Application of generalized singular value decomposition to ionospheric tomography

K. Bhuyan, S. B. Singh, and P. K. Bhuyan
Department of Physics, Dibrugarh University 786004, India

Received: 11 July 2003 – Revised: 3 May 2004 – Accepted: 18 May 2004 – Published: 3 November 2004

Abstract. The electron density distribution of the low- and mid-latitude ionosphere has been investigated by the computerized tomography technique using a Generalized Singular Value Decomposition (GSVD) based algorithm. Model ionospheric total electron content (TEC) data obtained from the International Reference Ionosphere 2001 and slant relative TEC data measured at a chain of three stations receiving transit satellite transmissions in Alaska, USA are used in this analysis. The issue of optimum efficiency of the GSVD algorithm in the reconstruction of ionospheric structures is being addressed through simulation of the equatorial ionization anomaly (EIA), in addition to its application to investigate complicated ionospheric density irregularities. Results show that the Generalized Cross Validation approach to find the regularization parameter and the corresponding solution gives a very good reconstructed image of the low-latitude ionosphere and the EIA within it. Provided that some minimum norm is fulfilled, the GSVD solution is found to be least affected by considerations, such as pixel size and number of ray paths. The method has also been used to investigate the behaviour of the mid-latitude ionosphere under magnetically quiet and disturbed conditions.

Key words. Ionosphere (modelling and forecasting; equatorial ionosphere) – Radio science (ionospheric propagation)

1 Introduction

Computerized tomography (CT) technique has had a revolutionary impact in the field of medical diagnostic imaging and is now routinely used to produce high quality images of a section of the human body. The mathematical foundation of CT was laid as early as 1917 by Radon (Deans, 1983). The technique has also been applied to the study of ocean structures (Munk and Wunch, 1979) and geological formations (Dines and Lyttle, 1979). CT technique can be used for investigation of the vertical and horizontal structure of the ionosphere when the line integral of the parameter (e.g. electron density) is the measured data through the region of interest.

A radio signal passing through the ionosphere suffers time delay, refraction and polarization rotations during its passage through the ionized plasma. Since the beginning of the satellite era, radio beacons on board artificial satellites have been routinely used to measure the total ionospheric electron content (TEC) using techniques such as Faraday rotation, group delay, Doppler shift (de Mendonca and Garriott, 1962; Davies, 1980, 1989, etc.). TEC is a measure of the total number of electrons contained in a tube of unit cross sectional area along the line of sight from the transmitter to the receiver. Individual and coordinated TEC measurements provide information about the temporal and horizontal distribution of electron content on various time scales. However, since TEC is the line integral of the electron density along the ray path, it cannot be used for investigating the vertical structure of the ionosphere using conventional techniques. Austen et al. (1986, 1988) first proposed the application of computerized tomography technique for reconstruction of the electron density distribution below and above the F2 peak by utilizing TEC observations. However, for utilizing TEC obtained from absolute phase measurements, an additional procedure of initial phase estimation is necessary, which leads to significant difficulties and reduces the reconstruction accuracy. Pryse and Kersley (1992) reported preliminary experimental results based on the method proposed by Austen et al. (1986, 1988). Andreeva et al. (1990) and Kunitsyn and Tereshchenko (1992) proposed a phase-difference approach making use of relative TEC derivatives and obtained experimental ray tomographic (RT) images of the ionosphere for the first time. The genre of computerized ionospheric tomography (CIT) that followed Austen et al. (1986, 1988) basically adheres to the principle of measurement of total electron content at a meridional chain of ground receivers using a transit satellite and then inverting the same to reconstruct a two-dimensional electron density versus height profile. Raymund et al. (1990) applied the computerized ionospheric tomography (CIT) to realistically simulate the ionospheric electron density variations over 16° in latitude within height range of 50 to 1000 km.

