**Annales
Geophysicae**

Medium-scale 4-D ionospheric tomography using a dense GPS network

M. M. J. L. van de Kamp

Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland

Correspondence to: M. M. J. L. van de Kamp (kampmax@fmi.fi)

Abstract. The ionosphere above Scandinavia in December 2006 is successfully imaged by 4-dimensional tomography using the software package MIDAS from the University of Bath. The method concentrates on medium-scale structures: between 100 km and 2000 km in horizontal size. The input consists of TEC measurements from the dense GPS network Geotrim in Finland. In order to ensure sufficient vertical resolution of the result, EISCAT incoherent scatter radar data from Tromsø are used as additional input to provide the vertical profile information.

The TEC offset of the measurements is unknown, but the inversion procedure is able to determine this automatically. This auto-calibration is shown to work well.

Comparisons with EISCAT radar results and with occultation results show that the inversion using EISCAT data for profile information is much better able to resolve vertical profiles of irregular structures than the inversion using built-in profiles. Still, with either method the intensities of irregular structures of sizes near the resolution (about 100 km horizontal size) can be underestimated. Also, the accuracy of the inversion worsens above areas where no receivers are available.

The ionosphere over Scandinavia in December 2006 often showed a dense E-layer in early morning hours, which generally disappeared during midday when a dense F-layer was present. On 14 December, a strong coronal mass ejection occurred, and many intense irregularities appeared in the ionosphere, which extended to high altitudes.

Keywords. Ionosphere (Auroral ionosphere; Ionospheric irregularities; Instruments and techniques)

1 Introduction

Ionospheric tomography is a method of reconstructing an image of the ionosphere based on measured signals of integrated parameters in various directions. This is a very useful source of information for the development of global navigation satellite systems (GNSS) such as GPS, GLONASS and Galileo, as these systems are affected by variations in ionospheric electron density, particularly those variations that do not follow a regular predictable pattern.

In ionospheric tomography, the integrated electron density (or total electron content, "TEC") is measured along many mutually crossing satellite signal paths, and subsequently an inversion procedure is applied to reconstruct the electron density distribution in the ionosphere. Albeit indirect, this method has the capability of giving a full multidimensional image, an advantage compared to direct measuring techniques, such as in-situ electron density measurements, which give only localised information at any one time and location.

Mitchell et al. (1995) successfully used ionospheric tomography to reconstruct two-dimensional images (in height and longitude) of the ionosphere from signals from polar LEO satellites received by five stations in Scandinavia. They were able to reconstruct features with a spatial resolution down to about 50 km. This showed the usefulness of the procedure, even though with this setup the result was limited in latitude (being only along one meridian) and time (giving only a single snapshot at each satellite pass).

The use of GNSS signals as input source allows signals received from many directions continuously over long periods of time, enabling the tomographic reconstruction of an ionospheric image in 4-D (i.e. 3 spatial dimensions and time). Bust et al. (2007) demonstrated this, presenting two different

tomography software packages with different inputs including GNSS and LEO satellite signals, and applied over various areas from the equatorial to the polar region. Verifying with in-situ measurements, they showed how tomography technique is able to resolve many features in the ionosphere in 4-D. However, because GNSS receivers can be sparsely and unevenly distributed, the resolution using this setup is almost always lower than that of the 2-D-setup, depending on the region.

Spencer and Mitchell (2007), Yin et al. (2008) and Pokhotelov et al. (2011) used ionospheric tomography to image the ionosphere over the polar area (the entire area above about 50° latitude) during a period of increased disturbance in October 2003 (known as the "Halloween storm"). This was challenging because, while the ionospheric structures can move and change at high velocities, the density of GNSS receivers in the polar area is low, making small-scale identification of structures difficult. They solved this by incorporating a-priori knowledge of plasma motion into the inversion procedure, which gave a good result for the motion of large-scale structures, but does not reveal structures smaller than about 500 km.

In this paper, a dense network of GPS measurements over Finland will be used, which allows the identification of smaller-scale plasma structures in the inversion with a horizontal resolution of about 100 km. Meanwhile, the spatially limited coverage of these measurements means that the inversion can only be performed over the region around Scandinavia (an area of roughly 2000 km across). No attempt to image the full-sized structure over the entire polar area is therefore taken.

The software package MIDAS (Mitchell and Spencer, 2003) developed at the University of Bath was used by Bust et al. (2007), Spencer and Mitchell (2007), Yin et al. (2008), and Pokhotelov et al. (2011), and version 2.0 of that software package is used in this setup. Section 2 will explain the differences between this setup and the one for larger-scale structures over the entire polar cap, and how the procedure needs to be adapted to this specific configuration.

2 Tomography technique

2.1 TEC measurement using GPS

The integrated electron density can be derived from the phase difference between the GPS signals L_1 and L_2 as follows. The phase φ_r of a signal at a frequency f, received after having passed along a certain path, is given by

$$\varphi_r = \varphi_t - \frac{2\pi f}{c} \int n \, ds \quad \text{(rad)}, \tag{1}$$

where φ_t is transmitted phase, $c =$ velocity of light in vacuum $= 299792458 \, \text{m s}^{-1}$, $n =$ refractive index of the medium (which in the ionosphere is dependent on frequency), and $s =$ distance along the ray path.

The phase difference between two signals at frequencies f_1 and f_2 is defined as

$$\Delta\varphi = \frac{f_1}{f_2}\varphi_2 - \varphi_1. \tag{2}$$

Substituting Eq. (1), this gives

$$\Delta\varphi_r = \Delta\varphi_t - \frac{2\pi f_1}{c} \int (n(f_2) - n(f_1)) \, ds. \tag{3}$$

If the signals are in phase at the transmitter, $\Delta\varphi_t = 0$.

In the troposphere, the refractive index n of the medium is independent of frequency, and therefore cancels out in Eq. (3). Therefore, although the integral is evaluated over the entire path between satellite and ground station, only the frequency-dependent refractive index of the ionosphere contributes to its result.

To find the frequency dependence of the refractive index n of an ionised medium, the Appleton-Hartree equation (Hargreaves, 1979) can be approximated for the case when collisions and the effect of the magnetic field are neglected as

$$n(f) = 1 - \frac{N_e r_e c^2}{2\pi f^2}, \tag{4}$$

with $N_e =$ electron density (m^{-3}), and $r_e =$ radius of an electron $= 2.81794 \times 10^{-15} \, \text{m}$.

Using this, Eq. (3) becomes

$$\Delta\varphi_r = r_e c f_1 \left(\frac{1}{f_2^2} - \frac{1}{f_1^2} \right) \int N_e ds. \tag{5}$$

The integral $\int N_e ds$ is equal to the number of free electrons in a column of unit cross-section area from the satellite to the receiver. This is referred to as the "total electron content" or TEC.

Summarising, the TEC can be determined from the phase difference $\Delta\varphi$, as

$$\text{TEC} = \frac{f_1}{r_e c \left(f_1^2/f_2^2 - 1 \right)} \Delta\varphi \quad \text{(electrons m}^{-2}). \tag{6}$$

However, because of the inevitable 2π-ambiguity of phases, the absolute phase difference $\Delta\varphi$ between the two signals cannot be measured. Therefore, only the increments in phase difference from one sample to the next, i.e. $\Delta\varphi(t) - \Delta\varphi(t - dt)$, can be measured (as long as these are within $-\pi$ and π radians). Hence, the differential TEC, or dTEC, as in

$$\text{dTEC}(t) = \text{TEC}(t) - \text{TEC}(t - dt) \tag{7}$$

can be measured, and from this, TEC can be reconstructed apart from an unknown offset value. For the tomography input, it would be best to determine this offset in order to have the absolute TEC.
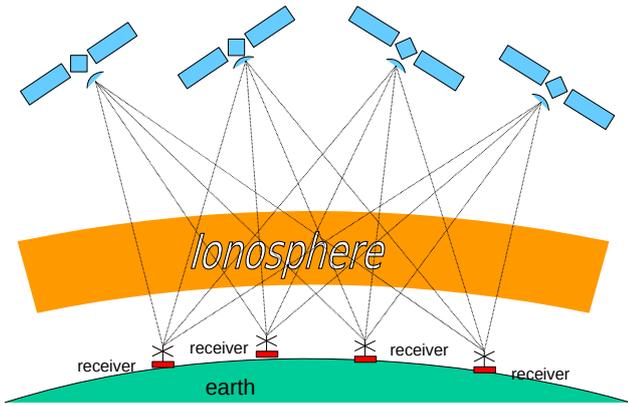
Fig. 1. Configuration with TEC measurements for ionospheric tomography.

This offset might be determined using "time stamp" codes transmitted by the GPS satellites, which, when received, allow determination of the absolute delay of the signal. The satellite's range determined this way is called its "pseudorange". Although this signal is generally too noisy to be used directly for TEC determination, it might be used to determine the offset of the TEC measured using the phase. Unfortunately, it was found that the pseudorange measurements in the database used in this study were of insufficient quality to be used for tomography purposes.

