



GNSS-retrieved precipitable water vapour in the Atlantic coast of France and Spain with GPT3 model

Raquel Perdiguier-Lopez¹ · José Luis Berne Valero² · Natalia Garrido-Villén²

Received: 27 February 2023 / Accepted: 26 September 2023
© The Author(s) 2023

Abstract

Water vapour is a critical atmospheric parameter to understand the Earth's climate system and it is characterized by a complex variability in time and space. GNSS observations have become an important source of information of the water vapour, thanks to its high temporal and spatial resolution. However, the lack of meteorological sites collocated with the GNSS site could hamper water vapour retrieval. The empirical blind models can fill this gap. This study analyses the temporal and spatial distribution of the water vapour using nine GNSS sites located on the Atlantic coast of Spain and France, with the empirical blind model GPT3 as the source of meteorological information. The observations were processed with Bernese 5.2 software on a double difference approach and validated with Zenith Total Delay EUREF REPRO2 values. Consequently, four-years series of water vapour was determined and validated using two matched radiosonde sites. The characterization of the water vapour on the area shows clear seasonal characteristics that the technique captures, using an empirical blind model for the whole process. Maximum values are observed in summer season and minimum in winter. The PWV tends to decrease with increasing latitude in the area of the study. The short-term variations can be reproduced by the high temporal resolution of the GNSS-retrieved water vapour and show a different behaviour over the area, but a similar pattern with a peak in the afternoon and minimum at night was found. Also, less variability is observed in winter season and higher in summertime.

Keywords GNSS · Precipitable water vapour · Radiosonde · GPT3

✉ Raquel Perdiguier-Lopez
rperdiguier@uniovi.es

José Luis Berne Valero
jlberne@upvnet.upv.es

Natalia Garrido-Villén
ngarrido@cgf.upv.es

¹ Cartographic Engineering, Geodesy and Photogrammetry Area, University of Oviedo, Mieres, Asturias, Spain

² Cartographic Engineering, Geodesy and Photogrammetry Department, Polytechnic University of Valencia, Valencia, Spain

1 Introduction

Water vapour in one of the most important constituents of the atmosphere. Apart from its obviously connection to the hydrological cycle, water vapour is crucial to the regulation of temperature on earth, and it is the most important natural greenhouse, accounting for about 60% of this effect (<https://gcos.wmo.int/en/essential-climate-variables/upper-vapour> accessed in November 2022). In fact, water vapour is defined by the Global Climate Observing System (GCOS) as an Essential Climate Variable because it critically contributes to the characterization of the Earth's climate system. The water vapour amount, as well as its temporal and spatial variation, therefore, are needed to understand and predict global climate change.

Water vapour can be expressed using different terms. One of these terms is the Integrated Water Vapour (IWV) which is defined as the total amount of water vapour present in a vertical atmospheric column, using $\text{kg}\cdot\text{m}^{-2}$ as its measurement unit. Another commonly used term is Precipitable Water Vapour (PWV) defined as the height to which water would stand if completely condensed and collected in a vessel with a cross section of 1 m^2 . In this paper, water vapour is expressed as PWV, using millimetres as its measurement unit. Water vapour proportion in the atmosphere is between 1 and 4%; of this percentage between 45 and 65% is concentrated from the surface to the 850 hPa level (approximately 1.5 kms) (J. Wang et al. 2007). One of the principal characteristics of atmospheric water vapour is its high variability on time and space that renders its study rather complex.

Because of its great importance, many different types of instruments, based on different techniques, had been used to retrieve the water vapour present in the atmosphere. Some of these techniques are Water Vapour Radiometers (Köpken 2001), Very Long Base Interferometry (Niell et al. 2001), LiDAR (Hicks-Jalali et al. 2020), sun photometers (Firsov et al. 2013), as well as different sensors on satellites (Gao & Kaufman 2003).

Furthermore, for decades, the traditional technique that had proved its capacity to retrieve the water vapour is the radiosonde (Ross and Elliott 2001). This technique allows to obtain the vertical profile of different atmospheric meteorological parameters from which Precipitable Water Vapour can be calculated. The long-term historic data and its high precision and good vertical resolution are some of its most valuable characteristics. However, most of the radiosonde stations retrieve just one or two profiles per day, so radiosonde is considered a low temporal resolution technique (Deblonde et al. 2005; Liu et al. 2005). Moreover, the number of stations is not high, with an inhomogeneous distribution characterised by a higher number of stations on the northern hemisphere and a very limited number of radiosondes sites on the oceans. Therefore, the spatial resolution is not high.

