Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 68 (2021) 4794–4805

ADVANCES IN SPACE RESEARCH (a COSPAR publication)

www.elsevier.com/locate/asr

Towards tropospheric delay estimation using GNSS smartphone receiver network

Tiago De Oliveira Marques [a,*], Maija Mäkelä [b], Leslie Montloin [a], Terhi Lehtola [c], Sarang Thombre [b], Ville Lehtola [b,d]

[a] AIRBUS Defence and Space, 31 Rue des Cosmonautes, 31400 Toulouse, France
[b] Finnish Geospatial Research Institute, National Land Survey, PO Box: 84, 00521 Helsinki, Finland
[c] Vaisala Oyj, PB 26, 00421 Helsinki, Finland
[d] EOS department, ITC faculty, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

Received 16 January 2020; received in revised form 20 September 2020; accepted 23 September 2020
Available online 9 October 2020

Abstract

Information on the water vapor distribution of the troposphere is useful for weather monitoring and forecast. Water vapor distribution can be estimated from tropospheric delays produced by high-grade Global Navigation Satellite System (GNSS) receivers. This type of techniques is currently used in the data assimilation process of Numerical Weather Prediction (NWP) models, especially, for the limited areas covered by these high-grade GNSS networks. We consider a new collaborative crowdsourcing-based alternative for obtaining these GNSS meteorology measurements. It relies on a GNSS smartphone receiver network, and hence promises to expand the use of GNSS meteorology techniques into areas not covered by high-end receiver networks. In order to assess the feasibility of estimating the troposphere water vapor distribution using such receiver networks, it is proposed a system architecture that supports the troposphere water vapor distribution estimation using a smartphone network. Next, it is presented the simulator test-bed that has been developed to emulate the proposed system in a representative way and to assess the system performances. The main motivation behind the simulator is that it provides a controlled environment for testing our method.

Keywords: GNSS; Troposphere; Crowdsourcing; Smartphones

1. Introduction

In order to improve weather monitoring and forecast, Numerical Weather Prediction (NWP) models currently assimilate data that describe the water vapor distribution of the troposphere. The accuracy of the water vapor distribution estimate itself, but also the resolution in horizontal and vertical directions, is important to precisely estimate the atmospheric processes. World Climate Research Programme (Kley et al., 2000) has assessed the different atmospheric water vapour measurement techniques with respect to accuracy and vertical resolution.

In general, satellite-based methods require atmospheric calibration, while ground-based and in-situ methods (Kämpfer, 2013) may be handicapped by the lack of ground networks. Council (2009) identified vertical accuracy of atmospheric humidity measurements as a gap for weather and climate. In the Authors' experience, commercial off the shelf (COTS) radiosondes equipped with

* Corresponding author.
E-mail addresses: tiago.marques@airbus.com, tiagomomarques@msn.com (T. De Oliveira Marques), maija.makela@maanmittauslaitos.fi (M. Mäkelä), leslie.montloin@airbus.com (L. Montloin), terhi.lehtola@vaisala.com (T. Lehtola), sarang.thombre@nls.fi (S. Thombre), v.v.lehtola@utwente.nl (V. Lehtola).

https://doi.org/10.1016/j.asr.2020.09.041
0273-1177/

hygrometers represent the commonly used state of the art for the best accuracy (2–4% relative humidity) and the best vertical resolution (about 1 Hz measurement frequency while airborne and raising with a 6 m/s speed on average). While providing good accuracy and vertical resolution, the sparse horizontal resolution and the operational costs of radiosondes have sparked needs to develop other techniques, e.g. LIDARs (Hicks-Jalali et al., 2019), ground-based radiometric profilers and techniques based on Global Navigation Satellite System (GNSS).

The meteorological application of GNSS techniques is well-established in areas, where a dense network of GNSS base stations exists, e.g. Europe (Guerova et al., 2016). Additionally, GNSS-based techniques are powerful in limited area NWP models that require the prediction of local phenomena, such as convective storms or fog-events (Haji-Aghajany and Amerian, 2017). In these meteorological applications, however, the estimation of tropospheric water vapor distribution relies on networks of high-grade (geodetic) GNSS receivers (Brenot et al., 2014; Bender et al., 2011). Thus, the state of the art in the meteorological application of GNSS techniques is limited to areas with high-end measurement networks, which is not the case for most of the world.

GNSS smartphone receivers, however, could provide additional data that would – in the best scenario – significantly improve the coverage of GNSS-based meteorology techniques. The number of smartphones equipped with chipset GNSS receivers and meteorological sensors, such as barometers, is currently experiencing strong growth. The most recent chipsets are multi-constellation receivers (Samsung S8 and Huawei P10), and the first multi-frequency smartphone is on the market (Xiaomi Mi8) is equipped with a Broadcom BCM47755 chip.

In order to assess the feasibility of estimating the water vapor distribution of the troposphere using GNSS and meteorological observables from a high number of GNSS smartphone receivers, this paper proposes an innovative algorithm that estimates tropospheric wet delays and outputs zenith total delay (ZTD) estimates for NWP purposes. The ZTD estimation is based on a collaborative crowdsourcing-based approach. This algorithm is called Weather Monitoring Collaborative Crowdsourcing (WMCC) algorithm in the following. Next, this paper assesses the performances of this algorithm. For this purpose, a simulator test-bed has been developed to test the WMCC system performances in a representative way. This test-bed is presented in the paper.

The novelty of this paper lies in the development of a WMCC simulator test-bed that enables performing experiments in a controlled environment and running scenarios that are hard to represent using a fully experimental approach. As examples, the test-bed enables controlling the emergence of new receivers and disappearance of old receivers, to simulate a so-called dynamic crowdsourcing receiver network.

This paper aims to describe both the WMCC system architecture and the simulator test-bed. Section 2 introduces the state of the art on the GNSS-based troposphere water vapor distribution estimation techniques. It justifies the approach selected to support the collaborative crowdsourcing troposphere estimation technique, namely the tomographic approach. Section 3 presents the WMCC system architecture and the simulator test-bed developed to assess the WMCC system performances. Section 4 describes the models and algorithms implemented in the test-bed to simulate the WMCC system in a representative way and in a controlled environment. Finally, Section 5 provides the first WMCC system performance results.

2. GNSS-based tropospheric estimation

GNSS signals received by the user equipment are delayed by the neutral atmosphere (hydrostatic component), and by the local weather (wet component). The wet component corresponds to the water vapor refraction experienced by the radio signals. Turning this phenomenon around, if the contribution of these effects in the received GNSS signals can be determined accurately, a direct mapping to the local weather conditions can be performed. Hence, numerical weather prediction (NWP) systems utilized in metrological organizations are assimilating the ZTD obtained from the GNSS ground stations for improving a model of the troposphere over a specific domain. These NWP models are then used for now-casting and forecasting, but also to improve GNSS positioning accuracy. Böhm et al. (2012) writes that the NWP augmentation using Vienna Mapping Functions increases the prediction accuracy of the conventional approach with up to 10%.