A number of theoretical as well as experimental CIT schemes have been reported so far in literature (Raymund et al., 1990; Fremouw et al., 1992; Na and Lee, 1990, 1991,
In this paper, we investigate the electron density distribution of the low- and mid-latitude ionosphere through tomographic reconstruction using the Generalized Singular Value Decomposition algorithm reported by Bhuyan et al. (2002). The TEC data used in this analysis for reconstruction of the low- and mid-latitude ionosphere through tomography can be found in Raymund (1994), Pryse and Kersley (1992), Kunitsyn and Tereshchenko (2003).

In our notation, the difference between the observed TEC data and TEC data computed from the reconstruction is based on the difference between the initial guess and measured data. Of course, the difference of MART with ART is that the changes are multiplicative rather than additive. More detailed discussion of algorithms used in tomography can be found in literature (Kunitsyn et al., 1994, 1995). The different orders of approximations for the function sought, that was studied by Kunitsyn et al. (1994), are the piecewise constant, piecewise planar, the product of cubic approximations, the product of linear approximations, the product of cubic approximations, etc. However, it should be noted that as the order of approximation increases, the matrix properties for solving the inverse problem worsen.

2 Method of reconstruction

2.1 Theory

In a CIT experiment, the phase difference suffered by two coherent radio signals transmitted from radio beacons on board transit satellites is measured using a meridional chain of ground receivers. For example, a number of CIT experiments (Kersley et al., 1993, 1997; Foster et al., 1994, etc.) have been performed using 150 MHz and 400 MHz transmissions from NNSS, Cicada satellites. The dispersive phase \( \phi \) (phase difference between the two signals reduced to a reference frequency) is proportional to the Total Electron Content (TEC) along the path S connecting the satellite and receiver and is given by (Andreeva et al., 1992, 2001)

\[
k \lambda r_e \int_S N d\sigma = \phi,
\]

where \( \lambda \) is the probing wavelength, \( r_e \) is the classical electron radius, \( k \) is proportionality constant and \( \int_S N d\sigma \) is the TEC along the path S. \( k \) is of the order of unity and is connected to the reference frequency at which the dispersive phase is measured.

This absolute phase approach to the ray tomography of the ionosphere suffers from the problem of initial phase determination known as phase ambiguity or \( 2\pi \) ambiguity. This ambiguity in phase measurement causes substantial error in the reconstruction process (Kunitsyn and Tereshchenko, 2003). However, the problem of initial phase determination can be avoided, if the phase-difference tomography as suggested by Andreeva et al. (1992) is adopted.

Subject to appropriate discretization, the ray tomography problem transforms into a problem of solving a system of linear equations. The system of linear equations can be written as

\[
y_{M \times 1} = A_{M \times N} x_{N \times 1} + e_{M \times 1},
\]

where \( A_{M \times N} \) is the operator matrix for transition of the sought function (i.e. the function to be reconstructed, \( x_{N \times 1} \) in our notation) to the linear integrals \( y_{M \times 1} \). The matrix \( A \) takes different forms depending on whether the approach is based on phase or phase difference. Its form also depends on the approximation used for the function sought. A detailed discussion can be found in literature (Kunitsyn et al., 1994, 1995). The different orders of approximations for the function sought, that was studied by Kunitsyn et al. (1994), are the piecewise constant, piecewise planar, the product of linear approximations, the product of cubic approximations, etc.
In this paper, the matrix $\mathbf{A}$ has been calculated using piecewise planar approximation of the function, sought following Andreeva et al. (1992). The piecewise constant approximation has been avoided, as it leads to the largest error in reconstruction (Andreeva et al., 2001). The discretization process and the intersection of a single ray path with only a very small fraction of the total number of the pixels makes the geometry matrix largely sparse and hence highly ill conditioned. Kunitake et al. (1995) used a method known as Modified Truncated Singular Value Decomposition (MTSVD) for solving such ill-conditioned problems. This method is based on the truncation of singular values causing larger perturbations in the solution. The choice of the truncation parameter is arbitrary. Zhou et al. (1999) derived an optimal truncation criterion for using SVD for reconstruction purposes. The approach is based on the inversion of a weighted geometry matrix constructed by incorporating a priori information in the form of a priori and a posteriori covariance for both the data and the model parameters (function sought). In the inversion of the weighted geometry matrix, the series of the weighted singular values has been truncated at unity. The truncation criterion has been chosen due to the fact that singular values greater than unity decrease the trace of the a posteriori criterion has been chosen due to the fact that singular values has been truncated at unity. The truncation has been avoided, as it leads to the largest error in reconstruction. Tikhonov regularization is one such method (Tikhonov, 1963).

