

A 4D tomographic ionospheric model to support PPP-RTK

G. Olivares-Pulido, M. Terkildsen, K. Arsov, P.J.G. Teunissen, A. Khodabandeh, V. Janssen

Abstract

Successful implementation of integer ambiguity resolution enabled precise point positioning (aka PPP-RTK) algorithms is inextricably linked to the ability of a user to perform near real-time positioning by quickly and reliably resolving the integer carrier-phase ambiguities. In the PPP-RTK technique, a major barrier to successful ambiguity resolution is the unmodelled impact of the ionosphere. We present a 4D ionospheric tomographic model that computes in real time the ionospheric electron density as a linear combination of basis functions, namely B-splines. The results show that when the ionospheric estimates are provided as atmospheric corrections for a PPP-RTK end-user, the time to fix its horizontal position below 10 cm is around 20 epochs (the sample rate is 30 s) at 90% of the cumulative distribution function (CDF), as opposed to the time it takes when no external corrections are provided, which is around 80 epochs at 90% of the CDF.

Keywords: Ionosphere, Tomography, B-splines, GNSS, PPP-RTK.

1 Introduction

The standard precise point positioning (PPP) technique can achieve horizontal position accuracy below 10 cm. It relies on the use of the carrier-phase observable, which is a hundred times less noisy than the code (pseudo-range) observable, and satellite precise orbit and clock errors estimates. On the other hand, the unknown initial value of the number of cycles makes the carrier-phase ambiguous. The time needed to achieve cm-level accuracy is usually above 15 minutes, or even up to around an hour, thus preventing standard PPP to be used in real time (Zumberge et al. 1997).

Ambiguity fixing is achieved much quicker by real-time kinematic (RTK) techniques, which have been used since the early 1990s. Nevertheless, as opposed to PPP, RTK cannot be applied by users located very far away from a reference station, thus typically limiting its performance to up to 20 km in single-baseline RTK, and up to 70 km baseline when a reference network is used [see Teunissen and Montenbruck (2017) and references therein].

PPP-RTK is integer ambiguity resolution-enabled PPP. PPP-RTK extends the PPP concept by providing users, along with the needed satellite orbit and clock corrections, information about satellite phase and code biases, and the atmosphere. This information, when accurately provided, enables recovery of the integer nature of the user ambiguities, thus enabling single-receiver ambiguity resolution and thereby reducing convergence times and improving positioning as compared to standard PPP (Odijk et al. 2016; Teunissen and Montenbruck 2017). Another important aspect is that PPP-RTK networks may provide global service, whereas the coverage area of RTK networks is regional. Uncertainty in the ionospheric signal delays is currently the main bottleneck for PPP-RTK in the sense of hindering fast successful ambiguity resolution. For more background on how the ionosphere affects the global navigation satellite system (GNSS) signals see, for example, Jakowski (2008) and Hernandez-Pajares et al. (2011).

The integration of the electron density along the line of sight between the satellite and the receiver, i.e. the slant total electron content or STEC, accounts for more than 99.99% of the total ionospheric delay. Higher-order ionospheric delays may have impacts on satellite orbit estimates and clock errors (Petrie et al. 2010, 2011; Hernandez-Pajares et al. 2014). However, such errors are usually smaller than 2 cm

or even at the mm-level (Hadas et al. 2017); therefore, they are usually considered negligible for real-time GNSS-based positioning. Hence, we have computed the STEC term only for this work.

The choice of the ionospheric model has an impact on the accuracy of the ionospheric estimates and their uncertainties. In this regard, a widely used approach for ionospheric modelling is the thin-shell layer, aka single-layer model (SLM), whereby it is considered that the ionosphere is projected onto the 2D SLM usually located between 350 and 450 km. However, 2D models are not accurate enough to support navigation and positioning techniques in real time at the cm-level (Odijk 2002; Banville 2014).

Higher accuracy and precision are achieved by considering the ionosphere as a grid of three-dimensional pixels, i.e. voxels, e.g. the tomographic method. One of the first GNSS-based data-driven tomographic models was presented by Hernandez-Pajares et al. (1999), whereby the electron density within each voxel is computed as a random walk time series. Such approach also yields a mismodelling due to the discontinuity of the electron density from one voxel to adjacent ones. Usually, this is overcome by assuming that the electron content within one voxel is equal to the average electron content of the voxels in the neighbourhood. This is an application of the mean value theorem (MVT) (Flanigan 1983; Hernandez-Pajares et al. 1999), which is a valid approach provided there are neither local maxima nor minima of the electron density (e.g. ionospheric plasma bubbles, large-scale travelling ionospheric disturbances) in the voxels where MVT is applied.

An alternative to the MVT is to estimate the electron density as a linear combination of smooth and continuous basis functions, aka B-functions, e.g. spherical harmonics (SPH) (Schaer 1999), spherical cap harmonics (SCH) (Haines 1988), B-splines and trigonometric B-splines (Schmidt et al. 2008). Moreover, such approach does not impose any constraints regarding the existence of local electron density maxima and minima. All B-functions models present advantages and inconveniences. For example, SPH and SCH are functions with global support; therefore, they provide estimates at all locations of the area of interest. However, the presence of data gaps may have an impact on the estimates of the coefficients of the linear combination (Schmidt et al. 2011).

On the other hand, B-splines, having local support, do not suffer from the same mismodelling as SPH and SCH do when data gaps are present. Nevertheless, only those coefficients constrained by data can be properly estimated. Ionospheric coefficient estimates in empty voxels can be constrained by using a background model, e.g. the International Reference Ionospheric (IRI) model (Schmidt et al. 2007) or global ionospheric maps (GIMs) from the International GNSS Service (IGS) (Hernandez-Pajares et al. 2011). However, climatological or empirical models and GIMs yield typical vertical TEC (VTEC) RMS errors from around 2 TECU ($1 \text{ TECU} = 10^{16} \text{ e}^- / \text{m}^2$; $\text{e}^- = 1 \text{ electron}$) (Hernandez-Pajares et al. 2017), to even 10 TECU in some cases (Hernandez-Pajares et al. 2011), thence they are not accurate enough to support GNSS-based navigation users at the cm-level.

