiver (e.g., Equation 2 in Montenbruck et al. (2014)). Both these effects need to be accounted for before the STEC can be estimated. In this paper we show how bending angles are related to the vertical derivative of STEC, $\partial S/\partial a$, where a is the impact parameter of the ray path. Taking the derivative removes the sensitivity to any constant offsets (the DCBs) and simplifies the problem.

The standard Abel transform, Equation 1, relates profiles of refractivity and bending angle that range from the surface to infinity. Real observations do not of course extend that far, and bending angles above the maximum height of the Metop-SG measurements, 600 km, are not small enough to be neglected. There are two standard Abel transform methods of generating electron density profiles from such truncated radio occultation signals (e.g., Schreiner et al., 1999). Neither requires absolute TEC, but each faces a difficulty. The first uses bending angles, and requires knowledge of the refractive index at the LEO. The second uses the vertical derivative of the slantwise TEC (or excess phase) (e.g., Equation 3 of Lei et al. (2007)). Unfortunately this necessarily diverges as impact heights approach the LEO, and the (integrable) singularity must be handled somehow. By contrast, in the method proposed in this paper the electron density at any height, including that of the LEO, can be derived from the parameters of the presumed VaryChap layer(s). This also means that electron densities can be inferred above the maximum height of the observations, which makes the method ideally suited for use with truncated bending angles. By construction, the method also enforces positive electron densities, thereby avoiding a failing that can be suffered by Abel inversion in the presence of horizontal gradients (see, e.g., Case 4 in Section 6.1). Finally, the variational method automatically generates measures of the quality of any particular retrieval (number of iterations, cost function at convergence, solution error covariance matrix, etc.)—see Section 2. Note, however, that the retrieval method described in this paper, being one-dimensional, cannot handle horizontal gradients because it assumes spherically symmetric refractivity/electron density fields. This drawback is shared by Abel transform methods.

We note that bending angles are assimilated in most global numerical weather prediction (NWP) data assimilation systems without bias correction, and they are considered "anchor measurements" because they constrain the bias corrections applied to other measurements (e.g., Poli et al., 2010). However, in the ionospheric data assimilation literature, radio occultation STECs with bias corrections (DCBs) are still usually assimilated, for example, Angling and Cannon (2004), Yue et al. (2011), and Angling et al. (2021).

The purpose of this paper is to describe the theoretical basis of the new ionospheric 1D-Var retrieval approach. This includes a description of the multiple Vary-Chap layer electron density profile, the relationship between the derivative of the slant TEC and bending angle, and the 1D-Var assimilation system. A companion paper, Elvidge et al. (2023), presents a full statistical validation of the method.

2. Data Assimilation Preliminaries

In a variational retrieval system the estimate of the state vector is obtained using the observations combined with a priori information. Since the observations, and a priori information, are only known up to certain levels of confidence these are usually described using probability density functions (PDFs). By assuming Gaussian PDFs, and that the observation and background errors are uncorrelated, Bayes' theorem can be used to show that this problem can be framed as the minimization of a cost function J (e.g., Talagrand, 2010):

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y}_o - \mathcal{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}_o - \mathcal{H}(\mathbf{x})). \tag{3}$$

Here, $\mathbf{x}_b$ is the background state (specifically, the parameters that define the vertical electron density profile), $\mathbf{B}$ is the background error covariance matrix, $\mathbf{y}_o$ is the set of observations (specifically, bending angles as a function of impact parameter), $\mathbf{R}$ is the observation error covariance matrix, and $\mathcal{H}$ is the forward operator, which generates pseudo observations from a particular state.

The retrieval solution, $\mathbf{x}_a$, is the state which minimizes the cost function, and it should be consistent with both the background $\mathbf{x}_b$ and the observations $\mathbf{y}_o$, to within their expected error statistics. The bending angle profile provides useful information about all the ionospheric variables in the state vector $\mathbf{x}$, and the uncertainty in the estimate of $\mathbf{x}$ is significantly reduced as a result of making the measurement. In this context, a useful feature of a 1D-Var retrieval is that it provides the following estimate of the theoretical solution error covariance matrix, $\mathbf{A}$ in terms of $\mathbf{B}$ and $\mathbf{R}$

$$\mathbf{A} = \left(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}\right)^{-1} \tag{4}$$

where $\mathbf{H}$ is the matrix of partial derivatives of the forward model with respect to the state vector elements: $\mathbf{H} = \partial \mathcal{H}/\partial \mathbf{x}$. In well-posed problems the diagonal elements of the $\mathbf{A}$ matrix are significantly smaller than the corresponding diagonal elements of the $\mathbf{B}$ matrix ($A_{i,i} \ll B_{i,i}$).

The 1D-Var approach should be more robust to measurement noise than the Abel transform because it is a weighted least-squares approach; it will not over-fit the measurements if the assumed observation error statistics, $\mathbf{R}$, are a reasonable approximation of the actual observation error statistics. The retrieval also produces useful diagnostics of the quality of the solution. These include the number of iterations required to converge, and the cost at convergence, $J(\mathbf{x}_a)$, the expectation value of which is half the size of the observation vector, $E[2J(\mathbf{x}_a)] = m \pm \sqrt{2m}$ (e.g., Rodgers (2000)), if the assumed error statistics are well specified.

Practical details of the solution method are discussed in Section 5.

3. Theory and Observations

3.1. Theory

To overcome the need for estimation of the Differential Code Biases (DCBs) in this work the derivative of slant-wise TEC S, with respect to the impact parameter a, $\partial S/\partial a$, is assimilated. The latter is the quantity used in the Abel transform solution for refractivity (see Equation 14 in Schreiner et al. (1999)), and we will show that, to a good approximation, it is proportional to the difference between the L2 and L1 bending angles, plus a term that involves the electron density at the LEO.

The STEC is the integrated electron density (N_e) along the path P between the GNSS and the LEO satellites (see Equation 2). To a first approximation the delay ϕ_i (in m) in the phase of the carrier wave (of frequency f_i), relative to the vacuum, accumulated along the path P is given by

$$\begin{aligned} \phi_i &= \int_P (n_i - 1)\, ds \\ &= \int_P -\kappa N_e / f_i^2\, ds \\ &= -\kappa S / f_i^2 \end{aligned} \tag{5}$$

where n_i is the refractive index at frequency f_i, which is approximately given by (e.g., Schreiner et al., 1999)

$$n_i - 1 \approx -\kappa N_e/f_i^2, \tag{6}$$

in which the proportionality constant $\kappa \approx 40.3 \text{ m}^3 \text{ s}^{-2}$.

If the electron density is assumed to be only a function of height (i.e., is spherically symmetric), the STEC S between a GNSS satellite at radius r_G and a LEO satellite at radius r_L is easily shown, from Equation 2 to be given approximately by

$$
\begin{aligned}
S &= \int_a^{r_L} \frac{r N_e(r)}{\sqrt{r^2 - a^2}} \, dr + \int_a^{r_G} \frac{r N_e(r)}{\sqrt{r^2 - a^2}} \, dr \\
&= \left[\int_a^{r_L} + \int_a^{r_G} \right] \frac{r N_e(r)}{\sqrt{r^2 - a^2}} \, dr
\end{aligned}
\tag{7}
$$

in which $N_e(r)$ is the electron density at radius r, and the impact parameter a is nearly equal to the radial distance to the tangent point, r_T. (For example, even for a high electron density of $3 \times 10^{12} \text{ m}^{-3}$ at a height of 300 km, which is appropriate for the F2 peak in daytime, solar maximum conditions, $n_i - 1 \approx -8 \times 10^{-5}$ at the L2 frequency of about 1.2 GHz. This means that a and r_T differ by less than 600 m.)