2.2 The algorithm

Huyan et al. (2002) reported a new CIT algorithm based on the Generalized Singular Value Decomposition (GSVD). For the sake of completeness of the paper the algorithm is discussed below in brief.

Ionospheric tomography is a part of the family of ill-posed problems, which can be written in the form of a system of linear equations given by Eq. (2). A common way to solve such ill-conditioned problems is to use the least-square solution to the set of equations

$$ \mathbf{x}_{ls} = \arg \min_{\mathbf{x}} \left( || \mathbf{A} \mathbf{x} - \mathbf{y} ||^2 \right), $$

(3)

where $|| \cdot ||$ denotes $l_2$ norm of a vector. When $\mathbf{A}$ is ill conditioned, the least-square reconstructed object $\mathbf{x}_{ls}$ obtained through Eq. (3) will be corrupted by amplified noise. There are many methods for regularizing such problems in order to generate reasonable estimate. Tikhonov regularization is one such method.\textsuperscript{2}

The GSVD can be used to solve the damped least-squares problem, as proposed by Tikhonov (1963). This approach amounts to finding out the $\mathbf{x}$ that solves

$$ \min_{\mathbf{x} \in \mathbb{R}^N} \left( || \mathbf{A} \mathbf{x} - \mathbf{y} ||^2 + \alpha^2 || \mathbf{L} \mathbf{x} ||^2 \right), $$

(5)

where $\alpha$ is the regularization parameter and $\alpha > 0$. $\mathbf{L}$ is a positive definite matrix. $\mathbf{L}$ takes different forms in accordance with the order of regularization. For zero order Tikhonov regularization, $\mathbf{L}=\mathbf{I}$, the identity matrix and for first order Tikhonov regularization

$$ \mathbf{L} = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{pmatrix}, $$

(6)

where $\mathbf{L}$ is a $P \times N$ matrix.

Under GSVD technique, the matrices $\mathbf{A}_{M \times N}$ and $\mathbf{L}_{P \times N}$ with $M \geq N \geq P$, and rank ($\mathbf{L}$)=P, can be written as

$$ \mathbf{A} = \mathbf{U} \begin{pmatrix} \Lambda & 0 \\ 0 & \mathbf{I}_{N-P} \end{pmatrix} \mathbf{B}^{-1}, $$

(7)

and

$$ \mathbf{L} = \mathbf{V} \begin{pmatrix} \mathbf{C} & 0 \\ 0 & \mathbf{I}_{N-P} \end{pmatrix} \mathbf{B}^{-1}, $$

(8)

where $\mathbf{U}$ is $M \times N$ and orthogonal, $\mathbf{V}$ is $P \times P$ and orthogonal, and $\mathbf{B}$ is $N \times N$ and non-singular. $\Lambda$ is $P \times P$ diagonal matrix with elements $\lambda_i$ such that

$$ 0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \lambda_4 \ldots \leq \lambda_P \leq 1 $$

(9)

$\mathbf{C}$ is a $P \times P$ diagonal matrix with elements $\mu_i$ such that

$$ 1 \geq \mu_1 \geq \mu_2 \geq \mu_3 \geq \mu_4 \ldots \geq \mu_P > 0. $$

(10)

The values of $\lambda_i$ and $\mu_i$ are normalized so that

$$ \lambda_i^2 + \mu_i^2 = 1, i=1, 2, \ldots, P $$

(11)

The generalized singular values are

$$ \gamma_i = \frac{\lambda_i}{\mu_i} $$

(12)

such that $0 \leq \gamma_1 \leq \gamma_2 \leq \gamma_3 \leq \ldots \leq \gamma_P$.

