Simultaneous multiplicative column-normalized method (SMART) for 3-D ionosphere tomography in comparison to other algebraic methods

T. Gerzen and D. Minkwitz
German Aerospace Center (DLR), Institute of Communications and Navigation, Kalkhorstweg 53, 17235 Neustrelitz, Germany

Correspondence to: T. Gerzen (tatjana.gerzen@dlr.de)

Received: 7 August 2015 – Revised: 16 December 2015 – Accepted: 22 December 2015 – Published: 22 January 2016

Abstract. The accuracy and availability of satellite-based applications like GNSS positioning and remote sensing crucially depends on the knowledge of the ionospheric electron density distribution. The tomography of the ionosphere is one of the major tools to provide link specific ionospheric corrections as well as to study and monitor physical processes in the ionosphere.

In this paper, we introduce a simultaneous multiplicative column-normalized method (SMART) for electron density reconstruction. Further, SMART+ is developed by combining SMART with a successive correction method. In this way, a balancing between the measurements of intersected and not intersected voxels is realised. The methods are compared with the well-known algebraic reconstruction techniques ART and SART. All the four methods are applied to reconstruct the 3-D electron density distribution by ingestion of ground-based GNSS TEC data into the NeQuick model.

The comparative case study is implemented over Europe during two periods of the year 2011 covering quiet to disturbed ionospheric conditions. In particular, the performance of the methods is compared in terms of the convergence behaviour and the capability to reproduce sTEC and electron density profiles. For this purpose, independent sTEC data of four IGS stations and electron density profiles of four ionosonde stations are taken as reference. The results indicate that SMART significantly reduces the number of iterations necessary to achieve a predefined accuracy level. Further, SMART+ decreases the median of the absolute sTEC error up to 15, 22, 46 and 67 % compared to SMART, SART, ART and NeQuick respectively.

Keywords. Ionosphere (mid-latitude ionosphere; modelling and forecasting)

1 Introduction

The ionosphere is the upper part of the atmosphere extending from about 50 to 1000 km and going over into the plasmasphere. The characteristic property of the ionosphere is that it contains sufficient free electrons to affect radio wave propagation. The electron density distribution is driven mainly by solar radiation, particle precipitation and charge exchange; it varies widely in both space and time. Thus, real-time determination of the ionospheric electron density distribution becomes important from the satellite applications perspective as well as for understanding ionospheric dynamics. Global Navigation Satellite System (GNSS) observations, which provide the total electron content (TEC) along a receiver-to-satellite ray path, have become one of the major tools for ionospheric sounding.

The ionosphere community carries out several activities that are aimed at describing the ionospheric behaviour by developing electron density models, based on historical GNSS data and other ionospheric measurements. For instance, the International Reference Ionosphere model (IRI; see Bilitza, 2001; Bilitza and Reinisch, 2008) describes empirically monthly averages of the electron density and temperature, based on historical ground- and space-based data. NeQuick (see Nava et al., 2008) is also an empirical model driven mainly by solar activity level and ionospheric F2 layer pa-
rameters, which are computed based on historical vertical sounders data (see ITU-R 1995, Sect. 3.3).

Since those models represent a median ionospheric behaviour, ingestion of actual ionospheric measurements is essential to update them. Several approaches have been developed and validated for ionospheric reconstruction by a combination of actual observations with an empirical or a physical background model. Galkin et al. (2012) present a method to update the IRI coefficients, using vertical sounding observations of a 24 h sliding window. Bust et al. (2004) use a variational data assimilation technique to update the background, combining the observations and the associated data error covariances. Also, other techniques, taking advantage of spatial and temporal covariance information, such as optimal interpolation, Kalman filter and kriging, have been applied (see e.g. Angling and Cannon, 2004; Angling and Khattatov, 2006; Angling et al., 2008; Gerzen et al., 2015; Minkwitz et al., 2015; Pérez, 2005) to update the modelled electron density distributions. Moreover, there are approaches based on physical models that combine the estimation of electron density with physical related variables, such as neutral winds or oxygen/nitrogen ratio (see Schunk et al., 2004; Wang et al., 2004).

In the beginning and even now, when looking for computer resource-saving approaches, algebraic iterative methods have been used to ingest data into background models, e.g. derivatives of the Algebraic Reconstruction Technique (e.g. ART, MART), column-normalized methods (e.g. SART) and the successive correction method (SCM) (see Daley, 1991; Heise et al., 2002; Wen et al., 2007, 2008; Li et al., 2012; Pezopane et al., 2013). Those methods are working without the modification of the model coefficients but by updating the background in the area surrounding the available measurements. Lorenc (1986) discusses the differences in mathematical framework and implementation for the majority of the above-mentioned methods.

In this paper, we introduce a multiplicative column-normalized method, called SMART. Further to this, SMART+ is developed as a combination of SMART and 3-D SCM, assuming a Gaussian covariance model. Both these methods are applied to reconstruct the electron density distribution from the measured ground-based GNSS slant TEC, using the NeQuick model as the background. A comparative study of the SMART and SMART+ approaches, in terms of convergence speed and accuracy, is carried out over the European area with the well-known SART and ART methods. The accuracy is tested by a case study comparing the reconstructed slant TEC values with independent GNSS sTEC and the reconstructed 3-D electron densities with vertical sounding data. The investigated periods cover quiet and disturbed ionospheric conditions within the year 2011.

2 Methods

Information about the total electron content, along the receiver-to-satellite ray path \( s \), can be obtained from the dual-frequency measurements permanently transmitted by the GNSS satellites (see e.g. Jakowski et al., 2011a, b and Sect. 3.4). This measured slant TEC is related to the electron density \( N_e \) by

\[
\text{TEC} = \int N_e (h, \lambda, \varphi) \, ds,
\]

where TEC is the slant TEC measurement in TECU, \( s \) is the ray path along which the corresponding TEC value was measured and \( N_e (h, \lambda, \varphi) \) is the unknown function describing the electron density values depending on altitude \( h \), geographic longitude \( \lambda \) and latitude \( \varphi \).