Assuming the electron density is zero at the GNSS, integrating Equation 7 by parts gives

$$S = N_e(r_L)\sqrt{r_L^2 - a^2} - \left[\int_a^{r_L} + \int_a^{r_G} \right] \sqrt{r^2 - a^2} \frac{d N_e(r)}{dr} \, dr \tag{8}$$

The derivative of Equation 8 with respect to a is:

$$\frac{\partial S}{\partial a} = -\frac{N_e(r_L)a}{\sqrt{r_L^2 - a^2}} + a \left[\int_a^{r_L} + \int_a^{r_G} \right] \frac{d N_e(r)/dr}{\sqrt{r^2 - a^2}} \, dr. \tag{9}$$

Equation 9 is singular at $a = r_L$, as noted in Lei et al. (2007). This singularity is necessary for the Abel inversion using slantwise TECs to be well-behaved.

Note that the integrals in Equation 9 are closely proportional to the standard Abel transform approximation to the ionospheric bending angle at impact parameter a and frequency f_i, $\alpha_i(a)$, namely (Angling et al., 2018; Kursinski et al., 1997; Vorob'ev & Krasil'nikova, 1994):

$$
\begin{aligned}
\alpha_i(a) &= -2a \int_{r_T}^{\infty} \frac{dn_i(r)/dr}{n_i \sqrt{(n_i r)^2 - a^2}} \, dr \\
&\simeq \frac{\kappa a}{f_i^2} \left[\int_a^{r_L} + \int_a^{r_G} \right] \frac{d N_e(r)/dr}{\sqrt{r^2 - a^2}} \, dr,
\end{aligned}
\tag{10}
$$

using $\log n_i \approx n_i - 1 = -\kappa N_e/f_i^2$, as in Equation 6, $n_i \approx 1$ (and therefore $r_T \approx a$) where appropriate, and replacing the infinite upper limits of the integrals by distances to the satellites.

The vertical gradient of the phase delay in Equation 5 is $\partial \phi_i/\partial a = -(\kappa/f_i^2)\partial S/\partial a$. Therefore, given measurements at two frequencies, f_1 and f_2, and assuming straight line paths with the same impact parameter, a, we can use Equation 9 to write the difference in the phase delay gradients as:

$$
\begin{aligned}
\partial(\phi_1 - \phi_2)/\partial a &= -\kappa \left(\frac{1}{f_1^2} - \frac{1}{f_2^2} \right) \frac{\partial S}{\partial a} \\
&= \kappa \left(\frac{f_1^2 - f_2^2}{f_1^2 f_2^2} \right) \left(a \left[\int_a^{r_L} + \int_a^{r_G} \right] \frac{d N_e(r)/dr}{\sqrt{r^2 - a^2}} \, dr - \frac{N_e(r_L)a}{\sqrt{r_L^2 - a^2}} \right).
\end{aligned}
\tag{11}
$$

There is a fortuitous cancellation of errors relevant to this problem. In the processing of real measurements, bending angle values are derived from the Doppler shift values assuming that the refractive index at the LEO satellite is unity: $n(r_L) = 1$. For circular orbits this assumption does not affect the impact parameter, a, but it introduces

a frequency dependent negative bias, b_i, in the observed bending angles. (The proportionality between Doppler shift and impact parameter for circular orbits is also exploited in the Full Spectral Inversion (FSI) technique (Jensen et al., 2003).) This is equal to (see Equation A6 of Schreiner et al. (1999)):

$$b_i = -\frac{\kappa}{f_i^2} \frac{N_e(r_L)a}{\sqrt{r_L^2 - a^2}},$$

(12)

which, being inversely proportional to the square of the frequency, cancels out in the usual dual frequency ionospheric correction of the bending angles (Vorob'ev & Krasil'nikova, 1994). The observed ionospheric bending angles, $\tilde{\alpha}_i(a)$, will therefore be related to the true bending angles, $\alpha_i(a)$, (see Equation 10), according to

$$\tilde{\alpha}_i(a) = \alpha_i(a) + b_i(a)$$
$$= \frac{\kappa a}{f_i^2} \left[\int_a^{r_L} + \int_a^{r_G} \right] \frac{dN_e(r)/dr}{\sqrt{r^2 - a^2}} \, dr - \frac{\kappa a}{f_i^2} \frac{N_e(r_L)}{\sqrt{r_L^2 - a^2}}.$$

(13)

Comparison with Equation 11 therefore reveals that

$$\frac{\partial(\phi_1 - \phi_2)}{\partial a} \simeq \tilde{\alpha}_2(a) - \tilde{\alpha}_1(a).$$

(14)

Therefore, the ionospheric retrieval can use either observed bending angle differences or the derivative of phase delay differences with respect to impact parameter.

3.2. Observations

Accurate estimation of the uncertainty of the observations is crucial to ensure the observations are not over-fitted. In general it is difficult to estimate the uncertainty of observations, since in most cases there is no additional, independent, reference truth.

As well as errors in the observations themselves, there are also errors arising from the use of a 1D-retrieval rather than a more realistic 3D-retrieval. Such errors obviously depend on the electron density distribution, which varies with location, time of day and solar conditions. To estimate these errors, electron density distributions appropriate to 143 Metop-SG occultations have been generated with 1D and 3D versions of the climatological ionosphere electron density model NeQuick (Nava et al., 2008), at four different solar activity levels, represented by the solar radio flux at 10.7 cm ($F_{10.7}$). Specifically we use $F_{10.7}$ values of 80, 130, 180, 230 solar flux units (sfu) (where 1 sfu = 10^{-22} Wm^{-2} Hz^{-1}).

Define $N_i^{3D}(h)$ as the electron density at height h that results from calculating the phase delays incurred by signals propagating through the ith occultation plane, using 3D NeQuick electron densities, and then passing the delays through an Abel transform to infer an electron density profile at the tangent point. $N_i^{1D}(h)$ is the same, except that it uses a spherically symmetric electron density field, which matches the 1D NeQuick field at the occultation tangent point. The difference between the two is a measure of the errors in an inferred electron density distribution incurred by the use of a 1D rather than a 3D retrieval. For each altitude h the absolute mean electron density error ϵ across all $4 \times 143 = 572$ simulated occultations is calculated thus:

$$\epsilon(h) = \frac{1}{572} \sum_{i=1}^{572} \left| N_i^{3D}(h) - N_i^{1D}(h) \right|.$$

(15)

