IONOSPHERIC TOMOGRAPHY

Esa Vilenius
Invers Oy
Tähteläntie 54 A, FIN-99600 Sodankylä, Finland
esa.vilenius@invers.fi
Tuomo Nygrén
Department of Physical Sciences, University of Oulu
P.O. Box 3000, FIN-90014 University of Oulu, Finland
Markku Lehtinen, Markku Markkanen and Antero Väänänen
Sodankylä Geophysical Observatory
Tähteläntie 112, FIN-99600 Sodankylä, Finland

ABSTRACT

The development of inexpensive, high speed computers has made tomography a readily accessible tool for remote sensing of the near-Earth space. The aim of ionospheric tomography is to produce a two-dimensional image of electron density. The image is reconstructed from line integrals of electron density along rays intersecting one another. The total electron content integrals are measured using the differential Doppler technique. The most widely used traditional iterative reconstruction methods and a non-iterative method based on stochastic inversion are presented. The iterative methods need an initial profile and a pre-defined criterion for stopping the iterations. There is no systematic way for error analysis. Stochastic inversion is a non-iterative method, which does not use any initial profile. It gives the most probable values of the electron densities and provides a formalism for error analysis.

1. INTRODUCTION

The purpose of tomography is to determine the internal structure of an inhomogeneous object using waves which are influenced by the object. Tomographic methods may be used in all applications where a set of projections can be measured in several directions. For example, in the two-dimensional case the projections are line-integrals of the space-dependent quantity of interest along the lines of measurement. For reconstructing the two-dimensional spatial structure of this quantity from a set of measured projections, an inversion problem must be solved. Tomographic methods have been used in medical and industrial applications as well as in studies of the solid earth and the oceans. During the last decade algorithms for the tomographic study of the ionosphere, which is the ionized component of the atmosphere, have been under development. The same tomographic reconstruction techniques have also been applied in optical tomography to find out the brightness of the aurora and the airglow.

In ionospheric tomography, a plot of electron density in a latitude-versus-altitude grid is produced. The altitude range of interest is typically from 100 km to 1000 km above the ground. The structure of the ionosphere includes large-scale inhomogeneities of electron density and isolated inhomogeneities of relatively small size as well as travelling ionospheric disturbances. Spatial and temporal variations of the electron density are closely related to the dynamics of the ionosphere, which are influenced by solar wind, magnetosphere-ionosphere

coupling and the interaction between the neutral atmosphere and the ionosphere. At high latitudes the ionosphere is strongly influenced by precipitating particles and by large-scale electric fields of magnetospheric origin. The electron density plots may be used in space weather predictions and in estimating propagation errors of radio signals traversing the ionosphere. Radio propagation forecasts and accurate satellite orbit determination require information about the electron densities. In space-based radio navigation systems such as the Global Positioning System (GPS) the range error contributions of the ionosphere may be up to 50 m for single-frequency users *[Jakowski et al., 1996]*. The total electron content (TEC) is measured along rays crossing each other in various directions within the ionosphere. This is normally done using radio signals from navigational satellites. The transmitters are aboard beacon satellites orbiting the earth below 2000 km altitude and a chain of receivers is located on the ground. Satellite-to-satellite set-ups are also possible. The trajectory of a satellite flying over the receivers provides each receiver a scan of the ionosphere from horizon to horizon. The tomographic reconstruction is calculated in the orbital plane of the transmitting satellite. Because of the measurement geometry, horizontal scans are missing. This requires more clever reconstruction algorithms than those traditionally used in medical and industrial applications, where a full scan around the object is available.

Ionospheric radiotomography was first suggested in 1986 *[Austen et al., 1986]*, but the first experimental results were not published until the next decade *[Kunitsyn et al., 1990; Pryse and Kersley, 1992]*. Meanwhile several iterative algorithms for tomographic reconstruction were studied with simulations *[e.g. Austen et al., 1988, Raymund et al., 1990]*. An alternative approach using stochastic inverse theory was introduced in 1992 *[Fremouw et al., 1992]*. Tomographic data analysis methods and a computer software package based on Bayesian stochastic inverse theory were developed during the following years *[Lehtinen et al., 1994; Markkanen et al., 1995; Nygrén et al., 1996a]*.

In auroral tomography, the result of inversion is a map of volume emission rate in the ionospheric region where the precipitating particles produce both ionisation and auroral light. The signal strength of scanning photometers and cameras is proportional to the integral of volume emission rate. The first studies of auroral tomography were made using satellite-borne scanning photometers *[Solomon et al., 1984]*. Two-dimensional *[Vallance-Jones et al., 1991; Nygrén et al., 1996b]* and three-dimensional *[Frey et al., 1996]* maps of volume emission rate have been obtained from ground-based measurements. The small number of optical instruments and their limited fields of view give limitations to the reconstruction methods.

Tomographic methods used in radiotomography can be divided into ray and diffraction tomography. Also statistical modelling of the ionosphere using tomographic measurements has been done. Most papers on ionospheric tomography use the assumption that the scale sizes are so large that ray approximation is applicable, because on sufficiently high frequencies the ray paths are essentially linear. In diffraction tomography, the scale sizes of the structures of the object are of the order of the wavelength of the probing wave, so that diffraction effects have to be taken into account. In statistical modelling of the ionosphere, the structure of the object is so complicated that only the statistical spatial spectrum is determined. If the ionosphere contains small-scale inhomogeneities, diffraction may sometimes be important. Examples of diffraction tomography and statistical modelling are presented by Kunitsyn et al. *[1995]*.