By discretization of the ionosphere into a 3-D grid and assuming the electron density function to be constant within a fixed voxel, we can transform Eq. (1) to a linear system of equations (LSE):

\[
\text{TEC} \approx \sum_{i=1}^{N_v} N_{e|i} \cdot a_{si} \Rightarrow \mathbf{y} = \mathbf{A}\mathbf{x},
\]

where \( \mathbf{y} \) is the vector of the sTEC measurements, \( N_{e|i} \) is the electron density in the voxel \( i \) and \( a_{si} \) is the length of the ray path \( s \) in the voxel \( i \). An important step therefore is the calculation of the whole ray path and voxel intersection geometry.

In the following chapters, algebraic iterative methods are presented to solve this LSE. All the methods work with an initial guess \( x^0 \) for the unknown electron density vector \( \mathbf{x} \), usually provided by a background electron density model. Within this study, the initial guess is calculated by the NeQuick model (see Sect. 3.3). The complexity of the methods is given by \( O(n^2 \cdot m) \) per iteration step, where \( m \) is the number of observations, and \( n \) the number of voxels.

2.1 ART

Originating with Kaczmarz (1937), the ART algorithm was suggested for medical computerised tomography by Gordon et al. (1970). ART works iteratively, starting with the initial guess \( x^0 \). The \((k+1)\)th iteration step is then given by

\[
x^{k+1} = x^k + c_k \cdot \frac{y_j - (x^k \cdot a_j)}{\sum_{l=1}^{n} a^2_{jl}} \cdot a_j \quad \text{with } j = 1, \ldots, m.
\]

In Eq. (3) \( a_j \) is the \( j \)th row of the matrix \( \mathbf{A} \), \( c_k \) is the relaxation parameter between 0 and 1, \((x^k a_j)\) is the dot product between the estimation of \( \mathbf{x} \) after the \( k \)th iteration and \( a_j \), \( m \) is the number of observations and \( n \) is the number of voxels. The current iterate \( x^k \) is renewed to \( x^{k+1} \) by considering each time just a single ray path \( j \) and changing only the electron density values of the voxels, which are intersected by the ray \( j \). The electron densities of all those voxels, which
are not intersected by a ray path, remain equal to the initial guess. We consider one iteration step of ART as performed, when Eq. (3) is applied to all the available ray paths.

The relaxation parameter $c_k$ plays an important role in practical realisation of algebraic methods, because it helps to overcome the instability problems resulting from measurement errors. When using noisy data the quality of the reconstruction can be improved with the proper choice of $c_k$ (see Kunitsyn and Tereshchenko, 2003; Austen et al., 1988), even when it slows down the convergence speed. In this study, we set the relaxation parameters for ART, SART and SMART to one.

When applied to a consistent LSE, ART was shown to converge to the minimum-norm least-squares solution for relaxation parameters equal to 1. The behaviour of this algorithm for inconsistent systems, when relaxation parameters are allowed, is studied by Censor et al. (1983) and Eggermont et al. (1981). Atkinson and Soria (2007) compare the ART method to other algebraic methods (e.g. with MART, AART, SIRT).

### 2.2 SART

The simultaneous Algebraic Reconstruction Technique (SART) is a kind of refinement of ART towards a column-normalized method. SART is successfully used for tomographic problems (see Andersen and Kak, 1984; Kunitsyn and Tereshchenko, 2003). The $(k+1)$th iteration step for the $i$th voxel is given by

$$x_i^{k+1} = x_i^k + c_k \sum_{r=1}^m \frac{y_j - (x_j^k, a_{ij})}{\sum_{l=1}^n a_{ij}}(a_{ji}),$$

(4)

In the above equation, $x_i^k$ is the estimated electron density in the voxel $i$, after the $k$th iteration. The remaining notation is the same as the one in Eq. (3). Again, only those voxels that are intersected by at least one measurement are innovated. Contrary to ART, SART takes into account all available measurements simultaneously. In the case that all the coefficients of matrix $A$ are non-negative, SART was shown (see Jiang and Wang, 2003) to converge to a solver of the minimisation problem:

$$\min ||y - Ax||_W,$$

where $W = \text{diag} \left( \frac{1}{\sum_{l=1}^n a_{ij}} \right)$. (5)

The use of the weighted mean of the deviations is the major refinement of the column-normalized methods in comparison with the classical row action methods, such as ART, which innovate separately for each ray path.

### 2.3 SMART

In this chapter, the simultaneous multiplicative column-normalized method SMART is introduced. The $(k+1)$th iteration step for $i$th voxel is then given by

$$x_i^{k+1} = x_i^k + \frac{c_k}{\sum_{r=1}^m a_{ri}} \sum_{j=1}^m \left( y_j - (x_j^k, a_{ji}) \right) a_{ji},$$

(6)

with the same notation as in Eq. (4). One iteration of SMART is performed, when Eq. (6) is applied once to all voxels.

Equation (6) can be interpreted as follows: for a voxel $i$, the multiplicative innovation is given by a weighted mean of the ratios between the measurements and the current estimate of the measurements. As for SART, the weights are given by the length of the ray path corresponding to the measurement in the voxel $i$ divided by the sum of lengths of all rays crossing voxel $i$. Thereby, again only voxels intersected by at least one measurement ray path are innovated during the procedure.

Until now we have studied the convergence behaviour of SMART empirically but have not proved the convergence. An advantage of the multiplicative methods, such as SMART, is that they automatically guarantee non-negative estimates of the $x$ components.

### 2.4 SMART+

We developed SMART+ as a combination of SMART and a 3-D successive correction method. First, SMART is applied to distribute the integral measured TEC among the local electron densities in the ray-path intersected voxels. In other words, Eq. (6) is applied for all voxels until the maximum iteration number (chosen as 100 for this work) is reached.