To ensure that "rare" geometries, such as occultations that cross the day/night terminator, do not distort the statistics, any outliers at each altitude are removed. Outliers are defined as any value which is less than 1.5 times the interquartile range (IQR) below the first quartile (Q1) of the data, or greater than 1.5 times the IQR above the third quartile (Q3), that is, only data, d, in the range

$$Q1 - 1.5IQR \le d \le Q3 + 1.5IQR$$

(16)

are kept at each altitude. The bending angle profile resulting from passing the 1D electron density difference $\epsilon(h)$ in Equation 15 through the Abel transform (which is *linear* in N_e) is shown in Figure 1. The blue vertical line at 2.10 μrad shows the average mean absolute error, and shows that 2.0 μrad is an excellent single value
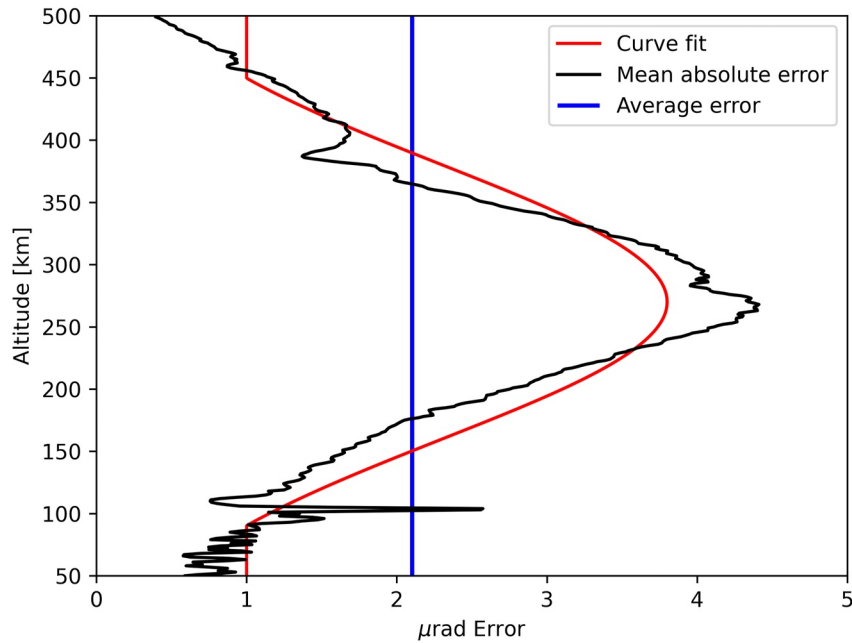
Figure 1. Mean absolute difference in bending angles derived from electron density differences $\epsilon(h)$ defined by Equation 15. Vertical average in blue. Fitting function in Equation 17 in red.

error to use. It is comparable to the value used in neutral atmosphere applications in the middle/upper stratosphere. For example, ECMWF uses 3.0 µrad above about 32 km when assimilating GNSS-RO data. It is clear, however, that the error is height-dependent. Rather than directly using the average errors from Figure 1, which are subject to sampling errors themselves, the following Gaussian fit to the data could be used in a 1D-Var retrieval:

$$\text{bending angle error} = 3.8 \ \ \mu\text{rad} \times \exp\left[-\frac{1}{2}\left(\frac{h - 270}{110}\right)^2\right] \tag{17}$$

where h is the altitude in km. A minimum value of 1.0 µrad is imposed. Equation 17 is plotted in red in Figure 1. This function provides a similar residual sum of squares error as using a fifth order polynomial fit, but without the associated numerical instabilities.

4. Electron Density Model

To reconstruct the ionosphere from the observations described in Section 3.2 a model electron density state is needed for the retrieval. Here we assume that the ionosphere can be modeled as a collection of one-dimensional 'Vary-Chap' electron layers.

The standard description (e.g., Wang et al., 2021) of a Vary-Chap electron density profile (described in Reinisch et al. (2007) as a generalization of an α-Chapman profile based on the work by Rishbeth and Garriott (1969)) with a height-dependent scale height, $H(h)$, is given by

$$N_e(h) = N_m \left(\frac{H(h)}{H_m}\right)^{-\frac{1}{2}} \exp\left(\frac{1}{2}\left[1 - u - \exp(-u)\right]\right), \tag{18}$$

where

$$u = \int_{h_m}^{h} \frac{1}{H(h)} \, dh. \tag{19}$$

In these equations N_m is the peak electron density, which occurs at $h = h_m$. H_m is the scale height at h_m.
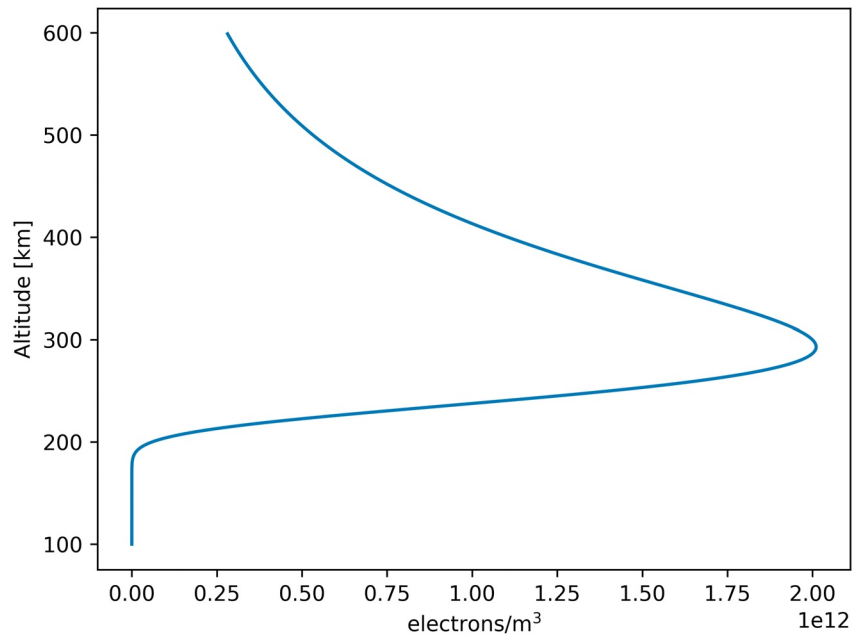
Figure 2. Example of a Vary-Chap profile with "standard" parameters given by Equation 26.

The estimation of the Vary-Chap parameters in AVHIRO is based on minimizing a least squares cost function, with two terms (Lyu et al., 2019). The first term is the fit to a previous state estimate (rather than a "background" constraint); the second is the fit to the observed STEC measurements, which can be calculated from the L1 minus L2 phase delay measurements.

In practice the height-dependent scale height, $H(h)$, in Equation 18 can be difficult to determine (Kutiev et al., 2009; Nsumei et al., 2012; Wang et al., 2021). Throughout this work the scale height used in the Vary-Chap profile varies linearly with height (above the peak). This means that Equations 18 and 19 reduce to

$$N_e(h) = N_m \left(\frac{H(h)}{H_m} \right)^{-\frac{1}{2}} \exp\left(\frac{1}{2}[1 - u - \exp(-u)] \right), \quad \text{where} \tag{20}$$

$$u = \log(H/H_m)/k \quad \text{and} \tag{21}$$

$$H(h) = H_m + k(h - h_m). \tag{22}$$

These equations apply if $h > h_m$. If $h \leq h_m$, or if $k \leq 10^{-3}$, the 'standard' Chapman layer approximation, which is given by the limit as k tends to 0 of the above, applies. This is given by:

$$N_e(h) = N_m \exp\left(\frac{1}{2}[1 - u - \exp(-u)] \right), \quad \text{where} \tag{23}$$

$$u = (h - h_m)/H_m \quad \text{and} \tag{24}$$

$$H(h) = H_m. \tag{25}$$

The Vary-Chap profile defined by Equations 20–22, using the "standard F2" values of

$$N_m = 2 \times 10^{12} \text{ electrons m}^{-3}$$
$$h_m = 300 \text{ km}$$
$$H_m = 50 \text{ km} \tag{26}$$
$$k = 0.15$$

is plotted between 100 and 600 km in Figure 2.