The 4D real-time ionospheric model presented in this work aims to support PPP-RTK algorithms. Such ionospheric model is envisaged to support real-time ambiguity resolution across Australia. In this regard, and considering the above-mentioned advantages and disadvantages, polynomial B-splines are chosen as the basis functions for the ionospheric modelling in 3D and in real time. In this contribution, the goal of the ionospheric modelling is not to provide a full map of the ionosphere, but to provide ionospheric corrections to PPP-RTK end-users which are consistent with other network corrections (e.g. clock errors, satellite phase bias). Note that in PPP-RTK, the user integer ambiguity depends on the external double-differenced (DD) STEC (Khodabandeh & Teunissen 2015), not on the un-differenced STEC; therefore, the following two aspects must be considered regarding the ionospheric modelling:

1. The DD STEC estimates must be accurate enough so as to ensure successful ambiguity resolution.
 - Considering that the STEC data set is not spatially homogeneous, the global support feature of SPH and SCH functions makes their estimates less accurate than the B-splines estimates (Schmidt et al. 2011).
2. The ionospheric corrections must be provided in real time or near real time. (By real time we mean updating the estimates as the observations and corrections are provided, typically less than 1 min.)
 - The decisive factor is the computational speed. Both SCH and B-splines can provide the same resolution as SPH with less parameters; therefore, the computational speed with SCH and B-splines is higher than using SPH.

Non-homogeneous data sets yield higher impact on SPH models than on B-splines-based models; hence, the estimates from the latter model are more accurate than those from SPH models (Schmidt et al. 2011). Additionally, B-splines models require less parameters than SPH-based models, consequently the computational speed is higher for B-splines models than for SPH. Altogether considered, those are the main reasons to use B-splines as basis functions for the 4D ionospheric model. Moreover, although not considered in the present work, B-splines provide tools for data compression (Mautz et al. 2005; Zeilhofer 2008), a property that might help to increase the resolution of ionospheric corrections broadcast in a fashion that is compliant with the Radio Technical Commission for Maritime Services (RTCM) State Space Representation (SSR) standards.

Regardless of the B-function approach, the computation in real time requires a computational method that updates the estimates and corrections on the fly. For this, we have implemented a simple Kalman filter.

Another important aspect is the STEC retrieval approach. If only code (pseudo-range) observables are used, then the ionospheric combination removes all terms that do not depend on the frequency, thus providing a direct observation of the STEC, although biased by satellite and receiver differential hardware delays. This method is the simplest but also the noisiest. An alternative is the carrier phase, which is 100 times less noisy than the code (pseudo-range). However, it is ambiguous. A standard procedure to estimate such ambiguity is the carrier-phase to pseudo-range levelling, which increases the noise as well. On the other hand, the un-differenced and uncombined approach of PPP-RTK provides a STEC estimate of higher accuracy and less noise than the levelling method (Zhang 2015). Altogether considered, we have computed the STEC from GPS observables using an uncombined and un-differenced PPP-RTK network platform developed and implemented by the Global Navigation Satellite Systems Research Centre at Curtin University [for further details see Odijk et al. (2016)].

The rest of this paper is organized as follows: Section 2 describes the method for the tomographic ionospheric computation based on B-splines. Section 3 presents the data set and the network used across New South Wales (Australia) in quiet atmospheric conditions. Then, Section 4 presents the results in terms of positioning accuracy. Section 5 summarizes the conclusions and proposes further work to be considered in the future.

2 Tomographic ionospheric model

The STEC is the integration of the ionospheric electron density along the line of sight (LOS) from the receiver position to the satellite position as follows:

$$S_r^s = \int_{\mathbf{r}_r}^{\mathbf{r}^s} N_e(\mathbf{r}, t) dl \quad (1)$$

where $S_r^s \equiv \text{STEC}_r^s$, i.e. the STEC from the receiver r to the satellite s ; \mathbf{r}_r and \mathbf{r}^s are the position vectors for the receiver and the satellite, respectively; $N_e(\mathbf{r}, t)$ stands for the electron density at position \mathbf{r} and time t ; and d_l is the differential path along the LOS.

However, the input STEC for the ionospheric model is not S_r^s , but ionospheric delays biased by satellite and receiver hardware delays. Then, Eq. (1) is modified as follows:

$$\tilde{S}_r^s = S_r^s + d_{r,\text{GF}} - d_{,\text{GF}}^s \quad (2)$$

where $d_{r,\text{GF}}$ and $d_{,\text{GF}}^s$ are the geometry-free receiver and satellite hardware delays (see Table 1 for their definition). Therefore, the estimator must decorrelate the physical STEC and the hardware delays. One way to do so is by considering that satellite and receiver hardware time variability is negligible during the computation, as opposed to the time-scale variability of the physical STEC.

Table 1: Estimable parameters for the *Common Clocks S*-system.

Estimable Parameter	Interpretation	Conditions
Receiver clock	$d\tilde{t}_r = (dt_r + d_{r,IF}) - (dt_1 + d_{1,IF})$	$r \geq 2$
Satellite clock	$d\tilde{t}^s = (dt^s + d_{,IF}^s) - (dt_1 + d_{1,IF})$	$s \geq 1$
Receiver phase bias	$\tilde{\delta}_{r,j} = \left(\delta_{r,j} - \frac{1}{\lambda_j} [d_{r,IF} - \mu_j d_{r,\text{GF}}] + z_{r,j}^1 \right) - \left(\delta_{1,j} - \frac{1}{\lambda_j} [d_{1,IF} - \mu_j d_{1,\text{GF}}] + z_{1,j}^1 \right)$	$j \geq 1; r \geq 2$
Satellite phase bias	$\tilde{\delta}_{,j}^s = \left(\delta_{,j}^s - \frac{1}{\lambda_j} [d_{,IF}^s - \mu_j d_{,\text{GF}}^s] \right) - \left(\delta_{1,j} - \frac{1}{\lambda_j} [d_{1,IF} - \mu_j d_{1,\text{GF}}] \right) - z_{1,j}^s$	$j \geq 1; s \geq 1$
STEC	$\tilde{S}_r^s = S_r^s + d_{r,\text{GF}} - d_{,\text{GF}}^s$	$r \geq 1; s \geq 1$
Phase ambiguity	$\tilde{z}_{r,j}^s = (z_{r,j}^s - z_{r,j}^1) - (z_{1,j}^s - z_{1,j}^1)$	$j \geq 1; r \geq 2; s \geq 2$
S-basis	$dt_1, d_{r,1}, d_{r,2}, d_{,1}^s, d_{,2}^s, \delta_{1,j}, z_{r,j}^1, z_{1,j}^s$	$j \geq 1; r \geq 1; s \geq 1$
$d_{,IF} \equiv \frac{\mu_2}{\mu_2 - \mu_1} d_{,1} - \frac{\mu_1}{\mu_2 - \mu_1} d_{,2}$		
$d_{,\text{GF}} \equiv \frac{1}{\mu_2 - \mu_1} [d_{,2} - d_{,1}]$		

The 4D tomographic ionospheric model, also called 4DB-Tomion, assumes that the electron density is a linear combination of B-splines, namely

$$N_e(\mathbf{r}, t) = \sum_{ijk} c_{ijk}(t) \phi_i(\lambda) \cdot \phi_j(\theta) \cdot \phi_k(h) \quad (3)$$

where the term $\phi_i(\lambda) \cdot \phi_j(\theta) \cdot \phi_k(h)$ comes out from the tensor product of basis functions for each dimension, i.e. longitude, latitude and height, respectively, and $c_{ijk}(t)$ is the basis function coefficient that varies over time. Henceforth, unless stated otherwise, $c_{ijk} \equiv c_{ijk}(t)$.