Thereafter, assuming electron densities covariance between the ray path intersected voxels and those not intersected by any TEC ray path, an extrapolation is done from intersected to not intersected voxels. For this purpose, exactly one iteration of a 3-D SCM (see Kalnay, 2011; Daley, 1991) is applied:

$$x_q^{\text{SMART+}} = x_q^0 + \frac{\sum_{i \in \text{[cut voxels]}} N_i \cdot w_{qi} \cdot (x_i^{\text{SMART}} - x_i^0)}{\sum_{i \in \text{[cut voxels]}} N_i \cdot w_{qi} + \varepsilon^2},$$

(7)

where $q$ is an arbitrary voxel number, $x^0$ is the initial guess (calculated in this study also by the NeQuick model), $x^{\text{SMART}}$ is the final estimation of the electron density by SMART, $\varepsilon$ is the estimated ratio between the electron density error variance reconstructed by SMART and the error variance of the background used to calculate $x^0$, and $N_i$ is the number of ray paths intersecting the $i$th voxel. The weights are defined, assuming a Gaussian covariance model for the electron densities, by

$$w_{qi} = \begin{cases} \frac{1}{\sqrt{2\pi}(\text{dist}_{\lambda,\phi}(q,i))^2 + (\text{dist}_{\lambda,\phi}(q,i))^2}}, & \text{for } \text{dist}_{\lambda,\phi}(q,i) \leq \text{RAD} \\ 0, & \text{for } \text{dist}_{\lambda,\phi}(q,i) > \text{RAD}. \end{cases}$$

(8)
In the above formula, $\text{dist}_{\lambda, \phi}(q, i)$ denotes the horizontal great circle distance between voxels with numbers $q$ and $i$; $\text{cor}^2_{\lambda, \phi}$ the square of the horizontal correlation length between the voxels; $\text{RAD}$ the influence radius of a voxel on its neighbourhood and $L_h$ the vertical correlation length, which increases with increasing altitude. Compared to the classical formulation of the SCM method (see Kalnay, 2011; Daley, 1991), in our application not the measurements (TEC) itself, but the SMART reconstructed electron densities in the intersected voxels, act as observations.

The horizontal and especially the vertical correlation lengths of the ionospheric electron densities have not been completely known until now. In the algorithm developed here, these key parameters are chosen empirically. But, we are currently working on methods that facilitate a better estimation of the correlation lengths for the 3-D electron densities (see Minkwitz et al., 2015). The ratio of error variances $\varepsilon$ is chosen as 0.5 (see e.g. Gerzen et al., 2015); the horizontal correlation length $\text{cor}^2_{\lambda, \phi}$ as $4^\circ$; the vertical correlation length starts from $30$ km in the E and F regions and gradually increases to $500$ km in the regions above the topside ionosphere (see e.g. Bust et al., 2004 and the references there); $\text{RAD}$ is chosen as $20^\circ$.

3 Tomography setup

The methods outlined here are developed and tested for the ionosphere tomography during two contrasting periods of the year 2011, one with quiet and the other with disturbed ionospheric conditions. In the following sections the chosen periods and reconstruction data base are described in detail.

3.1 Periods

Two periods of the year 2011 are selected for assimilation and case studies: DOYs 009–022 (January) and 294–298 (21–25 October). The geomagnetic and solar activities during these two periods are indicated in Fig. 1. The right-hand panel shows the global planetary 3 h index Kp as a measure of geomagnetic activity. The left-hand panel presents the variation of the solar radio flux at $10.7$ cm wave length ($F10.7$ index), which serves as indicator of solar activity.

For a fuller understanding of the temporal evolution of the ionosphere, the panels cover not just the targeted periods, but also a few more days before and after these periods. The days investigated within this study are displayed in bold font. The data have been acquired from the Space Physics Interactive Data Resource of NOAA’s National Geophysical Data Center (SPIDR) and the World Data Center for Geomagnetism (WDC) Kyoto. According to Suard et al. (2011), the ionospheric conditions can be assessed as quiet during DOY 009–022 (blue line) and as disturbed during DOY 294–298 (black line) with $F10.7$ between 130 and 170 and a severe geomagnetic storm on DOY 297–298 with a Kp index above 7. Also the geomagnetic index DST indicates a geomagnetic storm during the night from DOY 297 to DOY 298, with DST values below $-130$ nT.

3.2 Reconstructed area

In this study we apply the described methods to reconstruct the electron density in the extended European region covering the geographic latitudes $-90$ to $90^\circ$ N and $-100$ to $110^\circ$ E. The spatial resolution is $2.5^\circ$ along both latitude and longitude. The altitude resolution is $30$ km for altitudes from $60$ to $1000$ km and decreases exponentially with increasing altitude above $1000$ km altitude. In total, we get 54 altitude steps and thus 326,592 unknowns in Eq. (2). The time resolution is set to $30$ min.

3.3 Background model

To regularise the inverse problem in Eq. (2), the initial guess for the ionosphere tomography by algebraic methods is usually calculated by a background model. Therefor an arbitrary electron density model can be deployed. In this study we apply the three-dimensional NeQuick model version 2.0.2, released in November 2010.

The NeQuick model was developed at the International Centre for Theoretical Physics (ICTP) in Trieste/Italy and at the University of Graz/Austria (see Hochegger et al., 2000; Radicella and Leitinger, 2001; Nava et al., 2008). The vertical electron density profiles are modelled by parameters such as peak ionisation, peak height and semi-thickness, deduced from the ITU-R models (see ITU-R, 1995). We use the daily $F10.7$ index to drive the NeQuick model.