Fortunately, as will be shown in the next section, the inversion procedure for the tomography can also work with dTEC, so determination of the offset is not vital.

2.2 Inversion principle

The electron density through the entire 3-dimensional ionosphere can be estimated by the TEC measurements using multiple satellites and receivers, as illustrated in Fig. 1. This is done by looking for the electron density $N_e(\boldsymbol{x}, t)$ as a 4-D function of location $\boldsymbol{x}$ and time t, which fulfills the condition

$$\int N_e(\boldsymbol{x}, t) ds = \text{TEC}(t) \qquad (8)$$

for each measurement of TEC, i.e. for each sample on each path between any satellite and any receiver for which a measurement has been made. The integral is performed along the entire path from satellite to receiver.

The collection of all measurements provides a system of equations like Eq. (8), conditioning the 4-D solution which is sought. Furthermore, extra information on N_e can be provided by imposing additional conditions, which can, for instance, minimise the gradients of N_e in certain dimensions (temporal and/or spatial). These extra conditions also "fill the gaps of information" for locations where not enough measurements are available.

Because the measurement setup lacks a means of determining the TEC offset for all measurements, the absolute

Fig. 2. Geotrim GPS receiver network over Finland.

TEC is not available, as explained in the previous section. Because of this, dTEC is used, and the equations to be satisfied are modified from Eq. (8):

$$\int_{\text{path}_l(t)} N_e(\boldsymbol{x}, t) ds - \int_{\text{path}_l(t-dt)} N_e(\boldsymbol{x}, t-dt) ds = \text{dTEC}_l(t),$$

$$(9)$$

where $\text{path}_l(t)$ and $\text{path}_l(t-dt)$ are the paths from the satellite to the receiver of link l at the two sample times.

Obviously, the accuracy of the result of the inversion procedure increases with the number of available measurements on crossing paths. Hence, it is beneficial to use as many satellites and receivers as possible. In the case of this paper, the group of satellites are the GPS constellation, with currently 31 transmitting satellites, of which usually 8–10 are simultaneously visible at any given time from most locations on Earth. The receivers are the 86 receivers of the Geotrim network, distributed all over Finland, as shown in Fig. 2. Data from this measurement configuration are available for this study for the whole month of December 2006. The density of the receivers, and hence of crossing paths, allows a horizontal resolution of about 100 km for the inversion.

2.3 Mathematical setup

A detailed description of the mathematical technique used in this study is given by van de Kamp (2011). The principle is explained in the following.

In practice, the inversion procedure numerically looks for a four-dimensional solution within a certain finite 4-D space. In both horizontal dimensions, this area needs to be large enough to contain all measured paths that are taken into account. In altitude dimension it needs to contain all heights at which any ionisation is expected. In addition, the area also needs to extend in the time dimension to be able to include a gradient constraint, as mentioned in the previous section, not only in spatial but also in temporal dimension (i.e. a constraint on temporal variability of electron density). The necessary size of the area in the temporal dimension depends on the correlation time of N_e.

Because the problem is solved numerically, the integration over the paths in Eq. (9) is performed over voxels of a certain size, modifying the equation as

$$\sum_{\text{path}_l(t)} N_e(\boldsymbol{x}, t) P_l(\boldsymbol{x}, t) - \sum_{\text{path}_l(t-dt)} N_e(\boldsymbol{x}, t-dt) P_l(\boldsymbol{x}, t-dt)$$
$$= \text{dTEC}_l(t), \qquad (10)$$

where $P_l(\boldsymbol{x}, t)$ is the path length segment of link l at time t through the voxel at location $\boldsymbol{x}$.

The number of 4-D voxels in the grid ($=$ the number of terms in each equation) will be referred to as J. The grid used in the inversion results presented in this paper is defined as

- between 55° and 74° latitude with a resolution of 1° (20 latitudes);

- between 0° and 50° longitude with a resolution of 2° (26 longitudes);

- between 60 km and 780 km in altitude with resolution of 40 km (19 altitudes);

- for time lengths of 40 min with a resolution of 5 min (9 times).

This results in $J = 20 \times 26 \times 19 \times 9 = 88\,920$.

In horizontal dimension, this grid has a resolution of about 100 km. Its size is roughly 2000 km across, covering all of mainland Scandinavia. This is significantly larger than Finland, which is possible because the GPS ray paths cross the ionosphere (at low elevations) over a much larger area than just the area containing the receivers. In vertical dimension, with this altitude range the troposphere (below 60 km) is neglected, which is justified as it does not contribute to the TEC measurement (see Sect. 2.1). On the other hand, it should be noted that also the plasmasphere (above 780 km) is neglected, which in reality might contribute to the TEC observed in small amounts.

There is one equation of the form of Eq. (10) for each measurement (i.e. each combination of a satellite, a receiver and two points in time for which a dTEC was measured) within the time frame of one inversion. Hence, if this number of measurements is K, there is a set of K equations with J unknowns. This set of equations can be described as a matrix equation:

$$\mathbf{R}N = Y. \qquad (11)$$

Here, $\mathbf{R}$ is a $K \times J$ matrix with each row containing the coefficients of one equation of the form of Eq. (10). The vector N contains the J electron densities, i.e. the unknowns in the equations, and the vector Y contains the K measured values of $\text{dTEC}_l(t)$.

However, the inversion software package requires that the system of equations be square, i.e. it should have an equal number of equations and unknowns. Because of this, in the next step, both sides of the matrix equation are left-multiplied by a weight factor matrix $\mathbf{W}$ of size $J \times K$. The content of $\mathbf{W}$ is in principle arbitrary as long as the linear span of its J rows is the K−dimensional space (assuming $J > K$). Here, for this matrix is used the transpose of a matrix containing all the absolute values of the elements of $\mathbf{R}$, and with all K columns normalised by their sum. The multiplication with $\mathbf{W}$ results in a new set of equations described as

$$\mathbf{P}N = Z. \qquad (12)$$

Here, $\mathbf{P} = \mathbf{W}\mathbf{R}$ is a $J \times J$ matrix, and $Z = \mathbf{W}Y$ is a $J \times 1$ vector.

As mentioned in the previous section, the solutions N will also be confined by the condition to avoid sudden changes, both in horizontal direction and in time. This can be expressed in a correlation coefficient equation. As an example considering only one dimension, to minimise the first derivative it can be demanded that the difference between one value and the average of its two neighbours should be minimised. Hence, for the electron density in voxel i, the equation

$$-N_{i-1} + 2N_i - N_{i+1} = 0 \qquad (13)$$

should be fulfilled to an error as small as possible. To achieve the same for the second derivative in the same dimension, the equation can be convolved with itself:

$$N_{i-2} - 4N_{i-1} + 6N_i - 4N_{i+1} + N_{i+2} = 0. \qquad (14)$$

This equation is expressed for adjacent voxels in all dimensions where this is desired – in latitude, longitude, and time – and these various correlation equations are convolved with each other, resulting in one equation for each voxel, representing its correlation in various dimensions. All of those J correlation equations together can also be expressed as a matrix equation:

$$\mathbf{C}N = \mathbf{0}. \qquad (15)$$

The two equations, Eq. (12) and Eq. (15), are then combined:

$$\mathbf{P}N + F\mathbf{C}N = Z. \qquad (16)$$

The constant weight factor F determines how much weight is given to matching the measurements (Eq. 12) with respect

to the smoothness of the solution (Eq. 15). Here, $F = 1$ is used.

Although J equations with J unknowns can be solved entirely, solving such a large system of equations is a very challenging task. This is why the software does not look for the exact solution, but instead approximates the solution to a specified tolerance.

In the current setup, the equations are not solved for the electron densities themselves, but for their deviation from a background electron density. For this background density, the International Reference Ionosphere model "IRI-2007" is used (http://iri.gsfc.nasa.gov/) (Bilitza and Reinisch, 2008). This model gives the ionisation pattern that can be expected according to known theory, depending on latitude, longitude, altitude and time, using input information such as the solar activity. The condition of smoothness is imposed on these deviations from the background electron density, not on the electron densities themselves. Writing $N = N_0 + N'$, where N_0 is the background density and N' is the deviation, this means that the matrix equation is modified as

$$\mathbf{P}(N_0 + N') + F\mathbf{C}N' = \mathbf{Z}$$
$$(\mathbf{P} + F\mathbf{C})N' = \mathbf{Z} - \mathbf{P}N_0. \tag{17}$$

2.4 Time dependence

The combination of the inversion as a function of time (in the time-dimension of the grid defined in Sect. 2.3) and the continuity with long-term GPS measurements over the entire month of December 2006 is managed as follows. For every time point, the inversion is performed over a period of 40 min centred on the time point of interest. Of the result, only the values for this central time point are stored as output. Meanwhile, the entire 40-min result, shifted by 5 min, is used as a starting solution for the inversion for the next point in time.