Table 1
Default Vary-Chap Parameters for up to 5-Layers

	Parameters			
	N_m (m^{-3})	h_m (km)	H_m (km)	k
D-Region	2×10^8	70	5	0.05
E-Region	5×10^{10}	110	20	0.05
F1-Region	5×10^{11}	205	30	0.05
F2-Region	2×10^{12}	300	50	0.15
"Topside"	3×10^{11}	500	250	0.50

This one layer Vary-Chap profile provides a good approximation to the standard electron density profile. However, a better approximation can be formed through the addition of multiple Vary-Chap profiles, for example, one for each ionospheric layer (D, E, F1, and F2). It can also be beneficial to introduce a "topside" layer, to try to account for the systematic underestimation of electron densities above the F2 peak height given by the Vary-Chap model. (Prol et al. (2019) show that an F2 scale height that varies quadratically, rather than linearly, with height gives a better fit to topside sounder measurements). Various multi-layer Vary-Chap profiles using the parameters in Table 1 are shown in Figure 3.

The resulting five-layer Vary-Chap model provides an excellent approximation to the ionosphere, with a realistic looking E and F1 region. Using a fifth layer for the topside addresses the underestimation as highlighted by Prol et al. (2019). In practice the addition of the D-region layer has very little impact on the overall profile. Figure 4 shows the difference between a five-layer (F2+F1+E+D+Topside) and four-layer (F2+F1+E+Topside) version, that is, the difference in electron density by including the D-layer. As expected the impact of including a D-Region in the model increases with decreasing altitude, although the maximum difference is approximately 2×10^7 m^{-3}, which is over three orders of magnitude smaller than the absolute density at that altitude. The differences are insignificant, yet they add considerable time to the calculations. (See Section 6.2).

5. Practical Considerations

The equations of Sections 3.1 and 4 are encoded numerically as follows. The forward model calculates the electron density profile defined by the four Vary-Chap parameters $\{N_m, h_m, H_m, k\}$ according to Equations 20–25. The integrals needed to calculate the slantwise TEC, S, according to Equation 8 are estimated numerically by Simpson's rule, using at least 30 points between the tangent point and the LEO (counted twice), and 10 points between the LEO and the GNSS (counted once). The integrable singularity at $r = a$ is handled by working in terms of $\cosh^{-1}(r/a)$ rather than r. The procedure is repeated for each Vary-Chap layer. The resulting $\partial S/\partial a$ is related to the vertical gradient of the excess phases according to Equation 11, and this in turn is equated to the bending angle differences according to Equation 14. The resulting forward modeled bending angles $\mathcal{H}(\mathbf{x})$ are then compared to the observed bending angle differences $\mathbf{y}_o$ over the range 500–175 km, which is considered broad enough to
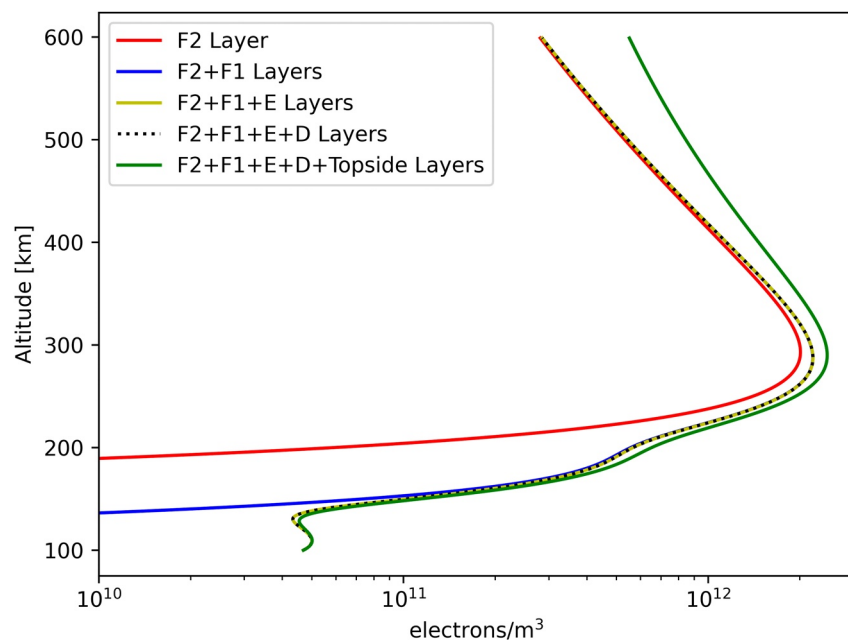

Figure 3. Examples of multi-layer Vary-Chap profiles with the parameters given in Table 1.
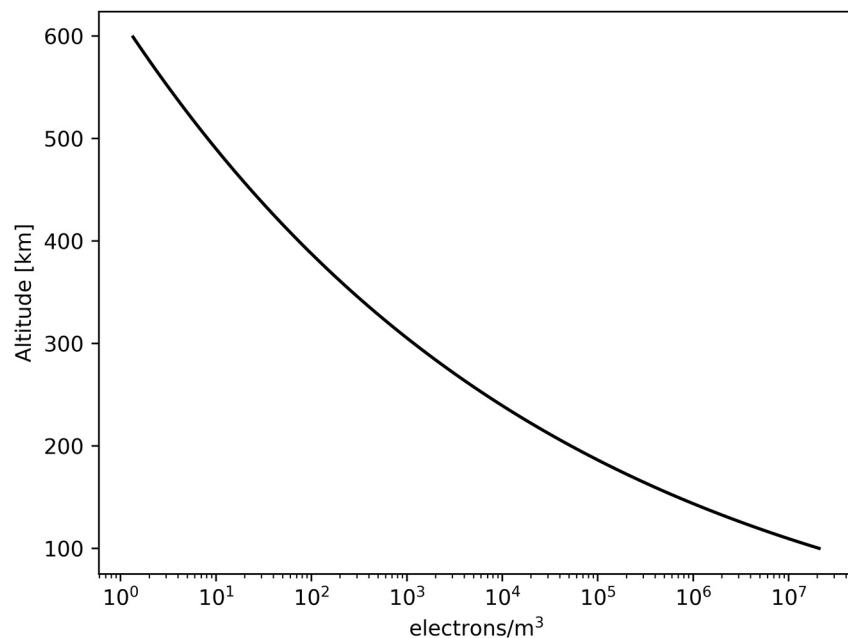
Figure 4. Electron density difference between using a five-layer (F2+F1+E+D+Topside) and four-layer (F2+F1+E+Topside) Vary-Chap model.

capture the key features of the ionosphere, without complicating the solution by trying to model smaller scale structures at lower altitudes.

The tangent linear model, $\mathbf{H}$, which is needed to minimize the cost function $J(\mathbf{x})$ of Equation 3, and the solution error covariance matrix $\mathbf{A}$ via Equation 4, is calculated by evaluating manually differentiated counterparts of the non-linear model.

The cost function $J(\mathbf{x})$ is minimized by means of a standard version of the Levenberg-Marquardt minimization algorithm (Press et al., 1992), in which the 'diagonal weighting factor' (usually called λ), which determines how closely the change in $\mathbf{x}$ follows the path of steepest descent, is multiplied by 0.1 if the cost function is decreasing, and by 100 if the cost function starts increasing. Iterations are deemed to have converged when changes to the state vector or to the cost function are sufficiently small. In addition, 'unphysical' state vector components, such as negative peak electron densities, are handled during the iterative phase of the minimization, usually by resetting them to a few percent of the associated errors, that is, the square roots of the diagonal elements of $\mathbf{B}$. These are deliberately chosen to be rather large—typically around $\{5.0 \times 10^{11}$ m^{-3}, 100 km, 20 km, 0.05$\}$ for $\{N_m, h_m, H_m, k\}$ respectively.

