Limited constraint, robust Kalman filtering for GNSS troposphere tomography

W. Rohm¹,², K. Zhang², and J. Bosy¹

¹Institute of Geodesy and Geoinformatics, Wroclaw University of Environmental and Life Sciences, Wroclaw, Poland
²Satellite Positioning for Atmosphere, Climate and Environment (SPACE) Research Centre, RMIT University, Melbourne, VIC, Australia

Correspondence to: W. Rohm (witold.rohm@up.wroc.pl)

Received: 14 August 2013 – Published in Atmos. Meas. Tech. Discuss.: 24 October 2013
Revised: 24 March 2014 – Accepted: 11 April 2014 – Published: 27 May 2014

Abstract. The mesoscale variability of water vapour (WV) in the troposphere is a highly complex phenomenon and modelling and monitoring the WV distribution is a very important but challenging task. Any observation technique that can reliably provide WV distribution is essential for both monitoring and predicting weather. The global navigation satellite system (GNSS) tomography technique is a powerful tool that builds upon the critical ground-based GNSS infrastructure (e.g. Continuous Operating Reference Station – CORS – networks) that can be used to sense the amount of WV. Previous research shows that the 3-D WV field from GNSS tomography has an uncertainty of 1 hPa. However, all the models used in GNSS tomography heavily rely on a priori information and constraints from non-GNSS measurements. In this study, 3-D GNSS tomography models are investigated based on a limited constrained approach – i.e. horizontal and vertical correlations between voxels were not introduced, instead various a priori information were added into the system. A case study is designed and the results show that proposed solutions are feasible by using a robust Kalman filtering technique and effective removal of linearly dependent observations and parameters. Discrepancies between reference wet refractivity data derived from the Australian Numerical Weather Prediction (NWP) model (ACCESS) and the GNSS tomography model using both simulated and real data are 4.2 ppm (mm km⁻¹) and 6.2 ppm (mm km⁻¹), respectively, which are essentially in the same order of accuracy.

1 Introduction

The distribution and dynamics of water vapour (WV) is closely associated with meteorological phenomena, such as long persistent rainfalls, tropical cyclones, mid-latitude cyclonic storms and thunder storms that are ongoing challenges for synoptic meteorology (Ahrens and Samson, 2010). These severe weather phenomena can potentially cause destructive damage to society and the economy and hence play a critical role in weather forecasting. Improving the understanding of WV distribution is important (Le Marshall et al., 2010), not only for meteorology, but also for an effective usage of the global navigation satellite system (GNSS) technology for precise positioning. For example, tropospheric effects are one of the important atmospheric errors that need to be removed or mitigated in both high-accuracy differential positioning and precise point positioning (Wielgosz et al., 2012).

GNSS tomography is based on the inverse Radon transform theory and it has been intensively investigated by a number of research groups and universities across the globe (Bender et al., 2011; Perler et al., 2011; Brenot et al., 2012; Flores et al., 2000). A special working group on troposphere tomography model integration has been set up recently by the International Association of Geodesy (i.e. IAG WG4.3.2) to coordinate this IAG initiative (Rohm et al., 2012). In addition, the GNSS tomography is an extended service of the Ground-Based Augmentation System (GBAS). The standard approach to establish GNSS tomography models is to divide the troposphere into a 3-D voxel structure over the area of the GNSS CORS network coverage. The intercepted distance of a GPS ray passing through the voxel of concern is used in the
The design matrix for the calculation of the refractivity (Fig. 1). The design matrix is then inverted to estimate unknown wet refractivity values (Flores et al., 2000).

The major challenge of the tomography is to obtain a stable solution, in the presence of the ill-conditionedness (high condition number) and ill-posedness of the inverse process. Possible solutions can be broadly divided into two categories: (1) increasing the number of pseudo-observations (and decreasing the condition number, i.e., reducing the impact of the observation noise on final results) by adding horizontal and/or top layer constraints (Rohm and Bosy, 2009; Perler et al., 2011; Hirahara, 2000; Flores et al., 2000; Bender et al., 2011); (2) extending the observation system with additional observations in the zenith direction, point observations, and radiosonde and radiometer profiles (Perler et al., 2011). As a consequence, all solutions listed in approach (1) are similar in the sense that refractivity values are given for all voxels in the model even though not all are intercepted by GNSS signals (Perler et al., 2011).

In addition to the studies of tomography observation system, the 3-D model structure has been investigated (e.g. by ETH Zurich tomography research group). Perler et al. (2011) recently showed that it is feasible to indirectly calculate the coefficients of a wet refractivity trilinear spline function instead of the wet refractivity inside each voxel. The most common inversion technique applied in GNSS tomography is based on a singular value decomposition (SVD) technique that allows for complete matrix inversion (Rohm and Bosy, 2009). A slightly different approach was proposed by Bender et al. (2011) whereby an algebraic reconstruction technique is applied to estimate refractivity in an iterative way. Each iteration step updates wet refractivity only in voxels that are actually intercepted by the GNSS signals.