2. THE DIFFERENTIAL DOPPLER TECHNIQUE

In ionospheric ray tomography, the projections that give information about the internal structure of the object, are integrals of electron density, called TEC (total electron content):

$$\mathrm{TEC} = \int_P N(s)ds. \tag{1}$$

The TEC represents the total number of free electrons in a column with cross-sectional area of 1. $N(s)$ is the electron density along the line of propagation P between the transmitter and the receiver. There are two methods for determining the TEC: the Faraday rotation technique and the differential Doppler technique, also called the differential carrier phase method. The latter one is described in this paper. A differential Doppler measurement gives the phase difference of coherent radio waves at two frequencies. The phase difference is proportional to the TEC. Typical frequencies used are 150 MHz and 400 MHz. The phase of the received single-frequency signal is

$$\phi = \omega \int \frac{n}{c} ds + \omega t + \psi, \tag{2}$$

where ω is the transmitted angular frequency, n the refractive index, c the velocity of light in vacuum, ds the phase path element, t the signal time of flight and ψ the phase constant. The ionospheric plasma is considered to be cold, that is the thermal motion of the particles of the plasma is neglected.[1] For frequencies higher than 100 MHz both the angular electron gyrofrequency and the collision frequency between electrons and other particles are small. Hence, the isotropic approximation for the refractive index may be used:

$$n^2 = 1 - \left(\frac{\omega_p}{\omega}\right)^2. \tag{3}$$

Since in ionospheric tomography the angular plasma frequency of the medium ω_p is smaller than the transmitted frequency ω, it follows that n is real. Because the influence of neutral gas on the refractive index is non-dispersive, the influence of plasma can be calculated by comparing the phases of two signals. If we have two phase-coherently transmitted signals with different angular frequencies ω_1 and ω_2, then the difference of scaled phases[2] is

$$\Delta\phi = \frac{\omega_2}{\omega_1}\phi_1 - \phi_2 = \frac{1 - (\frac{\omega_2}{\omega_1})^2}{2\omega_2 c} \int {\omega_p}^2 ds + \frac{\omega_2}{\omega_1}\psi_1 - \psi_2 \tag{4}$$

assuming that ray bending can be neglected and n is close to unity. The plasma frequency depends on electron density:

$${\omega_p}^2 = \frac{e^2}{m\varepsilon_o} N, \tag{5}$$

where e is the electron charge, m is the electron mass, ε_o is the permittivity in vacuum and N is the number density of free electrons[3]. Now the phase difference, also called the

[1] This simplification is valid for UHF and lower frequencies.

[2] In practice, both frequencies are coherently mixed to a common, new frequency.

[3] Ion plasma frequencies are negligible compared to the electron plasma frequency.

differential Doppler effect, takes the form

$$\Delta\phi = \frac{(1-(\frac{\omega_2}{\omega_1})^2)e^2}{2\varepsilon_o m\omega_2 c}\int N(s)ds + C, \tag{6}$$

where C is an unknown constant. Since phases are measured modulo 2π, the phase has experienced some unknown number of full revolutions on its phase path. The transmitter and the receiver also cause some phase shifts. Therefore, phase changes can only be measured starting from some constant value. The constant is different for each satellite pass and for each receiver. Each break during the measurement also creates an extra unknown phase constant. The constants are either left as additional unknowns in the reconstruction process or estimated by "calibration methods," which make use of additional information. For example, comparison with ionospheric electron content from the Faraday effect on the signals of geostationary satellites has been tried and a so-called two station method has been developed. Unfortunately, the error of these methods is at least 10% *[Leitinger et al., 1984]*.

The ionosphere is assumed to be stable during the satellite pass, which takes about 20 minutes for a satellite flying at the height of 1000 km. Changes in the ionosphere during a pass cause inconsistencies in the measurements. This is a problem especially at geophysically active times, e.g. during an aurora.

3. RECONSTRUCTION METHODS

The image generated from a set of measured projections is called a reconstruction. The ionospheric region of interest is divided into a two-dimensional rectangular grid[4]. For each observation the ray traverses certain rectangles of the grid. In a zero order interpolation the rectangles ("pixels") have a constant electron density. Thus, the TEC values are weighted sums of the pixel values:

$$\mathrm{TEC}_j = \int_{P_j} N(s)ds \approx \sum_{i=1}^{\text{no. of pixels}} N_i s_{ij}, \qquad j = 1, ..., M \tag{7}$$

with N_i being the electron density in pixel i, s_{ij} the length of the section of the ray j in pixel i and M the number of measurements during a satellite pass. $s_{ij} \neq 0$ only for the pixels which the ray traverses. The disadvantage of constant pixel values is that the reconstructed image of the ionosphere is only piecewise continuous. A continuos image is achieved by means of bilinear interpolation. The electron densities at the grid points are taken as the unknown quantities and the TEC integral is approximated using bilinear interpolation within every rectangle.