By inserting Eq. (3) into Eq. (1) and then replacing S_r^s in Eq. (2), the following expression is obtained:

$$\tilde{S}_r^s = \int_{\mathbf{r}_r}^{\mathbf{r}^s} \sum_{ijk} c_{ijk}(t) \phi_i(\lambda) \cdot \phi_j(\theta) \cdot \phi_k(h) dl + d_{r,\text{GF}} - d_{,\text{GF}}^s \quad (4)$$

wherein the B-splines coefficients do not depend on the position, and thus, Eq. (4) can be further expanded as follows:

$$\tilde{S}_r^s = \sum_{ijk} c_{ijk} \int_{r_r}^{r^s} \phi_i(\lambda) \cdot \phi_j(\theta) \cdot \phi_k(h) dl + d_{r,GF} - d_{,GF}^s \quad (5)$$

where the integral can be numerically computed by Gauss-Legendre, e.g. see Zeilhofer (2008).

A particular challenge for the tomographic approach is handling empty voxels in the grid. If the selected grid is too fine for the available data, some voxels may not be crossed by any GPS signal, in which case the B-splines coefficients cannot be constrained by observables. This is particularly problematic in regions not that densely covered by ground GNSS CORS infrastructure, such as in the western part of NSW.

If the dynamic model of the Kalman filter was a physical model, the B-splines could still be constrained. However, the dynamic model selected is a random walk process, which means that the coefficients of empty voxels evolve randomly in time. In this regard, there are several ways to minimize or avoid this problem:

1. Adding more GNSS constellations, thus increasing the sample and providing higher resolution by reducing the size of the voxels.
2. Adding a background model.
3. By interpolation on the end-user side.

Adding a second constellation is currently work in progress; we have used GPS data only for this work. The second option, a background model, e.g. the IRI or GIMs from IGS, may yield typical VTEC RMS errors around 2 TECU (Hernandez-Pajares et al. 2017), thence they are not accurate enough to support GNSS-based navigation users at the cm-level.

The last option is the one presented in this work, a hybrid model, which is a combination of B-splines modelling on the network side and spatial interpolation of STEC on the end-user side. Such hybrid method is based on the following spatial interpolation formula:

$$S_{\text{user}}^s = \sum_a w_a S_a^s \quad (6)$$

where S_a^s corresponds to the STEC for each network receiver and it is computed by integration of the B-splines coefficients along the receiver-satellite line of sight as follows:

$$S_a^s = \int_{r_a}^{r^s} \sum_{ijk} c_{ijk} \phi_i(\lambda) \cdot \phi_j(\theta) \cdot \phi_k(h) dl \quad (7)$$

Note that the only coefficients c_{ijk} in Eq. (7) are those located in illuminated voxels; hence, they are all constrained by observations. Thus, instead of computing the B-splines coefficients of the voxels along the line of sight of the satellite-user pair, the user STEC is obtained by interpolation of the network STECs for the same satellite. Should all the voxels be crossed by GNSS data, only then S_{user}^s may be computed by direct integration of the B-splines coefficients, thereby avoiding the interpolation method in Eq. (6), and increasing the accuracy of the estimate. Such scenario may be expected with four simultaneously fully operational GNSS constellations.

The coefficients w_a can be computed in many different ways. For example, provided there are neither local maxima nor minima for the STEC in the region encompassed by the network, we can assume that

the contribution of each STEC on the end-user STEC decays as the square of the inverse of the distance from the user to each network receiver (Flanigan 1983). Thus, we can compute normalized weighted coefficients as follows:

$$w_a \equiv \frac{d_a^{-2}}{\sum_b d_b^{-2}} \quad (8)$$

where d_a (d_b) is the distance from the a th (b th) network receiver to the end-user and \sum_b runs over all receivers.

Note that Eq. (7) guarantees that all B-splines coefficients used by the end-user are constrained by data. Thus it is possible to increase the spatial resolution as much as we need, whereas the direct integration along the LOS of the end-user and the satellite would require larger voxels in order to constrain all B-splines coefficients, thereby decreasing the spatial resolution.

The hybrid approach can be compliant as well with the RTCM-SSR standards. Indeed, it would only require for the RTCM message to broadcast the B-splines coefficients (which are the state-space parameters) estimated by the network. The satellites coordinates are already broadcast by the navigation message. As for the network receiver coordinates, they could be provided by an external data set that would only require low-rate update.

The GNSS network is located in NSW (Australia); therefore, in order to account for the boundary conditions of the tomographic grid, endpoint-interpolating B-splines are used. For further details as to how to compute B-splines and their properties please see Zeilhofer (2008).

For the rest of the paper, the following notation is used: scalars are denoted by lower case, column vectors by bold face lower case, and matrices by bold face upper case. In addition, any vector with the hat symbol at the top (e.g. $\hat{\mathbf{a}}$) is an estimate of its corresponding stochastic variable.

The B-splines coefficients and hardware delay estimates are computed by a simple Kalman filter, which consists in a linear least squares (LLS) solution for a system of measurement and dynamic equations as follows:

$$\begin{cases} \mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \boldsymbol{\epsilon}_k \\ \mathbf{f}(\hat{\mathbf{x}}_{k-1}) = \mathbf{x}_k + \boldsymbol{\delta}_k \end{cases} \quad (9)$$

where \mathbf{y}_k stands for the vector of measurements at epoch t_k , \mathbf{x}_k represents the state-space vector at t_k , \mathbf{H}_k is the observation model or design matrix that maps any vector from the state space onto the measurements space, $\boldsymbol{\epsilon}_k$ is the measurement noise at t_k , $\mathbf{f}(\hat{\mathbf{x}}_{k-1})$ is the dynamic model that forecasts the value of the last estimate of the state-space vector, and $\boldsymbol{\delta}_k$ the system prediction noise at t_k .

The LLS solution for Eq. (9) is as follows (see the ‘‘Appendix’’ for the demonstration):

$$\hat{\mathbf{x}}_k = \mathbf{f}(\hat{\mathbf{x}}_{k-1}) + \mathbf{K}_k [\mathbf{y}_k - \mathbf{H}_k \mathbf{f}(\hat{\mathbf{x}}_{k-1})] \quad (10)$$

and the covariance of $\hat{\mathbf{x}}_k$ is

$$\mathbf{C}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{Q}_{\mathbf{f}(\hat{\mathbf{x}}_{k-1})} \quad (11)$$

where \mathbf{I} is the identity matrix, and \mathbf{K}_k , the gain matrix, is defined as follows:

$$\mathbf{K}_k \equiv \mathbf{Q}_f(\hat{\mathbf{x}}_{k-1}) \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{Q}_f(\hat{\mathbf{x}}_{k-1}) \mathbf{H}_k^T + \mathbf{E}_k \right)^{-1} \quad (12)$$

with \mathbf{E}_k being the covariance of the measurements \mathbf{y}_k .