In this study, to overcome the ill-conditionedness of the inverse tomography problem without applying explicit constraints, the SVD method proposed by Xu (1998) and Lynch (2005) is used. The novelty in this approach is that the information provided in the observation matrix is used exclusively and singular values in the design matrix are sensibly selected. This paper aims to extend the previous research (e.g. Rohm and Bosy, 2011; Bosy et al., 2012; Rohm, 2012, 2013) by the new concept of a limited constraint, robust Kalman filter. Unlike all predecessors, the solution of the tomography model presented in this paper is not affected by some of the usually applied implicit constraints (i.e. no horizontal and no vertical constraints are applied), and it does not rely on additional observations (i.e. there is no need for NWP observations). It delivers solutions only for voxels that are intercepted by GNSS signals with a full variance–covariance matrix. The robust Kalman filter allows for variations of the refractivity field in time and reduces the noise propagation from the data into the output parameters. This technique is discussed in Sect. 4.

GNSS signal delay and the tomography model structure is presented in Sects. 2 and 3, respectively. Case study investigation using real and simulated data is performed in order to assess the quality and effectiveness of the new approach proposed. Conclusions and summary are given in the closing section of the paper.

2 GNSS signal delay

GNSS carrier frequencies reside in a microwave spectrum reserved for navigation (L-band, 1–2 GHz) (Hofmann-Wellenhof et al., 2008). Such spectrum’s allocation is designed to minimise signal attenuation in the atmosphere, and hence allows for all-weather operation of the system. The microwave signal’s refractivity in the neutral atmosphere is given as in Böhm and Schuh (2013), adopted after Essen and Froome (1951):

$$N_0 = k_1 \frac{R}{M_d} \cdot \rho + \left( k_2 \frac{e}{T} + k_3 \frac{e}{T^2} \right) \cdot Z_v^{-1} = N_h + N_w. \quad (1)$$

where $\rho$ is the density of air (mixed, dry + wet) (Kleijer, 2004) (kg m$^{-3}$), $R$ is an universal gas constant (J mol$^{-1}$ K$^{-1}$), $M_d$ is a molar mass of dry air (mol kg$^{-1}$), $e$ is the water vapour partial pressure (hPa), $T$ is the temperature (K), $k_1$, $k_2$ and $k_3$ are known empirical coefficients (Rüeger, 2002a, b), $Z_d^{-1}$ is an inverse compressibility factor for dry air and $Z_v^{-1}$ is an inverse compressibility factor for water vapour, respectively (both values are assumed to be 1 in this study). The analysis in this paper is focused only on the phase speed changes of the GNSS signals’ propagation (delay) in the neutral atmosphere, hence signal bending is neglected, and no effects on the signal’s energy are considered.

Tomography processing assumes that a signal is sufficiently modelled as a straight line between the satellite and
the receiver and that the slant total delay (STD) in neutral atmosphere is given by the following equation (Kleijer, 2004):

\[
STD = \int N_0 \cdot ds = \int N_0 \cdot ds + \int N_w \cdot ds = SHD + SWD, \quad (2)
\]

where SHD is the slant dry delay, and SWD is the slant wet delay. Usually tomography models utilise the SWD to reconstruct the water vapour distribution. The SWD is retrieved from GNSS troposphere estimates ZTD (e.g. Bosy et al., 2012) using Saastamoinen (Saastamoinen, 1972) dry delay model, fed with pressure values interpolated (Bosy et al., 2012) using Saastamoinen (Saastamoinen, 1972) dry

\[
SWD_n = 10^{-6} \sum_{i=1}^{m} N_{wm} d_{mn}, \quad (3)
\]

where \(N_{wm}\) is the wet refractivity in the voxel \(m\) (for the exemplary model structure see Fig. 2) and \(d_{mn}\) is the intercepted distance in the voxel \(m\) of signal \(n\).

The full functional model of the tomography solution in a matrix form is given as

\[
\begin{bmatrix}
SWD \\
N_{wpriori}
\end{bmatrix} =
\begin{bmatrix}
A_{inner} & A_{outer} \\
A_{apriori} & A_{apriori, outer}
\end{bmatrix}
\begin{bmatrix}
N_w \\
N_{ao}
\end{bmatrix} + v, \quad (4)
\]

The observations in Eq. (4) (SWDs) are appended with an a priori value of refractivity \(N_{wpriori}\). The design matrix \(A\) consists of four blocks: \(A_{inner}\) is the matrix containing distances in the inner model and \(A_{outer}\) is the equivalent matrix for the outer model (Fig. 2). \(A_{apriori}\) is the matrix containing value 1 when there are external observations and value 0 when there are no external observations in the voxel for inner model. \(A_{apriori, outer}\) is a similar matrix for the outer model. The unknowns \(N_w\) and \(N_{ao}\) are the wet refractivities in the inner model and the water vapour refractivity in the outer model, respectively.

The general equation for tomography using relation Eq. (4) in the matrix notation is given as

\[
SWD = A \cdot N_w + v. \quad (5)
\]

Equation (5) is ill-conditioned (no explicit constraints, with a limited number of observations) and ill-posed (a limited number of observations). Therefore, an effective inversion of matrix \(A\) is a central problem of all GNSS tomography applications.

The unconstrained tomography solution studies (Rohm, 2012) show that it is feasible to obtain precise wet refractivity exclusively from the GNSS SWD observations using stacked observations from at least ten 1 h consecutive epochs. However, ten hours is too long a time period to be represented by a single refractivity field. It is therefore convenient to use a Kalman filter which makes it possible to include a dynamic model of the troposphere. This is the focus of the next section.