It can be easily seen that

$$ \mathbf{B}^T \mathbf{A}^T \mathbf{A} \mathbf{B} = \begin{pmatrix} \Lambda^2 & 0 \\ 0 & \mathbf{I} \end{pmatrix}, $$

(13)

and

$$ \mathbf{B}^T \mathbf{L}^T \mathbf{L} \mathbf{B} = \begin{pmatrix} \mathbf{C}^2 & 0 \\ 0 & 0 \end{pmatrix}, $$

(14)

when $P < N$, the matrix $\mathbf{L}$ has a nontrivial null space. The vectors $\mathbf{B}_p, \mathbf{B}_{p+1}, \mathbf{B}_{p+2}, \ldots, \mathbf{B}_N$ form a basis for the null space of $\mathbf{L}$.

Equation (3) can now be written as

$$ \min \left( || \begin{pmatrix} \mathbf{A} & \mathbf{y} \end{pmatrix} \mathbf{x} - \begin{pmatrix} \mathbf{y} \\ 0 \end{pmatrix} ||^2 \right). $$

(15)
The normal equations are

\[
(A^T \alpha L^T) (a) x = (A^T \alpha L^T) (y)
\]

or,

\[
(A^T A + \alpha^2 L^T L) x = A^T y.
\]

The Tikhonov regularized solution can be written from the above equation as

\[
x_{\alpha,L} = (A^T A + \alpha^2 L^T L)^{-1} A^T y.
\]

The solution \(x\) is a function of the regularization parameter \(\alpha\), as well as \(L\). The solution varies with the regularization parameter quite strongly as compared to \(L\) and unfortunately, in the damped least-square approach, the value of \(\alpha\) is not specified. A key element of any Tikhonov regularized formulation is the proper choice of the regularization parameter \(\alpha\). If \(\alpha\) is chosen to be too small, the reconstruction will be dominated by large, high frequency components. If \(\alpha\) is chosen to be too large, important information in the data will be suppressed. The indeterminacy of \(\alpha\) can be eliminated by the method of Generalized Cross Validation (GCV).

Let

\[
G(\alpha) = \frac{||Ax_{\alpha,L} - y||^2}{\text{Trace}(I - AA^T\alpha,L))^2} = \frac{V(\alpha)}{T(\alpha)}.
\]

Here \(V(\alpha)\) measures the misfit. As \(\alpha\) increases, \(V(\alpha)\) also increases. \(T(\alpha)\) is a slowly increasing function of \(\alpha\). A regularization parameter \(\alpha\) is selected, for which \(G(\alpha)\) has a minimum value.

Bhuyan et al. (2002) proposed the determination of the regularization parameter \(\alpha\) using the method of Generalized Cross Validation (GCV). GCV has been used for the purpose of blur estimation and resolution enhancement in image processing (Nguyen et al., 2001). The solution also varies with the form of \(L\), i.e. the order of regularization. But, as the regularization approach tries to find a solution which is vertically as well as horizontally smooth, one has to choose an optimum regularization. In case of second order regularization, the large gradients are heavily penalized (Kamalabadi et al., 2002), thus tending to hide important aspects of the image of a highly disturbed ionosphere. A first order regularization should be a better option in such cases. In this paper, first order regularization has been used in all the cases.

In contrast to the MTSVD technique, the GSVD approach is based on the method of regularization of the problem, which eliminates the need for truncation of the singular values. It is noteworthy that the matrix \(L\) incorporates information regarding the smoothness of the solution. \(Lx\) for first order regularization is a finite difference approximation to the first derivative of \(x\) and for the second order; it is a finite difference approximation to the second derivative of \(x\).