The dynamic model considered for the B-splines coefficients is a random walk model, thence its covariance is $\mathbf{Q}_f(\hat{\mathbf{x}}_{k-1}) = \mathbf{C}_{k-1} + \mathbf{D}_{\delta_k}$, where \mathbf{D}_{δ_k} is the covariance of the noise of the dynamic model, which is considered zero mean and normally distributed.

As for the hardware delays estimates, they are considered constant during the computation, and with no process noise.

3 Data

The network receiver set, which is presented in Figure 1 and listed in Table 2, is formed by 21 GNSS continuously operating reference stations (CORS) belonging to CORSnet-NSW in New South Wales, Australia. CORSnet-NSW is a network of more than 200 CORS, which is built, owned and operated by Spatial Services, a unit of the NSW Department of Finance, Services and Innovation (DFSI).

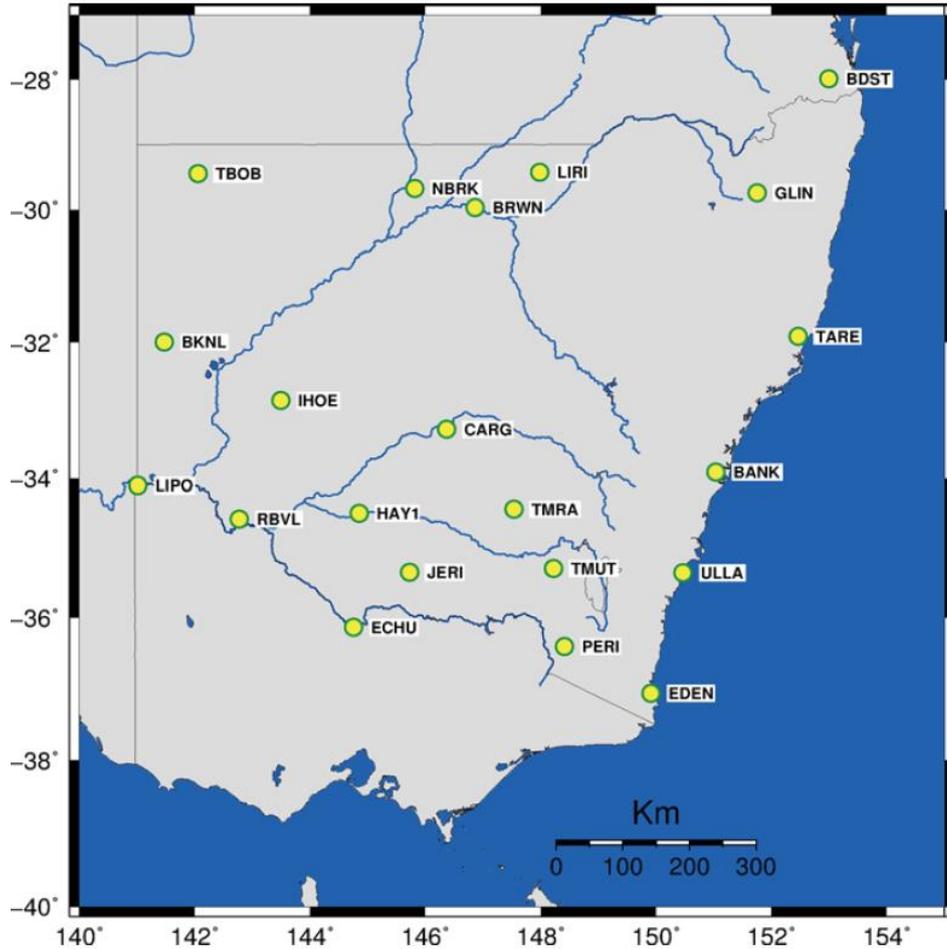


Figure 1: Distribution of the 21 network receivers in NSW, Australia.

Table 2: List of CORSnet-NSW network receivers used for ionospheric modelling in the post-fit analysis.

REC	1	2	3	4	5	6	7	8	9	10	
ID	BANK	BDST	BKNL	BRWN	CARG	ECHU	EDEN	GLIN	HAY1	IHOE	
REC	11	12	13	14	15	16	17	18	19	20	21
ID	JERI	LIPO	LIRI	NBRK	PERI	RBVL	TARE	TBOB	TMRA	TMUT	ULLA

The period considered goes from day of year (doy) 211 to doy 212 in 2016. The ionospheric conditions in those days were quite mild, with no strong activity in the previous days, namely the minimum disturbance storm time (Dst) for doys 211 and 212 was -14 and -12, respectively (values obtained from the *World Data Center for Geomagnetism* website).

The ionospheric conditions during these doys are representative of quiet conditions; hence, the conclusions of this paper cannot be extrapolated to storm conditions nor to areas near the equator which are affected by the Equator Anomaly.

The RINEX files of the stations for both doys were used as input data for the PPP-RTK platform. PPP-RTK platform computes the estimables shown in Table 1, i.e. tropospheric delay, satellites and receivers clock errors, phase biases, biased STEC and integer ambiguities. Further details on the PPP-RTK process can be found in Odijk et al. (2016).

We have only used GPS signals that pierce the grid from the top layer, thus guaranteeing that the STEC contains as much energy as possible from the ionosphere and plasmasphere. Otherwise, if the signals piercing from sideways of the grid were used for the ionospheric modelling, the electron density distribution would be shifted downwards.

3.1 Accuracy of the retrieved STEC data set

The ionospheric model is a data-driven computation; therefore, the accuracy of its ionospheric estimates depends on the quality of the input data, i.e. the biased STEC, S_r^s . Accordingly, we have conducted a first assessment on S_r^s .

For any given satellite simultaneously in view for two stations, the inter-station biased STEC difference, ΔS , provides a simple computation to assess the quality of S_r^s as follows:

$$\Delta S \equiv \tilde{S}_a^s - \tilde{S}_b^s = S_a^s - S_b^s + d_{a,GF} - d_{b,GF} \quad (13)$$

where subindexes a and b stand for each receiver, respectively, and $d_{a,GF}$ and $d_{b,GF}$ are their hardware delays. The satellite hardware delays cancel each other when computing the inter-station biased STEC. Moreover, for short baselines, $S_a^s \approx S_b^s$, thus yielding

$$\Delta S \approx d_{a,GF} - d_{b,GF} \quad (14)$$

which holds across all satellites.