The results of this paper have been produced by code written in Python 3.0. Another, widely available version, which is written in Fortran 95, has been part of the Radio Occultation Processing Package (ROPP) since version 11.0 (released January 2022). ROPP is a collection of software, build scripts, test scripts and documentation, which is intended to help users to process RO data. It is maintained, developed and supported by EUMETSAT, through the Radio Occultation Meteorology Satellite Applications Facility (ROM SAF), and freely available to download from its website (ROM SAF, 2023). Users should bear in mind that ROPP is under constant development, and that its code may not therefore exactly match that described in this paper. The differences, however, should be small.

6. Results

6.1. Example Retrievals

The 1D-Var retrieval results are compared with the AVHIRO retrieval (Lyu et al., 2019), and an Abel transform solution. The Abel transform retrievals are not absolute electron density values, because we do not add the electron density at the LEO to the profile. In addition, we have not corrected (calibrated) the bending angles to
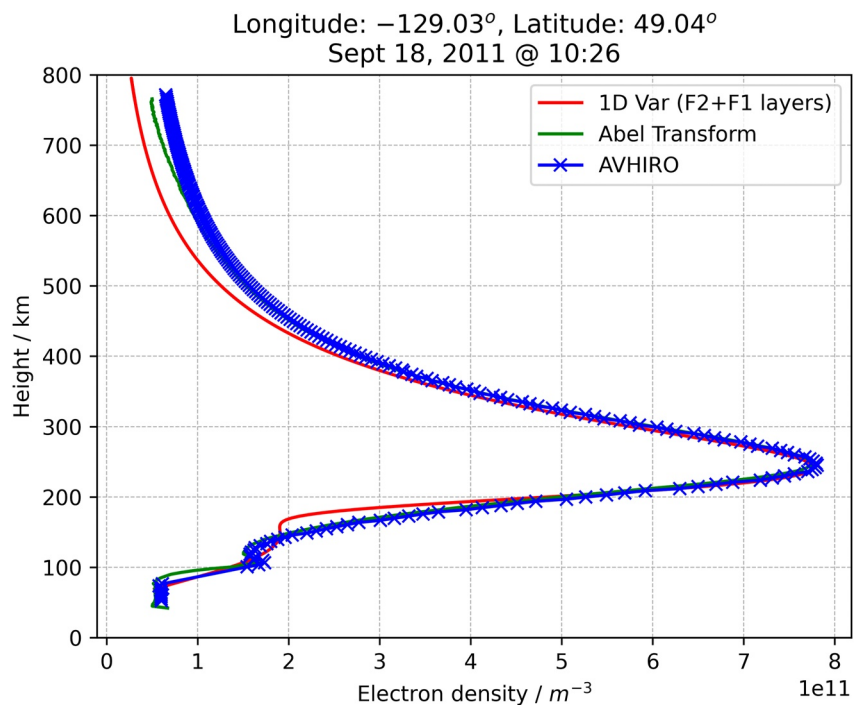
Figure 5. Comparison between the 1D-Var electron density profile retrieval with the Abel transform, and AVHIRO model (Lyu et al., 2019).

only include the section of the path below the LEO satellite, by subtracting positive elevation values with the same impact parameter (see Section 3.1 of Schreiner et al. (1999)). However, the Abel solution should indicate whether or not the 1D-Var results look reasonable. Our implementation of the Abel transform assumes that the L2–L1 bending angle differences vary linearly in the vertical between observations. This means that the Abel transform is linear, so the retrieved electron density profile can be computed as a matrix multiplied by a vector of (differenced) bending angles.

The performance of the 1D-Var is illustrated with four cases (see Elvidge et al. (2023) for the complete analysis), which have been chosen for their differing retrieval characteristics and convergence properties.

6.1.1. Case 1

The first case, shown in Figure 5, can be considered a good retrieval, and is the same example as shown in the top right panel of Figure 4 in Lyu et al. (2019). The 1D-Var retrieval uses two layers, an F2 layer and F1 layer, two layers being a reasonable compromise between accuracy (which increases with number of layers) and robustness and CPU time (which decrease with it)—see Section 6.2. (Also see Elvidge et al. (2023), which contains an extensive study of the impact of adding more and more layers to the 1D-Var retrieval method described in this paper). A one-off Abel Transform solution has also been added, to replicate Figure 4 of Lyu et al. (2019) more completely. The three approaches agree very similarly around the profile peak (between approximately 200 and 400 km) but the 1D-Var solution differs from the AVHIRO and the Abel transform solutions outside this region.

6.1.2. Case 2

The second case, shown in Figure 6, compares the 1D-Var retrieval with those from COSMIC-2 (UCAR Cosmic Program, 2019) and a closely located ionosonde. Ionosondes step through a range of HF frequencies transmitted vertically upwards and measure the return echoes, which enables them to image the vertical profile of the ionosphere up to the peak density. (Ionosonde density profiles above the peak are less reliable, because they depend on a background model, which is why they are plotted as dashed lines in this paper.) Ionosonde observations are widely used as reference observations for comparative studies (e.g., Feltens et al., 2011; Elvidge et al., 2014; Scherliess et al., 2011). Here an ionosonde observation (profile) from within 200 km of the location of the occultation have been used to demonstrate the 1D-Var retrieval (A full statistical study is undertaken in Elvidge
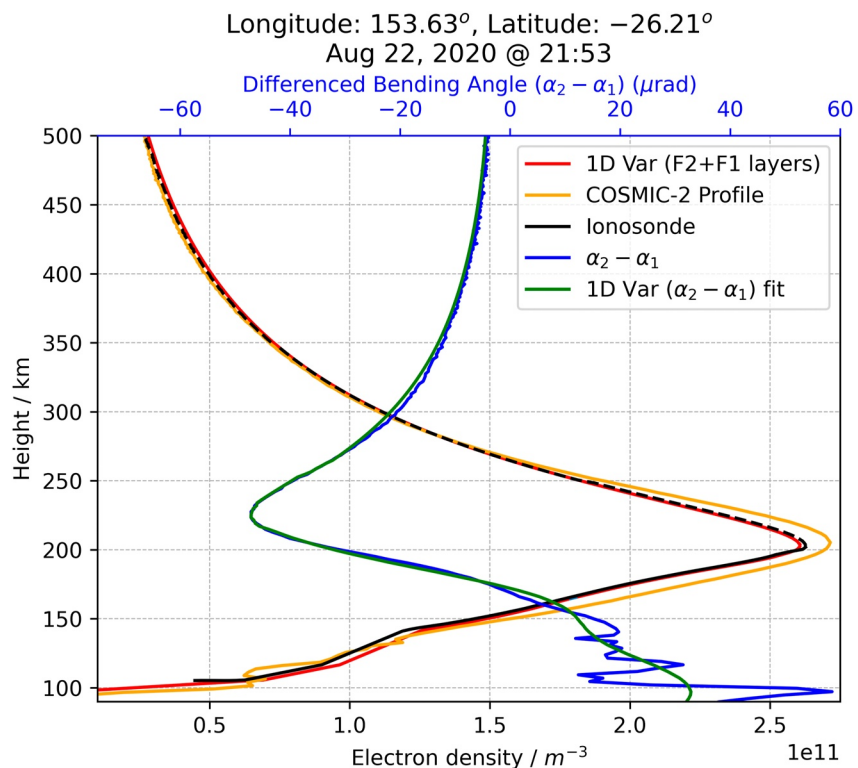
Figure 6. Comparison between the 1D-Var electron density profile retrieval with the Abel transform retrieval of a COSMIC-2 profile and a nearby ionosonde. Also shown are the observations and the forward modeled solution in bending angle space (i.e., $\alpha_2 - \alpha_1$). (Ionosonde profiles above the peak appear as dashed lines, as they are less reliable there).

et al. (2023)). Also shown is the convergence in bending angle space, that is, the observed (blue) and forward modeled solution (green) bending angle differences $\alpha_2 - \alpha_1$.