3.4 TEC measurements

As mentioned in Sect. 2, we use the ground-based absolute sTEC as input for the tomography approaches and also for the validation. The unambiguous relative sTEC is derived by the combination of GPS dual-frequency carrier-phase and code-pseudorange measurements. Then, the absolute sTEC and the receiver and satellite inter-frequency biases are separated by a model-assisted technique. The Neustrelitz TEC model (Jakowski et al., 2011a), together with a single-layer mapping function (assuming shell height of $400$ km), is applied for the calibration procedure. For more details, we refer to Jakowski et al. (2011b). For this study, the GNSS data of the global International GNSS Service (IGS) $1$ s high rate receiver network were acquired via ftp://cddis.gsfc.nasa.gov/pub/gps/data/highrate. Within the calibration procedure only the receiver-satellite link geometries with elevation angles not less than $20^\circ$ are used in order to avoid the usage of observations with multipath (see e.g. Yuan et al., 2008a, b). Thereafter, the geometry of the data is checked and only those calibrated sTEC measurements whose ray paths lie within the described reconstructed area (see Sect. 3.2) are used in the next step for reconstruction. For validation purposes, the cal-
Figure 1. The solar radio flux index F10.7 (left panel) and the global planetary 3 h index Kp (right panel) for the periods DOY 000–039/2011 (blue) and DOY 282–319/2011 (black). The investigated periods are highlighted in bold font.

Table 1. The independent IGS stations used for validation purposes.

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Country</th>
<th>Lat. (° N)</th>
<th>Long. (° E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ffmj</td>
<td>Germany</td>
<td>50.09</td>
<td>8.67</td>
</tr>
<tr>
<td>pado</td>
<td>Italy</td>
<td>45.41</td>
<td>11.90</td>
</tr>
<tr>
<td>ajac</td>
<td>France</td>
<td>41.93</td>
<td>8.76</td>
</tr>
<tr>
<td>mas1</td>
<td>Spain</td>
<td>27.76</td>
<td>−15.63</td>
</tr>
</tbody>
</table>

ibrated sTEC data of the four IGS stations, listed in Table 1, are excluded from the reconstruction procedure.

For one reconstruction epoch, the available sTEC data are collected within a 10 min interval and averaged regarding the ray path geometry. On average, around 80–90 stations and 600–700 averaged sTEC measurements become available in the reconstructed area. Comparing the measurement number with the number of unknowns in Eq. (2), we get a strongly underdetermined inverse problem with extremely limited angle geometry (see also Garcia and Crespon, 2008). Therefore, to regularise this inverse problem, we decided to use the corresponding vertical vTEC data, in addition to the sTEC measurements. Indeed, we validated the investigated tomography methods also without the additional use of vTEC values (i.e. assimilating only the ground-based sTEC) and detected a slight increase of the residuals statistics. This motivated us to concentrate on the results of assimilation, where both slant and vertical TEC is applied.

4 Results of the case study

This section is organised as follows: first, we present, by way of an example, the reconstructed 3-D electron density. Subsequently, the investigated tomography methods are validated for the two periods of the year 2011 by comparing the following:

1. the convergence behaviour;
2. the ability to reproduce the assimilated TEC;
3. the reconstructed sTEC with independent ground-based sTEC data;
4. the reconstructed electron densities with ionosonde electron density profiles.

The results obtained by different methods are colour-coded as follows: NeQuick model, orange; ART, light blue; SART, blue; SMART, red; SMART+, green.

4.1 Reconstructed electron densities

Figure 2 presents the 3-D electron densities at different altitudes for DOY 009/2011 at 12:00 UTC. The top panel is calculated using the pure NeQuick model; the middle panel depicts the SMART reconstruction and the bottom panel the SMART+ reconstruction.

The figures deduced from SART and ART are similar to those deduced from SMART and hence are not presented here. It is notable that the SMART result is rather patchy, which is usual for locally working reconstruction methods applied to sparse, unevenly distributed data. The application of 3-D SCM within SMART+ manages a balance between the neighbouring voxels.

4.2 Comparison in terms of the convergence behaviour

To compare the convergence behaviour of the investigated methods, we count the number of iterations needed by the
Figure 2. Electron density layers calculated by the NeQuick model (top), SMART (middle row) and SMART+ (bottom).

methods ART, SART and SMART to achieve a predefined threshold of ΔTEC, the mean deviation between the measured TEC values used for the reconstruction and the corresponding reconstructed TEC values, calculated by

\[
\Delta \text{TEC}_k^j = \frac{1}{m} \sum_{j=1}^{m} \left( | \text{TEC}^{\text{measured}}_j - \text{TEC}^{\text{reconstructed},k}_j | \right).
\]

(9)

In the above equation, \( k \) is the iteration step; \( m \) is the number of available measurements; \( \text{TEC}^{\text{measured}}_j \) is the \( j \)th measured TEC (sTEC or vTEC, used for reconstruction) and \( \text{TEC}^{\text{reconstructed},k}_j \) the corresponding TEC along the \( j \)th measurement ray path calculated from the reconstructed 3-D electron density distribution after \( k \) iterations. As accuracy threshold we set \( \Delta \text{TEC} \) equal to 1.5 TECU. In terms of the notation used for Eqs. (2)–(6), the above formula for the mean deviation can be stated as

\[
\Delta \text{TEC}_k^j = \frac{1}{m} \sum_{j=1}^{m} \left( | y_j - \langle x^k, a_j \rangle | \right).
\]

(10)

Figure 3 shows the decrease in the mean deviation \( \Delta \text{TEC}_k^j \) for ART, SART and SMART methods in dependency on the iteration step for the DOY 009 at 12:00 UTC. In general, the decrease is much faster for the SMART method: already after five iteration steps, \( \Delta \text{TEC} \) is around 0.5 TECU. After 100 SMART iterations, the \( \Delta \text{TEC}^{100} \) value is around 0.33 TECU. However, the subsequent realisation of 3-D SCM (see Eq. 7) within the SMART+ method introduces a \( \Delta \text{TEC} \) increase from 0.33 to 0.94 TECU.

The left-hand panel of Fig. 4 illustrates the number of iterations \( k \) needed to achieve the threshold \( \Delta \text{TEC}_k^j \leq 1.5 \) TECU for the quiet period. We see that SMART needs, in all epochs, the least number of iterations to achieve the threshold. Additionally, it is observable that, whereas with SART and SMART it was possible to reach the threshold of 1.5 TECU at all epochs, this is not the case with ART, as indicated by interruptions in the light blue curve.