2.5 Vertical profile information

The inversion procedure works well if the measurements are made on many crossing paths in all directions. However, in the case of satellite measurements the directions of the measured paths are not evenly distributed, as already noted by Mitchell et al. (1997) for LEO satellite measurements. In the current setup using GPS measurements in a relatively small area at high latitudes, the situation is even worse: the paths are concentrated in certain directions, and very sparse in others, as the example of Fig. 3 shows. As a result, for instance, the inversion has a poor vertical resolution at altitudes above about 200 km. Because of this, just as for Mitchell et al. (1997), the inversion result benefits from extra information about the vertical profile of electron density.

In order to provide this vertical profile information, in MIDAS version 2.0, the grid is not entered into the inversion procedure as it is. Instead, the height dependence of the grid is convolved with a number of orthogonal functions of height, each representing a component of the altitude pro-
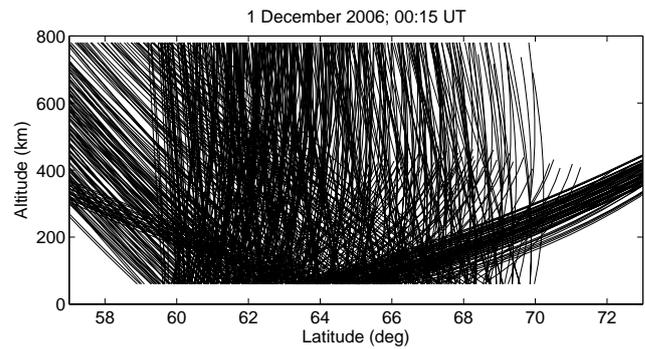
Fig. 3. Example of the measurement paths inside the grid of one inversion, projected onto the latitude–altitude plane.

file of electron density. It is assumed that a particular altitude profile can be approximated as a linear combination of these orthogonal functions. Then, for each latitude, longitude, and time, the modified (convolved) grid consists of a number of profiles (instead of a number of heights), and on this grid the inversion is performed. The electron density resulting from the inversion contains a coefficient for each profile; this resulting solution is again convolved inversely with these orthogonal functions to obtain the electron density as a function of altitude.

The convolution with these orthogonal functions provides the necessary profile information. As long as enough horizontal measurement paths are available to resolve a combination of the orthogonal functions at some (e.g. the low) altitudes, these orthogonal functions make sure the profile is appropriately extended at other (higher) altitudes.

An additional advantage of this extra mathematical operation is memory usage reduction. Modification of the grid so that, in the dimension of altitude, it does not represent separate altitudes but instead separate orthogonal functions can reduce the memory usage in the inversion significantly if the number of orthogonal functions is smaller than the number of altitudes.

There are various ways to obtain the orthogonal functions necessary for this. MIDAS has a flexible input for these functions, but as a default uses functions based on the Chapman or Epstein profile models. In MIDAS version 2.0, also used by Spencer and Mitchell (2007), the orthogonal functions are assumed proportional to the formula of electron production rate q according to the Chapman profile (Ratcliffe, 1972):

$$q(h) = q_m e^{1-z-e^{-z}} \tag{18}$$

$$z = \frac{h - h_m}{h_s}, \tag{19}$$

where $h =$ height above sea level (km); $h_m =$ height where production rate is maximised (km); $q_m =$ production rate at height h_m (q_m and h_m considered at the same incidence angle of solar radiation); and $h_s =$ scale height (km) (the scale
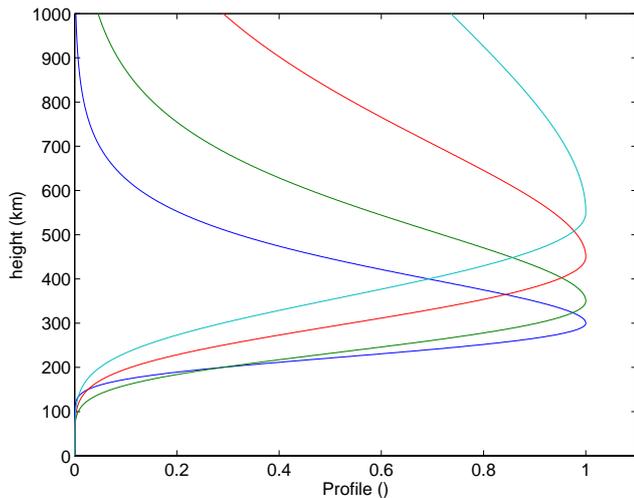
Fig. 4. The basic Chapman profiles used as default for vertical profile information in MIDAS.

Fig. 5. Some examples of electron density profiles measured by the EISCAT incoherent scatter radar in Tromsø.

height can generally be different for $h < h_m$ and $h > h_m$, which are then referred to as "bottom scale height" h_{sB} and "top scale height" h_{sT}).

A large number of Chapman profiles is generated, with each one's parameters varying loosely around the parameters of four chosen "typical" profiles. Figure 4 shows the four "average" profiles used in this random generation. Because all of these profiles have maxima at 300 km altitude or higher, these are well suitable to describe various parts of the F-layer ionosphere. However, this characterisation does not allow for any E-layer. This is presumably because this configuration of the software was used for large-scale imaging of the ionosphere over the polar area, where no significant E-layer is expected.

Another approach is to use, rather than a modelled vertical profile, a measured input. In the case of Scandinavia, this can be provided by the EISCAT incoherent scatter radar in Tromsø, Norway (http://www.eiscat.se/) (Folkestad et al., 1983). This radar site conducts various measurement campaigns on demand. Many of these consist of vertical profiling of the ionosphere, and the results of most campaigns are publicly available (http://www.eiscat.se/madrigal/). These measurements can provide the input with good vertical resolution that the GPS receiver setup otherwise lacks. Although the measurements in Tromsø correspond to only one location within the grid, it is assumed that the typical shapes of regularly occurring vertical profiles of the ionosphere are similar within the entire Scandinavian area. The orthogonal functions can be based on many measured profiles, and as a result, the height dependence of the inversion output will correspond to a typical shape of such measured profiles.

Figure 5 shows a selection of typical profiles measured by the EISCAT radar for different hours of the day on 10 December 2006. It can be seen that on this day there was a strong E-layer at about 100 km altitude in the early morning.
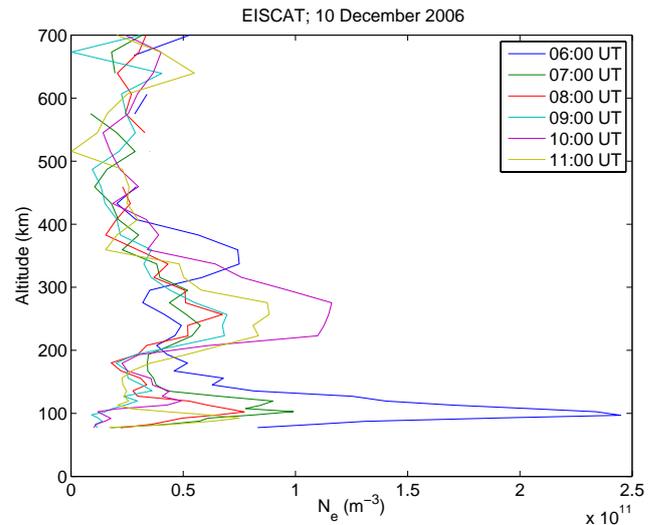
Later in the day, when the Sun was up, this E-layer decreased and the ionisation was predominantly present in an F-layer between 200 and 300 km height. This variation between E- and F-layers was also found in many other days in the EISCAT results for December 2006. Especially because of this, it is expected that this input for the vertical profiles will be a better representation for the Scandinavian ionosphere than the built-in Chapman profiles in MIDAS 2.0, which do not cater for an E-layer.

In this paper, three approaches are used for the vertical dependence of the inversion:

A. The default option in the MIDAS software is used. This means that orthogonal functions are based on Chapman profiles, similarly as was done by Spencer and Mitchell (2007). Five orthogonal functions are derived from a large collection of profiles; these are the left-singular vectors associated with the five largest singular values of the matrix made up of all profiles together. The vertical dependence of the result of the inversion (for any latitude, longitude and time) will be a linear combination of these orthogonal functions, and as a consequence, will have the general shape of a linear combination of several Chapman profiles.

B. Orthogonal functions based on EISCAT measurements are used. The typical shapes of vertical profiles are considered to be not very dependent on location within Scandinavia or on the date within the month of December. However, they are assumed to be dependent on the hour of the day. EISCAT measurements are not available for every day of December 2006, and many were performed on/around 14–15 December. Typically, measurement campaigns lasted several hours, and during such campaigns typically one profile was measured

every minute. As a result, for every hour of the day, at least one day in December 2006 can be found where many measurements at that hour have been performed.