However, the reliance on GNSS base stations has its shortcomings. Due to the uneven distribution of such monitoring stations in different parts of the world, there can be considerable spatial distances between neighbouring data sources. In other words, the horizontal resolution is hampered by the lack of density of the receiver network. For example, in France, the permanent GNSS stations are about 50 km from each other (Champollion et al., 2005). In such cases, the accuracy of the predicted weather deteriorates as one moves further away from a monitoring station. This research activity, therefore, aims to assess the feasibility of increasing the spatial resolution by crowdsourcing GNSS observables from a high-density network also containing end-user smartphones.

Increasing the network density is, however, not enough. The ZTD-based conventional GNSS meteorology technique runs into trouble by smoothing out local-scale fluctuations of the water vapour (Oigawa et al., 2018). This is because all the received satellite signals are simplified into one ZTD value for each receiver, by using a parametrized model e.g. VMF1 or Foelsche & Kirchengast (F&K) MF (Foelsche and Kirchengast, 2002), see Fig. 1(a). In other

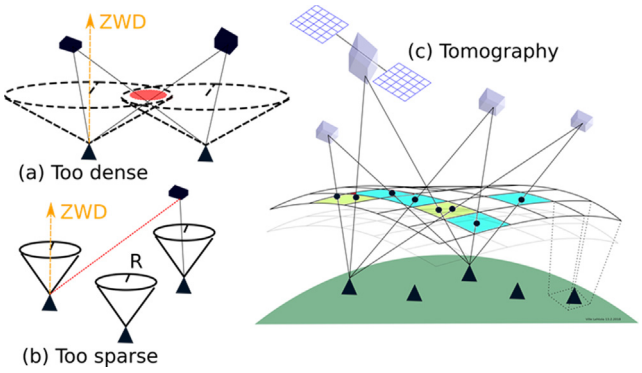

Fig. 1. (a) Conventional GNSS meteorology technique using e.g. 10 degrees elevation cut-off angle. (b) The signal from the highest elevation angle satellite only for each receiver. (c) Tomographic approach uses all data from all receivers.

words, each receiver can be thought to average measurements from an inverted cone with a base radius of R and a typical elevation cut-off of e.g. 10 degrees. With a dense network, this leads to overlapping of the cones, as data is captured partly from the same spatial domain. Lindskog et al. (2017) reduces the available data, i.e., they use a thinning distance to down-sample the station network. But their approach is not favourable when the idea is to increase both the accuracy and the resolution of the estimate.

One attempt to increase the local-scale accuracy and avoid this overlapping problem has been to exclude a major part of the data so that only the signals from the highest elevation angle satellite(s) are utilized (Bender et al., 2011), see Fig. 1(b). However, this is only an ad-hoc solution, as then most of the data is discarded and problems related to horizontal accuracy may persist. For example, if the spatial distribution of the GNSS receivers is inhomogeneous, the inverse cones may be too sparsely spread in some areas.

If a dense network of receivers is employed, it would be optimal to retrieve Slant Wet Delays (SWD) from GNSS data (Kawabata et al., 2013) which contain more information about 3-dimensional distribution of the water vapour and is, therefore, more accurate. This information is more relevant to local weather prediction. The ZTDs can be extracted without overlap, following the resolution of the NWP. Furthermore, SWDs that are especially useful for predicting thunderstorms and/or rainbands (Kawabata and Shoji, 2018) can be separately extracted for those purposes. Regardless of the assimilation details, the 3D distribution of the water vapor can be estimated using tomographic techniques. If this estimation is done probabilistically, the altering spatial accuracy should also be correctly represented in the posterior distributions. In other words, the estimate contains the knowledge of its own uncertainty.

Furthermore, as we are considering crowdsourcing, this uncertainty needs also to include the errors originating from the measurement devices, i.e., GNSS receivers (see Tables 1–3). These uncertainties have been studied in Lehtola et al. (2019) and the references there-in.

The best way is to utilize all the data, in a way that does not run into problems regardless of whether the GNSS receiver network is dense, sparse, or inhomogeneous. One such technique is tomography. The tomographic grid makes use of all data by distributing the measured water vapor refraction in 3D via ray-tracing (Haji-Aghajany and Amerian, 2017). The expected GNSS receiver quality is a key parameter of the proposed tomography-based algorithm. GNSS data characterized by a poor expected quality is still used in the tomographic algorithm, but it is underweighted. This allows for the computation of slant wet delays for new receivers that are added into the network, for precise point positioning (PPP) purposes. The discretization of the tomographic grid is adjustable and defines the spatial resolution.

For a sparse receiver network, the tomographic approach resembles the conventional GNSS meteorology but outperforms it for dense networks. Inhomogeneous networks can be handled with a regularization that balances the measurements also to grid cells that lack observations.

Our end goal is to obtain an accurate estimate that describes the tropospheric water vapour distribution. In this paper, we utilize a feed-forward implementation of PPP and tomography. In other words, the position of each receiver is calculated first, and the estimates for each receiver are given as input to the tomographic module. This implementation is detailed in Section 4.

The novelty of our paper is neither in the PPP, which is well known, nor in the tomography per se, as several tomography studies exist as well (Brenot et al., 2014; Bender et al., 2011; Oigawa et al., 2018; Bokoye et al., 2003; Flores et al., 2000; Rohm and Bosy, 2011). Rather, the novelty lies in performing experiments in a controlled environment to study the impact of different receiver densities with respect to tomographic grid resolution. Additionally, simulations allow for the control of the network homogeneity both in static and dynamic sense. We can control the emergence of new receivers and the disappearance of old receivers, to simulate a so-called dynamic crowdsourcing receiver network. Our output is given in the form of ZTD for NWP purposes.

3. Simulator TEST-BED

We describe here the design and implementation of the test-bed that simulates the WMCC system. The proposed WMCC system architecture consists of a unique Central Processing Facility (CPF) that collects the GNSS carrier

Table 1
Satellite clock and ephemeris error model parameters.

σ_{radial}	σ_{cross}	σ_{along}	$\sigma_{clocksat}$
5 cm	5 cm	5 cm	3 ns

Table 2
GNSS receivers used for error model development (Lehtola, 2019).

Simulated receiver grade	Receiver	Multipath environment	Magnitude of carrier phase noise + multipath error
Low grade (Smartphone)	Samsung S8	Open sky + semi-urban	dm
Medium grade	uBlox NEO-M8T with Patch antenna	Open sky + semi-urban	cm
High grade	Septentrio POLARX5S with Leica antenna	Open sky + semi-urban	mm

Table 3
Barometer pressure noise model parameters from Bosh BME 280 datasheet.

Model parameters	α_1	β_1	α_2	β_2	$\sigma_{p,\eta}(T = 25\ ^{\circ}\mathrm{C})$
Parameter values	0.1429	−3.5714	0.5	−12.5	0.2 Pa

phase measurements from all receivers of the network. Part of the pre-processing tasks, such as the GNSS error mitigation (e.g. the ionospheric delay) or the meteorological data (pressure/temperature) estimation at the GNSS receiver antenna, can be performed individually by the smartphones. The troposphere water vapor distribution estimation is made by the CPF.

The test-bed developed to assess the performances of the WMCC system is depicted in Fig. 2. The true GNSS ranges from the GNSS satellites to the GNSS receivers of the network for three constellations (GPS, GALILEO, GLONASS) are set in GNSS data sets.