The right-hand panel of the same figure illustrates that, during the disturbed period, ART could not reach the thresh-
Figure 4. Number of iterations needed to achieve $\Delta$TEC $\leq$ 1.5 TECU applying the methods ART (light blue), SART (blue dashed) and SMART (red); left-hand panel is during the quiet period; right-hand panel is during the disturbed period.

Figure 5. $\Delta$TEC$^{100}$ for ART (light blue stars), SART (blue points) and SMART (red pluses). Left-hand panel during the quiet period; right-hand panel during the disturbed period.

old at any epoch. Using SART, the threshold is achieved only for approximately half of the processed epochs and even SMART has four short time intervals missing the threshold.

4.3 Plausibility check by comparison with assimilated TEC

When applying the methods ART, SART and SMART to the LSE (2), the iteration process is stopped after performing 100 iteration steps. To check how well the methods work, we consider the mean deviation $\Delta$TEC$^{100}$, reached after 100 iteration steps. This allows us to assess the ability of the methods to reproduce the assimilated TEC.

Additionally, we look at the percentage reduction of the mean deviation achieved by the tomography methods after 100 iteration steps, in comparison to $\Delta$TEC$^0$:

$$\text{Improvement} = \frac{\Delta\text{TEC}^0 - \Delta\text{TEC}^{100}}{\Delta\text{TEC}^0} \times 100\%.$$  

(11)

Thereby, $\Delta$TEC$^0$ values are calculated from the NeQuick model providing the initial guess $x^0$:

$$\Delta\text{TEC}^0 = \frac{1}{m} \sum_{j=1}^{m} \left( |y_j - \langle x^0, a_j \rangle | \right)$$  

(12)

The $\Delta$TEC$^{100}$ values of ART, SART and SMART are depicted in Fig. 5 against the minutes of the considered periods. The performance of SMART is found to be the best, followed by SART and then ART.

As expected, the subsequent application (after 100 iterations with SMART) of the 3-D SCM method within SMART+, in general increases the $\Delta$TEC values. A comparison of the mean TEC deviation values for SMART (the same as in Fig. 5) and those of SMART+ is shown in Fig. 6. Also presented are the $\Delta$TEC$^0$ values deduced from the NeQuick model. For the quiet period $\Delta$TEC$^0$ ranges between 2 and 5 TECU. SMART could reduce the deviation to values around 0.5 TECU. SMART+ provides mean TEC deviation values of around 1 TECU.

At the beginning of the disturbed period high $\Delta$TEC$^0$ values are produced by NeQuick. This induces of course
corresponding high values of $\Delta$TEC$^{100}$ for all tomography methods (see Figs. 5 and 6, right-hand panels). Notably, the peak of $\Delta$TEC occurs abruptly and then decreases gradually. We assume that these high values are caused by some outliers still present in the assimilated TEC data regardless of the TEC data filtering applied. This assumption might explain the abrupt appearance of the peak. The subsequent slow decrease of $\Delta$TEC$^{100}$ and $\Delta$TEC$^0$ is probably due to the receiver and station hardware calibration of the GNSS ground-based sTEC measurements, performed according to the method in Jakowski et al. (2011b). The calibration process applies a time forecasting method leading to the slow vanishing of possible errors in the sTEC calculation.

Except for the two $\Delta$TEC peaks at the beginning of the disturbed period, all the other $\Delta$TEC values are below 15 TECU. Therefore, Fig. 7 presents only the values below 15 TECU, so that distinguishing between the tomography methods becomes easier. We observe that, for the disturbed period, $\Delta$TEC$^0$ values range between 4 and 15 TECU, ART $\Delta$TEC$^{100}$ values between 2 and 10 TECU, SART $\Delta$TEC$^{100}$ values between 0.5 and 3, SMART between 0.24 and 2.1 TECU and SMART+$\Delta$TEC$^{100}$ values between 1–3.8 TECU.

Figure 8 displays the extent of $\Delta$TEC reduction achieved by the methods ART, SART and SMART after 100 iterations in relation to $\Delta$TEC$^0$ values (see Eq. 11). During the quiet period (left-hand panel) ART decreases the mean deviation by $\approx 35\%$, SART by $\approx 75\%$ and SMART by $\approx 90\%$. Similarly, during the disturbed period (right-hand panel) ART reduces the mean deviation values by $\approx 40\%$, SART by $\approx 80\%$ and SMART by $\approx 90\%$.

### 4.4 Validation with independent ground-based sTEC data

The reconstruction outcomes are highly dependent on the quality and availability of data and on the accuracy of the background. Therefore, for this first comparison, we concentrate on the European region covering the geographic latitudes 20 to 60° N and −20 to 30° E, because the availability of the IGS stations in this region is relatively dense. Further,
the performance of the NeQuick model is expected to be better for mid-latitude regions.

For validating the outlined methods regarding their capability to estimate independent sTEC, four IGS stations are chosen. They are listed in Table 1. These stations are not used for tomography. For each station, the measured sTEC (namely sTEC\text{measured}) is compared to the reconstructed sTEC (sTEC\text{reconstructed}) derived from the reconstructed 3-D electron densities according to the measurements geometry. Additionally, the sTEC of the NeQuick model, sTEC\text{model}, is analysed to assess the background model errors.

For each IGS validation station, the residuals between the reconstructed values and the measured TEC values are calculated as \( \Delta \text{TEC} = \text{sTEC}\text{measured} - \text{sTEC}\text{reconstructed} \). Further, the absolute values of the residuals (|\( \Delta \text{TEC} \)|) and the relative residuals (\( \frac{\Delta \text{TEC}}{\text{sTEC}\text{measured}} \times 100\% \)) are considered. Equivalent to these, the NeQuick model residuals \( \text{sTEC}\text{measured} - \text{sTEC}\text{model} \) are computed.

Figure 9 depicts the sTEC\text{measured} values for the southernmost validation station mas1 (27.76° N, −15.63° E), for the quiet period in the left-hand panel and for the disturbed period in the right-hand panel. The figure gives an impression of the magnitude of the compared values. For the quiet period, the sTEC measurements vary between 0 and 90, and for the second period between 0 and 300 TECU.