Therefore, for each inversion, all EISCAT measurements are selected which were performed on any day in December 2006, within the time frame of two hours centred around the central time of the inversion. From this group of measurements five orthogonal functions are derived (by determining the singular values, as above). As a consequence, the vertical dependence of the result of the inversion (for any latitude, longitude and time) will have the general shape of a linear combination of several typical measured profiles.

C. For comparison, to verify the usefulness of the vertical profile convolution, the inversion is also run without any vertical profile convolution. This should demonstrate the poor vertical resolution of the measurement setup at high altitudes, as described above.

3 Results

The inversion was performed on the data measured by the Geotrim network over the month of December 2006, for a time resolution of 5 min and for a grid over Scandinavia as specified in Sect. 2.3.

The results of the inversion consist of the 4-dimensional electron density pattern over the full spatial dimensions of the grid, and over the whole month of December 2006. Unfortunately it is impossible to present all of these results in the paper. Many results can be found online at http://www.space.fmi.fi/MIRACLE/Geotrim. In the following sections, results will be presented for several special cases, where they can be tested against other data.

3.1 Comparison to calibrated TEC measurements

First, the inversion results are verified using the TEC measurements which were used as inputs for the inversion. As explained in Sect. 2.1, the TEC measurement used in the inversion contains an unknown offset, and the inversion automatically adjusts this offset, "calibrating" the data. This automatic calibration is tested in this section against TEC measurements that have been calibrated using another method.

Leitinger et al. (1975) described a method of calibrating TEC measured at two stations simultaneously, which has become a well-established technique. However, that method uses measurements from a single satellite whose orbit is lined up with the array of the stations, and assumes that the ionosphere is stationary during the time of the pass of the satellite. The method described in this paper is a variation on Leitinger's method, and uses multiple satellites, does not require a particular geometry of the stations, and does not make any assumption on the temporal stationarity of the ionosphere (outside of a certain correlation time).
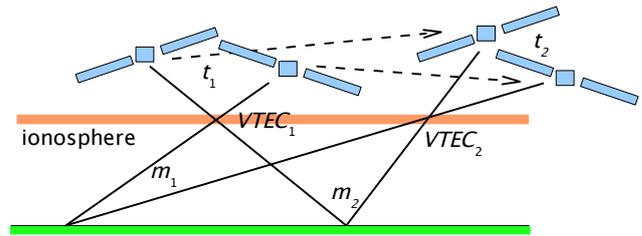
Fig. 6. Situation where TEC calibration is possible.

The offset of GPS TEC measurements can be determined for any set of measurements for which the following conditions are met:

– The ionosphere is quiet (no scintillation), though not necessarily horizontally uniform or temporally stationary.

– The electron density can be approximated as concentrated in an infinitely thin slab at a known height, presumably the height of the F-layer.

– Two satellite-receiver links cross this layer with horizontal IPP (ionospheric pierce point) locations and time points close enough together – in practice, within certain distance and time delay limits within which the ionosphere is assumed stationary.

– At a later time, the same two links again have two IPP locations and times close enough together.

– Both links stay in lock continuously between the two time points.

– The ratio of sines of elevation angles of the two links at the first point in time is sufficiently different from that at the second point in time.

Figure 6 depicts this situation schematically. In practice, within the Geotrim network, such cases happen numerous times per day.

From the assumptions follows that the TEC offset of each of the two links is constant between the two time points t_1 and t_2. Furthermore, at any of the two time points, both links m_1 and m_2 measure the same electron density (assuming a thin F-layer). These four measurements can then be expressed as

$$\begin{cases} \text{TEC}_{m1}(t_1) = \text{TEC}_{ofs1} + \text{VTEC}_1(t_1)\big/\sin\varepsilon_1(t_1) \\ \text{TEC}_{m2}(t_1) = \text{TEC}_{ofs2} + \text{VTEC}_1(t_1)\big/\sin\varepsilon_2(t_1) \\ \text{TEC}_{m1}(t_2) = \text{TEC}_{ofs1} + \text{VTEC}_2(t_2)\big/\sin\varepsilon_1(t_2) \\ \text{TEC}_{m2}(t_2) = \text{TEC}_{ofs2} + \text{VTEC}_2(t_2)\big/\sin\varepsilon_2(t_2) \end{cases} \quad (20)$$

where $\text{TEC}_{mi}(t) = $ TEC measured on link i as a function of time, $\varepsilon_i(t) = $ elevation angle (in the ionosphere) of link i as a function of time, $\text{TEC}_{ofsi} = $ the (constant) TEC offset of link
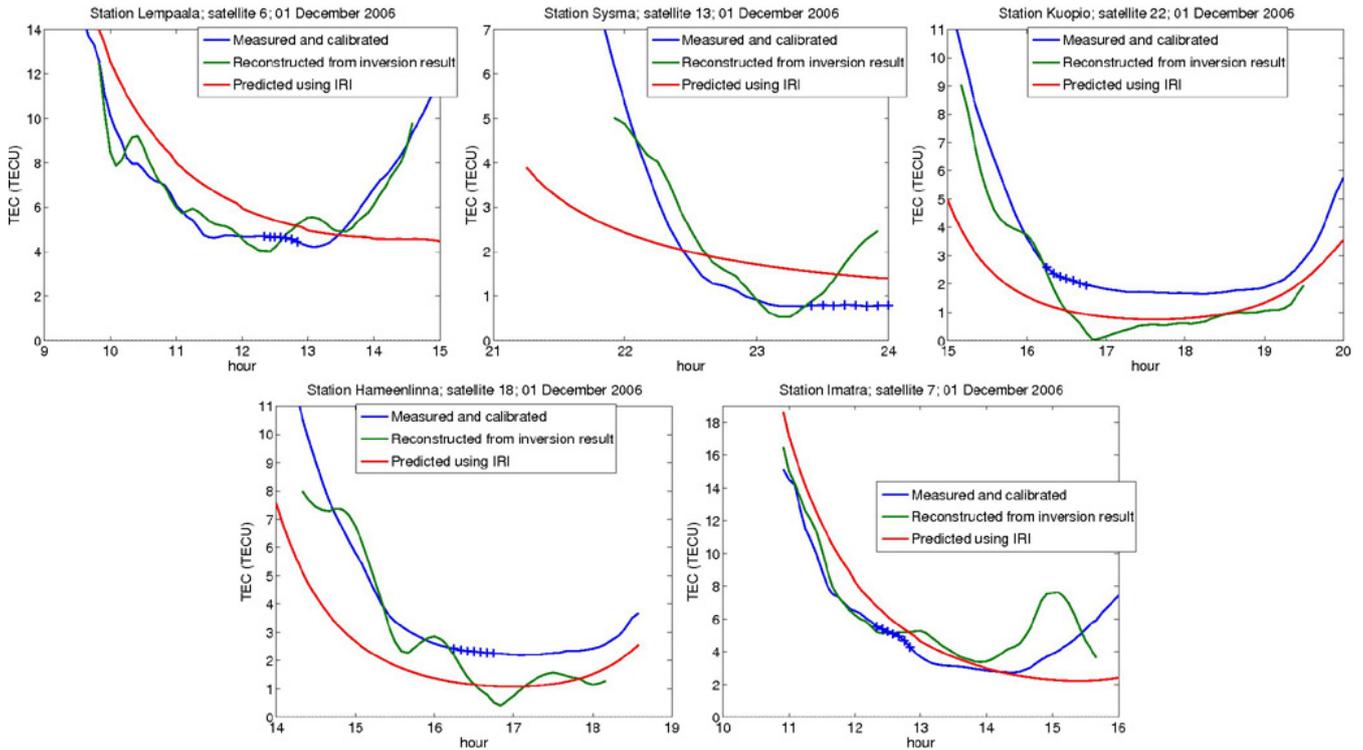
Fig. 7. Calibrated TEC measurements of several receivers and satellites on 1 December (blue curves). The samples at which calibration was performed are indicated as "+". Also included: the inversion results integrated for the same satellite-receiver paths (green curves), and the TEC along the same paths according to the IRI model for the same date and time (red curves).

i, and $\text{VTEC}_j(t) = \text{VTEC}$ in the ionosphere at location j as a function of time.

Here, VTEC denotes "vertical TEC", i.e. the electron density integrated over a vertical column of unit cross-section. Under the mentioned assumption of a thin ionosphere at a known height, VTEC is equal to TEC $\sin \varepsilon$ and is independent of measurement.