Equations (6) and (7) show that each measurement is a linear combination of the unknown density values and phase constants. In matrix form the tomography measurement is gived by the equation

$$\mathbf{m} = \mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon}. \tag{8}$$

Here $\mathbf{m}$ is a column vector containing the phase difference measurements from all receivers at discrete instants of time during a single pass, $\mathbf{x}$ is a column vector containing the unknown electron densities N_i and the unknown phase constants C, and $\boldsymbol{\varepsilon}$ is a column vector

[4]in geocentric spherical coordinates

containing the stochastic errors of the measurements. $\mathbf{A}$ is a matrix containing the coefficients of N_i and C, since the measured phase values are, after the bilinear interpolation, linear combinations of the unknown electron densities and the unknown phase constants. The problem is to find the exact or the best value of the unknown vector. When $\mathbf{x}$ is determined by the inversion of equation (8), both the density values and the phase constants are determined simultaneously. Because of the lack of measurements in the horizontal direction the measurements do not contain much information on the vertical electron density profile. There are an infinite number of solutions with different vertical variations. Therefore, some additional information is needed to determine the solution.

3.1 Iterative methods

There are many iterative methods for solving equation (8), most of which were originally developed for medical or technical tomography applications. The main difference is that in ionospheric tomography there are only a few receivers and the measurement geometry does not allow a full scan *around* the object under study.

Iterative algorithms start from an initial guess, which is derived from a model ionosphere or independent incoherent scatter radar or ionosonde measurements. The iteration stops when a predefined criterion is met, e.g. the result no longer changes or a least-squares minimum is achieved. The unknown phase constants C need to be estimated by some separate method before the iteration can start.

In Algebraic Reconstruction Technique (ART) a correction term depending on the difference between $\mathbf{m}$ and $\mathbf{m}^k$, where $\mathbf{m}^k$ is obtained by using the current estimate of the solution $\mathbf{x}^k$, is added in every iteration:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda_k \frac{m_j - \sum_{i=1}^{N} A_{ji} x_i^k}{\sum_{i=1}^{N} {A_{ji}}^2} \mathbf{A}_j^T. \tag{9}$$

Here N is the number of unknowns, $\mathbf{A}_j$ is the jth row of $\mathbf{A}$ and $j = (k \bmod N) + 1$. The sequence of relaxation parameters λ_k are usually confined to the interval $0 < \lambda_k < 2$. The relaxation parameter keeps the algorithm stable when the problem is underdetermined. ART seeks to minimize the root mean squared (rms) difference between observed data and data computed from the reconstruction. The rms difference is computed at each iteration step and additive changes are made, in order to minimize the difference. The minimum solution contains artifacts and strong gradients and the reconstructed image tends to be noisy, all of which are characteristic properties of the minimum norm type reconstruction. Therefore, the iteration has to be stopped before the gradients grow too large.

The Simultaneous Iterative Reconstruction Technique (SIRT) algorithm computes the difference according to the right-hand term of the ART equation (9), but only updates the $\mathbf{x}^{k+1}$ vector after all ray paths are calculated, i.e. $j = 1...M$. ART and SIRT do not usually preserve the overall shape of the initial guess.

Multiplicative Algebraic Reconstruction Technique (MART) uses root mean squared minimization like ART does, but the changes are applied multiplicatively from iteration to iteration:

$$x_l^{k+1} = x_l^k \left(\frac{m_j}{\sum_{i=1}^{N} A_{ji} x_i^k} \right)^{\lambda_k A_{jl}}, \tag{10}$$

where x_l^{k+1} is the lth element of the vector of unknowns and the exponent is normalized and bounded so that $0 < \lambda_k A_{jl} \leq 1$. The algorithm takes one measurement at a time and modifies the image $\mathbf{x}$ each time. Large changes are made where the initial guess had large density and small changes are made where the initial guess had small density. An iteration step is finished when all the measurements have been used. The order in which the measurements j are taken is random. The ratio of m_j to the inner product of $\mathbf{x}$ and $\mathbf{A}_j$ represents the ratio between the measurements and the values calculated from the image. If the electron density in the image is too low, this ratio will be greater than one, and the density will be increased. The weighting of the exponent by A_{jl} scales the modification of individual pixels to be proportional to the length of the ray-pixel intersection of that pixel. MART preserves the large scale structure of the start profile.

3.2 Matrix inversion methods

The inversion of equation (8) can be carried out in a single step. The inversion problem can be converted to a smaller one by constructing the ionosphere from a set of model base functions and taking the pseudoinverse of the matrix describing the mapping from the function space to the measurement space *[Raymund et al., 1995]*. A proper choice of the base functions is a crucial point in this method.

In stochastic inversion the measurements $\mathbf{m}$, their errors $\boldsymbol{\varepsilon}$ and the unknowns $\mathbf{x}$ are treated as random variables. In the inversion method by Fehmers *[Fehmers, 1994]* a regularizing function is minimized within a set consisting of positive solutions of the equation $||\mathbf{Ax} - \mathbf{m}|| = E$ where the norm is Eucledian and E is an estimated length of the error vector. The regularizing function is constructed to include both point-to-point horizontal flatness, vertical smoothness, and boundary constraint at the bottom and on the top of the ionosphere.