Figure 2 shows ΔS for all satellites. Arcs tagged with different colours correspond to different satellites. For this example, two stations in Sydney were chosen, BANK and VLWD, on doy 061 in 2016, under quiet ionospheric conditions, and with a baseline below 7 km. Their inter-station hardware delays stay around -13 and -11 TECU. Despite the variability, it is much smaller as opposed to the inter-station-differenced STECs obtained by the widely used Phase-to-Code levelling method as shown in Figure 3. Indeed, the inter-station biased STEC varies across satellites from -13 to -8 TECU when the Phase-to-Code levelling is used.

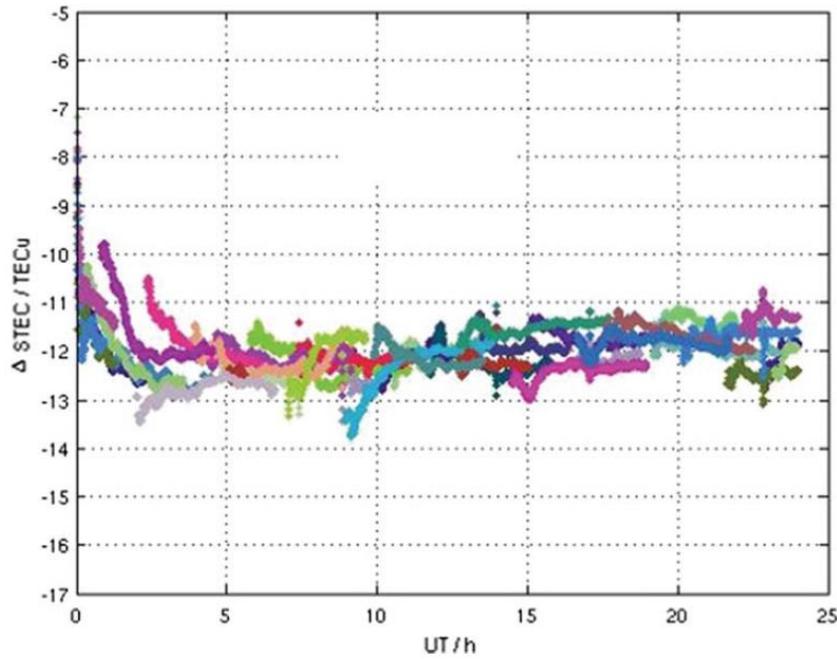


Figure 2: Inter-station biased STEC, Δ STEC, computed by PPP-RTK. Each colour represents the Δ STEC from a different satellite.

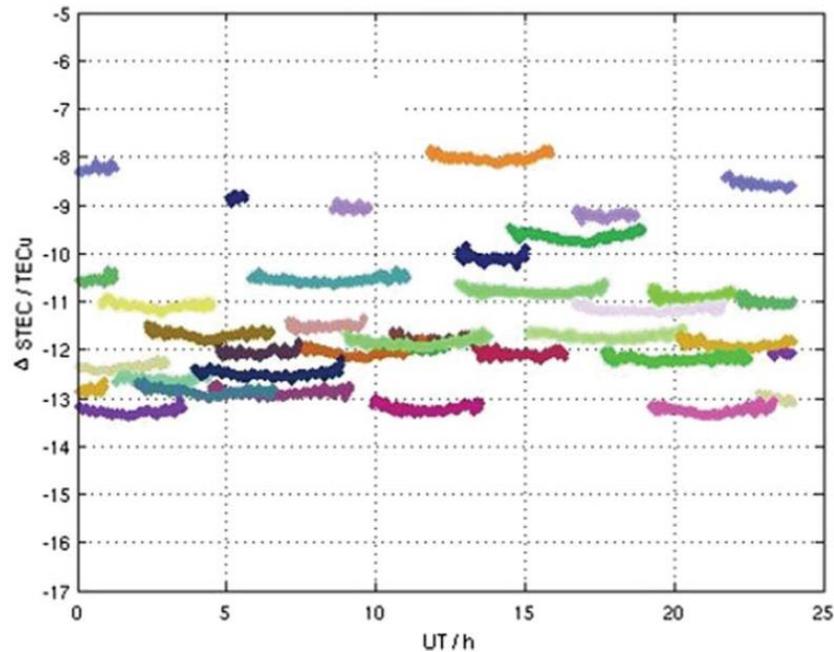


Figure 3: Inter-station biased STEC, Δ STEC, computed by the Phase-to-Code technique. Each colour represents the Δ STEC from a different satellite.

The reason for such difference in both STEC retrieval methods is the integer ambiguity fixed performed by the PPP-RTK method, thereby filtering out the noise from the code (pseudo-range) observable and using the phase only for STEC retrieval, which is two orders of magnitude less noisy than the former. On the other hand, the Phase-to-Code levelling uses phase and code (pseudo-range) observables to estimate the float ambiguity; hence, the higher variability is shown in Figure 3. Other authors have recently reported similar conclusions regarding both ambiguity fixing methods, e.g. Nie et al. (2018). Furthermore, we will see in the following sections that the dimension chosen for the ionospheric model plays as well an important role for the accuracy of the outputs in positioning.

4 Results and discussion

4DB-Tomion computes the B-splines coefficients, alongside with satellites and receivers hardware delays (namely, their geometry-free counterparts). The performance of the tomographic model has been assessed in terms of time to fix position (TTFP) for 28 rovers in kinematic mode across NSW. TTFP has been defined as the time it takes for the user to achieve position error RMS below 10 cm. For this, the Kalman filter of the PPP-RTK user has been reinitialized every epoch and then run for over 400 epochs.

For the computation, there are several parameters that must be previously defined and set-up, for example the size of the knots grid, the resolution level, i.e. the size of the voxels (and the distance between knots) and the order of the B-splines (e.g. Haar, linear, quadratic, cubic). The number N of B-splines depends on the resolution level n and the order m of the functions according to the following formula (Stollnitz et al. 1995):

$$N = (2^{n_\lambda} + m_\lambda - 1) \cdot (2^{n_\theta} + m_\theta - 1) \cdot (2^{n_h} + m_h - 1) \quad (15)$$

and the total number of voxels is $2^{n_\lambda} \cdot 2^{n_\theta} \cdot 2^{n_h}$, where the subindexes λ , θ , and h stand for longitude, latitude and height, respectively.

The tomographic grid is a set-up from $+100^\circ$ to $+200^\circ$ longitude, from -40° to $+10^\circ$ latitude, and from 50 to 1500 km height. In order to have the same resolution in latitude and longitude, thirty-two and sixteen cubic B-splines functions were evenly set-up for the longitude and latitude components, respectively. Therefore, according to Eq. (15), $n_\lambda = 5$, $n_\theta = 4$, and $m_\lambda = m_\theta = 4$. Due to the location of the ground stations, the tomographic method cannot provide high vertical resolution; hence, only two evenly distributed Haar functions were set-up, i.e. $n_h = m_h = 1$. Table 3 summarizes the set-up of the tomographic set and B-splines parameters.