These four equations contain four unknowns: TEC_{ofs1}, TEC_{ofs2}, $\text{VTEC}_1(t_1)$ and $\text{VTEC}_2(t_2)$. Resolving these for TEC_{ofs1} gives

$$\text{TEC}_{ofs1} = \frac{\text{TEC}_{m2}(t_2) - \text{TEC}_{m2}(t_1) - \text{TEC}_{m1}(t_2)\frac{\sin \varepsilon_1(t_2)}{\sin \varepsilon_2(t_2)} + \text{TEC}_{m1}(t_1)\frac{\sin \varepsilon_1(t_1)}{\sin \varepsilon_2(t_1)}}{\frac{\sin \varepsilon_1(t_1)}{\sin \varepsilon_2(t_1)} - \frac{\sin \varepsilon_1(t_2)}{\sin \varepsilon_2(t_2)}}$$

$$(21)$$

With this, not only the measured samples $\text{TEC}_{m1}(t_1)$ and $\text{TEC}_{m1}(t_2)$ can be calibrated, but the entire pass of link m_1. In the same way, TEC_{ofs2} is also found from this information and link m_2 can be calibrated.

This calibration has been performed for all measurements of the whole day of 1 December 2006, with a sampling time of 5 min. For this day, the inversion results showed that the ionosphere was quiet, and the altitude of maximum ionisation was ~300 km. Therefore, the thin layer of ionisation was assumed to be at 300 km. Any two paths were considered to be close enough together to be calibrated if they crossed this height at points no more than 20 km apart at the same time

sample (i.e. within 5 min). All satellite passes that contained at least two such co-located IPPs with another pass were calibrated.

Figure 7 shows the result for several passes on 1 December (blue lines). For verification of the tomography result, the TEC for the same passes has also been calculated from the inversion (the one using method B described in Sect. 2.5), as follows: for every sample, the electron density was integrated along the momentary path between the satellite and the receiver through the grid, obtaining a "virtual TEC measurement". These have also been included in Fig. 7 (green lines). Furthermore, for reference, the electron density according to the IRI model was also integrated along the same satellite-receiver paths and included in Fig. 7 (red lines).

Note that while TEC as calculated in e.g. Eq. (6) has the unit of electrons m^{-2}, in Fig. 7 it is expressed in TEC-units or "TECU". This is defined as 1 TECU = 10^{16} electrons m^{-2}.

These graphs reveal many different situations. On the top row of Fig. 7, the left-hand graph (Lempäälä – satellite 6) shows a good agreement between the inversion result and the measured and calibrated TEC. Both differ significantly from the IRI prediction. Given that the inversion tends to approach the IRI model for any case where the measurements do not give any better information, this means that the measurement results were well used in this case.
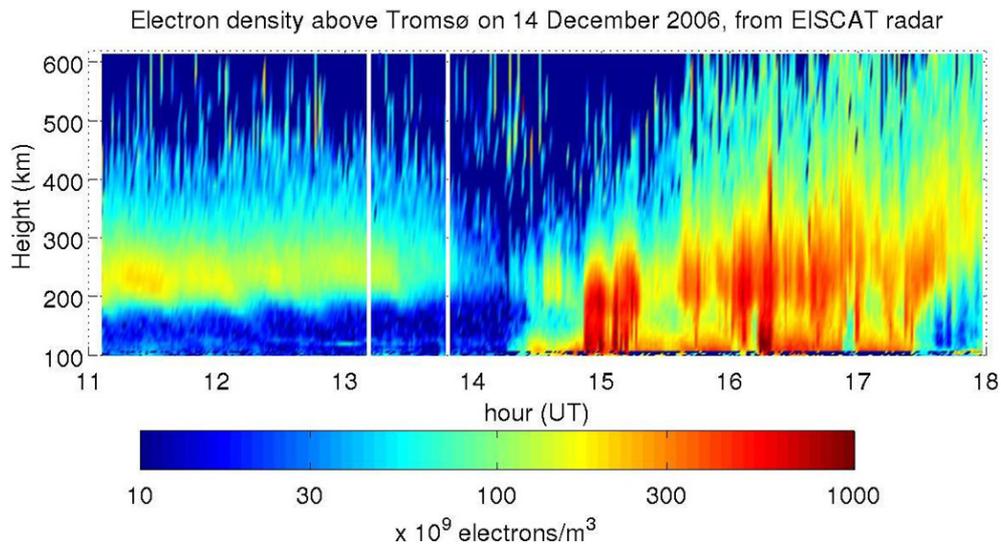
Fig. 8. Electron density measured by EISCAT at Tromsø on 14 December.

It should also be emphasised that the inversion does not use the offset information obtained using Eq. (21) from the measurements; it uses only dTEC. Yet, not only the shape but also the offset of the inversion is very similar to that of the measurements. This means that the inversion is well able to automatically calibrate the measurements.

The top middle graph (Sysmä – satellite 13) shows a similar result for part of the satellite pass; after 23:00 the inversion result increases more than the measured TEC. This must be due to other measurements, nearby in horizontal direction or in time, which affect the result through the smoothing condition (either at the same altitude, or at a different altitude, in which case they affect the result through the vertical orthogonal profiles). A similar thing happens in the top right graph (Kuopio – satellite 22), where the inversion result is lower than the measured TEC after 16:15 and stays close to the IRI model.

Similar effects are also seen on the bottom row of graphs in Fig. 7: in the left-hand graph (Hämeenlinna – satellite 18), the inversion result agrees with the calibrated measurements for part of the pass, and stays lower after that; and in the right-hand one (Imatra – satellite 7), the inversion after 14:00 increases even more away from the IRI model than the calibrated measurements, which is probably because of influence from other measurements, as the one from Lempäälä (top row, left), which is in the vicinity. This shows that different measurements can have a combined effect on the same voxels, so that it should not be expected that the inversion always exactly matches the measurements.

These examples show that, even though the calibrations of the measured results using Eq. (21) were done independently of the inversion, they arrived very often at similar offsets. This shows that the inversion procedure is well able to "auto-calibrate" the dTEC results.

3.2 Comparison with EISCAT

Comparison to incoherent scatter radar returns is a useful verification of ionospheric tomography results, as it proved to be already to e.g. Mitchell et al. (1995). The verification in this section concentrates on the height profile of electron density above Tromsø, measured by the EISCAT radar on 14 December between 11:00 and 18:00 UT.

Figure 8 shows measured profiles from the EISCAT radar. The resolution of EISCAT data depends on the campaign; the data of Fig. 8 have a time resolution of 1 min, and an altitude resolution varying from about 3 km in the E-layer to about 30 km at the highest altitude. This figure shows that on this day, according to the radar results the "regular" ionisation was present at daytime, at an F-layer height between 200 and 300 km, and disappeared around 14:00 UT. After this, from 14:30 some strong disturbances occurred: they started first at the E-layer at low altitudes (100 km), but at 15:00 many areas of intense ionisation appeared around the F-layer, which after 16:00 spread gradually even higher.

This event was most likely driven by an interplanetary coronal mass ejection (ICME). It is not the purpose of this paper to go into detail about the physics of this particular magnetospheric event, but because of the combination of "regular" and "incidental" occurrences, this measurement period serves as a good test case for the inversion procedure.