4 Kalman filter application

A classic Kalman filter formulation follows the notation given in Grewal et al. (2001) whereby observations and process are separated. In this study, the process is set to be a wet refractivity field \(N_{w_{k+1}}\) with the time evolution given as a linear dynamic system (Yang, 2010):

\[
N_{w_{k+1}} = \Phi_k \cdot N_{w_k} + w_k, \quad (6)
\]

where \(\Phi_{k+1}\) is a state transition matrix (in this study it is an identity matrix \(\Phi_{k+1} = I\)). The \(w_k\) parameter is the noise with the characteristics of mean \(E(w_k) = 0\) and covariance \(E(w_k w_k^T) = Q_k\), which is called the dynamic disturbance noise matrix. The observation linear model for epoch \(k\) is given by Eq. (5):

\[
SWD_k = A_k \cdot N_{w_k} + v_k, \quad (7)
\]

where \(SWD_k\) are uncorrelated normally distributed observations, and measurement noise \(v_k\) has mean \(E(v_k) = 0\) and covariance \(E(v_k v_k^T) = R\). However, in a robust Kalman filter observations are assumed to be of a normal distribution contaminated with outliers (Yang, 2010), therefore \(E(v_k v_k^T) = R^R\) and the observation covariance matrix needs to be iteratively adjusted. The prediction step of Kalman filtering is given as a set of equations:

\[
\hat{N}_{w_k}(-) = \Phi_k \cdot \hat{N}_{w_{k-1}}(+)
\]

\[
P_k(-) = \Phi_k \cdot P_{k-1}(+) \cdot \Phi_k^T + Q_k,
\]

Figure 2. A skeleton of the exemplary horizontal (a) and vertical (b) structures of the TOMO2 model.
where $\hat{N}_{w_k}(\cdot)$ and $\hat{N}_{w_{k-1}}(\cdot)$ are the predicted and the corrected estimates of wet refractivity in the voxels of the GNSS troposphere tomography model. The matrix $P_k(-)$ is the prediction and $P_{k-1}(\cdot)$ the correction of the covariance matrix of the estimated state.

The Kalman gain matrix $K$ is

$$K = P_k(-)A_k^T \left( A_kP_k(-)A_k^T + R_k^R \right)^{-1}. \quad (10)$$

The covariance matrix $R_k^R$ of the robust Kalman filter is calculated using the following equation:

$$R_k^R = (\text{diag} (p_1, \ldots, p_m))^{-1}, \quad (11)$$

where

$$p_i = p \text{ for } |r_i| \leq \frac{\sigma \cdot c}{\sqrt{\eta}}, \quad (12)$$

$$p_i = \frac{c \sqrt{\eta}}{|r_i|} \text{ for } |r_i| \geq \frac{\sigma \cdot c}{\sqrt{\eta}}, \quad (13)$$

where $p$ is the weight of the observation, and the parameter $r_i$ is a posteriori residual value calculated via

$$r_k = A_k \cdot \hat{N}_{w_k}(\cdot) - SWD_k, \quad (14)$$

where $c = 1.5$ is a scaling factor and $\sigma$ is a reference variance (usually 1 mm).

Usually, a robust Kalman filter is applied to observations contaminated with outliers, to minimise or remove their impact by increasing the selected observations’ variances in the estimation process. Therefore, the process of estimating $R_k^R$ is iterative and might need to be repeated several times.

In the paper by Koch and Yang (1998) downweighting is applied on the parameters. However, in this study, to be consistent with the previously developed SVD methodology (Rohm and Bosy, 2011; Rohm, 2012, 2013) the downweighting of the parameters is not used. In this paper the structure of a design matrix $A$ is evaluated to reveal and remove linearly dependent observations (in a numeric sense). This technique reduces matrix $A$ condition number and improves inversion stability. The design matrix $A$ filtering process follows the methodology developed by Rohm (2013), based on the work of Xu (1998) and Lynch (2005). In essence, matrix $A$ is decomposed into three matrices (Strang and Borre, 1997):

$$A = U \Sigma V^T = \begin{bmatrix} u_1 & \ldots & u_k & \ldots & u_m \\ 0 & \ddots & \vdots & \ddots & \vdots \\ 0 & \ldots & \vdots & \ddots & \vdots \\ 0 & \ldots & \vdots & \ddots & u_m \end{bmatrix} \begin{bmatrix} s_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & s_m \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_k \\ \vdots \\ v_m \end{bmatrix}^T \quad (15)$$

where $U$ and $V$ are a set of orthonormal bases and $\Sigma$ is a set of singular values ($s_k$). A condition number of any matrix (e.g. matrix $A$) is calculated as (Anderson et al., 1999)

$$\text{cond}(A) = \frac{s_k}{s_1} \quad (16)$$

where $x$ is the rank of matrix $A$.

Previous investigation by Rohm (2013) shows that the uncertainty expressed as a covariance of wet refractivity is linked with the singular values on the diagonal of matrix $\Sigma$. The smaller the singular value considered in the design matrix ($A$) inversion (Eq. 10) the stronger the amplification of observation uncertainty $R_k^R$. The last, useful singular value ($s_k$) considered in the processing is found using functional analysis of singular values function (Xu, 1998; Lynch, 2005; Hansen and O’Leary, 1993), this method to improve the matrix condition number is named truncated singular values decomposition (TSVD). The new filtered $\hat{A}$ matrix is obtained by composing the $A$ matrix back according to the equation:

$$\hat{A} = \begin{bmatrix} u_1 & \ldots & u_k \\ \vdots & \ddots & \vdots \\ s_1 & \ldots & s_m \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}^T. \quad (17)$$