Table 3: Input parameters for 4DB-Tomion software for the TTFP test. Each number in n and m corresponds to the resolution and the order parameters for longitude, latitude and height, respectively. $\Delta\lambda$, $\Delta\theta$ and Δh correspond to voxel dimensions in longitude, latitude and height, respectively.

Parameter	Hybrid Model
n	(5, 4, 1)
m	(4, 4, 1)
Δt	30 s
λ	[100° , $+200^\circ$]
θ	[-40° , $+10^\circ$]
h	[50 km, 1500 km]
$\Delta\lambda$	3.13°
$\Delta\theta$	3.13°
Δh	725 km

Time resolution is another parameter to consider. This is a set-up by the update rate for the Kalman filter, Δt . The time resolution depends very much on the ionospheric conditions. Indeed, during periods of strong space weather activity there may be enhancement of ionospheric gradients and time variability of the ionospheric electron density. Moreover, even in mild ionospheric conditions, the time variability of the ionospheric electron density may increase influenced by the equator anomaly. For this work, we have considered only the area in New South Wales in quiet conditions, thus neither the equator anomaly nor the aurorae activity (at high-latitude regions) has a strong influence. In this regard, since the ionospheric conditions are quite mild during the days analyzed, we have set up $\Delta t = 30$ s for all models and computations.

Finally, the time and spatial resolution determine the time needed for the Kalman filter to converge, which may require several hours of input data to achieve the steady state (for GNSS time series of 30 s update rate). Therefore, the results shown in this work correspond to the second day (doy 212), when the Kalman filter has finally achieved the steady state. The main goal of the ionospheric model is to support PPP-RTK end-users. In this regard, the metric for the performance assessment of the tomographic model is the time needed for the user to achieve a position residual RMS lower than 10 cm.

Figure 4 shows the cumulative distribution function (CDF) for the time to fix position (TTFP) in the horizontal plane for the *closed-loop* (aka *network* – green solid lines), the rovers with ionospheric corrections from 4DB-Tomion (blue solid lines), and the rovers with no external ionospheric corrections (aka *float* solution – red solid line).

The *closed-loop* solution corresponds to a scenario where the PPP-RTK user is given its own observed ionospheric corrections, in other words, it is the ideal scenario; therefore, it is taken as ground truth and reference for the assessment of the ionospheric corrections provided by the ionospheric model.

In the horizontal component, 90% of the computations achieve the 10-cm positioning threshold within 5 and 20 epochs for the *closed-loop* and ionospheric models, respectively.

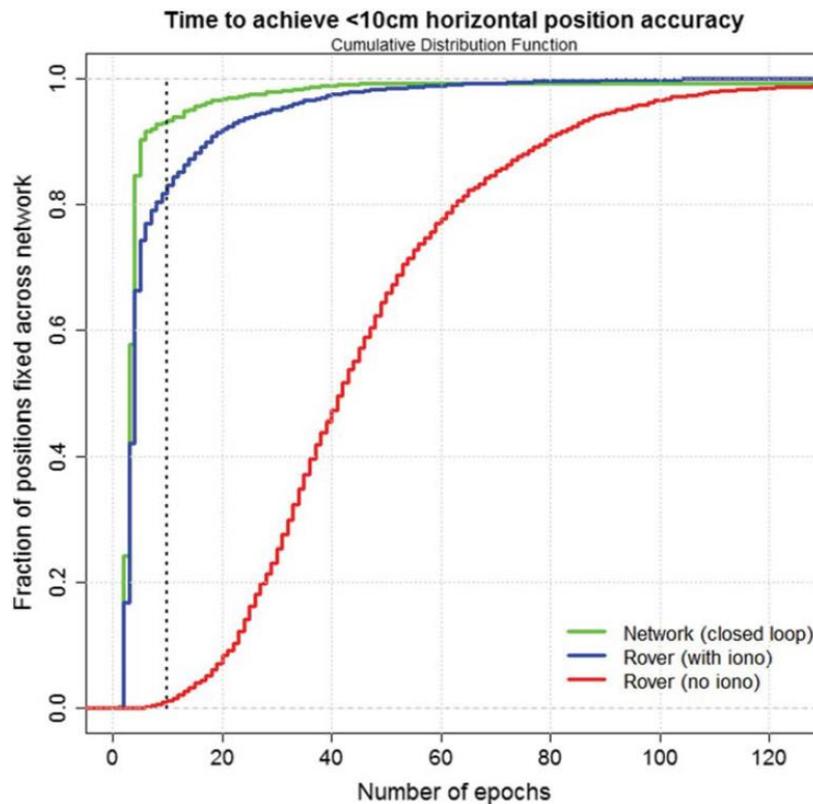


Figure 4: Cumulative Distribution Function of time to achieve horizontal position accuracy below 10 cm for the network or closed-loop solution (green solid line), for the rover with the ionospheric model (blue solid line), and the rover with no ionospheric corrections (red solid line).

Table 4 summarizes the results for TTFP in horizontal for all three approaches. Although not shown in this paper, we did not find any significant correlation between the locations of the rovers within the network and the CDFs for the TTFP.

Table 4: Results in epochs for Time To Fix (horizontal) Positioning (TTFP) for the *closed-loop* model, the ionospheric hybrid model, and the float model. Each epoch is 30 s.

	90% TTFP (epochs)
Network	5
3DB-Tomion	20
Float	80

Finally, we have also computed the TTFP using ionospheric corrections estimated by a 2D ionospheric model. Such 2D ionospheric model consisted of a B-splines set in longitude and latitude located at 450 km height. With such corrections, the PPP-RTK user was not capable of computing its position with an accuracy of less than 10 cm in less than 400 epochs. These results indicate that the mismodelling induced by the thin-shell approach may hamper highly precise positioning computations in real time.

5 Conclusions

We have presented a new four-dimensional tomographic ionospheric model, known as 4DB-Tomion, that supports PPP-RTK positioning and navigation. Such model uses a Kalman filter that computes in real time the ionospheric electron density as a linear combination of B-splines. The biased STEC input data are previously computed by a PPP-RTK network platform that decorrelates the integer ambiguity from the STEC. Such STEC output from the PPP-RTK network is still biased by satellite and receiver hardware delays. Then, the Kalman filter decorrelates the hardware delay from the physical STEC under the hypothesis that the former does not change as quickly as the ionospheric electron density within a period of several hours.

We have used the time to fix position for 28 PPP-RTK users in kinematic mode as metrics to assess the performance of the ionospheric model in quiet ionospheric conditions using a network of 21 receivers belonging to CORSnet-NSW in New South Wales, Australia.