Figure 9 shows the results of the inversion for this day, time and location, for the three configurations as described in Sect. 2.5. The upper graph shows the result for method A, where the vertical profile information was obtained according to the default built-in method in MIDAS using the orthogonal functions derived from Chapman profiles. The middle graph shows the result for method B, where the vertical profile information was obtained by orthogonal functions
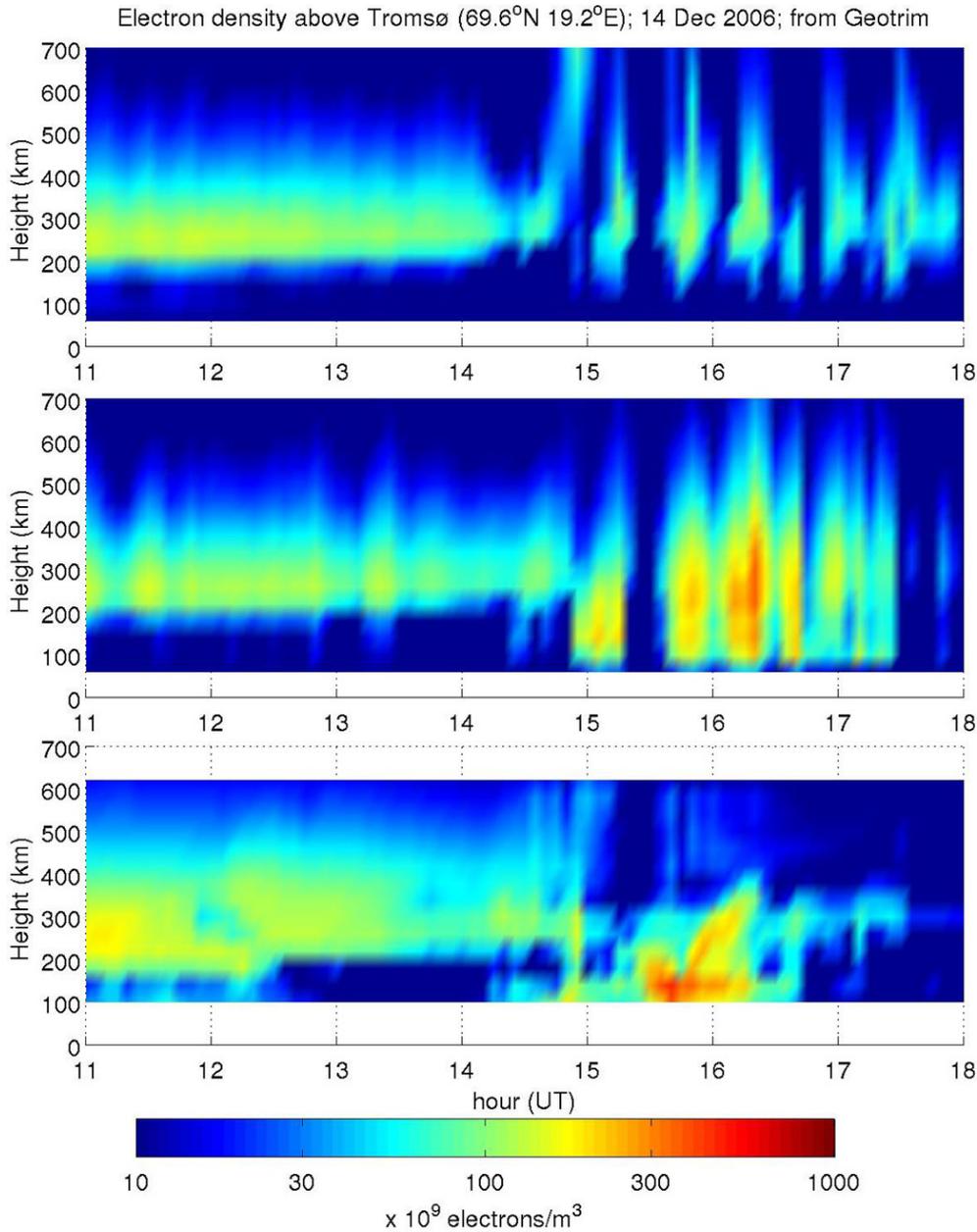
Fig. 9. Inversion result on December 14 above Tromsø (69.59° N; 19.23° E). Upper: method A, using the orthogonal functions from Chapman profiles. Middle: method B, using the orthogonal functions from the EISCAT measured profiles. Lower: method C, without using any vertical profile information.

derived from the profiles measured by the EISCAT incoherent scatter radar. For comparison, the lower graph shows the result for method C, where no vertical profile convolution was used.

The different graphs of Fig. 9 show that all of the methods are able to retrieve the "regular" F-layer up to 14:00 and the disturbances in the afternoon/early evening of this day. However, significant differences between the methods are also found.

Method A places all peaks of ionisation at at least 300 km height. Obviously this was expected, considering that the default Chapman profiles all have peak altitudes of at least 300 km; hence, the inversion would not be able to retrieve any peak ionisation at lower altitudes. Furthermore, the method does show the moments of increased ionisation in the afternoon above Tromsø at the correct times; however, the ionisation is much weaker than measured by the EISCAT radar.

Table 1. RMS errors with respect to EISCAT incoherent scatter radar data

Method	RMS error N_e (electrons m^{-3})	RMS error $\log_{10} N_e$ ()
Inversion A (Chapman profiles)	11.29×10^{10}	0.824
Inversion B (EISCAT profiles)	8.41×10^{10}	0.596
Inversion C (no profiles)	10.64×10^{10}	0.708

The result of method B is clearly more in agreement with the radar profiles. To some extent this could be expected, given that these measured profiles were used in the procedure. Still, the agreement is less self-evident than it might seem, as the used information from the measured profiles only consists of possible shapes of the profiles, depending on the hour of the day.

This method is better able to retrieve irregularities in the daytime F-layer that were seen in the EISCAT measurement, e.g. between 11:00 and 12:00 UT. Furthermore, compared to method A, the results for the disturbances in the afternoon/early evening show stronger ionisation areas, including the E-layer at 100 km altitude, with values and shapes more similar to those in the EISCAT measurements. Still, the level of ionisation found during the disturbances from also this method is lower than observed by EISCAT, especially in the E-layer. This is presumably due to the small horizontal size of many of these irregularities, which is near the resolution of the inversion (about 100 km).

For a quantitative comparison, the EISCAT data were interpolated in altitude, and averaged in time, to arrive at the same resolution in both these dimensions as the inversion results. Next, the RMS difference between electron density N_e of the inversion results and that of the EISCAT data over the entire period 11:00–18:00 and all altitudes was calculated; the result for all inversion methods are shown in Table 1. The RMS differences of $\log_{10} N_e$, which reveal the differences also in lower electron density levels, are also included. Both values are for method B clearly smaller than for method A.

Obviously, another possible method to obtain better results than using the default built-in Chapman profiles would be to adapt these profiles. Ideally, these profiles could be modelled on the profiles measured by the EISCAT radar, dependent on the hour of the day. However, this option is not used in this paper, since at this point there is no reason to expect that the results will be any different from those using the EISCAT measured profiles directly. On the other hand, when tomography is applied over a wider area and a longer time scale, the built-in modelled profiles could be made dependent on other parameters as well, such as date, location, and space weather parameters. This could provide a considerable improvement to the inversion procedure for general use.

Finally, method C without using any profile convolution results in Fig. 9 in a more irregular pattern than the other two methods. The pattern of ionisation observed here, as a function of time and altitude, is not in good agreement with the EISCAT measurements, particularly at altitudes above 200–300 km. To its credit, however, below 200 km this method is able to retrieve the strongest E-layer ionisation of all three, similar to that found by the EISCAT radar.

The RMS difference of N_e with the EISCAT data for this case is included in Table 1. According to these figures, this method performs slightly better than method A.

The observation that method C is best able to retrieve strong low-altitude ionisation, while method B is much better at higher altitudes, gives rise to the idea that these two methods might be combined: the profile convolution could be used only at higher altitudes, where not enough crossing measurement paths are available, while the height-dependence of the measurements would be used directly when they are. However, the application of this idea is not straightforward, and the feasibility of this should be a subject of further study.

3.3 Comparison to occultation measurements

A useful test of the quality of the inversion results in terms of vertical profile is to compare them with a source of data that has been measured in a similar way as the GPS measurements, but which have a better vertical resolution. Such measurements are radio occultation (also called "limb-sounding") measurements. In these experiments, TEC is measured not between a satellite and the ground, but between two satellites, whose line-of-sight path crosses the atmosphere in horizontal direction.

As can be understood from Sect. 2.5, these horizontal paths are just what is missing from the setup using the GPS ground stations. Therefore, occultation measurements provide the vertical resolution which the inversion otherwise lacks (and which is compensated by including separate vertical profiles). Indeed, it would have been good to include occultation measurements in the input for the inversion procedure, were it not that those measurements are too sparse to provide a continuous source of information, as they provide only localised and short-term measurements. Nevertheless, these measurements are very suitable as test cases for the inversion results.

The geometry and parameter calculation in occultation measurements are described in detail by Hajj et al. (2002), and can be summarised as follows. As the ray between two satellites traverses the atmosphere, it is delayed and significantly bent. From the resulting phase shift and Doppler shift of the dual-frequency signal, the vertical profile of atmospheric refractivity is determined for the point where the path between the satellites is closest to Earth. From this refractivity, tropospheric temperature and water vapour profiles and
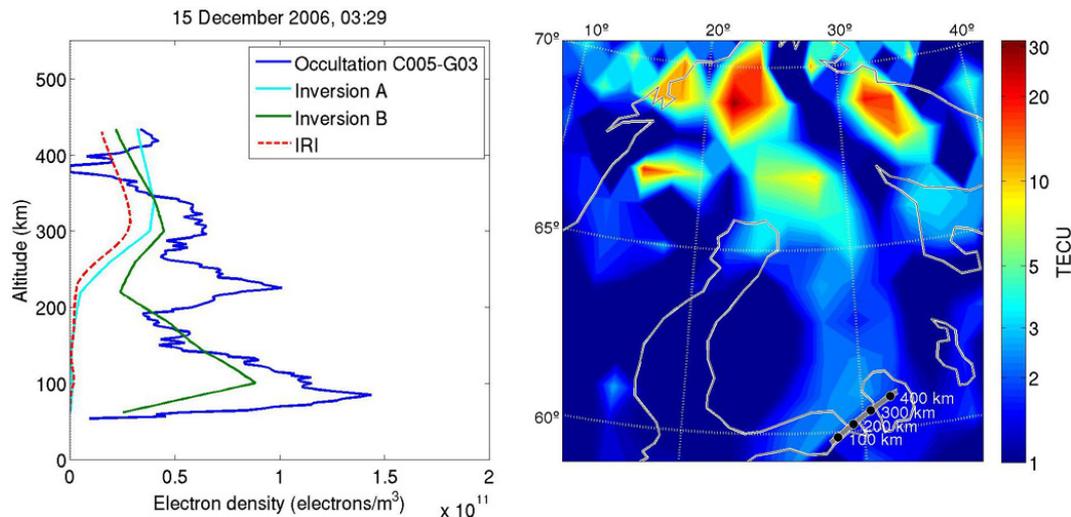
Fig. 10. Comparison between the inversion result and occultation measurement on 15 December, 03:29, between COSMIC satellite 5 and GPS satellite 3. Left: vertical electron density profiles, including also the IRI-model. Right: VTEC map from the inversion result, with the location of the occultation measurement included.

ionospheric electron density profiles can be derived. In that estimate, the assumption is used that the ionosphere is horizontally uniform and does not change within the time of each profile measurement (usually a few minutes). This estimate is reasonable in the situations when horizontal variations over the length of the path through the ionosphere are small. However, if some irregularities exist on the path, occultation measurements may integrate those over the path.