Presented hereafter the list of the main configurable parameters:

- Simulation time
- Multipath environment (open-sky or urban)
- Meteorological model (AROME/ARPEGE)
- GNSS data rate
- GNSS constellations (GPS, GALILEO, GLONASS or a combination of these constellations)
- GNSS frequencies
- Data failures (duty cycle or network failure)
- Receiver grade

The test-bed main blocks are the following:

- The Configuration Generator block is in charge of the definition (using the configuration file), and of the generation of all system configuration parameters;
- The GNSS Errors and Meteorological Data Generator block is in charge of the generation of the GNSS errors. The GNSS errors, combined with its associated true GNSS ranges from the data set, reproduce the estimated observables (phase pseudo-range) such as estimated by the GNSS smartphone receiver. Moreover, the module is responsible of the generation of the meteorological data (temperature, pressure) such a measured by the GNSS smartphone receiver;
- The Network Simulator is in charge of the simulation of the data transfer between the receivers and the CPF;
- The Central Processing Facility (CPF) is in charge of the estimation of the tropospheric delay;

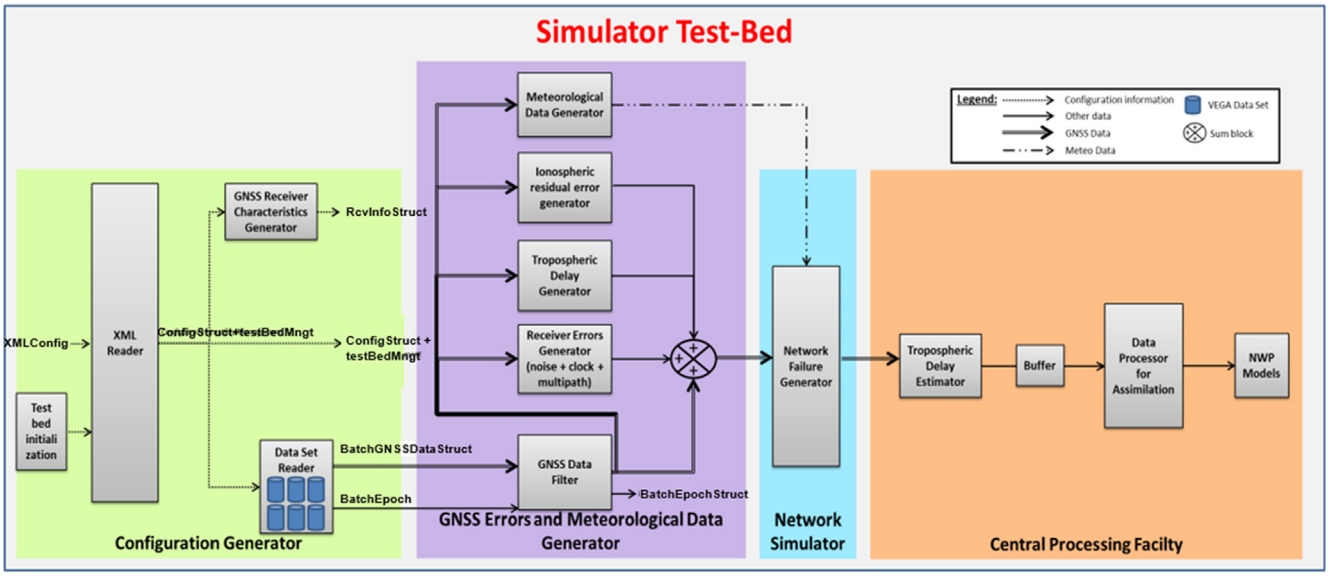

Fig. 2. Simulator Test-bed Architecture.

- The External block represents all the elements that the test-bed needs but which do not belong to the test-bed (such as GNSS data sets, meteorological files and NWP real data to generate GNSS tropospheric delays).

4. Test-bed models and algorithms

This section describes the algorithms implemented in the test-bed to simulate GNSS carrier phase measurements and meteorological data estimated at the smartphone position, and to estimate the tropospheric delays using these data.

4.1. Receiver spatial distribution

The receiver position spreading replicates a local and regional irregular distribution of several hundreds of receivers inside areas of diameter ranging between 10 km and 250 km.

The generated location considers the Earth's shape. The locations are projected on the reference ellipsoid and next from the reference ellipsoid to the terrain model. For the reference ellipsoid the World Geodetic System 1984 (WGS84) model is used and for the terrain model the digital elevation model of the AROME/ARPEGE model is used, as represented in Fig. 3.

The principle of the algorithm can be summed up in three steps: generation of the different simulation geographic areas, generation of the receiver locations and projection of the locations on the Earth surface model.

4.2. GNSS errors

One of the objectives of the test-bed is to simulate GNSS carrier phase measurements from the network of smartphone receivers in a representative way. The GNSS carrier phase measurements are reconstructed in the test-bed as follows:

$$\begin{aligned}\phi^{s_i}_{Lk,r_j}(t_k) = {} & \rho^{s_i}_{0,r_j}(t_k) + c{*}dt_{r_j}(t_k) - c * dt^{s_i}(t_k) \\ & + STD^{s_i}_{r_j}(t_k) + Eph^{s_i}_{r_j}(t_k) - I^{s_i}_{Lk,r_j}(t_k) \\ & - \lambda_{Lk} A^{s_i}_{Lk,r_j}(t_k) + \eta^{s_i}_{Lk,\phi,r_j}(t_k) \end{aligned} \quad (1)$$

where:

- $\phi^{s_i}_{Lk,r_j}(t_k)$ is the carrier phase measurement between receiver r_j and satellite s_i at time t_k for frequency band Lk [m],
- c is the speed of light [m/s],
- $dt_{r_j}(t_k)$ is the receiver clock offset [s],
- $dt^{s_i}(t_k)$ is the ranging error due to the satellite clock offset error (after correction) [s],
- $STD^{s_i}_{r_j}(t_k)$ is the tropospheric Slant Total Delay [m],
- $I^{s_i}_{Lk,r_j}(t_k)$ is the ionospheric delay estimation error (after correction) [m],
- $\eta^{s_i}_{Lk,\phi,r_j}(t_k)$ is the receiver thermal noise + multipath on the carrier phase measurement [m],
- $Eph^{s_i}_{r_j}(t_k)$ is the ranging error induced by ephemeris error (after correction) [m],
- $A^{s_i}_{Lk,r_j}(t_k)$ is the (integer) carrier phase ambiguity [cycle],
- λ_{Lk} is the wavelength on frequency Lk [m].
- $\rho^{s_i}_{0,r_j}(t_k)$ is the true range between receiver r_j and satellite s_i [m].

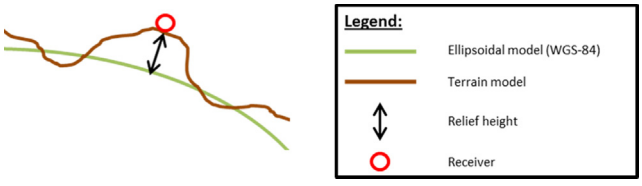

Fig. 3. Earth models used to generate receiver locations.