To reflect the changes in the design matrix $A$ observation matrices $\text{SWD}$ and $R$ have to be converted to the matrices $\text{SWD}$ and $R$ to eliminate the linearly dependent observations (operator $T$ in Eq. 20). The identification of linearly dependent rows ($f$) is based on comparing rows from matrices $A$ and $\hat{A}$ according to the following equation:

$$f, \text{ such that } \begin{bmatrix} A_{i,1,j} - \hat{A}_{i,1,j} \end{bmatrix}^T \geq 2 \cdot \text{SD} \begin{bmatrix} A_{i,1,j} - \hat{A}_{i,1,j} \end{bmatrix}^T. \quad (18)$$

Consequently, the Kalman filter sequence as shown in Koch and Yang (1998) for filtering observations will be transformed to the following sequence:

$$\hat{A} = \text{TSVD}(A) \quad (19)$$

$$\text{SWD} = T\{\text{SWD}\}, \quad \hat{R}_k^R = T\{R_k^R\} \quad (20)$$

$$\hat{N}_{w_k}(\cdot) = \hat{N}_{w_k}(\cdot) + \overline{K}_k \left( \text{SWD}_k - \overline{A}_k \hat{N}_{w_k}(\cdot) \right) \quad (21)$$

$$r_k = \hat{A}_k \cdot \hat{N}_{w_k}(\cdot) - \text{SWD}_k \quad (22)$$

$$\hat{R}_k^R = (\text{diag} (p_1, \ldots, p_m))^{-1} \quad (23)$$

$$\overline{K} = P_k(-) \hat{A}_k^T \left( \hat{A}_k \overline{P}_k(-) \hat{A}_k^T + \hat{R}_k^R \right)^{-1} \quad (24)$$

$$P_k(\cdot) = P_k(-) + \overline{K}_k \hat{A}_k \overline{P}_k(-) \quad (25)$$

$$\hat{N}_{w_k}(\cdot) = \Phi_k \cdot \hat{N}_{w_{k-1}}(\cdot) \quad (26)$$

$$P_k(-) = \Phi_k \cdot P_{k-1}(\cdot) + \overline{Q}_k. \quad (27)$$

The robust estimation process of wet refractivity starts with the filtering of matrix $A$ to produce $\hat{A}$ (Eq. 19), as well as truncating observations $\text{SWD}$ to obtain $\text{SWD}$ (Eq. 20). Initial estimates of parameters $\hat{N}_{w_k}$ are obtained $\text{SWD}$ (Eq. 20). Afterwards, residuals $r_k$ are derived, which form the base for downweighting of the outliers (Eqs. 22, 12, 13) and calculation of $\hat{R}_k^R$ matrix. The following step consist of calculating the Kalman gain $\overline{K}$ (Eq. 19). Equations (21)–(24) are repeated several times to remove outliers from observations. This operation is followed by an update step (Eq. 25), and the propagation of covariance and parameters to the next epoch.
The initial covariance $P_k(-)$ was calculated using errors estimates published in one of the previous papers (Rohm, 2012).

The matrix $Q_k$ adds a noise to each voxel in a covariance matrix $P_{k-1}$. The amount of noise in each element of matrix $Q_k(q_{m,i}$ for the inner domain and $q_{o,i}$ for the outer domain) is driven by three factors: (1) location in the model $h$ (height); (2) time since last update $t$; and (3) location in inner or outer model, according to the following formula:

$$Q = \text{diag}(q_{1,i} \ldots q_{m,i} q_{1,o} \ldots q_{l,o})$$

(28)

where

$$q_{m,i} = \left(1 - \exp^{-\frac{t}{T}}\right) \cdot \sigma^2_{h,i}$$

(29)

$$q_{l,o} = \sigma^2_o$$

(30)

where exp is the base of the natural logarithm.

In this study, uncertainty parameters $\sigma^2_{h,i}$ and $\sigma^2_o$ are calculated from NWP model outputs, interpolated to the tomography model voxels. The performance analysis takes into account vertical variability of the wet refractivity as well as time autocorrelations of these parameters. Therefore, the input parameters for Eqs. (29) and (30) are essentially anticipated wet refractivity variations in the model space and in the time domain. In a more general case (without access to NWP data) the uncertainty parameters may be derived from climatological data.

5 Case study

To demonstrate the capability of this new GNSS tomography model TOMO2 (introduced in this paper), a case study based on simulated and real data is performed and the results are validated against NWP model outputs.

5.1 NWP model

The meteorological data covers ACCESS-R model outputs (analysis run) with the time resolution of 6h and spatial outline covering Australia and a ~ 20° buffer zone. The model is based on the UK Meteorological Office Unified Model, and a number of data sources are used to produce forecasts (e.g. COSMIC, AIRS, SYNOP) (Le Marshall et al., 2010). The model in the horizontal plane contains 229 nodes, with the grid spacing of 0.375° (~37.5 km) and the model utilises in the vertical direction terrain following hybrid (pressure/height) coordinates with 50 levels. This study, from all possible NWP model parameters, considers only pressure, temperature and WV partial pressure (given as a mixing ratio) (Fig. 3).