In the Bayesian approach to the inversion problem

$$\mathbf{m} = \mathbf{A}_m \mathbf{x} + \boldsymbol{\varepsilon}_m. \tag{11}$$

the solution is the conditional probability distribution $D(\mathbf{x}|\mathbf{m})$ called the *a posteriori* distribution. Here $\mathbf{m}$ is known and $\boldsymbol{\varepsilon}_m$ is not known, but its components are random variables with a known joint distribution function. In the absence of systematic errors the mean value of each of the components of $\boldsymbol{\varepsilon}_m$ is zero. The Bayesian approach assumes that we have some *a priori* information about the unknown. This information is given as the marginal distribution of $\mathbf{x}$. If it is flat, the mean value of $D(\mathbf{x}|\mathbf{m})$ is very unstable, that is point-to-point variations are very large. These variations have to be limited by using some kind of prior information. A usual way of avoiding numerical instability in matrix inversion is to use regularization by adding small numbers to the diagonal elements. In the present method, loose *a priori* information corresponding to a modified version of regularization is used. If the difference between the electron densities at two neighbouring grid points g_i and g_j is assumed to be zero, then some error ε_r^{ij} is made:

$$0 = x_i - x_j + \varepsilon_r^{ij}. \tag{12}$$

When this condition is written for all horizontally and vertically neighbouring grid points the equations can be written as

$$\mathbf{0} = \mathbf{A}_r \mathbf{x} + \boldsymbol{\varepsilon}_r. \tag{13}$$

Here $\mathbf{A}_r$ is a non-diagonal matrix with elements $0, -1$ and 1 such that each of the equations in (13) is of the form (12). In this method the point-to-point steps of electron density, both in vertical and horizontal directions, are treated as random variables with zero expectation values. Suitable variances are given to control the step sizes. If the variance is large, both small and large positive and negative steps are probable, whereas in the case of a small variance large steps are unlikely to happen. Matrix equation (13) also allows us to put the electron density very close to zero at the bottom and on the top of the grid to get physically reasonable results. The mean value of $D(\mathbf{x}|\mathbf{m})$ is guided towards a general layer-shaped ionosphere by selecting a variance profile with a certain peak value and peak height and thickness. Close to the peak height the variances are large, and large electron density steps from one grid point to another are possible. Although a certain variance profile is used, the resulting electron density does not have the same shape, and the peak height of the density is usually different from that of the variance profile. Equations (11) and (13) can be combined to a single matrix equation:

$$\begin{pmatrix} \mathbf{m} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_m \\ \mathbf{A}_r \end{pmatrix} \mathbf{x} + \begin{pmatrix} \boldsymbol{\varepsilon}_m \\ \boldsymbol{\varepsilon}_r \end{pmatrix}. \tag{14}$$

Since $\mathbf{m}$ and $\mathbf{0}$ have the same role in this equation, giving prior information is equivalent to adding fictitious measurements. The total number of these 'measurements' and the true measurements is greater than the number of unknowns in the inversion problem. It is possible to have dense grids with the number of unknowns exceeding the number of true measurements.

If all errors in equations (14) are Gaussian with zero mean, the density function of the *a posteriori* distribution of the unknowns is

$$D\left(\mathbf{x}|\mathbf{m},\mathbf{0}\right) \propto D_{pr}\left(\mathbf{x}\right) \exp\left[-\frac{1}{2}\left(\mathbf{A}_m\mathbf{x}-\mathbf{m}\right)^T \boldsymbol{\Sigma}_m^{-1}\left(\mathbf{A}_m\mathbf{x}-\mathbf{m}\right)\right], \tag{15}$$

where

$$D_{pr}\left(\mathbf{x}\right) = \exp\left[-\frac{1}{2}\left(\mathbf{A}_r\mathbf{x}\right)^T \boldsymbol{\Sigma}_r^{-1}\left(\mathbf{A}_r\mathbf{x}\right)\right] \tag{16}$$

is the *a priori* density of the unkowns. The matrices $\boldsymbol{\Sigma}$ are the covariance matrices of the measurement errors and the *a priori* variances: $\boldsymbol{\Sigma}_m = \langle \boldsymbol{\varepsilon}_m \boldsymbol{\varepsilon}_m^T \rangle$ and $\boldsymbol{\Sigma}_r = \langle \boldsymbol{\varepsilon}_r \boldsymbol{\varepsilon}_r^T \rangle$. The result we are usually interested in is the most probable point of the *a posteriori* distribution of $\mathbf{x}$:

$$\hat{\mathbf{x}} = \mathbf{C}\mathbf{A}_m^T \boldsymbol{\Sigma}_m^{-1} \mathbf{m}, \tag{17}$$

where $\mathbf{C}$ is the *a posteriori* covariance matrix

$$\mathbf{C} = (\mathbf{A}_m^T \boldsymbol{\Sigma}_m^{-1} \mathbf{A}_m + \mathbf{A}_r^T \boldsymbol{\Sigma}_r^{-1} \mathbf{A}_r)^{-1}. \tag{18}$$

The point $\hat{\mathbf{x}}$ is both the maximum likelihood point and the mean value of $D\left(\mathbf{x}|\mathbf{m},\mathbf{0}\right)$. The matrix $\mathbf{C}$ gives the error estimates of the unknowns. The variance of the component x_i is $\sigma_i^2 = C_{ii}$. The reliability of the error estimates depends on the correct size of the *a priori* variances.

It can be noted that the point $\hat{\mathbf{x}}$ is a result of minimizing $\left(\mathbf{A}_m\mathbf{x}-\mathbf{m}\right)^T \boldsymbol{\Sigma}_m^{-1}\left(\mathbf{A}_m\mathbf{x}-\mathbf{m}\right) + \left(\mathbf{A}_r\mathbf{x}\right)^T \boldsymbol{\Sigma}_r^{-1}\left(\mathbf{A}_r\mathbf{x}\right)$. The same minimization is done when regularization is used to solve the inversion problem.