This section shows some results of vertical electron density profiles from the radio occultation mission using the FORMOSAT-3/COSMIC satellites. This constellation consisting of six small LEO satellites was designed to measure tropospheric and ionospheric profiles by radio occultation using signals received from GPS satellites. Between April 2006 and October 2007, the satellites were stepwise being moved from initial 516 km single-plane orbits to their final 800 km six-plane orbits (Fong et al., 2009). Measured data can be obtained upon registration from http://www.cosmic.ucar.edu/. In December 2006, there were on average about 20 profile measurements per day within Scandinavia.

The blue line in the left-hand graph of Fig. 10 shows the occultation profile measured on 15 December in the early morning. To compare the inversion results with this, the 4-D inversion results have been interpolated at the locations and times of the occultation profile. These results have been included in the figures: the inversion using method A (cyan) and B (green) as described in Sect. 2.5, as well as the IRI model for the same location and time (red). The occultation result shows that at this time, there were some strong enhancements, not predicted by the IRI model, in the E- and F-layers. The inversion using method B also shows these anomalies, though slightly lower in value. The inversion obviously is not able to retrieve the irregularities at an equally

Table 2. RMS errors with respect to COSMIC occultation measurements (electrons m^{-3}).

Date/time	Inversion B (EISCAT profile)	Inversion A (Chapman profile)	IRI
15 Dec, 03:29	2.97×10^{10}	5.76×10^{10}	6.08×10^{10}
20 Dec, 00:59	2.76×10^{10}	4.95×10^{10}	6.42×10^{10}
9 Dec, 02:07	3.40×10^{10}	3.45×10^{10}	4.23×10^{10}

high resolution as the occultation measurements, with the current grid voxel size and GPS receiver density. The inversion result using method A shows similar values as that using method B from 250 km upward, but completely misses the ionisation present below this height.

For a quantitative comparison, the RMS of the difference between the occultation and all of the other curves is given in Table 2. Also these figures show that the inversion using method B agrees by far best with the occultation results.

The right-hand graph of Fig. 10 shows a colour map of VTEC calculated from the inversion result, over an area around Finland (not the entire inversion grid). In this map, the location where the occultation profile was measured is indicated by a thick grey line. The labels show the altitudes of the profile at locations indicated by the black dots. This shows that the occultation measurement was made at a location that, according to the inversion, was close to the edge of a region of irregularities in the southeast of Finland.

At the time of measurement, GPS satellite 3 was at a location of coordinates 43.1° N/−171.2° E, and was seen from the locations marked in Fig. 10 toward the north, at azimuths between 14° and 17°. This means that, assuming the
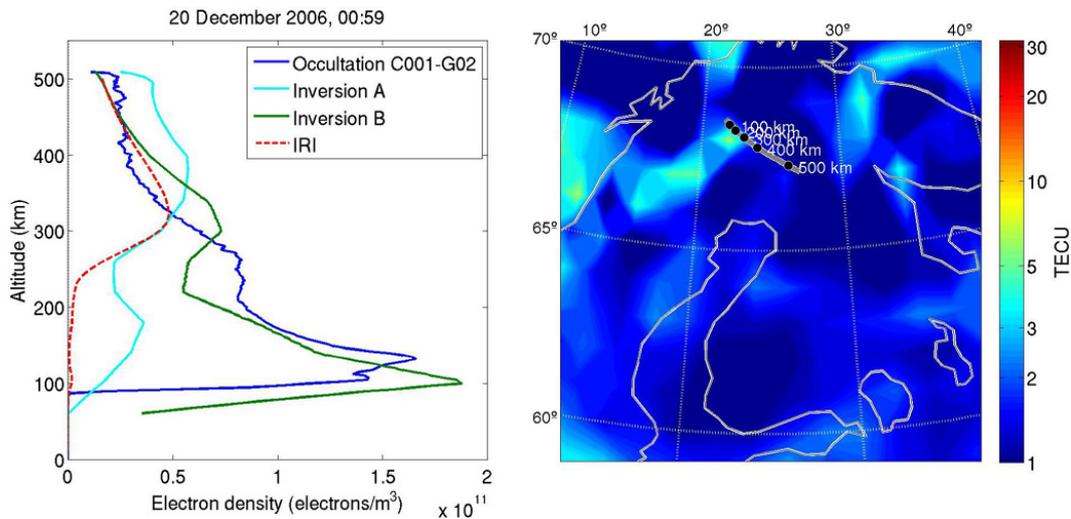
Fig. 11. Comparison between inversion result and occultation measurement, similarly as in Fig. 10, for 20 December, 00:59, between COSMIC satellite 1 and GPS satellite 2.
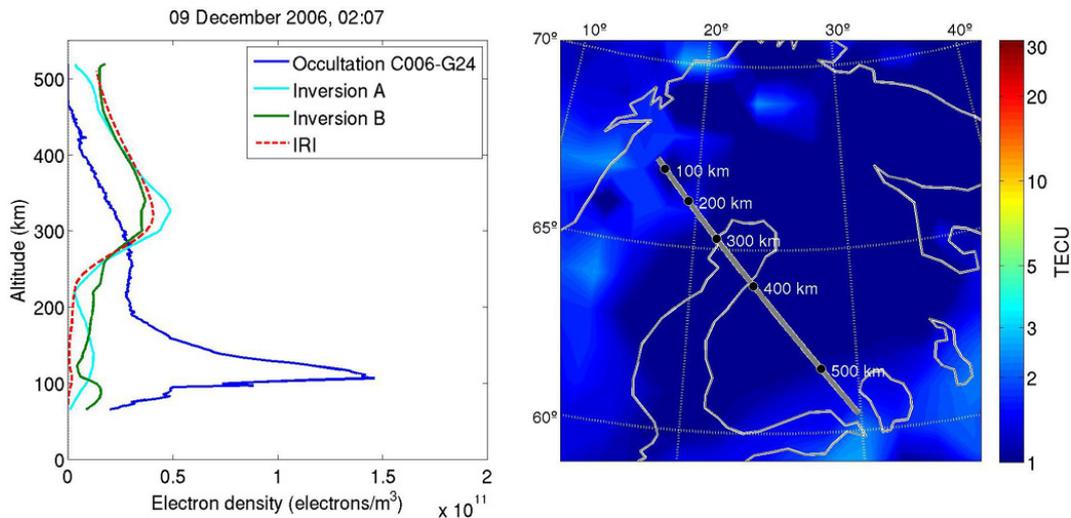
Fig. 12. Comparison between inversion result and occultation measurement, similarly as in Fig. 10, for 9 December, 02:07, between COSMIC satellite 6 and GPS satellite 24.

inversion result is correct, the occultation has likely measured TEC on a path which also crossed the local centre of the ionisation region (especially for the measurements at lower altitudes), which may have caused the occultation measurement to overestimate the electron density at the location assigned. Also, it can be assumed that if the assigned locations of the occultation had been slightly further north, its results might have been similar to the ones shown, but the interpolation from the inversion would have given higher values. Seen in this light, it can be concluded that the occultation and inversion results agree well within the expected accuracy.

Figure 11 shows another example from 20 December, just after midnight. The GPS satellite 2 was at 1.0° N/−29.4° E, and was seen from the measurement location to the west, at

azimuths between 270° and 278°. Here, the occultation and the inversion method B both detect an intense ionisation region in the E-layer above northwestern Finland not predicted by the IRI model. In this example, both results agree very well.