5.2 GNSS stations network

Two ZTD data sets, with the same time, terrain and identical receiver network settings are prepared, one is simulated and the other one is real data. The same tomographic model setup is used. The time span covered by this case study is limited to 325 epochs of ZTD estimation between 3 March 2010 and 9 March 2010, whereby each ZTD epoch covers an interval of 30 minutes. GNSS observations from 45 stations were taken during the development, transition and dissipation of a heavy hail storm (Choy et al., 2011). The GNSS network employed in this study (GPSnet) (Fig. 3) is owned and operated by the Victorian Government Department of Sustainability and Environment (Victoria, Australia). Standard GPSnet stations are equipped with Trimble NetR5 receivers and high-quality antennas (mostly TRM55971.00). A few International GNSS Service (IGS) stations were also processed, with receivers/antennas from other manufacturers (e.g. Leica, Ashtech). The inter-receiver distance is roughly 50 km.
in the investigated case. Terrain undulation, especially in the east mountainous part of Victoria, is in favour of tomography because of large height differences (the voxels close to the ground are crossed by slant delays from other stations), some receivers are located on the mountains’ peaks (e.g. MTBU – Mount Buller 1600 m). The troposphere in the vertical direction, above a network of the GNSS receivers, is divided into a number of cuboids (in this study called voxels, see Fig. 2), from ground level to 10 km with increasing layer thickness (varying between 500 and 1700 m), whereas the last voxel spans from 10 to 15 km and has 5 km thickness. The horizontal plane of the tomography model consists of six voxels in the north direction and 12 voxels in the east direction (to accommodate the model outline to the GPSnet shape). As a consequence of the vertical and horizontal settings, the size of each voxel is approximately 75 km × 45 km × 0.5–5 km, which roughly represents the average inter-station distance.

Pressure, temperature and water vapour produced by the ACCESS-R model (1) are utilised to separate wet and dry delays (pressure), and (2) act as a reference value for tomography model reliability investigation (water vapour). To separate wet and dry delays, pressure values from eight NWP model nodes (four below, four above) surrounding GNSS stations are interpolated to the antenna reference point. The NWP acts as a supplementary pressure data source as ground-based pressure observations at the GNSS stations are not available. The second use of NWP model outputs is realised via interpolating water vapour partial pressure and temperature values from NWP model nodes to the tomography’s voxel centre points (Bosy et al., 2010, 2012). Alternatively the pressure and temperature values are obtained from the global pressure and temperature (GPT) model (Boehm et al., 2007) and water vapour pressure is calculated from UNB3m (Leandro et al., 2008). At the time of experiments the GPT2 (Lagler et al., 2013) model had not yet been introduced, so the consistency between pressure, temperature and water vapour content from the two different models (UNB3m and GPT) was assumed.

5.3 Simulated slant delays

The first data set comprises wet refractivities $N_R$ derived from NWP temperature and WV interpolated to the centre of each voxel of the tomography model. Then, using simple analytical ray tracing (Rohm, 2013; Rohm and Bosy, 2009) the signal intercepting distances in each voxel of the inner $A_{\text{inner}}$ and the outer models $A_{\text{outer}}$ (Eq. 4) are calculated, along with observations SWD$_S$ (Eq. 7). Uncertainty of the ray tracing through the model of unknown precision (there is no impartial measure of weather model parameter precision) remains unidentified. Therefore, all elements of the covariance matrix $R_R$ (Eq. 11) are of equal weight (1 mm). The NWP model data and the simulated slant delays also constitute the reference data.

The simulated data were reprocessed to take into account random noise and bias (0.025 and 0.007 m, respectively). The proposed values are based on previous tests comparing the simulated and observed slant delays (Bosy et al., 2012). The degradation is distributed randomly regardless of the satellite elevation angle, the receiver position and the time of the day.

5.4 Real GNSS data

The second data set consists of the real observations from a GNSS network processed with the Bernese GPS Software 5.0 (Dach et al., 2007). Only GPS observations are considered. The wide/narrow lane (L5/L3) GPS processing procedure is applied (Dach et al., 2007). The ambiguities are solved with the wide-lane L5 (98% success rate) strategy. Final coordinates are estimated with the minimum constraint conditions imposed on the translation parameters of coordinates and velocities of IGS reference stations (MOBS, HOB2, STR1 and CEDU). The mean accuracy of the solutions in the horizontal directions and in the vertical direction, based on repeatability score, are 1.5 and 4 mm, respectively.

The troposphere estimates in 30 min resolution are obtained in the next processing step by fixing the translation parameters of the network (the solution inherited from precise geodynamic studies) and pre-eliminating the velocities as well as the coordinates from weekly solution (removing from normal equations). In Bernese GPS software the standard approach to estimate the ZTD (Dach et al., 2007) is used. The ZTD parameters are estimated as corrections to an a priori standard atmosphere model using piecewise linear functions. The Niell (1996) mapping function is used to parameterise the mapping of troposphere delays to the vertical direction. The atmospheric gradients (Dach et al., 2007) are estimated at the same time resolution as the total delay. The adopted processing setup is not an optimal configuration (state of the art mapping functions are not used) to estimate troposphere parameters. However, it is a common approach used in Bernese 5.0 GPS Software. Output TRO and TRP files comprise the most important part of the second data set. This set also includes pressure parameters interpolated to the antenna heights from the NWP model, and final station coordinates as well as precise orbits from IGS. The dry part is subtracted from the total delay based on the Saastamoinen model of dry delay (Saastamoinen, 1972) with pressure values from the NWP model.