4. EXPERIMENTAL PROCEDURES

In November 1995 an experimental measurement campaign with four receiving stations was carried out. The stations of Tromsø (Norway), Esrange (Sweden), Kokkola (Finland) and Kärkölä (Finland) were installed in a chain approximately along a magnetic meridian. In this way it is possible to get a better picture of the field-aligned structures in the ionosphere which are caused by the geomagnetic field. The length of the chain was 1036 km and the site separations were 214, 458 and 364 km from north to south. The receivers carried out differential Doppler measurements using signals from Russian navigational satellites of the Tsykada system flying at 1000 km altitude. Since the inclination angle of the satellite orbit is about 83°, the satellites fly approximately along magnetic meridians above Scandinavia. The satellites transmit coherent waves at 150 and 400 MHz. On average, a suitable passage is encountered at 90-min intervals and a single passage takes about 18 min. Russian receivers made for this campaigne were used for receiving the signals. The 400 MHz signal was mixed down to 150 MHz, its amplitude was fixed and it was used as a reference in a hybrid detector. The sampling frequency used was 50 Hz. The complex output of the hybrid detector gives the phase difference of the two frequencies. The complex output together with 1-s pulses extracted from the satellite telemetry signal and a timing signal from an omega-clock were fed to a multiplexer, and the multiplexer output was fed to a PC computer. In the preliminary phase of data analysis the data were post-integrated to a lower time resolution. Phase curves free from ionospheric scintillation or local radio disturbance were selected for tomographic analysis.

The Tsykada system is equivalent to the American Transit system, which was phased out in 1997. The Russian system remains operational for still some time, new satellites replacing aging older ones. There are usually about 10 satellites active, giving new and also simultaneous sweeps reasonably often. Some new purely scientific American satellites will soon be available, and India is planning two different types of tomography satellites. The low-orbit satellites give a much faster sweep across the ionosphere than the slowly moving high-orbit GPS satellites. Also the ionospheric influence is much weaker at the 1.2 or 1.5 GHz GPS frequencies.

Sodankylä Geophysical Observatory is currently conducting a project of setting up a chain of five or six new receiving stations, together with Invers Oy of Sodankylä, University of Oulu, University of Tromsø, and Polar Geophysical Institute and Tekhvers Ltd from Murmansk. The project is funded by the above mentioned organizations together with the European Union, the Technology Devopment Centre of Finland (TEKES), and the Finnish Academy of Science and Letters. In the old receivers all the special signal processing like the elimination of Doppler shift due to the movement of the satellite along its orbit and the determination of phase differences has been done with the receiver electronics. In the new receivers the RF receiver and signal processing are separated so that the RF part only supplies raw signals to A/D converters, and everything essential is done with software in both the interface circuits and in a suitable computer. The new system can receive several satellites simultaneously. The TEC values are buffered in the local computers and transferred to the tomography analysis site using standard ways of connecting to the Internet.

5. RESULTS

5.1 Inversion of simulated data

In order to test the tomographic software package, simulated data were calculated from a model ionosphere constructed for this purpose and the results of tomographic inversion were compared with the original model. The upper panel of figure 1 shows the model, which contains a travelling ionospheric disturbance (TID). Disturbances of this type are created by a gravity wave propagating in the neutral atmosphere. The same software has also been used in studying airglow emission observations from the mesopause region at 80 - 100 km altitude, where gravity waves are important drivers of the mean winds. The middle panel shows the ionosphere reconstructed by the inversion of simulated measurements. Four receiving stations with equal separations of 333 km were used. The data were simulated by integrating the model electron density along rays from the satellite to the receiver. Each receiver made 200 measurements at angles above 10° above the horizon. In order to simulate the statistical error of measurement, random Gaussian noise with a standard deviation of 0.1 rad was added to the calculated phase angles. The peak altitude of the *a priori* variance profile was chosen to be 300 km. The shape was chosen to be bi-Gaussian with widths of 150 and 250 km in the lower and upper parts respectively. In the reconstruction there is an absence of the TID to north of the northernmost receiving station. This is due to the fact that the rays from a satellite to all the stations traverse the wave fronts and front-aligned rays are completely missing. This is a good example of limitations posed by the measurement geometry.

5.2 Inversion of observed phase curves

Results from a satellite passage during an experimental campaign are shown in figure 2. The upper panel shows the phase registered by each receiver as a function of the satellite vertical projection on the ground. The positive direction of the horizontal axis is northwards. The calculation of the phase starts from zero at the moment when the receiver is locked to the satellite signal. The satellite flew from north to south. Only observations collected at elevation angles higher than 10° are plotted and used in the analysis, which is the reason why the phase curves start at different levels. There is a maximum in the phase curves seen by the Tromsø and Esrange stations but not seen at all by the Kärkölä station. Kokkola sees a small maximum far north of the station. From this it can be concluded that there is a maximum in electron density somewhere north of Esrange.