Figure 10 displays the histograms of the sTEC residuals during the quiet period for the four reference stations, from top (north) to bottom (south): ffmj, pado, ajac, mas1. The distribution of the relative residuals for the methods SART, SMART and SMART+ at the stations ffmj, pado and ajac is almost symmetric. But this is not so in the case of ART residuals, which broadly follow the NeQuick residuals distribution. At the southernmost station mas1 the distributions of all residuals show the wide spread of the deviations. Ignoring the IGS station mas1, almost all the statistics of the absolute residuals decrease in magnitude from north to south. The mas1 station shows the highest medians regarding both the relative and the absolute residuals; whereas the northernmost station ffmj (followed by mas1) shows the highest RMS and SD values of the relative residuals. This behaviour of the mas1 residuals is caused most probably by the location of the station within the ionospheric equatorial crest region. The
Figure 10. Histograms of the absolute (left-hand panel) and relative (right-hand panel) sTEC residuals during the quiet period. For IGS validation stations from top to bottom: ffmj, pado, ajac, mas1 (north to south).
Figure 11. Histograms of the absolute (left-hand panel) sTEC residuals and sTEC residuals (right-hand panel) over all the four validation stations during the quiet period.

Table 2. The ionosonde stations used for validation purposes.

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Country</th>
<th>Lat. (°N)</th>
<th>Long. (°E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JR055</td>
<td>Germany</td>
<td>54.6</td>
<td>13.4</td>
</tr>
<tr>
<td>DB049</td>
<td>Belgium</td>
<td>50.1</td>
<td>4.6</td>
</tr>
<tr>
<td>EB040</td>
<td>Spain</td>
<td>40.8</td>
<td>0.5</td>
</tr>
<tr>
<td>GM037</td>
<td>Italy</td>
<td>37.6</td>
<td>14.0</td>
</tr>
</tbody>
</table>

high RMS and SD values of the ffmj relative residuals are probably explainable by the relative low sTEC values at this high-latitude station. Thus, the errors, possibly still present in the reference observations as well as in the assimilated observations, might outweigh the statistics.

The general behaviour of the residuals during the disturbed period is very similar to that during the quiet period. Thus, we just present the corresponding statistics of the absolute residuals in Table 3.

During both periods at all stations, the NeQuick model seems to overestimate the sTEC values visible in the negative relative residuals. A similar overestimation was observed by Nigussie et al. (2012). The authors assimilated the GNSS ground-based sTEC data into the NeQuick model with an alternative least square approach. Afterwards, the results obtained before and after the assimilation were compared with GNSS sTEC of four independent ground-based stations located in East Africa. They even detected a higher level of overestimation by the pure NeQuick model. This higher level can be explained probably by the low-latitude locations of the therein chosen validation stations and by the 10° elevation mask – lower than the 20° mask used in our study.

For the quiet period, the medians of the NeQuick relative residuals range between −38.2% at the southernmost mas1 station and −7.0% at the pado station. For the disturbed period, the values are between −21.8% (at mas1) and −15.6% (at pado). All the tested tomography methods succeed in reducing this overestimation. Regarding the relative residuals, SMART+ performs the job best, followed by the SMART method. For the quiet period, the median values of the SMART+ relative residuals range between −24.6% (at the mas1 station) and 0.1% (at the ajac station). The disturbed period SMART+ median values range between −11.1% (at mas1) and 1.1% (at pado).

During both periods, all the compared tomography methods could significantly decrease all the statistics of the absolute residuals at each validation station, as compared to the corresponding background values. Again, at each station, the reduction achieved by the SMART+ method is the highest,
followed by that achieved by SMART. SMART+ reduces the background absolute median value by up to 71\% during the quiet period (see pado station) and by up to 77\% during the storm period (see ajac station); the RMS value is decreased by up to 70\% during the quiet and up to 24\% during the storm period.

Figures 11 and 12 present the histograms and the statistics of the residuals and absolute residuals over all the four validation stations for the quiet and the storm period, respectively. Notably, the NeQuick model, once again overestimates the sTEC values. The overall statistics confirm that the performance of SMART+ is the best, followed by that of SMART, SART and ART, in that descending order. The RMS of SMART+ is about 3.48 TECU for the quiet period. This corresponds to a range error of about 0.56 m on the GPS L1 frequency. In comparison, Yuan et al. (2008a) developed a method to update the Klobuchar model coefficients by GPS observations. The approach was validated with independent ionospheric delays during the quiet period of 1–8 January 2001. As result, a RMS of 1.96 m was obtained over all stations and days. This points out the potential advantage of 3-D reconstructions over simple single-layer models concerning the accuracy of positioning.

A comparison of the overall statistics of quiet and storm conditions shows an increase of $d\text{TEC}$ and $|d\text{TEC}|$ values for NeQuick and all tomography methods. The median of the NeQuick absolute residuals increases by around 196\%, from 2.49 to 7.38 TECU. The SMART+ median increases by around 150\% (0.98 TECU for the quiet period and 2.45 TECU for the disturbed period) and the ART median by around 165\%. Considering the RMS values, the increase is up to 345\% for NeQuick, 597\% for SMART+ and around 450\% for SMART, SART and ART.

### 4.5 Validation with independent vertical sounding data

In this section the investigated tomography methods are compared in terms of their capability to estimate the vertical electron density profiles. Therefore, the 3-D reconstructions are validated with vertical sounding data of four ionosonde stations, listed in Table 2. The ionosonde profiles of these ionosondes are downloaded from SPIDR.

According to the ionosonde locations, the electron density profiles are deduced from the 3-D electron density reconstructions. Since the reconstructions are calculated with resolution of 30 km altitude (below 1000 km height), the derivation of the F2 layer characteristics, $NmF2$ and $hmF2$, from the reconstructed profiles would be inaccurate. But, computation of reconstructions with a higher altitude resolution would increase the computation time significantly. At the same time, anticipating the comparison results presented below, at the present state, we do not expect a better understanding of the performance of tomography methods in reconstructing the vertical behaviour of the ionosphere, based on comparisons of the reconstructed and ionosonde profiles, in terms of $NmF2$ and $hmF2$ values.