In this case the 1D Var solution is in very close agreement with the ionosonde observations throughout the whole profile. Above approximately 300 km and below 160 km there are some deviations from the observations, but these are small. The COSMIC-2 retrieval is also very good, but overestimates the peak electron density. The COSMIC-2 profile shows more structure below 150 km, presumably resulting from the variability in the $\alpha_2 - \alpha_1$ observations in that region. The 1D-Var solution is the result of fitting observations between 500 and 175 km (see Section 5), and therefore has no knowledge of this structure. Even so, the forward modeled bending angle differences give a reasonable fit to the observations.

6.1.3. Case 3

The third case, Figure 7, again compares the 1D-Var retrieval with COSMIC-2 and a nearby ionosonde (in the same manner as described in Section 6.1.2) and both the COSMIC-2 and 1D-Var retrievals are fairly similar. The peak heights of the profiles, at about 200 km, are similar and both are slightly higher than that reported by the ionosonde. (But note that autoscaled ionosonde observations should be treated carefully, and the errors in the F2 peak height can be reasonably large (as much as 10%, according to Themens et al. (2022)). The COSMIC-2 retrieval slightly overestimates the peak density whilst the 1D-Var slightly underestimates it. Between the peak of the electron density profile and about 150 km, the COSMIC-2 retrieval more closely matches the observations. Above about 275 km the two retrievals are very similar. (Differences from the ionosonde profile in this region should be treated skeptically, for reasons explained in Section 6.1.2).

6.1.4. Case 4

In the final case, shown in Figure 8, the COSMIC-2 retrieval is generally closer to the ionosonde observation than the 1D-Var retrieval. Again, both retrievals have similar heights of peak density, but the COSMIC profile more closely matches the observations above and below the peak. However the COSMIC-2 retrieval returns negative electron densities below 100 km, whereas the 1D-Var retrieval cannot, for reasons described in Section 5.
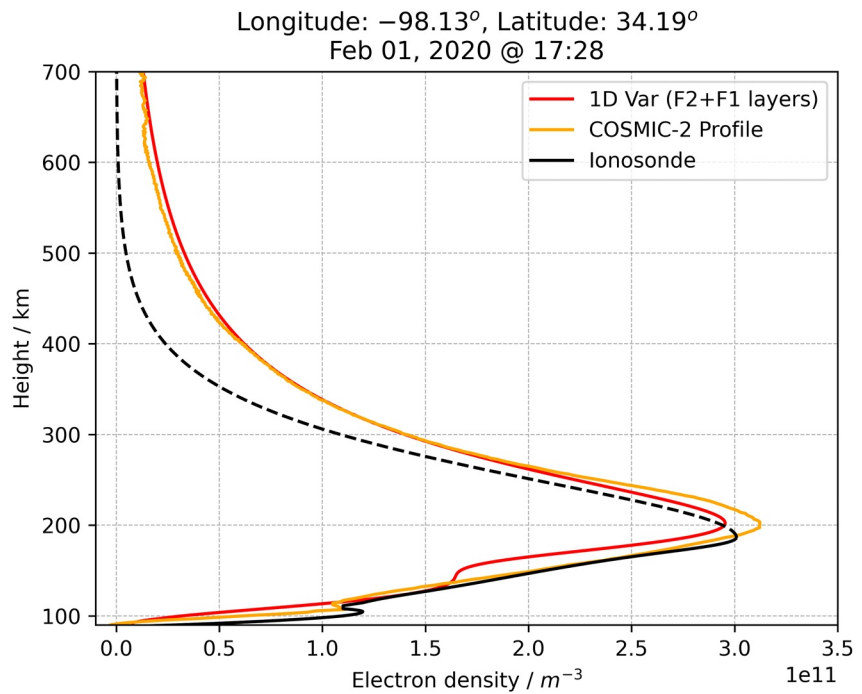
Figure 7. Comparison between the 1D-Var electron density profile retrieval with the Abel transform retrieval of a COSMIC-2 profile and a nearby ionosonde. (Ionosonde profiles above the peak appear as dashed lines, as they are less reliable there).

Figure 8. Comparison between the 1D-Var electron density profile retrieval with the Abel transform retrieval of a COSMIC-2 profile and a nearby ionosonde. (Ionosonde profiles above the peak appear as dashed lines, as they are less reliable there).
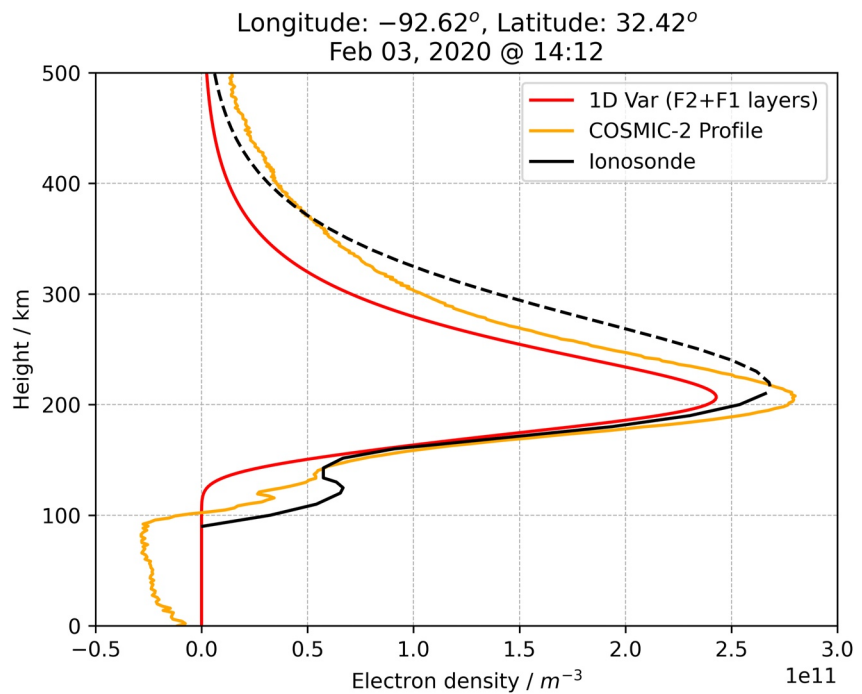
Table 2

Average Time Per Iteration, Percentage of Observations Which Converge, and Statistics of the Number of Iterations Needed for Convergence, for up to Five Vary-Chap Layers

	Time per iteration (seconds)	Observations that converged	Mean ± std dev no. of iterations
1 layer	0.99	98.6%	16.2 ± 7.1
2 layers	2.22	85.5%	34.1 ± 8.0
3 layers	4.06	66.7%	28.1 ± 9.1
4 layers	6.72	65.2%	28.7 ± 9.1
5 layers	9.98	58.7%	30.7 ± 11.4

6.2. Convergence and Computational Expense

Elvidge et al. (2023) provides a full statistical analysis of the results of the 1D-Var retrievals compared to Abel-transform-based COSMIC-2 retrievals. This section simply provides a flavor of the 1D-Var ionospheric retrieval method's performance on 143 bending angle profiles provided by the Institut D'Estudis Espacials de Catalunya (IEEC) in Barcelona, Spain. These are COSMIC-1 measurements from 18 September 2011. The data contains the "geometry free" ($\phi_1 - \phi_2$) phase differences as function of impact parameter, a, up to the COSMIC-1 satellite, which operated at an altitude of about 800 km.