The reconstruction shows an active region above Esrange and Tromsø at 200 to 400 km altitude. The shape of the region suggests that there might well be a visible aurora present. During an aurora the ionosphere is not stable during the satellite passage, which causes inconsistency in the measurement data. In the reconstruction this is seen as regions of negative densities. An equally probable reason is that the *a priori* variance profile is slightly inconsistent with the measurements. The starting and ending measurement lines for each station are shown. The lines are curved in this geometry because of the curvature of the Earth. The grid size was 43 km in the horizontal direction and 25 km in the vertical direction. A Chapman function was used as a profile for the *a priori* variances. The peak height and scale height were 300 km and 150 km, respectively. In this way the electron density is

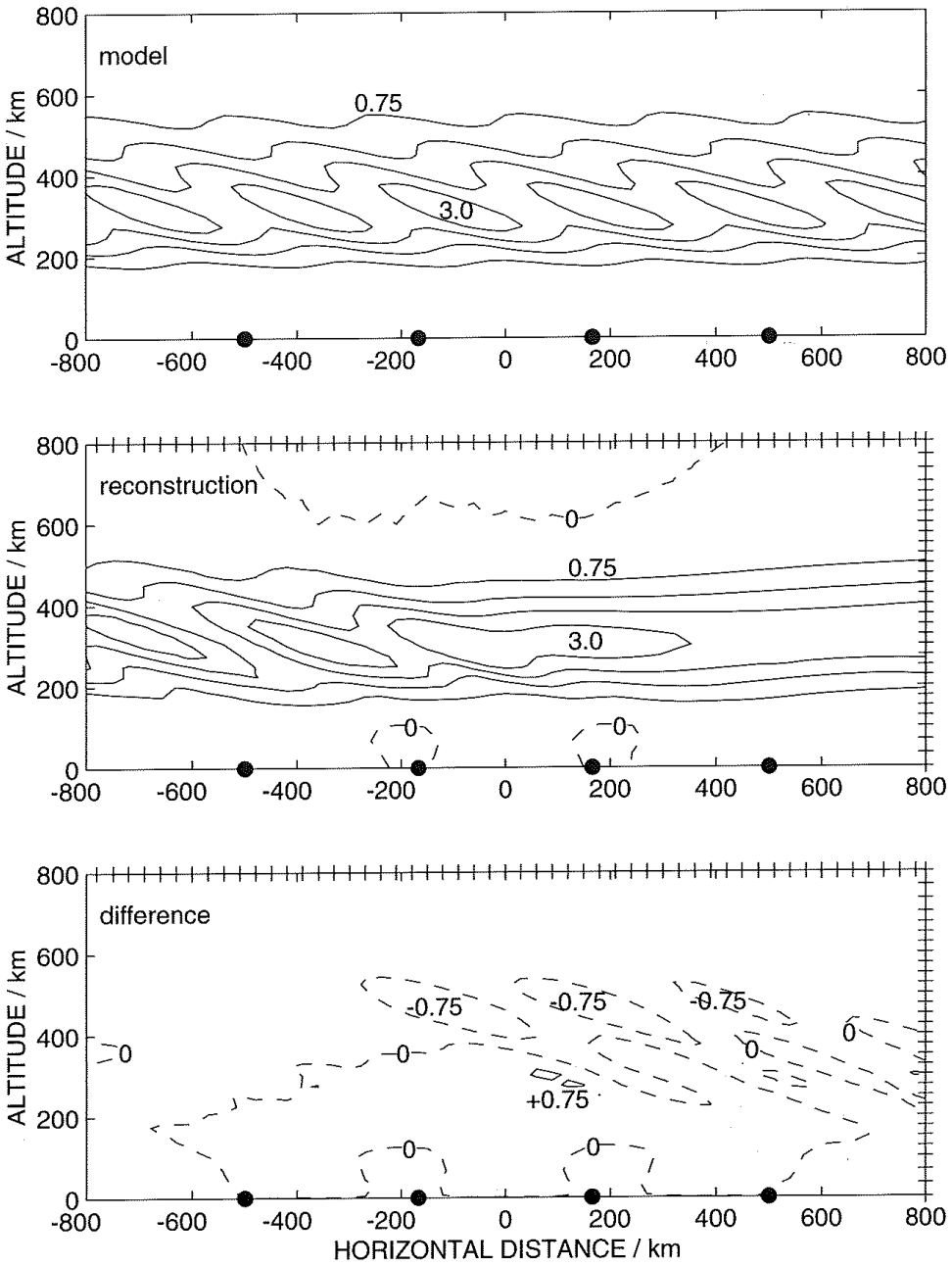

Figure 1. Model, reconstruction and their difference. The contour interval is 0.75.
The receiving stations are marked with black circles and edges of the grid with crosses.