The GNSS errors and meteorological data generator aims to compute $I^{s_i}_{Lk,r_j}$, $STD^{s_i}_{r_j}$, $Eph^{s_i}_{r_j}$, $\eta^{s_i}_{Lk,\phi,r_j} dt_{r_j}$. These computations are further detailed in this section.

4.2.1. Ephemeris and satellite clock error *Eph* and dt^{s_i}

$Eph^{s_i}_{r_j}(t_k)$is computed as follows:

$$Eph^{s_i}_{r_j}(t_k) = \boldsymbol{u}^{s_i}_{r_j}(t_k) * \left(\boldsymbol{p}^{s_i}_{brdc}(t_k) - \boldsymbol{p}^{s_i}_{true}(t_k)\right) \quad (2)$$

Where:

- $\boldsymbol{p}^{s_i}_{true}(t_k)$ is the satellite position (true position) at the reception time t_k in ECEF [m].
- $\boldsymbol{p}^{s_i}_{brdc}(t_k)$ is the broadcast satellite position (estimated position) at the reception time t_k in ECEF [m]. It is assumed that the broadcast satellite position is corrected using ultra-rapid precise orbit and clock products.
- $\boldsymbol{u}^{s_i}_{r_j}(t_k)$ is the unit line-of-sight vector defined by the true satellite and receiver positions. $\boldsymbol{p}_{true,r_j}(t_k)$ is the receiver position (true position) at the reception time t_k in ECEF [m].

The satellite clock error$dt^{s_i}(t_k)$ is computed as follows:

$$dt^{s_i}(t_k) = t^{s_i}_{brdc}(t_k) - t^{s_i}_{true}(t_k) \quad (3)$$

where:

- $t^{s_i}_{true}(t_k)$ is the satellite clock (true clock) at the reception time t_k in [s],
- $t^{s_i}_{brdc}(t_k)$ is the broadcast satellite clock (estimated clock) at the reception time t_k in [s].

In this application, it is assumed that the broadcast satellite clock and ephemeris are corrected using ultra-rapid precise orbit and clock products. Random (zero-mean Gaussian errors) time-varying errors are applied on the true satellite ephemeris (in radial, cross and along directions) and on the true clock to obtain broadcast satellite ephemeris and clock after ultra-rapid corrections. The values

of the standard deviations (SD) of the satellite position and clock errors are aligned with the assumption on the ultra-rapid precise orbit and clock product accuracy provided in (Sanz et al., 2013). Table 1 recaps these values.

Note that using cross-/along-track orbit errors higher than the radial component would have been more realistic.

4.2.2. Tropospheric delay ***STD***

Simulating realistic tropospheric delay on GNSS carrier phase measurements $STD_{r_j}^{s_i}$ is crucial to correctly assess the performances of the proposed tropospheric delay estimator. In order to represent small-scale and time-varying tropospheric events, such as local storms or local convective events, tropospheric delays are generated using AROME model database. AROME is a high spatial resolution NWP model developed by Météo France with a horizontal spatial resolution of 1–2 km. We use VMF1 Vienna mapping function (MF) to generate slant delays from zenith delays. The next sketch depicts the algorithm used to generate ZTD using AROME and VMF1 coefficient data.

Comparison with independent IGS ZTD data has shown that the accuracy of the AROME-based ZTD estimation is on the order of a few millimetres.

Note also that an optimal solution to compute the STD would be to use ray-tracing of 3D AROME data instead of using VMF1 coefficients. This approach would lead to simulate more realistic STD since it would represent the azimuthal variations of the troposphere delay, while the VMF1 mapping function assumes that the troposphere is symmetric around the GNSS receivers. Since the STD estimation algorithm does not assume that the STD is independent of the azimuth, this simplification is not expected to have a significant impact on the accuracy of the STD estimation algorithm developed in this paper. However, this must be further validated.

4.2.3. Ionospheric residual delay ***I***

For PPP using single frequency (SF) receivers, the ionospheric correction is required. Thus, simulating realistic ionospheric residual delay $I_{Lk,r_j}^{s_i}$ is fundamental to simulate the performance of the single-frequency receivers. It has been decided to simulate the ionospheric error residuals after the PPP correction (Zumberge et al., 1997). Being aware of the difficulty to capture the small scale and rapid ionospheric disturbances, the authors decided to use the Fast-PPP product (Rovira-Garcia et al., 2016), that offers a good residual error magnitude, <1 TECU root mean square (RMS), and global coverage.

The discrete time model of the ionospheric residual GM random process at the k-th epoch is expressed as follows:

$$b_{I,k} = exp\left(-\frac{\Delta t}{\tau}\right).b_{I,k-1} + \eta_I \tag{4}$$

- $b_{I,k-1}$ is the ionospheric residual Gauss-Markov (GM) random process corresponding to the previous epoch;
- τ is the ionospheric error correlation time that is set to 1800 s according to the GPS L1 and Galileo E1 MOPS (EUROCAE, 2019).
- η_I is a [N_k * 1] vector representing the noise error term. It follows a stationary, zero-mean normal distribution characterized as follows:

$$\eta_I \sim N\left(0, {\sigma_{I,\eta}}^2\right) \tag{5}$$

In the discrete time, the SD $\sigma_{I,\eta}(k)$ of the ionospheric delay as a function of the epoch k is given using the following relation:

$${\sigma_{I,\eta}}^2 = {\sigma_{I,k}}^2.\left(1 - exp\left(-\frac{2\Delta t}{\tau}\right)\right) \tag{6}$$

where the SD of the ionospheric residual $\sigma_{I,k}$ for the single frequency signals is obtained from the Fast-PPP absolute residuals RMS, that is considered in a conservative way to be 0,5 TECU in the vertical direction (σ_{VTEC}). The slant residual error can be determined using a single layer model-based MF depending on the satellite elevation, as it is used in the MOPS for civil aviation (EUROCAE, 2019).

$$\sigma_{I,k} = MF.\sigma_{VTEC} \tag{7}$$

With:

$$MF = \frac{1}{\left(\sqrt{1 - \frac{(RE.\cos(el))^2}{RE+h}}\right)} \tag{8}$$

where: RE is the mean Earth radius, h is the effective height of the ionosphere layer, and el is the elevation of a given satellite.

4.2.4. Receiver-related errors $\boldsymbol{dt_r}$ and $\boldsymbol{\eta_{Lk,\phi}}$

The test-bed generates the receiver clock offset dt_{r_j} and the phase lock loop (PLL) + multipath errors $\eta_{Lk,\phi,r_j}^{s_i}$ using models derived for L1 and E1 frequencies, but are also used for L2, L5 and E5a frequencies. Additional work is needed to estimate the model parameters for each signal frequency and elevation angle.

In order to perform experiments in a controlled environment in terms of receiver quality, three grades of receivers are used in the experiments, and the error models for each grade have been derived in two multipath environments (open-sky and semi-urban) using real measurement analyses. Table 2 maps the receiver used in measurements into a simulated receiver grade (see Figs. 4 and 5).

The developed carrier phase noise + multipath error model and receiver clock offset model have been developed for these three kinds of receiver grades and are fully described in a previous publication (Lehtola et al., 2019).