The simulated differenced bending angles $\alpha_2 - \alpha_1$ are computed with a finite difference approximation to $\partial(\phi_1 - \phi_2)/\partial a$. The vertical separation between the bending angles is typically 500 m. A common radius of curvature of 6,371 km is used for all cases since it was not included in the test data sets.

Although adding layers seems to improve the overall shape of the retrieved ionosphere, this comes at the cost of higher computational time and number of iterations required for the 1D-Var to converge. This is shown in Table 2. Note that the maximum number of iterations was set to 50. Any retrievals needing more iterations than this were recorded as convergence failures, and discarded from further analysis. On the basis of the 143 profiles studied here, two layers generate a reasonable fit to the observations at a moderate computational cost, which is suitable for illustrating the method. In the companion paper Elvidge et al. (2023), however, which uses a much better background model on a much larger sample, four layers (F2, F1, E and topside) work better, having a smaller RMS error with respect to ionosondes (albeit at a lower convergence rate). The correct number to use in any particular application will depend on the balance of the competing demands of speed, robustness and accuracy that the user feels appropriate for that application.

It can be seen from Table 2 that the average time per iteration increases as the number of layers increases. However after an immediate jump in the average number of iterations when using one- and two-layer models, the average number of iterations for converging observations remains fairly steady (at around 30). This leads to a monotonic increase in the average total length of time needed for an observation to converge using the 1D-Var, as shown in Figure 9. In addition, the percentage of observations that converge continually decreases as the number of layers increases, from 98.6% with just one layer to only 58.7% with five layers. More realistic retrieved electron density profiles therefore come at a cost in CPU computational time and robustness.
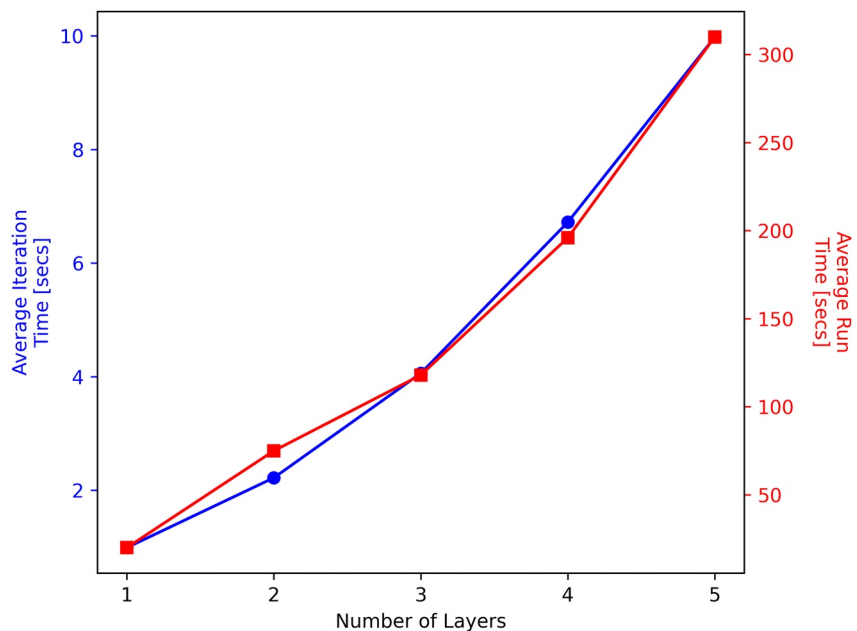

Figure 9. Average time taken per iteration (blue) and average length of time for an observation to converge (red).

7. Discussion and Conclusions

This paper has described a new 1D-Var ionospheric retrieval method. This approach can be applied to the Metop-SG measurement geometry, which will truncate the ionospheric GNSS-RO measurements around 600 km above the surface, as well as 'standard' GNSS-RO measurement geometry, in which measurements are assumed to be available at all heights.

The original plan was to use the difference between the slantwise TECs of the L1 and L2 radio signals, expressed as a function of impact parameter, but the L2-L1 observed bending angle differences have been used instead. This approach is closer to the way that GNSS-RO data is used in neutral atmosphere applications, such as operational NWP.

A new 1D ionospheric bending angle forward operator has been developed, which computes L2-L1 bending angle differences as a function of impact parameter. This operator assumes that the ionospheric electron density can be modeled as a collection of 'Vary-Chap' layers. As suggested by a comparison of bending angles generated by 1D- and 3D-NeQuick electron density models, the L2-L1 bending angle uncertainty is assumed to be a constant 2.0 µrad, and vertical error correlations are neglected. Some experiments were made using the more complicated error formula of Equation 17, but the results were found to differ little from the same experiments using 2.0 µrad. For simplicity and reproducibility, therefore, and to illustrate the method, the constant value was adopted.

We find that gradient-based minimization techniques can be successfully applied to this retrieval problem, and that the non-linear, nested exponential nature of the Vary-Chap profiles (see Equation 20) does not cause problems. This is in contrast to the assertion of Lyu et al. (2019). Typically, at least two out of three retrievals converge within 50 iterations, although this convergence rate decreases as more Vary-Chap layers are introduced. The iteration count, and the high final cost functions of the retrievals that fail to converge, provide useful diagnostic information, as does the automatically generated solution error covariance matrix **A**.

A few example 2-layer 1D-Var retrievals have been compared against the results of the AVHIRO method, Abel transform retrievals, and nearby ionosondes. In general the 1D-Var method performs at least as well as AVHIRO and the Abel transform, and produces a close fit to the observed bending angle differences $\alpha_2 - \alpha_1$ in the region where it is supposed to. By construction, 1D-Var also avoids one drawback of the Abel inversion technique, namely the generation of negative electron densities. Adding more layers can improve the fit, but at the cost of longer CPU times and a lower convergence rate.

The a priori information is essentially a first guess used to start the minimization, rather than a strong constraint on the final 1D-Var solution, because the background errors are (deliberately) rather large. Better a priori information might speed up convergence but would not necessarily improve the 1D-Var solution. It would, however, make it easier to screen out poor bending angles at the start of the minimization.

The key suggestion of this paper is that it may be possible to use differenced (L1-L2) bending angles in ionospheric data assimilation (DA) systems. Usually, such systems assimilate slantwise TEC values, but these require a correction for the Differential Code Biases (DCBs). Constant DCBs during the occultation, however, will not affect the bending angles derived from raw phase delays, and can therefore be ignored. This suggestion should at some point be tested in the context of a more formal data assimilation system.

Data Availability Statement

COSMIC-1 data are available from (UCAR, 2014). The processed files from the IEEC and AVHIRO output may be requested from Dr Hernández-Pajares, Universitat Politécnica de Catalunya (UPC), and are available from IEEC (2011).

References

Angling, M. J., & Cannon, P. S. (2004). Assimilation of radio occultation measurements into background ionospheric models. *Radio Science*, *39*(1), RS1S08. https://doi.org/10.1029/2002RS002819

Angling, M. J., Elvidge, S., & Healy, S. B. (2018). Improved model for correcting the ionospheric impact on bending angle in radio occultation measurements. *Atmospheric Measurement Techniques*, *11*(4), 2213–2224. https://doi.org/10.5194/amt-11-2213-2018

Angling, M. J., Nogués-Correig, O., Nguyen, V., Vetra-Carvalho, S., Bocquet, F.-X., Nordstrom, K., et al. (2021). Sensing the ionosphere with the Spire radio occultation constellation. *Journal of Space Weather and Space Climate*, *11*, 56. https://doi.org/10.1051/swsc/2021040

Space Weather

10.1029/2023SW003572

Dyrud, L., Jovancevic, A., Brown, A., Wilson, D., & Ganguly, S. (2008). Ionospheric measurement with GPS: Receiver techniques and methods. *Radio Science*, *43*(6), RS6002. https://doi.org/10.1029/2007RS003770

Elvidge, S., Angling, M. J., & Nava, B. (2014). On the use of modified Taylor diagrams to compare ionospheric assimilation models. *Radio Science*, *49*(9), 737–745. https://doi.org/10.1002/2014RS005435

Elvidge, S., Healy, S. B., & Culverwell, I. D. (2023). One-dimensional variational ionospheric retrieval using radio occultation bending angles: Part 2—Validation. *Journal of Geophysical Research: Space Physics*.