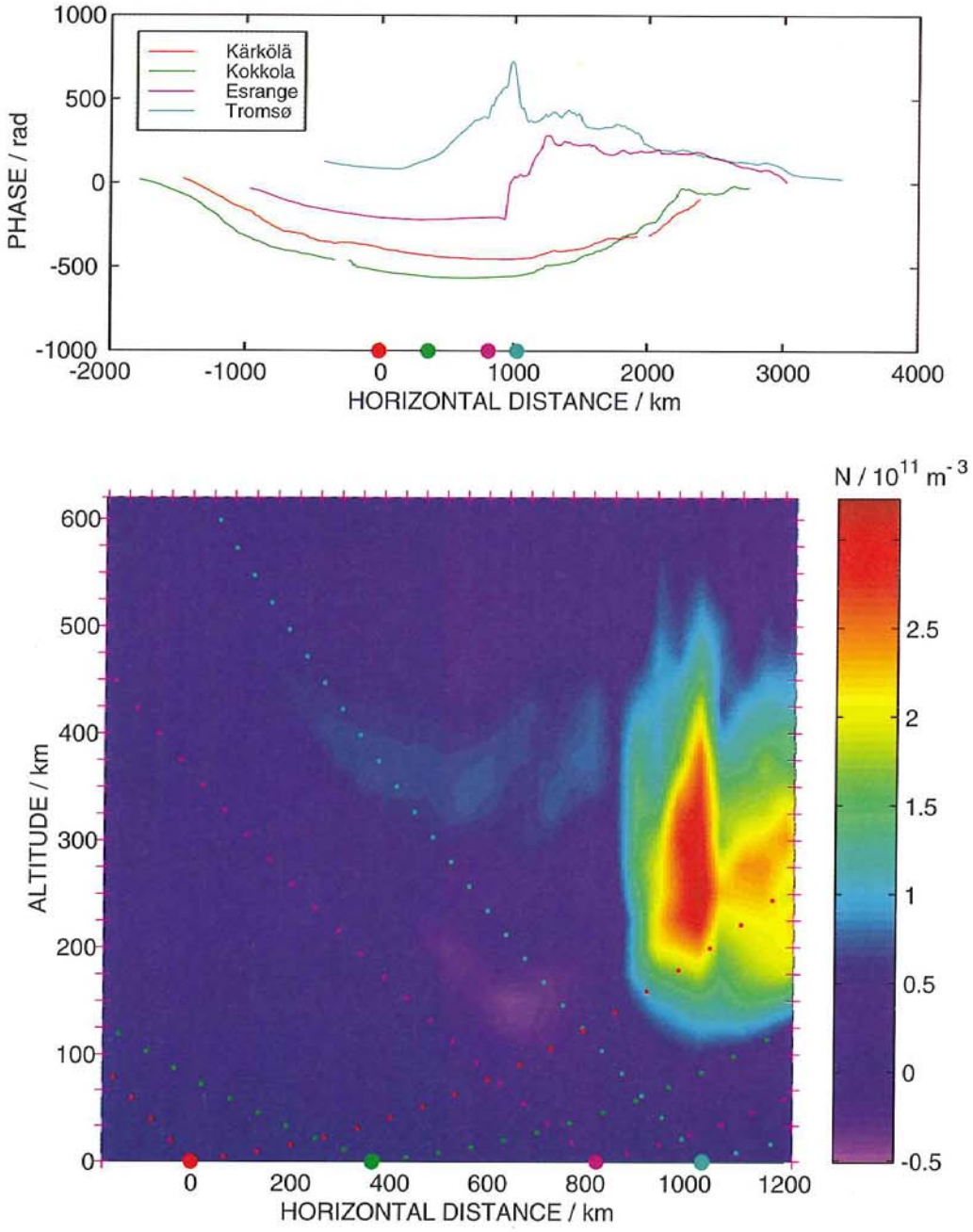

Figure 2. Observed phase curves and the calculated reconstruction on 17 Nov 1995 at 21:32-21:48 UT.

allowed to change more freely at heights where the density is expected to be large.

6. DISCUSSION AND CONCLUSIONS

Satellite tomography of the ionosphere extends the capabilities of research of the near-Earth space environment. It can provide additional information about structures like the mid-latitude ionization trough, the equatorial anomaly region, travelling ionospheric disturbances and large-scale blob-like features of the high-latitude ionosphere. Tomographic mapping takes place quickly within a wide region, which can be extended simply by adding more receivers to the measurement chain. Because of the measurement geometry, the most difficult case for tomography is a horizontally stratified layer. The problem gets easier with increasing horizontal gradients and the easiest structure is an isolated blob *[Nygrén et al., 1997]*.

Reconstruction of tomographic images from measured projections is a mathematical inversion problem. Several methods to solve the problem are available. Iterative methods are commonly used, but it is not well understood how good the solutions are. The convergence of the iterative method is not necessarily clear and it is not known how close to the convergence limit it has stopped. Depending on the start profile, the iteration may lead to different results which all satisfy the original measurements equally well. The quality of the result depends on how suitable the start profile is. Test results clearly indicate that most methods are not capable of reproducing an abrupt step in the peak altitude of the ionospheric layer *[Raymund, 1995]*, because the height is too strongly controlled by the start profile.

The solution of Bayesian stochastic inversion is well understood: it gives the most probable values of the unknowns once the measurements and prior information are known. *A priori* variance profile does not determine the peak altitude of the resulting profile, but it may make the width of the profile too wide or narrow if the variances are not well chosen. In the form presented in this paper, there are no positivity constraints. The reconstructions may lead to negative electron densities within regions of very small densities, where the absolute error is greater than the density value itself. Negative regions appear also if the measurement data are inconsistent. Also prior information has to be consistent with real measurements. The method provides means of including information or additional measurements provided by other instruments like ionosondes or incoherent scatter radars.

The results from Bayesian stochastic inversion of tomography measurements have been tested against independent incoherent scatter radar measurements in an experimental campaign in November 1995. The method is capable of reconstructing reliably the gross-scale structures of the ionosphere, although differences may occur in details *[Nygrén et al., 1996a]*.

Acknowledgements. The authors are grateful to E. D. Tereshchenko and B. Z. Khudukon from Polar Geophysical Institute, Murmansk, for their efforts and equipment during the November 1995 measurement campaign and their co-operation thereafter. The authors also wish to thank A. Ossipov and Th. Ulich for useful remarks and P. Piiroinen for tips in using LaTeX.