Thus, instead of comparing the profiles in terms of $NmF2$ and $hmF2$, we decided to analyse the residuals $dNe = Ne_{\text{measured}} - Ne_{\text{reconstructed}}$ and the relative residuals: $\frac{dNe}{Ne_{\text{measured}} \cdot 100\%}$ for each reconstruction altitude separately. The $dNe$ values are calculated for each altitude step of a reconstructed electron density profile up to the measured $F2$ layer peak height, $hmF2$, of the corresponding ionosonde. The ionosonde electron density profiles are usually provided also for several altitudes above the $hmF2$ value, but the quality of these values is considered rather poor (see e.g. McNamara, 2006; Davies, 1990). In the same way, also the residuals of the NeQuick model, $Ne_{\text{measured}} - Ne_{\text{model}}$, are considered.

Figure 13 presents the profiles for the DOY 009/2011 at 12:00 UTC at the four ionosondes. The different methods are colour-coded as has been done in the figures of foregoing sections. For Fig. 13, the reconstructed profiles are interpolated by the piecewise cubic Hermite interpolation to the higher resolution of the ionosonde profiles (usually 10 km).
At JR055 and DB049, the electron density values of the NeQuick model, and also of all the methods being compared, are smaller than the ionosonde measurements for altitudes below 180 km. Above this altitude, the electron density values of the NeQuick model, ART and SART at the DB049 station are higher than the ionosonde values, whereas those of SMART and SMART+ are lower. At JR055, the reconstructed density values around the F2 layer peak move even further away from the ionosonde values than the background.

At EB040, the E and F2 layer peak heights given by the NeQuick model are completely different from the ionosonde values. As a result, the NeQuick modelled electron densities and consequently all the reconstructed electron densities are smaller than those of the ionosonde at all altitudes.

At GM037, the estimated E layer peak height and also the densities, estimated by the investigated methods below 200 km altitude, match the measured values. The ionosonde profile of this station is provided with 1 km altitude resolution, but seems unsmoothed above the 200 km altitude.

Figure 14 points out the results of the comparison between the profiles. The altitude-dependent median values of the relative residuals at DB049 station are shown. During both periods, the electron densities estimated by the NeQuick at the lower altitudes from 120 to 180 km, are significantly lower than those of the ionosonde station. This can be explained most probably due to the difference in the estimation of the ionospheric layers peak heights (especially the E-layer seems to be problematic), which causes different shapes of the model and ionosonde electron density profiles.

To elaborate this point further, attention is invited to Fig. 15, which shows the $h_{m}F2$ values, measured by the ionosonde station DB049 in magenta colour and those calculated by the NeQuick model as orange coloured diamonds. During the disturbed period, NeQuick seems to overestimate the F2 layer peak height, except for the storm night between DOY 297 to 298. On this night, the differences between $h_{m}F2_{measured}$ and $h_{m}F2_{model}$ values are up to 150 km. For the quiet period, the differences are up to −40 km during the daytime and up to 70 km during the night-time. Such major discrepancies will inevitably lead to different estimations of the profile shape and thus, to huge differences between the estimated electron densities at the corresponding altitudes. Especially, when assimilating only ground-based GNSS sTEC data, the estimation of the vertical shape of the profile, particularly the ionospheric layer characteristics becomes a difficult task, because of limited vertical information in these data (see e.g. Minkwitz et al., 2015; McNamara et al., 2008, 2011).

It is important to realise here that such huge deviations between ionosonde and modelled profiles could be induced,
at least partly, by the inaccuracy of the ionosonde profiles themselves (see McNamara, 2006; Gerzen et al., 2015). McNamara (2006) addresses this topic in a comprehensive way, especially by pointing out the weakness in electron density estimation between E and F layers and in determination of layers height.

Because of these reasons, we restrict our further comparison to the area that usually provides the most reliable ionosonde data for altitudes ranging from 210 km to altitudes that are just above the corresponding ionosonde $hmF2$ value. The results are presented in Fig. 16 for the quiet period and in Fig. 17 for the disturbed period. The left column panels depict the 90 % bound of the $|dNe|$ values, and the right column panels the median values of the relative residuals. The 90 % bound values are computed, by sorting of the $|dNe|$ values and calculating the nine-tenths bound.

For the quiet period, the low median at 210 km altitude for JR055 is conspicuous. Probably, the 210 km altitude cut-off used is too low for this station, and thus we observe a similar behaviour as at DB049 in Fig. 14. Regarding the median values during the quiet period, SMART and SMART+ perform best at DB049 and worst at EB040 and JR055 at almost all altitudes. On average, the NeQuick model seems to underestimate the electron density at DB040, at all altitudes, and at JR055, EB040 and GM037 at lower altitudes (below 270–330 km, depending on the location). Considering the comparison presented in Fig. 15, this is most probably caused by the discrepancies between the ionosonde and the NeQuick estimations of the layers’ peak heights. Therefore, the subsequent results should also be considered very carefully. For the investigated periods, the tomography methods tend to increase the background $N_e$ values (contrary to the epoch depicted in Fig. 13) at almost all the chosen altitudes, at all four stations.

Regarding the 90 % bound, again during the quiet period, the results provided by ART and SART methods are very similar to those provided by NeQuick, except for the 210 km altitude. The bounds for SMART and SMART+ are even higher than the bound of NeQuick for altitudes below 330 km at EB040 and JR055. At all stations, except the southernmost GM037, SMART+ provides the lowest bound for the higher altitudes (above 300–390 km, depending on the station). The
Figure 16. The 90% values (left) and the medians (right) of the relative residuals versus the corresponding altitude during the quiet period. For the validation ionosonde stations from top to bottom: JR055, DB049, EB040, GM037 (north to south).
Figure 17. The 90% values (left) and the medians (right) of the relative residuals versus the corresponding altitude during the disturbed period.
90% bound results at GM037 are similar for the different methods, and by far the highest compared to other stations, caused by the low-latitude location of GM037.