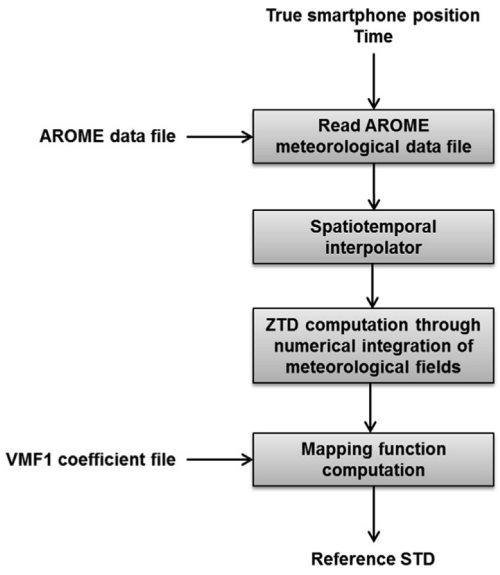

Fig. 4. Reference STD computation using weather forecast service data.

4.3. Pressure measurement errors

Surface pressure measurements at the smartphone positions are needed to extract the hydrostatic tropospheric delay from the total tropospheric delay in the tropospheric delay estimator (CPF). A quantitative analysis has shown that the performance level of the most recent smartphone barometers, such as the Bosh BME280 is sufficient to accurately extract the hydrostatic component from the total tropospheric delay for the weather monitoring application.

A surface pressure measurement at a given smartphone position is generated as follows in the test-bed:

$$p(t_k) = p_{true}(t_k) + b_p(t_k) + \eta_p(t_k) \tag{9}$$

- p_{true} is the true pressure at the smartphone position, extracted using a spatial high-resolution meteorological model (AROME/ARPEGE)
- b_p is the barometer measurement bias,
- η_p is the barometer measurement noise.

The bias is made of a constant part (that can be removed by a calibration process) and a time-varying part. The time-varying part is induced by several causes, including temperature effects. It is proposed to model it using a first-order Gauss-Markov model:

$$b_p(t_k) = \beta b_p(t_{k-1}) + \sqrt{1-\beta^2}\eta_b \tag{10}$$

where:

$$\eta_b \sim N\left(0, {\sigma_{p,b}}^2\right) \tag{11}$$

Using the Bosh BME280 datasheet we get:

$$\sigma_{p,b} = \frac{1}{3}hPa = 0.33hPa \tag{12}$$

- $\beta = exp\left(-\frac{\Delta t}{T_{corr}}\right)$
- Δt is the sampling rate of the pressure output [s].
- T_{corr} is the correlation time of the Gauss-Markov model [s]. Since the time-varying bias is highly correlated to the temperature changes, it is proposed to align the correlation time of the process to the time variation of the outdoor temperature: $T_{corr} = 3600$ s.

The noise term is computed as:

$$\eta_p \sim N\left(0, {\sigma_{p,\eta}}^2\right) \tag{13}$$

The pressure noise $\boldsymbol{\eta_p}$ depends on the temperature T. From Bosh BME 280 datasheet, we model $\sigma_{p,\eta}$ as:

$$\begin{aligned} \sigma_{p,\eta}(T) &= \left(1 + \tfrac{\alpha(T)}{100}\right)\sigma_{p,\eta}(T = 25\ ^\circ\mathrm{C}) \\ \alpha(T) &= \begin{cases} \alpha_1 T + \beta_1 & for\ \ T < 25\ ^\circ C \\ \alpha_2 T + \beta_2 & for\ \ T \geq 25\ ^\circ C \end{cases} \end{aligned} \tag{14}$$

Table 3 provides the values of α_1, β_1, α_2, β_2 and $\sigma_{p,\eta}(T = 25\ ^\circ\mathrm{C})$ for the Bosh BME280 sensor.

4.4. Link loss

To simulate the failures induced by the data transfer through the network between a given receiver and the CPF, the simulator uses statistic models of the occurrence of the link loss as well as the parameters that characterize the phenomenon: Mean Time Between Failure (MTBF), Mean Down Time (MDT).

First, it is randomly determined the epoch where the failure occurs, using the Mersenne Twister pseudorandom number generator, determining the epoch t_{init} when the first failure occurs for a given receiver.

The epochs with a simulated failure between the considered receiver and the CPF are determined as follows:

$$\begin{aligned} TimeFailure = [t_{init}; t_{init} + (MTBF + MTD); t_{init} \\ +2 * (MTBF + MTD; \cdots, t_{init} + (N-1) * (MTBF + MTD)] \end{aligned} \tag{15}$$

where N is the total number of failures for the considered receiver during the entire simulation.

Finally, the failure is applied to the GNSS observations matrix to disable the use of the identified GNSS receiver when a certain link loss occurs.

Similarly, duty cycles (that are periods in which the smartphone is not active to save battery) are generated randomly over the smartphone receivers to simulate the gaps on the GNSS signals acquisition that happens for the sake of power consumption.

4.5. Tropospheric delay estimator algorithm

The water vapor distribution in the atmosphere is estimated using a collaborative tomography approach. A grid