We have compared the CDFs of the TTFP for the ionospheric model to the *closed-loop* and *float* models in quiet ionospheric conditions. The results from the TTFP show that the ionospheric model clearly outperforms the *float* solution mode and performs slightly worse than the *network* or *closed-loop* solution.

In the future, a multi-GNSS scenario will allow to increase the spatial resolution, thereby improving the STEC accuracy and the estimates on the end-user side. Eventually, spatial interpolation might not be required, and direct integration of the B-splines coefficients would provide enough accuracy to support PPP-RTK in real time.

Additionally, a quantitative comparison between local support and global support ionospheric models (e.g. between B-splines and spherical cap harmonics) as tools to support PPP-RTK users would provide further insight into accurate ionospheric modelling for GNSS-based navigation and positioning algorithms.

Finally, note that the conclusions of this work cannot be extrapolated to storm conditions and hence it is still necessary to assess the performance of 4DB-Tomion in periods of strong space weather dynamics.

Acknowledgements

This work has been supported by the Cooperative Research Centre for Spatial Information, whose activities are funded by the Business Cooperative Research Centres Programme.

Appendix

Equation (9) can be rearranged into one single equation as follows:

$$\begin{bmatrix} y_k \\ \mathbf{f}(\hat{\mathbf{x}}_{k-1}) \end{bmatrix} = \begin{bmatrix} \mathbf{H}_k \\ \mathbf{I} \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ \boldsymbol{\delta}_k \end{bmatrix} \quad (16)$$

The optimal solution (in a LLS sense) for Eq. (16) corresponds to the argument \mathbf{x}_k that minimizes the error, i.e.

$$\hat{\mathbf{x}}_k = \arg \min_{\mathbf{x}_k} \|\mathbf{z}_k - \boldsymbol{\Omega}_k \mathbf{x}_k\|_{\boldsymbol{\Sigma}_k}^2 \quad (17)$$

whose solution is as follows:

$$\hat{\mathbf{x}}_k = \left[\boldsymbol{\Omega}_k^T \boldsymbol{\Sigma}_k^{-1} \boldsymbol{\Omega}_k \right]^{-1} \left[\boldsymbol{\Omega}_k^T \boldsymbol{\Sigma}_k^{-1} \mathbf{z}_k \right] \quad (18)$$

where $\mathbf{z}_k \equiv [\mathbf{y}_k^T, \mathbf{f}^T(\hat{\mathbf{x}}_{k-1})]^T$; $\boldsymbol{\Omega}_k$ is the matrix that maps the state-space estimates onto the observations and onto the a priori state-space estimates,

$$\boldsymbol{\Omega}_k = \begin{bmatrix} \mathbf{H}_k \\ \mathbf{I} \end{bmatrix} \quad (19)$$

which is always full rank due to the presence of the identity matrix.

The covariance of the state-space estimates is

$$\boldsymbol{\Sigma}_k \equiv \begin{bmatrix} \mathbf{E}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{\mathbf{f}(\hat{\mathbf{x}}_{k-1})} \end{bmatrix} \quad (20)$$

where \mathbf{E}_k is the covariance of the measurements y_k , and $\mathbf{Q}_{\mathbf{f}(\hat{\mathbf{x}}_{k-1})}$ is the covariance of the stochastic dynamic model $\mathbf{f}(\hat{\mathbf{x}}_{k-1})$.

The covariance of the estimates $\hat{\mathbf{x}}_k$ is as follows:

$$\mathbf{C}_k = \left[\boldsymbol{\Omega}_k^T \boldsymbol{\Sigma}_k^{-1} \boldsymbol{\Omega}_k \right]^{-1} \quad (21)$$

which, after developing the right-hand side, yields

$$\mathbf{C}_k = \left[\left(\mathbf{Q}_{\mathbf{f}(\hat{\mathbf{x}}_{k-1})} \right)^{-1} + \mathbf{H}_k^T \mathbf{E}_k^{-1} \mathbf{H}_k \right]^{-1} \quad (22)$$

The matrix inversion lemma, e.g. Stengel (1994), states that

$$\left[\left(\mathbf{Q}_{\mathbf{f}(\hat{\mathbf{x}}_{k-1})} \right)^{-1} + \mathbf{H}_k^T \mathbf{E}_k^{-1} \mathbf{H}_k \right]^{-1} = \mathbf{Q}_{\mathbf{f}(\hat{\mathbf{x}}_{k-1})} - \mathbf{Q}_{\mathbf{f}(\hat{\mathbf{x}}_{k-1})} \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{Q}_{\mathbf{f}(\hat{\mathbf{x}}_{k-1})} \mathbf{H}_k^T + \mathbf{E}_k \right)^{-1} \mathbf{H}_k \mathbf{Q}_{\mathbf{f}(\hat{\mathbf{x}}_{k-1})} \quad (23)$$

thence, after replacing Eq. (23) in Eq. (22), it leads to the following expression:

$$\mathbf{C}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{Q}_f(\hat{\mathbf{x}}_{k-1}) \quad (24)$$

where \mathbf{K}_k , the gain matrix, is defined as follows

$$\mathbf{K}_k \equiv \mathbf{Q}_f(\hat{\mathbf{x}}_{k-1}) \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{Q}_f(\hat{\mathbf{x}}_{k-1}) \mathbf{H}_k^T + \mathbf{E}_k \right)^{-1} \quad (25)$$

Finally, by using Eq. (24) in Eq. (18), and noting that

$$\mathbf{I} = \left(\mathbf{H}_k \mathbf{Q}_f(\hat{\mathbf{x}}_{k-1}) \mathbf{H}_k^T + \mathbf{E}_k \right)^{-1} \left(\mathbf{H}_k \mathbf{Q}_f(\hat{\mathbf{x}}_{k-1}) \mathbf{H}_k^T + \mathbf{E}_k \right) \quad (26)$$

the following expression is obtained:

$$\hat{\mathbf{x}}_k = \mathbf{f}(\hat{\mathbf{x}}_{k-1}) + \mathbf{K}_k [\mathbf{y}_k - \mathbf{H}_k \mathbf{f}(\hat{\mathbf{x}}_{k-1})] \quad (27)$$

which is Eq. (10) presented in Section 2.