For the disturbed period at DB040 and JR055, the behaviour of the median and the 90% bound for the tomography methods is similar to the NeQuick values. Also at EB040, it is hard to name any method as the best-performing one, because the methods perform differently at different altitudes. The high negative median values obtained by NeQuick, SART and ART at altitudes between 330 and 390 km are conspicuous. The sharp increase of the 90% bound for the SART, SMART and SMART+ methods at altitudes 330 and 360 km, especially when compared to the lower NeQuick values in this interval, is also notable. At GM037, at altitudes between 330 and 390 km, the negative median values produced by NeQuick and ART are even higher than at EB040. At this station SMART and SMART+ perform best regarding the median, but show the highest 90% values between 270 and 330 km.

From a comparison of the statistics of the quiet and the disturbed periods (especially at the lower altitudes), a significant increase in the 90% bound becomes visible for the residuals of NeQuick and all tomography methods at all stations, except GM037. At the low-latitude station GM037, the behaviour of the 90% bound differs for the quiet and disturbed periods in dependence on the altitude: during the quiet period, very high values of the 90% bound are obtained at altitudes above 360 km. In contrast, during the disturbed period, the highest bound values are obtained below 360 km, which is similar to the 90% bound behaviour at the other ionosondes. The increase in the ionosonde hmF2 values at the ionosondes DB049 and EB040, expressed in the presence of higher altitudes in Fig. 17, as compared to those in Fig. 16, is also noteworthy.

5 Conclusions

In the present work, our main goal has been to introduce the algebraic tomography methods SMART and SMART+ and to set the performance of them into the context of the well-known methods ART and SART carrying out the first case study in this regard.

The SMART method shows the best performance, in terms of convergence speed, especially visible during the storm period, followed by SART and ART. The reduction in the mean TEC deviation achieved by SMART, SART and ART, after 100 iterations, in comparison to the background (NeQuick model) initial mean deviation, is up to 90, 85 and 40 % respectively.

For the purpose of validation, we selected sTEC GNSS observations of four independent ground-based IGS stations and the vertical electron density profiles of four ionosonde stations in the European region. Two periods within the year 2011, one with quiet ionospheric conditions and the other with disturbed conditions, were investigated.

In summary, comparison of the sTEC results of this case study reveals that all the investigated tomography methods improve the background. During both periods and at each validation station, all the four methods could successfully reduce the median, RMS and SD values of the absolute sTEC residuals, in comparison to the background values.

SMART+ gave the best performance, decreasing the overall median, RMS and SD of |dTEC| (compared to the corresponding background values) by up to 67, 18 and 12 % respectively during the quiet period (see Fig. 11) and up to 61, 37 and 28 % respectively during the disturbed period (see Fig. 12). SMART is the second best method: for the quiet period, the differences between the SMART and SMART+ median, RMS and SD values are around 0.4, 1.7 and 1.2 TECU respectively. For storm days, the differences are around 0.2, 0.8 and 1 TECU respectively. The performance of SART is inferior to that of SMART.

The first validation with vertical sounding data reveals, on the one hand, the difficulties involved in correct characterisation of the electron density profile shapes, when only ground-based TEC is used for tomography. This is in agreement with the results deduced by similar studies (see e.g. Minkwitz et al. 2015; McNamara et al., 2008, 2011). On the other hand, the validation emphasises the need for careful treatment and filtering of ionosonde profile data. Here, bigger discrepancies between the background estimated and the true (or ionosonde) ionospheric layer heights cause significant differences in the electron density profile shape estimation and thus, huge differences between the modelled and true (or ionosonde) electron densities at the same altitudes. This problem seems to be difficult to solve by mere ingestion of ground-based data.

To get comprehensive 3-D reconstructions in the future, the step of assimilation of data providing more information about the vertical distribution, like ionosonde profiles and ionospheric radio occultation profiles, may prove important and promising (see McNamara et al., 2007; Angling 2008). Moreover, in order to improve the data coverage and measurement geometry we will assimilate space-based GNSS sTEC and further ground-based sTEC measurements available due to the recent development and modernisation of the different GNSS (e.g. BDS, Galileo and GLONASS – see e.g. Li et al., 2012, 2015). Further, adjustment of the background in terms of F2 layer characteristics (because the F2 layer dominates the shape of the whole profile) before starting the assimilation procedure seems to be helpful (see e.g. Bidaire and Warnant, 2010). In this context, because of the limitations of ionosonde profile estimation, filtering of data is a further important topic (see e.g. McNamara, 2006 and Gerzen et al., 2015).

Additionally, we are currently working on methods that enable better estimation of the correlation lengths and error bounds for the 3-D electron densities (see Minkwitz et
This information can be used to improve upon the SMART+ method by adopting the same approach as the one applied for the 2-D modified SCM by Gerzen et al. (2015).

Acknowledgements. We thank the IGS (ftp://cddis.gsfc.nasa.gov/pub/gps/data/highrate) for making available GNSS data. We also thank the SPIDR (http://spidr.ngdc.noaa.gov/spidr/) and WDC Kyoto (http://wdc.kugi.kyoto-u.ac.jp/wdc/Sec3.html) for making the ionosonde and geo-related data available. Additionally, we express our gratitude to the Aeronomy and Radiopropagation Laboratory of the Abdus Salam International Centre for Theoretical Physics Trieste/Italy for providing the NeQuick version 2.0.2 for scientific purposes. We are grateful to our colleagues from DLR IKN, especiallyVolker Wilken, for the useful discussions we had with them and for calibration of the GNSS data.

The article processing charges for this open-access publication were covered by a Research Centre of the Helmholtz Association.

The topical editor, K. Hosokawa, thanks Y. Yunbin and one anonymous referee for help in evaluating this paper.

References