In this study SWDs are calculated using either the zenith part of the delay or the zenith delay and horizontal gradients (Boehm and Schuh, 2007); in either case the double differentiated residuals (Manning et al., 2014) are not considered. The gradients show large variability in the zenith direction and are significant (statistically speaking). The north component varies between –2.2 and 2.5 mm, whereas the east component oscillates between –2.4 and 3.5 mm. The estimation formal errors are relatively small, the average value...
The TOMO2 tomography model voxel settings superimposed on the wet refractivity field of 6 March ... + no downweighting, OOO – no observation removal + no design matrix reconditioning + no downweighting). The uncertainty measures \( R \) are calculated based on the estimated error of particular ZTD value by applying the law of variance propagation (Rohm, 2012).

5.5 Tomography processing results discussion

A number of test runs of the tomography model are performed to precisely assess the impact of particular methodological improvements. The following major groups of settings are adopted with regard to observations: (1) simulated observations \( M \) (with and without a noise); (2) real observations with gradients \( R \); and finally (3) real observations without gradients \( Z \). Furthermore, experiments are grouped together according to the a priori models adopted. The following settings are considered: (1) NWP-derived outer model values for all epochs and NWP-derived inner model values for the first value (as an initial value) \( N \) and as an alternative with the same settings for outer model and all epochs for inner model \( W \); (2) NWP-derived outer model values for all epochs and UNB3m- and GPT-derived inner model for the first epoch \( G1 \) and with a reverse settings, first epoch inner NWP, all epochs outer GPT + UNB3m \( G0 \); (3) NWP-derived outer model values for all epochs and UNB3m- and GPT-derived inner model for all epochs \( G2 \) and fully independent from NWP data with all epochs populated using GPT + UNB3m data \( G1P \). To assess the impact of the innovative robust Kalman filter processing procedure, the following three levels of validation are adopted: (1) firstly all equations related to the observation selection criteria are applied including: SWD observation removal (Eq. 20) \( S \); reconditioning of matrix \( A \) (Eqs. 17 and 18) \( A \) and downweighting of selected observations (Eqs. 11, 23, 24) \( D \); (2) secondly the downweighting scheme is not applied \( O \) but observation removal \( S \) and reconditioning \( A \) is, (3) thirdly no robust improvements of the Kalman filter are considered \( OOO \), so the filter runs like a classic Kalman filter.

The experiment setup is shown in Table 1. In total 21 different settings are investigated, and the most significant results are presented in Table 2. The solutions for all 325 epochs are depicted in Figs. 4–15. Figures 4, 6, 8, 10, 12 present the mean bias of the tested solution against reference data, whereas Figs. 5, 7, 9, 11, 13 show the mean standard deviation of the tested solution against reference data. Figures 14 and 15 present detailed vertical structure of bias and standard deviation. The major outcomes of these experiments are summarised below.

The most important conclusion drawn from the set of experiments is that the a priori value \( N_{\text{apriori}} \) for inner model \( N, W, G0, G1, G2, G1P \) is the main factor in all processing...
Table 1. The list of tested tomographic solutions. The naming conventions explained in two bottom rows.

<table>
<thead>
<tr>
<th>Tested combinations</th>
<th>ZNASD, ZNASO, ZNOOO</th>
<th>M1NASD, M1NASO, M1NOOO</th>
<th>M2NASD, M2NASO, M2NOOO</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNASD, RNASO, RNOOO</td>
<td>RG1ASD, RG1ASO, RG1OOO</td>
<td>RG2ASD, RG2ASO, RG2OOO</td>
<td>RW ASD, RG0ASD, RG1PASD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Naming convention (example)</th>
<th>SWD type</th>
<th>A priori type</th>
<th>Reconditioning</th>
<th>SWD removal</th>
<th>Downweighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>N</td>
<td>A</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>Real observations with gradients (R), with a priori observations from NWP model in outer model (for all epochs) and inner model (initial epoch) (N), in Kalman filter A matrix reconditioning (A), SWD removal and downweighting (D) has been used.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observations</th>
<th>A priori type</th>
</tr>
</thead>
<tbody>
<tr>
<td>R, Real observations with gradients</td>
<td>N, NWP outer (all epochs) and inner (first epoch)</td>
</tr>
<tr>
<td>Z, Real observations without gradients</td>
<td>W, NWP outer and inner (all epochs)</td>
</tr>
<tr>
<td>M1, Simulated observations without noise</td>
<td>G1, NWP outer (all epochs) and GPT + UNB3m inner (first epoch)</td>
</tr>
<tr>
<td>M2, Simulated observations with realistic noise</td>
<td>G0, GPT + UNB3m outer (all epochs) and NWP inner (first epoch)</td>
</tr>
<tr>
<td></td>
<td>G2, NWP outer (all epochs) and GPT + UNB3m inner (all epochs)</td>
</tr>
<tr>
<td></td>
<td>G1P, GPT + UNB3m outer and inner (all epochs)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kalman filter improvements</th>
<th>Reconditioning</th>
<th>SWD removal</th>
<th>Downweighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A – Yes</td>
<td>S – Yes</td>
<td>D – Yes</td>
</tr>
<tr>
<td></td>
<td>0 – No</td>
<td>0 – No</td>
<td>0 – No</td>
</tr>
</tbody>
</table>

Figures 4–15 show that whenever the a priori value for all epochs and all voxels is set (compare RG2ASD and ZG2ASD to RNASD and ZNASD, in Table 1), even with a simple deterministic model such as UNB3m and GPT the quality of the reconstruction is much higher than in all other cases. The values in Table 2 show also higher accuracy of the G2 solution. The experiments using different combinations and “intensity” of a priori data (G0, G1P, G1, G2, N, W) show that there is very limited impact of quality of outer model data (RG0ASD and RG1ASD, Figs. 6 and 7), but clearly the retrieval quality increases with larger number of a priori data (RG1ASD and RG2ASD, Figs. 8 and 9; RNASD and RW ASD, Figs. 10 and 11). Comparing retrieval based on UNB3m and GPT only (RG1PASD in Figs. 12 and 13) with retrieval based on NWP only (RW ASD in Figs. 12 and 13), we may see that solution quality converges after 150 epoch and is essentially similar. Introducing initial values into the tomography system (Eq. 4) can effectively stabilise the tomography solution. In this study the initial wet refractivity field is a function of the day of the year, latitude, longitude and altitude. The variance of the a priori observations in Eq. (17) is set to rather large value (i.e. 30 mm km$^{-1}$). Hence, the results show that the quality of the a priori observations is not an issue for tomography reconstruction.