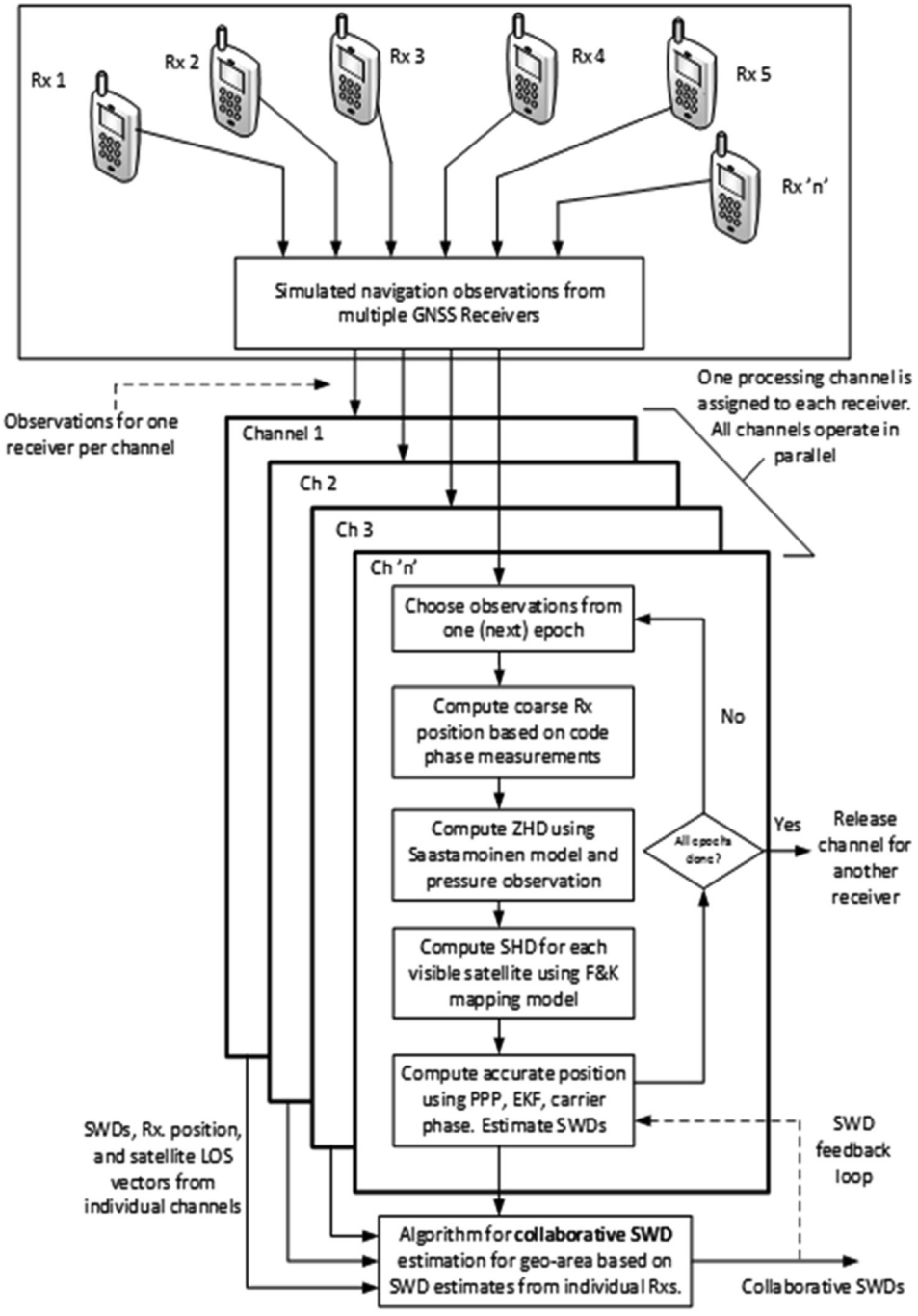

Fig. 5. Methodology for SWD estimation.

with one or several layers is defined over the area of interest, where each cell within the grid is bounded in latitude and longitude. The amount of delay in each range measurement caused by the water vapour is associated with one cell in each layer of the grid. In this crowdsourcing approach, instead of static high-quality receivers in known locations, we additionally utilize roving receivers. For this reason, we need to simultaneously estimate the locations of the receivers. This brings additional uncertainty to the problem, but this can be balanced by a large number of the receivers and by weighting the measurements based on their quality.

The position estimation is done using a standard PPP approach. The state estimate in the Extended Kalman Filter (EKF) contains the receiver position and velocity, as well as the receiver clock bias and carrier phase ambiguity for each satellite. The EKF also outputs the state uncertainty. This information is used in the tomography algorithm.

The delay caused by the water vapor, the wet residual, is extracted from the range measurements by removing the distance between the satellite and the receiver, and estimates of errors caused by receiver clock bias, carrier phase ambiguity, ionosphere and hydrostatic part of the troposphere. The clock bias and carrier phase ambiguity are estimated in the PPP.

Managing of the residual ionospheric error depends on whether dual-frequency measurements are available. In the case of dual-frequency measurements being available, the wet residuals are extracted from the ionosphere-free combination of the carrier phase measurements. As the ionosphere-free combination is used also in the PPP, the estimated phase ambiguity is multiplied with the ionosphere-free wavelength (van der Marel, 2012) when extracted from the range measurement. In the case of single-frequency measurements, the residual ionospheric error cannot be extracted but it is simulated as presented above.

The Zenith Hydrostatic Delay (ZHD) is estimated using the Saastamoinen model (Saastamoinen, 1972), and mapped to Slant Hydrostatic Delay (SHD) based on the elevation angle of each satellite by using F&K MF. F&K mapping function was chosen over the Niell mapping function due to computational simplicity. Furthermore, the Vienna mapping function is used in the testbed in generating the simulated tropospheric delay, so using it again in estimating the final ZWD would likely lead in too optimistic results.

The remainder of the observation is the Slant Wet Delay (SWD) and a residual error consisting, for example, of errors in satellite and receiver positions and clocks. The value is converted to Zenith Wet Delay (ZWD) using again F&K MF. Next, ZWDs are under-sampled to reach a sampling period of 15 min. This corresponds to the sampling rate of the GNSS-based tropospheric data currently assimilated in Météo France NWMs (Mahfouf et al., 2015).

We use numerous receivers of different grades in different locations, and thus assume measurement errors originating from environmental effects and position and clock errors to be independent of each other. For this reason, when combining all the measurements then the environmental error is assumed to approximately cancel from the final result. Furthermore, the estimated variance for the position and the clock error from the PPP is used to weight the measurements in the tomography.

In case of static receivers, errors resulting from antenna phase centre offset and variation depend on the antenna type and elevation and azimuth of the satellite. Usually, a correction term is given by the antenna manufacturer to mitigate the effect on the range measurement. If the phase centre variation is not corrected, the resulting error in position is on the level of few centimetres, mainly on the height component (El-Hattab, 2013). In the presented simulations, the phase centre offset and variations are not considered.

In, for example, IGS predicted Ultra-Rapid products the RMS error of orbits is 5 cm and the standard deviation of the clock error is 1.5 ns (IGS Products, 2020). As these errors are common to all receivers, there is a chance that there will be a common bias depending on the current satellite visibility in the ZWD measurements. This would affect especially the receivers close to each other, but if the receivers and pierce points are spread wider, the effect on the estimated ZWD would be smaller. In the simulations, we use constant standard deviations (zero bias) for the orbit and clock errors to weight the ZWD measurements.

The uncertainties in the simulations are used when estimating the standard deviations for the ZWD values, and thus each estimated ZWD will be associated with a respective confidence interval. However, for realistic implementations, the possible biases and variances need to be considered more carefully.

The distribution of ZWD, that results from water vapor in the atmosphere, is modelled as a thin layer at an appropriate height in a similar manner as is done in the Klobuchar ionosphere model (Klobuchar, 1987), for example. However, the height for the thin layer approximation is significantly lower, in the order of a few kilometres. Also, the location where the signal from the satellite to the receiver reaches the height of the thin layer, the pierce point, is computed in a similar manner as in the Klobuchar model. Based on the computed pierce point, the measurement is placed into one of the predefined cells in the grid. If a measurement falls outside the grid, it is not used. We apply Tikhonov regularization for the ZWD estimation in an iterative manner, where the ZWD estimate with associated covariance from the previous epoch is used as a priori estimate on the next epoch. The known (Lehtola et al., 2019) uncertainties and the covariances estimated in the PPP are used to weight the ZWD measurements such that the measurements with larger uncertainties have a smaller effect on the final estimate. The resulting estimate contains a ZWD value for each cell of the grid with the associated estimate of the uncertainty. These estimates are then assimilated to the NWP model.

4.6. Main test-bed limitations

This section discusses the main limitations related to the representativeness of the test-bed simulator. The ways to improve the representativeness of the test-bed are listed as follows:

- Generate TEC values on the GNSS measurements using ionosphere products, such as IONEX files, and to implement the correction technique aiming to mitigate GNSS measurement ionospheric delays for SF receivers.
- Generate tropospheric Slant Total Delay using AROME/ARPEGE and ray tracing instead of using VMF1 function to compute STD from ZTD.
- Add cycle slips on the carrier phase measurements.