The inversion using method A is able to retrieve a small peak of ionisation at 180 km height. Even though none of the orthogonal functions used in this method has a peak at such a low altitude, this may have been achieved by a linear combination (including negative coefficients) of a few of the profiles as those in Fig. 4. Nevertheless, this low-altitude peak is too small compared with the occultation measurement. The RMS differences with the occultation measurements are again shown in Table 2.
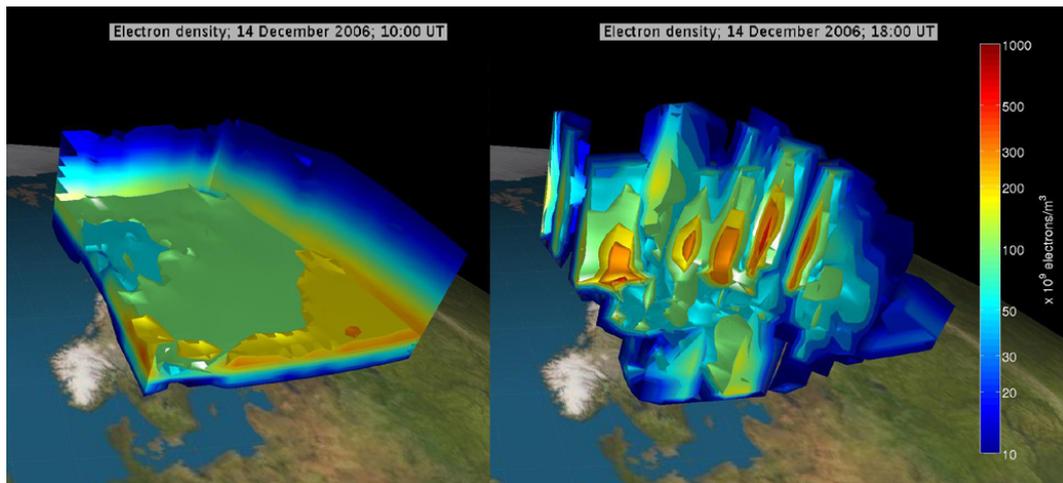
Fig. 13. Two snapshots of the 3-dimensional inversion results for 14 December.

Admittedly, there are also cases where the occultation and the inversion do not agree. Figure 12 shows an example for 9 December. The GPS satellite was located at $-0.9°$ N/$-30.7°$ E, and was seen from the measurement locations to the west, at azimuths between 267° and 273°. The seemingly higher vertical resolution of both inversion results than in the previous figures is only due to the fact that the occultation measurement location moved more rapidly in horizontal direction, so that the inversion trace traverses more grid voxels.

Here, the occultation measured an intense E-layer, which is all but missed by both inversions. Table 2 shows that the RMS of the differences with the occultation results is for inversion B hardly better than for inversion A, or for the IRI model.

The explanation for this failed result may well be that the occultation detected this E-layer above northern Sweden, outside the area where the Geotrim receivers are, and where hence no measurement paths traversing at low altitudes can be expected as inputs for the inversion. This example shows once more that reliable tomography results can only be expected in areas where a high density of crossing paths at all altitudes and in various directions is available.

At this point, method B is considered the best available method and will be used in the further graphs of this paper.

3.4 Three-dimensional results

Figure 13 shows two snapshots of the inversion results of method B as 3-dimensional visualisations. Both are from 14 December, the same day that was shown in Sect. 3.2. These graphs do not show the full grid used in the inversion (described in Sect. 2.3), but only the part for latitudes between 58° and 72° and longitudes between 11° and 43°. This is the area around Finland, where the measurement density is high enough to obtain high-resolution results.

The left-hand graph is for 10:00 UT, at which time it is close to solar midday in Finland. It shows that at this time, there was a significant F-layer, but not many irregularities. The right-hand graph is for 18:00, which corresponds to the right-hand end of the graphs in Fig. 9. It is seen that at this time, many strong medium-scale irregularities were present, which extended to high altitudes. These irregularities are the same ones as seen in the centre graph of Fig. 9, space-time converted. As the comparison with EISCAT in Sect. 3.2 showed, the size and general shape of these irregularities are well in agreement with reality; only their intensity can be somewhat underestimated by the inversion.

4 Conclusions

The ionosphere above Scandinavia in December 2006 is successfully imaged by 4-dimensional tomography using TEC measurements from the dense Geotrim GPS network in Finland, and inversion using the software package MIDAS version 2.0. The results have been tested by comparing with

– Geotrim TEC measurements calibrated independently;

– EISCAT incoherent scatter radar returns from Tromsø;

– FORMOSAT3/COSMIC radio occultation measurements.

These comparisons show the general good performance of the procedure. Specifically, the following observations are made.

Any measurement setup using GPS beacon receivers lacks sufficient vertical resolution for ionospheric tomography, especially at high altitudes. Because of this, the inversion requires extra input information concerning vertical ionisation profiles in order to model the vertical dependency of electron density properly. The software MIDAS version 2.0, which

was designed to image large-scale structures over the polar cap, used built-in Chapman profiles for this purpose, which modelled the F-layer but not the E-layer. This configuration is not suitable for medium-scale imaging over Scandinavia, and better results in terms of vertical profile are obtained by replacing this with profiles from EISCAT incoherent scatter radar measurements in Tromsø.

Because no reliable TEC offset determination is available, the inversion uses only dTEC, the difference in TEC from one sample to the next. In general, the inversion procedure is well able to calibrate these values (determine the TEC offset) automatically.

The comparison with EISCAT and FORMOSAT measurements shows that the tomography procedure is well able to image the ionosphere at a resolution as high as can be expected from the chosen grid resolution. Limitations are

– the accuracy of the inversion worsens above areas where no receivers are available;

– irregular structures of sizes near the resolution (around 100 km horizontal size) are resolved with correct shapes and sizes, but their intensities can be underestimated.

References

Bilitza, D. and Reinisch, B. W.: International Reference Ionosphere 2007: Improvements and new parameters, Adv. Space Res., 42, 599–609, doi:10.1016/j.asr.2007.07.048, 2008.

Bust, G. S., Crowley, G., Garner, T. W., Gaussiran II, T. L., Meggs, R. W., Mitchell, C. N., Spencer, P. S. J., Yin, P., and Zapfe, B.: Four-dimensional GPS imaging of space weather storms, Space Weather, 5, S02003, doi:10.1029/2006SW000237, 2007.

Folkestad, K., Hagfors, T., and Westerlund, S.: EISCAT: An updated description of technical characteristics and operational capabilities, Radio Sci., 18, 867–879, doi:10.1029/RS018i006p00867, 1983.

Fong, C. J., Yen, N. L., Chu, C. H., Yang, S. K., Shiau, W. T., Huang, C. Y., Chi, S., Chen, S. S., Liou, Y. A., and Kuo, Y. H.: FORMOSAT-3/COSMIC spacecraft constellation system, mission results, and prospect for follow-on mission, Terr. Atmos. Ocean. Sci., 20, 1–19, doi:10.3319/TAO.2008.01.03.01(F3C), 2009.

Hajj, G. A., Kursinski, E. R., Romans, L. J., Bertiger, W. I., and Leroy, S. S.: A technical description of atmospheric sounding by GPS occultation, J. Atmos. Solar-Terr. Phys., 64, 451–469, 2002.

Hargreaves, J. K.: The Upper Atmosphere and Solar-Terrestrial Relations, Van Nostrand Reinhold Co. Ltd., ISBN 0-442-30216-9, 1979.

Leitinger, R., Schmidt, G., and Tauriainen, A.: An Evaluation Method Combining the Differential Doppler Measurements from Two Stations that Enables the Calculation of the Electron Content of the Ionosphere, J. Geophys. (Zeitschrift für Geophysik), 41, 201–213, 1975.

Mitchell, C. N. and Spencer, P. S. J.: A three-dimensional time-dependent algorithm for ionospheric imaging using GPS, Annals of Geophysics, 46, 687–696, 2003.

Mitchell, C. N., Jones, D. G., Kersley, L., Pryse, S. E., and Walker, I. K.: Imaging of field-aligned structures in the auroral ionosphere, Ann. Geophys., 13, 1311–1319, 1995.

Mitchell, C. N., Kersley, L., Heaton, J. A. T., and Pryse, S. E.: Determination of the vertical electron-density profile in ionospheric tomography: experimental results, Ann. Geophys., 15, 747–752, doi:10.1007/s00585-997-0747-1, 1997.

Pokhotelov, D., Jayachandran, P. T., Mitchell, C. N., MacDougall, J. W., and Denton, M. H.: GPS tomography in the polar cap: comparison with ionosondes and in situ spacecraft data, GPS Solutions, 15, 79–87, doi:10.1007/s10291-010-0170-z, 2011.

Ratcliffe, J. A.: An introduction to the ionosphere and magnetosphere, Cambridge University Press, Great Britain, 1972.

Spencer, P. S. J. and Mitchell, C. N.: Imaging of fast moving electron-density structures in the polar cap, Annals of Geophysics, 50, 427–434, 2007.

van de Kamp, M.: Medium-scale 4-D Ionospheric Tomography using a Dense GPS Network, Research Report, Finnish Meteorological Institute, Helsinki, Finland, 2011.

Yin, P., Mitchell, C. N., Spencer, P., McCrea, I., and Pedersen, T.: A multi-diagnostic approach to understanding high-latitude plasma transport during the Halloween 2003 storm, Ann. Geophys., 26, 2739–2747, doi:10.5194/angeo-26-2739-2008, 2008.