### Table 1. The geometry used for low-latitude ionosphere reconstruction.

<table>
<thead>
<tr>
<th>Number of ray paths</th>
<th>220–250</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude range</td>
<td>100–1000km</td>
</tr>
<tr>
<td>Latitude range</td>
<td>18°N–28°N</td>
</tr>
<tr>
<td>Vertical dimension of one pixel</td>
<td>100 km</td>
</tr>
<tr>
<td>Horizontal dimension of one pixel</td>
<td>55 km (≈0.5° in Latitude)</td>
</tr>
<tr>
<td>Number of receivers</td>
<td>5</td>
</tr>
<tr>
<td>Receiver positions (°N)</td>
<td>20.7, 21.2, 23.4, 25.1, 27.3</td>
</tr>
</tbody>
</table>

### 3 Results

#### 3.1 CIT reconstruction of the low-latitude ionosphere

The well-known equatorial ionization or “Appleton” anomaly (EIA) over Indian low latitudes was investigated by reconstructing the same using the GSVD algorithm developed for this purpose and described in Sect. 2. The ionosphere within the altitude range of 100 km to 1000 km and latitude range of 18° N to 28° N was divided into pixels of 55 km×100 km dimensions (details in Table 1). Bhuyan et al. (2002) used lower resolution 110 km×100 km pixel size for reconstructing the low-latitude ionosphere within a latitude span of 6.5°. The TEC along 220 different ray paths within the region of interest was obtained from the International Reference Ionosphere (IRI-2001). Figure 1a illustrates the electron density distribution for 20 January 2004 at 12:00 LT along 75° E meridian. The reconstructed electron density distribution is plotted in Fig. 1b. The regularization parameter \(\alpha\) has a value of 5.85. It is seen that the equatorial ionization anomaly (EIA) is reconstructed particularly well. The tilt of the ionosphere is also well preserved. Andreeva et al. (2000), Yeh et al. (2001) have shown by radio tomographic investigations that EIA crest can be quite well pronounced. Also, the anomaly core was found to be oriented in the direction of the geomagnetic field. However, the IRI is a climatological model that gives monthly averages of ionospheric parameters under quiet solar activity conditions and since the TEC data used for the present reconstruction of the EIA, is obtained from the IRI, investigation of structural peculiarities of EIA as obtained by tomographic experiments such as the ones mentioned above, could not be undertaken.

#### 3.2 Effect of geometry on reconstruction

The computations carried out with the GSVD-based algorithm require large computer disc space, which, in turn, adversely affects the resolution one tries to achieve. In order to investigate the effect of the geometry on the resolution of the reconstructed image, another geometry with increased number of sampling ray paths was considered. The number of
Fig. 1. (a) Electron Density \( (10^6 \text{ cm}^{-3}) \) distribution obtained from IRI 2001 (b) GSVD reconstruction using 220 ray paths (c) reconstruction with 250 ray paths.

Fig. 2. (a) Model ionosphere with complicated irregularities (electron density in units of \( 10^6 \text{ cm}^{-3} \)) (b) Reconstruction of the model ionosphere of Fig. 2a.

3.3 Reconstruction of complicated irregularities

The ionosphere has a complicated structure where irregularities of different dimensions cohabit the regular structures, such as EIA, mid latitude trough, etc. For the reconstruction of such complicated ionospheric structures, a model ionosphere was first constructed by superimposing three irregular structures on a background ionosphere obtained from the Chapman function, as shown in Fig. 2a. Figure 2b shows the reconstructed ionosphere using the geometry described in Sect. 3.1. All the irregularities preserve their positions in latitude and altitude in the reconstructed image. It should be noted that the regularization approach introduces a smoothening effect in the gradients of electron density (Ka-malabadi et al., 2002). It has been noticed during the simulation of different test cases that in case of a highly disturbed ionosphere, first order regularization is optimum. Second order regularization has a much stronger smoothening effect. The regularization parameter \( \alpha \) has a value of 0.39131 in this case.