REFERENCES

Austen, J. R., S. J. Franke, C. H. Liu, K. C. Yeh, 1986, Application of computerized tomography techniques to ionospheric research, in Proceedings of the International Beacon Satellite Symposium, University of Oulu, Finland, Ed. A. Tauriainen, pp 25-35.

Austen, J. R., S. J. Franke, C. H. Liu, 1988, Ionospheric imaging using computerized tomography, Radio Science, Vol. 23, pp 299-307.

Censor, Y., 1983, Finite Series-Expansion Reconstruction methods, Proceedings of IEEE, Vol. 71, pp. 409-419.

Davies, K., 1969, Ionospheric Radio Waves, Blaisdell Publishing Company.

Fehmers, G., 1994, A new algorithm for ionospheric tomography, in Proceedings of the International Beacon Satellite Symposium, University of Wales, Aberystwyth, UK, 11-15 July 1994, Ed. L. Kersley, pp, 52-55.

Fremouw, E. J., A. Secan, B. M. Howe, 1992, Application of stochastic inverse theory to ionospheric tomography, Radio Science, Vol. 27, pp. 721-732.

Frey, S., H. U. Frey, D. J. Carr, O. H. Bauer, G. Haerendel, 1996, Auroral emission profiles extracted from three-dimensional reconstructed arcs, Journal of Geophysical Research, Vol. 101, pp. 21731-21741.

Jakowski, N., E. Sardon, E. Engler, A. Jungstand, D.Klähn, 1996, Relationships between GPS-signal propagation errors and EISCAT observations, Annales Geophysicae, Vol. 14, pp. 1429-1436.

Kunitsyn, V. E., E. D. Tereshchenko, E. S. Andreeva, A. V. Galinov, M. A. Filimonov, 1990, Radiotomography of global ionospheric structures (in Russian), Polar Geophysics Institute Press, Apatity, No. 90-10-78, pp. 1-30.

Kunitsyn, V. E., E. D. Tereshchenko, E. S. Andreeva, B. Z. Khudukon, Y. A. Melnichenko, 1995, Radiotomographic investigations of ionospheric structures at auroral and middle latitudes, Annales Geophysicae, Vol. 13, pp. 1242-1253.

Lehtinen, M., M. Markkanen, P. Henelius, E. Vilenius, T. Nygrén, E. D. Tereshchenko, B. Z. Khudukon, 1994, Bayesian approach to satellite radiotomography, in Proceedings of the International Beacon Satellite Symposium, University of Wales, Aberystwyth, UK, 11-15 July 1994, Ed. L. Kersley, pp. 80-83.

Leitinger, R., 1995, Tomography, in Modern Ionospheric Science, Ed. H. Kohl, R. Ruster, K. Schlegel, pp. 346-370.

Leitinger, R., G. K. Hartmann, F.-J. Lohmar, E. Putz, 1984, Electron content measurements with geodetic Doppler receivers, Radio Science, Vol. 19, No. 3, pp. 789-797.

Markkanen, M., M. Lehtinen, T. Nygrén, J. Pirttila, P. Henelius, E. Vilenius, E. D. Tereshchenko, B. Z. Khudukon, 1995, Bayesian approach to satellite radiotomography with applications in the Scandinavian sector, Annales Geophysicae, Vol. 13, pp. 1277-1287.

Nygrén, T., M. Markkanen, M. Lehtinen, E. D. Tereshchenko, B. Z. Khudukon, 1996a, O. V. Evstafiev, P. Pollari, Comparison of F region electron density observations by satellite radio tomography and incoherent scatter methods, Annales Geophysicae, Vol. 14, pp. 1422-1428.

Nygrén, T., M. Markkanen, M. Lehtinen, K. Kaila, 1996b, Application of stochastic inversion in auroral tomography, Annales Geophysicae, Vol. 14, pp. 1124-1133.

Nygrén, T., M. Markkanen, M. Lehtinen, E. D. Tereshchenko, B. Z. Khudukon, 1997, Stochastic inversion in ionospheric radiotomography, Radio Science, Vol. 32, No. 6, pp. 2359-2372.

Nygrén, T., M. J. Taylor, M. S. Lehtinen, M. Markkanen, 1998, Application of tomographic inversion in studying airglow in the mesopause region, Annales Geophysicae, Vol. 16, pp. 1180-1189.

Pryse, S. E., L. Kersley, 1992, A preliminary experimental test of ionospheric tomography, Journal of Atmospheric Terrestial Physics, Vol. 54, pp. 1007-1012.

Raymund, T. D., Comparison of several ionospheric tomography algorithms, 1995, Annales Geophysicae, Vol. 13, pp. 1254-1262.

Raymund, T. D., J. R. Austen, S. J. Franke, C. H. Liu, J. A. Klobuchar, J. Stalker, 1990, Application of computerized tomography to the investigation of ionospheric structures, Radio Science, Vol. 25, pp. 771-789.

Solomon, S. C., P. B. Hays, V. J. Abreu, 1984, Tomographic inversion of satellite photometry, Applied Optics, Vol. 23, pp. 3409-3414.

Vallance-Jones, A., R. L. Gattinger, F. Creutzberg, F. R. Harris, A. G. McNamara, A. W. Yau, E. J. Llewellyn, D. Lummerzheim, M. H. Rees, I. C. McDade, J. Margot, 1991, The ARIES auroral modelling campaign: characterization and modelling of an evening auroral arc observed from a rocket and a ground-based line of meridian scanners, *Planetary Space Science,* Vol. 39, pp. 1677-1705.