- Develop error models that depend on the signal elevation angle, and on the signal frequency.
- Simulate inter-frequency biases on the GLONASS pseudo-range measurements.

5. System performance assessment

The system performance has been assessed for nominal and perturbed conditions and for different receiver distributions, e.g. dense receiver concentration to simulate a city. Taking advantage of the test-bed flexibility to simulate different controlled environments, several simulations were performed. Keeping the tomographic grid in a single-layer mode, it is seen that the main parameters influencing the system performance are the following: the receiver density, the GNSS constellation mode, the GNSS frequency mode, the receiver grade and the system failures. The ZWD estimation error was obtained comparing to a reference ZWD obtained from AROME/ARPEGE models.

As an illustration, the impact of the receiver duty cycle on the ZWD estimation is analysed in this section. The next figures illustrate the maximal ZWD estimation error obtained without (Fig. 6) and with (Fig. 7) duty cycle during a 3 h simulation using a network composed of 1000 GNSS receiver spread randomly in a circle of 25 km diameter, representing Toulouse area, in France. Tests were conducted to assess the needed receiver density and it has been concluded that ~1 smartphone per km^2 is recommended. Given that convergence is achieved in a maximum of 30 min, a window of 3 h is enough to have a good number of assimilated ZWDs. In this simulation the receivers are dual-frequency, in case of high residual ionospheric errors when using single-frequency receivers, it would be expected to have a degradation of 25% in the number of assimilated ZWDs.

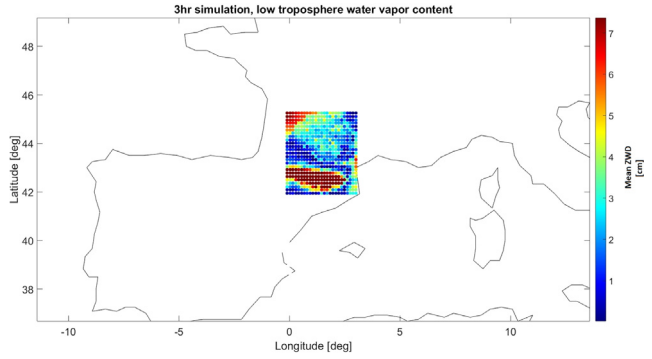

Fig. 6. Spatial evolution of the ZWD estimation error in the South of France, with no system failures. GPS + GAL, dual-frequency.

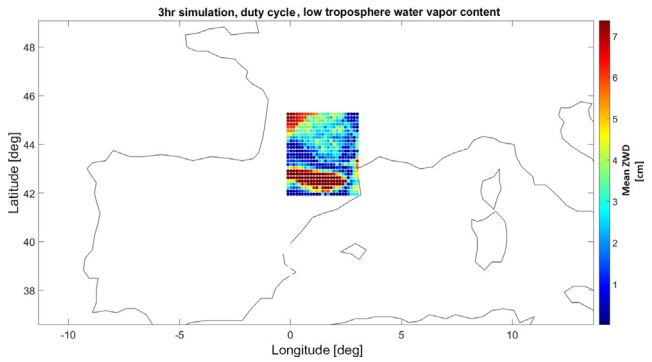

Fig. 7. Spatial evolution of the ZWD estimation error in the South of France, with duty cycle (30 min of MTBF, MTD of 15 min and 50% of the receivers affected by DC), GPS + GAL, dual-frequency.

Quantitatively, the mean ZWD estimation error in the centre of the grid is roughly 4 cm. When system failures are introduced, a performance degradation is noticed, as expected. The error reaches 5 cm when duty cycles are present. We observe that the tropospheric delay estimator is robust against duty cycles when 50% of the receivers are affected. Additionally, the ZWD error SD is on the order of 1 mm with and without duty cycle, because of the tomography process that acts as an averaging filter.

The required accuracy of ZWD for NWM assimilation is roughly 6 mm in terms of SD and below 3 mm in terms of residual bias (Bock et al., 2016). This corresponds to a SD of roughly 1.0 kg/m^2 and a residual bias of 0.5 kg/m^2 in terms of Integrated Water Vapor (IWV). Indeed $IWV = \rho \frac{ZWD}{k}$, where ρ is the density of water and k depends on the surface temperature and is taken here at 6.4 (Bevis et al., 1992). The proposed system leads to reach the required SD of the estimation error, but the ZWD residual bias (4 cm) is too high for the NWM assimilation. However, applying a bias removal algorithm (time domain averaging) on the ZWD series leads to decrease the obtained mean ZWD up to a few mm, which is close to the required ZWD residual error.

Although our method is not limited to this, we used only a single layer in vertical dimension for simplicity. Larger grids are prone to demand stronger regularization assumptions because the rover stations and satellites rarely offer homogeneous distribution of GNSS signals. Future work should explore the reliability of such grids with respect to highly inhomogeneous rover station distributions. Finally, note that the ZWD estimation is degraded at the border of the tomography grid. This is due to the low number of troposphere pierce points passing through grid cells at the border of the grid. In addition, degraded ZWD estimation is observed in the area of Pyrenees mountains. Further investigations are needed to improve the ZWD estimation over high altitude areas.

6. Conclusion

Exploiting GNSS smartphone receiver networks for meteorology applications is tempting. We propose a system architecture and its implementation, a simulation test-bed, for assessing the feasibility of estimating the tropospheric water vapor distribution using GNSS smartphone receiver networks.

This paper proposes a simulator that emulates the WMCC system in a representative way, and that enables providing a controlled environment for the performance analyses. In this context, it has been shown that the designed test-bed enables simulating a wide variety of configurations, such as:

- The system parameters, gathering the smartphone network parameters (receiver density), the smartphone parameters (receiver grade, constellation, ...), and the smartphone to CPF network parameters (data link loss).
- The environmental parameters correspond to the atmospheric conditions.

The main limitations of the test-bed have been also discussed in this publication, in Section 4.5.

The development of the test-bed enables testing the system performances in specific configurations, and assessing the robustness of the system against several perturbations, such as dynamic crowdsourcing receiver network conditions, that are the emergence of new receivers and disappearance of old receivers. The first test-bed results are analysed in this paper, and the full results are expected to be presented and analysed in a future publication.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The NAVISP-EL1-008 project has been supported by the European Space Agency (ESA). The authors also thank Centre National d'Etudes Spatiales (CNES) for supporting this project.

References

Bender, M., Dick, G., Ge, M., et al., 2011. Development of a GNSS water vapour tomography system using algebraic reconstruction techniques. Adv. Space Res. 47, 1704–1720. https://doi.org/10.1016/j.asr.2010.05.034.

Bevis, M., Businger, S., Herring, T.A., Rocken, C., Anthes, R.A., Ware, R.H., 1992. GPS meteorology remote sensing of atmospheric water vapor using the global positioning system. J. Geophys. Res. 97 (D14), 15787–15801.

Bock, O., Bosser, P., Pacione, R., Nuret, M., Fourrié, N., Parracho, A., 2016. A high-quality reprocessed ground-based GPS dataset for atmospheric process studies, radiosonde and model evaluation, and reanalysis of HyMeX special observing period. Q.J.R. Meteorol. Soc. 142, 56–71. https://doi.org/10.1002/qj.2701.