3.4 Reconstruction of mid-latitude ionosphere

Earlier simulation studies (Andreeva et al., 1992) have shown that a minimum of three receiving stations could form a good platform for performing reconstruction of the ionosphere from total electron content data. The GSVD algorithm is
further used for reconstruction of the mid-latitude ionosphere along the 145° W meridian in Alaska, USA. The data used has been made available by the International Ionospheric Tomography Community (IITC) through the Internet. Table 2 gives the coordinates of the three receiving stations of LEO satellite signals. Reconstruction has been performed on two days under different geomagnetic conditions. Figures 3a and 3b represent the variation of the relative TEC data with the angle of arrival of the signal at the three receivers on 26 February 2003 at 01:03 UT and 14 July 2003 at 11:03 UT, respectively. The geomagnetic condition of those days as characterized by the $D_{st}$ indices is 4 nT and $-33$ nT respectively. Figure 4a represents the reconstruction for 26 February. The midlatitude trough is reconstructed reasonably well. Figure 4b represents the reconstruction for 14 July. Departures from the average quiet day behavior of the midlatitude ionosphere are quite evident from the reconstruction and it matches with what is expected from Figure 3b. The variation of relative TEC with arrival angle of the signal at all the receivers suggests the presence of such a structure in the ionosphere. It should be noted that the latitudinal span of the two reconstructions is different, owing to the availability of slant TEC data.

### Table 2. Geographical coordinates of the three stations located in Alaska, USA.

<table>
<thead>
<tr>
<th>Receiver Number</th>
<th>Receiving Station</th>
<th>Latitude ° N</th>
<th>Longitude ° W</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cordova</td>
<td>60.495</td>
<td>145.476</td>
</tr>
<tr>
<td>2</td>
<td>Gakona</td>
<td>62.399</td>
<td>145.157</td>
</tr>
<tr>
<td>3</td>
<td>Delta</td>
<td>63.902</td>
<td>145.240</td>
</tr>
</tbody>
</table>

4 Discussion

The formulation of phase-difference tomography makes it free from the phase ambiguity problem. The reconstruction result is affected only by the presence of noise and random recording errors (Andreeva et al., 1992). These errors are averaged out during processing of data and as such, the reconstruction becomes more effective.

A reconstruction error can be characterized by relative errors $\rho_l$ and $\rho_c$ as being discrete analogue for the deviation
of the function being reconstructed ($x$ from the true function $x_{\text{true}}$) in metrics of $l^2$ and c-space. They are defined as

$$
\rho_{l^2} = \left( \frac{\sum (x_{\text{true}}^i - x^i)^2}{\sum (x_{\text{true}}^i)^2} \right)^{\frac{1}{2}}
$$

and

$$
\rho_c = \max_i \frac{(x_{\text{true}}^i - x^i)}{\max_i (x_{\text{true}}^i)}.
$$

The values of the parameters $\rho_{l^2}$ and $\rho_c$ for reconstruction of EIA for the two cases are 0.4251, 0.035 and 0.6148, 0.0093, respectively. In case of reconstruction of complicated irregularities, the respective values are 0.0789 and 0.019. Thus, comparison of $\rho_{l^2}$ and $\rho_c$ for the first two cases shows that while the overall reconstruction error is lower in the first test case compared to the second test case, the reconstruction error of the maximum value of the sought function ($x$) is higher. The low value of both the parameters in the last case points to better reconstruction.

The indeterminacy of the regularization parameter $\alpha$ has been removed by using GCV, which makes the algorithm self-consistent. It must be noted that the GSVD approach requires the number of ray paths to equal or exceed the number of pixels. The algorithm is capable of reconstructing regular as well as complicated structures.

## 5 Conclusion

A radio tomographic algorithm based on the Generalized Singular Value Decomposition technique is used to investigate the regular and irregular structure of the F region ionosphere at equatorial and mid latitudes. Model and measured TEC data are used to reconstruct the equatorial ionization anomaly and the mid-latitude ionosphere, respectively. The method has also been used to reconstruct complicated density irregularities embedded in a simulated regular ionosphere. Results indicate that this GSVD based CIT algorithm is capable of recreating the electron density distribution of the ionosphere satisfactorily. Another advantage of the method has been its ability to reconstruct without the help of any initial guess regarding the state of the ionosphere in the region of interest.

**Acknowledgements.** This work is partially supported by the Indian Space Research Organisation through the grant 10/2/258-NE. One of the authors (KB) is grateful to ISRO for providing him a fellowship under the grant. The authors thank NSSDC, USA for making the IRI 2001 available in the Internet. They are also thankful to the International Ionospheric Tomography Community for making the TEC data available in the Internet.

Topical Editor M. Lester thanks E. J. Fremouw and another referee for their help in evaluating this paper.

**References**