The second most important outcome of this research is that the robust filtering helps to reduce noise in outputs. Clearly, the solid red line showing the standard deviation of the real data solution in Fig. 5 (RNASD) is much lower than the one with partial robust algorithms (RNASO) and no robust procedures in place (RNOOO). However, the difference between the last two is not significant which in turn means that the most significant improvement in real-time data processing is due to the downweighting not reconditioning. The same effect is visible when the processing covers the real observations with a large number of a priori data (RG2ASD – the dark blue line in Fig. 9). However, the effect is not strong, at least in the solution scatter. In terms of systematic errors, the mean difference is effectively removed by the robust algorithm (RG2ASD) as depicted in Fig. 8.
of the realistic noise and bias to the observations (0.025 m of shown in Fig. 14 in the third panel). However, introduction of the realistic noise and bias to the observations (0.025 m of the random noise and 0.007 m of the bias) results in tomography quality degradation to 6.4 mm km\(^{-1}\) (Table 2, M1G2ASD shown in Fig. 14 in the third panel). However, introduction of the realistic noise and bias to the observations (0.025 m of the random noise and 0.007 m of the bias) results in tomography quality degradation to 6.4 mm km\(^{-1}\) (Table 2, M1G2ASD, Fig. 14, second panel). In comparison, the quality of the tomography retrieval based on real data (Table 2, RG2ASD) is 6.2 mm km\(^{-1}\), (dark blue line in Fig. 9). This suggests that both solutions (real, RG2ASD, and simulated, M2NASD) converge to the same solution with similar bias and standard deviation measures; hence all real data outliers are effectively filtered out. The retrieval quality presented separately for each layer (Figs. 14 and 15), measured as a mean and standard deviation solution departure from NWP based profile, shows reasonably high agreement in mid-troposphere (above 2 km) and significant bias and large scatter in the lower section of the profile (below 2 km). The wet refractivity converted to water vapour (Fig. 15) using inversion of Eq. (1) and temperature profile from NWP, shows that the standard deviation of retrieval is close to 2 hPa in the middle part of the troposphere. The obtained results confirm that station separation and cut-off angle limits the number of signal intersections in the troposphere boundary layer and hence the tomography model uses a priori data as a solution in this section of atmosphere.

Table 2. Set of the quality measures for investigated models; bias is a mean discrepancy between reference wet refractivity (from the ACCESS-R model) and refractivity retrieved from TOMO2 model, standard deviation (SD) is a measure of scatter for discrepancies between reference wet refractivity (from the ACCESS-R model) and refractivity retrieved from the TOMO2 model, a posteriori rms of SWD observations (SWD\(_k\) - \(A_k \cdot N_{wk}\)) and mean processing uncertainty as in Eq. (25). Statistics derived using only retrieval for inner model. The number of resolved voxels per layer varies between 33\% (bottom layer) to more than 80\% above 4000 m.

<table>
<thead>
<tr>
<th>Model and observation setup</th>
<th>Validation by NWP</th>
<th>A posteriori rms of SWD [mm]</th>
<th>Processing uncertainty SD [ppm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>RG1PASD</td>
<td>0.5</td>
<td>6.2</td>
<td>13</td>
</tr>
<tr>
<td>RWASD</td>
<td>−0.1</td>
<td>5.8</td>
<td>12</td>
</tr>
<tr>
<td>RG0ASD</td>
<td>−0.4</td>
<td>8.7</td>
<td>31</td>
</tr>
<tr>
<td>RG1ASD</td>
<td>−0.1</td>
<td>8.8</td>
<td>31</td>
</tr>
<tr>
<td>RG2ASD</td>
<td>0.5</td>
<td>6.2</td>
<td>13</td>
</tr>
<tr>
<td>ZG2ASD</td>
<td>0.4</td>
<td>6.7</td>
<td>12</td>
</tr>
<tr>
<td>ZNASD</td>
<td>−0.2</td>
<td>8.0</td>
<td>28</td>
</tr>
<tr>
<td>M1G2ASD</td>
<td>−0.1</td>
<td>4.2</td>
<td>5</td>
</tr>
<tr>
<td>M1NASD</td>
<td>−0.2</td>
<td>4.4</td>
<td>9</td>
</tr>
<tr>
<td>M2NASD</td>
<td>−0.3</td>
<td>6.4</td>
<td>13</td>
</tr>
<tr>
<td>RNASD</td>
<td>−0.5</td>
<td>8.8</td>
<td>31</td>
</tr>
<tr>
<td>RNOO</td>
<td>−1.2</td>
<td>12.4</td>
<td>10</td>
</tr>
<tr>
<td>UNB3mGPT</td>
<td>3.5</td>
<td>7.2</td>
<td>–</td>
</tr>
</tbody>
</table>