References

- Banville S (2014) Improved convergence for GNSS precise point positioning. Ph.D. dissertation, Department of Geodesy and Geomatics Engineering, Technical Report No. 294, University of New Brunswick, Fredericton, New Brunswick, Canada.
- Chambodut A, Panet I, Mandeau M, Diament M, Holschneider M, Jamet O (2005) Wavelet frames: An alternative to spherical harmonic representation of potential fields. *Geophys J Int* 163(3): 875-899. <https://doi.org/10.1111/j.1365-246X.2005.02754.x>.
- Feltens J, Angling M, Jackson-Booth N, Jakowski N, Hoque M, Hernandez-Pajares M, Aragon-Angel A, Orus R, Zandbergen R (2011) Comparative testing of four ionospheric models driven with GPS measurements. *Radio Sci* 46(6): 1-11.
- Flanigan FJ (1983) Complex variables. Dover books on mathematics. Dover Publications, Mineola.
- Hadas T, Krypiak-Gregorczyk A, Hernández-Pajares M, Kaplon J, Paziewski J, Wielgosz P, Garcia-Rigo A, Kazmierski K, Sosnica K, Kwasniak D, Sierny J, Bosy J, Pucilowski M, Szyszko R, Portasiak K, Olivares-Pulido G, Gulyaeva T, Orus-Perez R (2017) Impact and implementation of higher-order ionospheric effects on precise GNSS applications. *J Geophys Res Solid Earth*. <https://doi.org/10.1002/2017JB014750>.
- Haines GV (1988) Computer programs for spherical cap harmonic analysis of potential and general fields. *Comput Geosci* 14(4): 413-447.
- Hernandez-Pajares M, Juan JM, Sanz J (1999) New approaches in global ionospheric determination using ground GPS data. *J Atmos Solar Terr Phys* 61: 1237-1247.
- Hernandez-Pajares M, Juan JM, Sanz J, Aragon-Angel A, Garcia-Rigo A, Salazar D, Escudero M (2011) The ionosphere: Effects, GPS modeling and the benefits for space geodetic techniques. *J Geod*. <https://doi.org/10.1007/s00190-011-0508-5>.
- Hernandez-Pajares M, Juan JM, Sanz J, Orús R (2014) Second-order ionospheric term in GPS: Implementation and impact on geodetic estimates. *J Geophys Res Solid Earth* 112: B08417. <https://doi.org/10.1029/2006JB004707>.
- Hernandez-Pajares M, Roma-Dollase D, Krankowski A, Garcia-Rigo A, Orus-Perez R (2017) Methodology and consistency of slant and vertical assessments for ionospheric electron content models. *J Geod*. <https://doi.org/10.1007/s00190-017-1032-z>.
- Jakowski N (2008) Ionospheric impact on GNSS signals. *Fisica de la Tierra* 20: 1125.
- Khodabandeh A, Teunissen PJG (2015) An analytical study of PPP-RTK corrections: Precision, correlation and user-impact. *J Geod* 89(11): 1109-1132.

- Ma G, Maruyama T (2003) Derivation of TEC and estimation of instrumental biases from GEONET in Japan. *Annal Geophys* 21: 2083-2093.
- Mautz R, Ping J, Heki K, Schaffrin B, Shum C, Potts L (2005) Efficient spatial and temporal representation of global ionosphere maps over Japan using B-spline wavelets. *J Geod* 78: 660-667. <https://doi.org/10.1007/s00190-004-0432-z>.
- Nie W, Xu T, Rovira-Garcia A, Juan Zornoza JM, Sans Subiran J, González-Casado G, Chen W, Xu G (2018) Revisit the calibration errors on experimental slant total electron content (TEC) determined with GPS. *GPS Solut* 22: 85. <https://doi.org/10.1007/s10291-018-0753-7>.
- Odiijk D (2002) Fast precise GPS positioning in the presence of ionospheric delays. *Publications on geodesy*, vol 52. The Netherlands Geodetic Commission, Delft. ISBN-13: 978 90 6132 278 8.
- Odiijk D, Khodabandeh A, Nadarajah N, Choudhury M, Zhang B, Li W, Teunissen PJG (2016) PPP-RTK by means of S-system theory: Australian network and user demonstration. *J Spat Sci* 62: 3-27.
- Petrie EJ, King MA, Moore P, Lavallée DA (2010) A first look at the effects of ionospheric signal bending on a globally processed GPS network. *J Geod* 84(8): 491-499. <https://doi.org/10.1007/s00190-010-0386-2>.
- Petrie EJ, Hernandez-Pajares M, Spalla P, Moore P, King MA (2011) A review of higher order ionospheric refraction effects on dual frequency GPS. *Surv Geophys* 32(3): 197-253. <https://doi.org/10.1007/s10712-010-9105-z>.
- Schaer S (1999) Mapping and predicting the Earth's ionosphere using the global positioning system. Ph.D. dissertation, Astron Institute, University of Bern, Berne.
- Schmidt M (2007) Wavelet modelling in support of IRI. *Adv Space Res* 39: 932940. <https://doi.org/10.1016/j.asr.2006.09.030>.
- Schmidt M, Bilitza D, Shum C, Zeilhofer C(2008) Regional 4-D modelling of the ionospheric electron density. *Adv Space Res* 42: 782790. <https://doi.org/10.1016/j.asr.2007.02.050>.
- Schmidt M, Dettmering D, Mößmer M, Wang Y, Zhang J (2011) Comparison of spherical harmonic and B spline models for the vertical total electron content. *Radio Sci* 46: RS0D11. <https://doi.org/10.1029/2010RS004609>.
- Stengel RF (1994) Optimal control and estimation. Dover Publications Inc., New York.
- Stollnitz EJ, DeRose TD, Salesin DH (1995) Wavelets for computer graphics: A primer 2. *IEEE Comput Graph Appl* 15(4): 75-85.
- Teunissen PJG, Khodabandeh A (2015) Review and principles of PPP-RTK methods. *J Geod* 89(3): 217240.
- Teunissen PJG, Kleusberg A (1998) GPS for geodesy. Springer, Berlin.
- Teunissen PJG, Montenbruck O (2017) Handbook of global navigation satellite systems. Springer, Berlin. ISBN 978-3-319-42926-7. <https://doi.org/10.1007/978-3-319-42928-1>.
- Themens DR, Jayachandran PT, Langley RB, MacDougall JW, Nicolls MJ (2012) Determining receiver biases in GPS-derived total electron content in the auroral oval and polar cap region using ionosonde measurements. *GPS Solut* 17: 357.
- Zeilhofer C (2008) Multi-dimensional B-spline modeling of spatiotemporal ionospheric signals, Reihe A123, Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften.
- Zhang B (2016) Three methods to retrieve slant total electron content measurements from ground-based GPS receivers and performance assessment. *Radio Sci* 51: 972988. <https://doi.org/10.1002/2015RS005916>.
- Zumberge JF, Heflin MB, Jefferson DC, Watkins MM, Webb FH (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. *J Geophys Res* 102: 50055017.

Author Posting. © Journal of Geodesy, 2019.

This is the authors' version of the work. The definitive version was published in the *Journal of Geodesy*, Volume 93, Number 9, September 2019, pp. 1673-1683. <https://doi.org/10.1007/s00190-019-01276-4>.