Feltens, J., Angling, M. J., Jackson-Booth, N., Jakowski, N., Hoque, M., Hernández-Pajares, M., et al. (2011). Comparative testing of four ionospheric models driven with GPS measurements. *Radio Science*, *46*(6), RS0D12. https://doi.org/10.1029/2010RS004584

Hajj, G. A., & Romans, L. J. (1998). Ionospheric electron density profiles obtained with the Global Positioning System: Results from the GPS/MET experiment. *Radio Science*, *33*(1), 175–190. https://doi.org/10.1029/97rs03183

Healy, S. B., & Eyre, J. R. (2000). Retrieving temperature, water vapour and surface pressure information from refractive index profiles derived by radio occultation: A simulation study. *Quarterly Journal of the Royal Meteorological Society*, *126*(566), 1661–1683. https://doi.org/10.1002/qj.49712656606

IEEC. (2011). AVHIRO 2011 Day 261. Retrieved from https://zenodo.org/records/10000327

Jensen, A. S., Lohmann, M. S., Benzon, H.-H., & Nielsen, A. S. (2003). Full Spectrum Inversion of radio occultation signals. *Radio Science*, *38*(3), 1040. https://doi.org/10.1029/2002RS002763

Kursinski, E. R., Hajj, G. A., Schofield, J. T., Linfield, R. P., & Hardy, K. R. (1997). Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *Journal of Geophysical Research*, *102*(D19), 23429–23465. https://doi.org/10.1029/97jd01569

Kutiev, I., Marinov, P., Belehaki, A., Reinisch, B., & Jakowski, N. (2009). Reconstruction of topside density profile by using the topside sounder model profiler and digisonde data. *Advances in Space Research*, *43*(11), 1683–1687. https://doi.org/10.1016/j.asr.2008.08.017

Lei, J., Syndergaard, S., Burns, A. G., Solomon, S. C., Wang, W., Zeng, Z., et al. (2007). Comparison of cosmic ionospheric measurements with ground-based observations and model predictions: Preliminary results. *Journal of Geophysical Research*, *112*(A7), A07308. https://doi.org/10.1029/2006JA012240

Lyu, H., Hernández-Pajares, M., Monte-Moreno, E., & Cardellach, E. (2019). Electron density retrieval from truncated radio occultation GNSS data. *Journal of Geophysical Research: Space Physics*, *124*(6), 4842–4851. https://doi.org/10.1029/2019JA026744

Montenbruck, O., Hauschild, A., & Steigenberger, P. (2014). Differential code bias estimation using multi-GNSS observations and global ionosphere maps. *Navigation*, *61*(3), 191–201. https://doi.org/10.1002/navi.64

Nava, B., Coisson, P., & Radicella, S. (2008). A new version of the NeQuick ionosphere electron density model. *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*(15), 1856–1862. https://doi.org/10.1016/j.jastp.2008.01.015

Nsumei, P., Reinisch, B. W., Huang, X., & Bilitza, D. (2012). New Vary-Chap profile of the topside ionosphere electron density distribution for use with the IRI model and the GIRO real time data. *Radio Science*, *47*(4), RS0L16. https://doi.org/10.1029/2012RS004989

Palmer, P. I., Barnett, J. J., Eyre, J. R., & Healy, S. B. (2000). A non-linear optimal estimation inverse method for radio occultation measurements of temperature, humidity and surface pressure. *Journal of Geophysical Research*, *105*(D13), 17513–17526. https://doi.org/10.1029/2000jd900151

Poli, P., Healy, S. B., & Dee, D. P. (2010). Assimilation of global positioning system radio occultation data in the ECMWF ERA–interim reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *136*(653), 1972–1990. https://doi.org/10.1002/qj.722

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in FORTRAN: The art of scientific computing* (2nd ed.). Cambridge University Press.

Prol, F. S., Themens, D. R., Hernández-Pajares, M., de Oliveira Camargo, P., & de Assis Honorato Muella, M. T. (2019). Linear vary-chap topside electron density model with topside sounder and radio-occultation data. *Surveys in Geophysics*, *40*(2), 277–293. https://doi.org/10.1007/s10712-019-09521-3

Reinisch, B. W., Nsumei, P., Huang, X., & Bilitza, D. K. (2007). Modeling the F2 topside and plasmasphere for IRI using IMAGE/RPI and ISIS data. *Advances in Space Research*, *39*(5), 731–738. https://doi.org/10.1016/j.asr.2006.05.032

Rishbeth, H., & Garriott, O. K. (1969). In J. Van Mieghem (Ed.), *Introduction to ionospheric physics*. Academic Press.

Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: Theory and practice, atmospheric, oceanic and planetary physics* (Vol. 2). World Scientific Publishing.

ROM SAF. (2023). ROPP software deliverable. Retrieved from https://rom-saf.eumetsat.int/ropp/

Scherliess, L., Thompson, D. C., & Schunk, R. W. (2011). Data assimilation models: A new tool for ionospheric science and applications. In W. Liu & M. Fujimoto (Eds.), *The dynamic magnetosphere* (pp. 329–339). Springer.

Schreiner, W. S., Sokolovskiy, S. V., Rocken, C., & Hunt, D. C. (1999). Analysis and validation of GPS/MET radio occultation data in the ionosphere. *Radio Science*, *34*(4), 949–966. https://doi.org/10.1029/1999rs900034

Talagrand, O. (2010). Variational assimilation. In W. Lahoz, B. Khattatov, & R. Menard (Eds.), *Data assimilation: Making sense of observations*. Springer.

Themens, D. R., Reid, B., & Elvidge, S. (2022). ARTIST ionogram autoscaling confidence scores: Best practices. *URSI Radio Science Letters*, *4*. https://doi.org/10.46620/22-0001

UCAR. (2014). Reprocessed COSMIC-1 data. Retrieved from https://data.cosmic.ucar.edu/gnss-ro/cosmic1/repro2013/

UCAR Cosmic Program. (2019). COSMIC-2 data products. https://doi.org/10.5065/T353-C093

Vorob'ev, V. V., & Krasil'nikova, T. G. (1994). Estimation of the accuracy of the atmospheric refractive index recovery from the NAVSTAR system. *USSR Physics of the Atmosphere and Ocean (Eng. Trans.)*, *29*(5), 602–609.

Wang, S., Huang, S., & Fang, H. (2021). A novel method for deriving the vary-chap scale height profile in the topside ionosphere. *Journal of Geophysical Research: Space Physics*, *126*(2). https://doi.org/10.1029/2020JA028514

Yue, X., Schreiner, W. S., Lin, Y.-C., Rocken, C., Kuo, Y.-H., & Zhao, B. (2011). Data assimilation retrieval of electron density profiles from radio occultation measurements. *Journal of Geophysical Research*, *116*(A3), A03317. https://doi.org/10.1029/2010JA015980