The third conclusion is that the best achievable performance using this tomography model and simulated observations (without noise) is 4.2 mm km\(^{-1}\) (Table 2) (M1G2ASD shown in Fig. 14 in the third panel). However, introduction of the realistic noise and bias to the observations (0.025 m of the random noise and 0.007 m of the bias) results in tomography quality degradation to 6.4 mm km\(^{-1}\) (Table 2, M2NASD, Fig. 14, second panel). In comparison, the quality of the tomography retrieval based on real data (Table 2, RG2ASD) is 6.2 mm km\(^{-1}\), (dark blue line in Fig. 9). This suggests that both solutions (real, RG2ASD, and simulated, M2NASD) converge to the same solution with similar bias and standard deviation measures; hence all real data outliers are effectively filtered out. The retrieval quality presented separately for each layer (Figs. 14 and 15), measured as a mean and standard deviation solution departure from NWP based profile, shows reasonably high agreement in mid-troposphere (above 2 km) and significant bias and large scatter in the lower section of the profile (below 2 km). The wet refractivity converted to water vapour (Fig. 15) using inversion of Eq. (1) and temperature profile from NWP, shows that the standard deviation of retrieval is close to 2 hPa in the middle part of the troposphere. The obtained results confirm that station separation and cut-off angle limits the number of signal intersections in the troposphere boundary layer and hence the tomography model uses a priori data as a solution in this section of atmosphere.

Another important issue clearly visible is that there is not much difference between the tomography solution fed by the observations with and without gradients (Table 2, RG2ASD 6.2 mm km\(^{-1}\) standard deviation, ZG2ASD 6.7 mm km\(^{-1}\) standard deviation). The same level of bias has been also observed for both types of measurements (Table 2, RG2ASD 0.5 mm km\(^{-1}\) and ZG2ASD 0.4 mm km\(^{-1}\) bias). Using either data type leads to the same a posteriori errors of observations and uncertainties. Therefore, using gradients in the signal delay modelling does not improve the solution in this model setup.

Many authors (Bender et al., 2011; Manning et al., 2014; Perler et al., 2011; Rohm, 2012) report that the tomography quality varies between 4 to 10 mm km\(^{-1}\) and is lower for the bottom level of troposphere and increases with height until
the amount of water vapour is lower than the sensitivity of the method. This suggests that the effectiveness of the tomography method in resolving the vertical structure of troposphere needs further investigation. In this study to validate whenever this method has some advantage over a deterministic model, we simply subtract GPT and UNB3m wet refractivity from NWP derived refractivities and calculate the statistics such as standard deviation and bias. The results in Table 2 (UNB3mGPT) shows that the standard deviation is slightly higher than that of the tomography model (7.2 mm km\(^{-1}\)), but the bias is much higher (3.5 mm km\(^{-1}\)). Hence, the tomography processing has the advantage over the deterministic models; the question remains of whether the level of the obtained quality is satisfactory for meteorological and positioning applications.

6 Conclusions

In this paper, the new GNSS tomography model TOMO2 is presented. This model employs a robust Kalman filter to solve the limited constraint (i.e. the correlation between voxels are not applied) tomography problem. This study demonstrates that the real slant wet delay data set is affected by noise and outliers and the estimated zenith delay uncertainties are overly optimistic. Therefore, the real GNSS data require advanced processing beyond the ordinary Kalman filter. In this study, both the robust Kalman filter and a truncation of the design matrix (with TSVD method) are investigated to limit the noise impact on the model updates. The estimations of wet refractivities and their associated uncertainties in the troposphere above a network of GNSS receivers for selected voxels can be determined through these methods. Results show that the STD discrepancy between the reference wet refractivity and the tomography model outputs is of the order of 6.2 mm km\(^{-1}\) (or ppm), which is the equivalent to 2 hPa of WV. The results are in good agreement with GNSS tomography simulation studies with an intermediate level of noise of 5.2 mm km\(^{-1}\) (or ppm) (Bosy et al., 2012). The most important contribution of this paper is an effective GNSS tomography reconstruction without using implicit constraints, which allows for a quicker tomography model response to the changing environment conditions.
The limited constraint approach investigated in this research produces more realistic wet refractivity uncertainties that are unbiased by inner constraints. The results presented in this paper shows the current level of quality achievable with tomographic reconstruction. Further discussion with the meteorological community is needed to investigate an efficient way to assimilate the GNSS tomography outputs into NWP models.

Acknowledgements. This project is supported by the Australian Space Research Program (ASRP2), the Australian Research Council (ARC) Linkage (LP0883288) and the Australian Antarctic Scheme (4159) projects funded by the Australian Federal Government and the Polish Ministry of Science and Higher Education: research project No. NN526 197238. We also thank the Australian Bureau of Meteorology for providing the NWP data, and, the Department of Sustainability and Environment for providing the GPSnet data. The authors would like to express their gratitude to Brett Carter for enhancing the language of this paper. Special thanks to Sue Lynn Choy and Yurii Kuleshov for their valuable comments and discussions.

Edited by: S. Slijkhuis

References


Rohm, W., Geiger, A., Bender, M., Shangguan, M., Brenot, H., and Manning, T.: IAG WG4.3.2 Inter-Comparison and Cross-Validation of Tomography Models – Aims, Scope and Methods, in: IGS Workshop 2012, 23–27 July 2012, Olsztyn, Poland, 2012.