Böhm, J., Böhm, S., Nilsson, T., et al., 2012. The new Vienna VLBI Software VieVS. Geodesy for Planet Earth. Springer, pp. 1007–1011. doi: 10.1007/978-3-642-20338-1_126.

Bokoye, A.I., Royer, A., O'neill, N.T.O., et al., 2003. Multisensor analysis of integrated atmospheric water vapor over Canada and Alaska. J. Geophys. Res.: Atmos. John Wiley & Sons, Ltd, vol. 108, no. D15, 4480. doi: 10.1029/2002JD002721.

Brenot, H., Walpersdorf, A., Reverdy, M., et al., 2014. A GPS network for tropospheric tomography in the framework of the Mediterranean hydrometeorological observatory Cévennes-Vivarais (southeastern France). Atmos. Meas. Tech. 7 (2), 553–578. https://doi.org/10.5194/amt-7-553-2014.

Champollion, C., Masson, F., Boin, M., et al., 2005. GPS water vapour tomography: preliminary results from the ESCOMPTE field experiment. Atmos. Res. 74 (1), 253–274. https://doi.org/10.1016/j.atmosres.2004.04.003.

Council, N.R., 2009. Observing weather and climate from the ground up: a nationwide network of networks, Observing Weather and Climate from the Ground Up: A Nationwide Network of Networks. The National Academies Press, Washington DC. doi: 10.17226/12540.

EL-Hattab, A.I., 2013. Influence of GPS antenna phase center variation on precise positioning. NRIAG J. Astronomy Geophys., 2(2), 272-277.

EUROCAE, 2019. Minimum Operational Performance Standard for Galileo/Global Positioning System/Satellite-Based Augmentation System Airborne Equipment. EUROCAE ED-259, 01/02/2019 (EUROCAE, Paris 2019).

Flores, A., Ruffini, G., Rius, A., 2000. 4D tropospheric tomography using GPS slant wet delays. Ann. Geophys. 18 (2), 223–234. https://doi.org/10.1007/s00585-000-0223-7.

Foelsche, U., Kirchengast, G., 2002. A simple "geometric" mapping function for the hydrostatic delay at radio frequencies and assessment of its performance. Geophysical Research Letters. John Wiley & Sons, Ltd, 29(10), pp. 111–114. doi: 10.1029/2001GL013744.

Guerova, G., Jones, J., Douša, J., Dick, G., de Haan, S., Pottiaux, E., et al., 2016. Review of the state of the art and future prospects of the ground-based GNSS meteorology in Europe. Atmos. Meas. Tech. 9 (11), 5385–5406.

Haji-Aghajany, S., Amerian, Y., 2017. Three dimensional ray tracing technique for tropospheric water vapor tomography using GPS measurements. J. Atmos. Sol. Terr. Phys. 164, 81–88. https://doi.org/10.1016/j.jastp.2017.08.003.

Hicks-Jalali, S., Sica, R., Haefele, A., Martucci, G., 2019. Calibration of a water vapour Raman lidar using GRUAN-certified radiosondes and a new trajectory method. Atmos. Meas. Tech. 12 (7), 3699–3716. https://doi.org/10.5194/amt-12-3699-2019.

IGS Products. Online: https://www.igs.org/products (accessed: 29.4.2020).

Kämpfer, N., 2013. ISSI Scientific Report Series. Springer Science & Business Media, vol. 10, Springer-Verlag, New York.

Kawabata, T., Shoji, Y., 2018. Applications of GNSS Slant Path Delay Data on Meteorology at Storm Scales. In: Hashimov (Ed.), In Multifunctional Operation and Application of GPS, Ch. 7. IntechOpen: London, UK, doi: 10.5772/intechopen.75101.

Kawabata, T., Shoji, Y., Seko, H., Saito, K., 2013. A numerical study on a mesoscale convective system over a subtropical Island with 4D-Var assimilation of GPS slant total delays. J. Meteorol. Soc. Jpn. 91, 705–721. https://doi.org/10.2151/jmsj.2013-510.

Kley, D., Russell, J.M., Phillips, C., 2000. SPARC Assessment of Upper Tropospheric and Stratospheric Water Vapour. World Climate Research Organization, (1043).

Klobuchar, J.A., 1987. Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users. IEEE Transactions on Aerospace and Electronic Systems, AES-23(3), pp. 325–331. doi: 10.1109/TAES.1987.310829.

Lehtola, V.V., Söderholm, S., Koivisto, M., Montloin, L., 2019. Exploring GNSS crowdsourcing feasibility: combinations of measurements for modeling smartphone and higher end GNSS receiver performance. Sensors (Basel, Switzerland) MDPI 19 (13), 3018. https://doi.org/10.3390/s19133018.

Lindskog, M., Ridal, M., Thorsteinsson, S., Ning, T., 2017. Data assimilation of GNSS zenith total delays from a Nordic processing centre. Atmos. Chem. Phys. 17 (22), 13983–13998. https://doi.org/10.5194/acp-17-13983-2017.

Mahfouf, J.F., Ahmed, F., Moll, P., Teferle, F.N., 2015. Assimilation of zenith total delays in the AROME France convective scale model: a recent assessment. Tellus A: Dynamic Meteorol. Oceanography 67, 1. https://doi.org/10.3402/tellusa.v67.26106.

Oigawa, M., Tsuda, T., Seko, H., et al., 2018. Data assimilation experiment of precipitable water vapor observed by a hyper-dense GNSS receiver network using a nested NHM-LETKF system. Earth Planets Space 70 (1), 74. https://doi.org/10.1186/s40623-018-0851-3.

Rohm, W., Bosy, J., 2011. The verification of GNSS tropospheric tomography model in a mountainous area. Adv. Space Res. 47 (10), 1721–1730. https://doi.org/10.1016/j.asr.2010.04.017.

Rovira-Garcia, A., Juan, J., Sanz, J., et al., 2016. Accuracy of ionospheric models used in GNSS and SBAS: methodology and analysis. J. Geod. 90 (3), 229–240. https://doi.org/10.1007/s00190-015-0868-3.

Saastamoinen, J., 1972. Atmospheric Correction for the Troposphere and Stratosphere in Radio Ranging Satellites. In: Henriksen, S.W., Mancini, A., Chovitz, B.H. (Eds.), The Use of Artificial Satellites for Geodesy. Geophysical Monograph Series. pp. 247–251. AGU, Washington, DC., doi: 10.1029/GM015p0247.

Sanz, J., Juan, J.M., Pajares, M., 2013. GNSS Data Processing, Vol I: Fundamental and Algorithms. ESA Communications, Noordwijk, the Netherlands.

van der Marel, H., 2012. Single- versus dual-frequency precise point positioning. Inside GNSS, 30–35.

Zumberge, J.F., Heflin, M.B., Jefferson, D.C., Watkins, M.M., Webb, F. H., 1997. Precise point positioning for the efficient and robust analysis of GPS data from large networks. J. Geophys. Res. 102 (B3), 5005–5017. https://doi.org/10.1029/96JB03860.