Global Navigation Satellite Systems (GNSS) observations, originally designed to provided positioning, have become a powerful tool to retrieve the water vapour. There are some advantages on the use of GNSS observations to retrieve the atmospheric water vapour. On one hand, the network of GNSS sites continues to grow so the availability of data is increasing. In addition, temporal resolution is high, being able to retrieve the PWV values with a ratio of 15 min or even less, so the technique can be useful to study short-term water vapour variations. However, the distribution of GNSS stations is not homogenous. Moreover, the lack of meteorological stations collocated with the GNSS site makes it difficult, as will be explained later, to retrieve water vapour from GNSS measurements.

There have been various studies that proved the capacity of GNSS to retrieve the water vapour on a global, continental, or local scale (Bock et al. 2007; Vaquero-Martínez et al. 2018; J. Wang and Zhang 2009). The water vapour retrieved from GNSS observations have

been validated with other reference measurements such as radiosonde (Ohtani and Naito 2000), radiometers on satellites like MODIS (Z. Li et al. 2003) or numerical weather models (Heise et al. 2009), reaching a quality between 1 and 3 mm. A selection of some multi-comparisons can be found in Guerova et al., (2016) Table 3.

The retrieval of the water vapour from GNSS observations is based on the delay that the signals undergo on the last part of its way to the antenna. This is commonly known as troposphere delay because it mainly occurs in this part of the atmosphere. The total delay of the signal, once mapped onto the zenith, is known as Zenith Total Delay (ZTD). ZTD is the sum of the Zenith Hydrostatic Delay (ZHD) and Zenith non-hydrostatic or Zenith Wet Delay (ZWD). ZHD depends on the atmospheric pressure and temperature and ZWD depends on the partial water vapour pressure and temperature so, the calculation of the delay would imply in-situ measurements of these atmospheric parameters on which they depend. Unfortunately, not all the GNSS sites have collocated meteorological stations to measure the needed atmospheric parameters. However, using just a set of GNSS sites with collocated meteorological stations could reduce the valuable high spatial resolution of the technique. In these cases, an empirical model can be useful to obtain the atmospheric parameters. Moreover, the real time PWV retrieval, needed on the short-term rainfall forecasting, can be easily accomplished with an empirical model. However, it is necessary to verify the capacity of the empirical blind model, if it is the unique source of meteorological data in PWV retrieval, to contribute to the capturing of the variations of PWV. Thereafter, this study focuses on the characterization of temporal and spatial variations of water vapour using an empirical blind model in each step in which atmospheric parameters are required.

This paper is structured as follows. In Sect. 2, a description of the retrieval of Precipitable Water Vapour from GNSS observations is shown. In Sect. 3, data sets and methodology are described. In the results section, Sect. 4, firstly the validation process of the ZTD using the official EUREF values and the PWV with radiosonde data is provided. After which, an analysis of the temporal and spatial variations of Precipitable Water Vapour in the investigated area are explained. Finally, Sect. 5 shows the summary and conclusions of the study.

2 Precipitable water vapour retrieval from GNSS observations

The starting point of the retrieval of PWV is the estimation of the neutral or troposphere delay, that is commonly obtained from the processing of the GNSS observations as a time dependent parameter. The troposphere delay, T_D , can be expressed as the difference between the ray path and the geometrical vacuum distance:

$$T_D = \int (n - 1) ds \quad (1)$$

where n is the index of atmospheric refraction, ds is the differential increment in distance with respect to the line of sight. Often, the refractivity, N , defined by $N = 10^6(n - 1)$, is used instead of the atmospheric refraction, so Eq. 1 can be formulated as:

$$T_D = 10^{-6} \int N ds \quad (2)$$

The refractivity can be divided into two terms:

$$N = N_{HYD} + N_{WET} \quad (3)$$

where N_{HYD} is due to the hydrostatic part of the atmosphere and N_{WET} is due to the non-hydrostatic or wet part of the atmosphere. Then, the Eq. 2 can be rewritten as:

$$T_D = 10^{-6} \int N_{HYD} ds + 10^{-6} \int N_{WET} ds \quad (4)$$

The refractivity depends on the atmospheric temperature, pressure and water vapour pressure and can be calculated empirically (Essen and Froome, 1951; Thayer 1974). However, usually, the troposphere delay showed on Eq. 4 are calculated by considering the dependence of the delay on the satellite zenith distance z (Dach et al. 2015) so the integrals on Eq. 4 can be expressed as:

$$T_D = mf_{HYD}(z) \cdot ZHD + mf_{WET}(z) \cdot ZWD \quad (5)$$

where mf_{HYD} and mf_{WET} are the hydrostatic and wet mapping functions respectively and ZHD and ZWD are the Zenith Hydrostatic Delay and the Zenith Wet Delay respectively. Zenith Hydrostatic Delay is caused by the atmospheric gases, assuming the hydrostatic equilibrium, and accounts for 90% of the total delay (Hofmann-Wellenhof et al. 2007). Its low variability facilitates its estimation with enough accuracy from models. One of the most widely used is the Saastamoinen model (Saastamoinen 1972):

$$ZHD = \frac{0.0022767 \cdot P_0}{1 - 0.00266 \cdot (\cos 2\varphi) - 0.00028 \cdot H} \quad (6)$$

where P_0 is the atmospheric pressure in hPa, and φ (in degrees) and H (in kilometres) are the latitude and height of the site, respectively. With an error less than 0.30 hPa, ZHD can be modelled with a quality of around 1 mm (Duan et al. 1996). In those stations without a collocated meteorological station, the pressure could be obtained from empirical models like Global Pressure and Temperature 2 wet model (GPT2w) (Böhm et al. 2015) or Global Pressure and Temperature 3 (GPT3) (Landskron and Böhm 2017) or from reanalysis products like ERA5 (Hersbach et al. 2020).

The Zenith Wet Delay is defined as the integral of the wet refractivity along the vertical profile above the station and depends on the partial pressure of water vapour P_v and temperature T (Yao and Hu 2018). ZWD contributes the remaining 10% of the total delay. According to (Bevis et al. 1992), ZWD is given by:

$$ZWD = 10^{-6} \int \left(k^2 \cdot \frac{P_v}{T} + k^3 \cdot \frac{P_v}{T^2} \right) dz \quad (7)$$

where the values and uncertainties of constants are $k^2 = (22.1 \pm 2.2) \text{ K} \cdot \text{mb}^{-1}$ and $k^3 = (3.7390 \pm 0.0012) 10^5 \cdot \text{K}^2 \cdot \text{mb}^{-1}$ (Bevis et al. 1994). One of the characteristics of ZWD is its high variability, both temporal and spatial, that complicates its calculation from models so, if profiles of P_v and T are not available, ZWD is often obtained from the subtraction of ZTD and ZHD:

$$ZWD = ZTD - ZHD \quad (8)$$

Once ZWD is obtained, the last step to the retrieval of PWV from GNSS observations is the application of the dimensionless factor Π :

$$PWV = \Pi \cdot ZWD = \frac{10^6}{\rho \cdot R_V \left[\frac{k_3}{T_m} + k_2 - m \cdot k_1 \right]} \cdot ZWD \quad (9)$$

where ρ is the density of liquid water, R_V is the specific gas constant for water vapour, m is the ratio of the molar masses of water vapour and dry air, and T_m is the water-vapour-weighted mean temperature. The values and uncertainties of the physical constants are $k_1 = (70.60 \pm 0.05) \text{ K} \cdot \text{mb}^{-1}$, $k_2 = (70.40 \pm 2.2) \text{ K} \cdot \text{mb}^{-1}$ and $k_3 = (3.7390 \pm 0.0012) 10^5 \cdot \text{K}^2 \cdot \text{mb}^{-1}$ (Bevis et al. 1994). The approximate value of factor Π is around 0.15, with a variation of 20% related to the location, height, weather season and time (Bevis et al. 1994). Notably, other authors consider a factor value of 0.16 (Foelsche and Kirchengast 2001). Taking the value of 0.16 as a reference then, approximately 6.3 mm of ZWD would correspond to 1 mm of PWV.

The water-vapour-weighted mean temperature, T_m , can be obtained by:

$$T_m = \frac{\int \left(\frac{P_v}{T} \right) dz}{\int \left(\frac{P_v}{T^2} \right) dz} \quad (10)$$

where P_v is the partial pressure of water vapour, T is the absolute temperature and dz is the integral path. Relative error of conversion factor Π is closely related to the relative error of T_m so this magnitude is crucial to improve the quality of the conversion factor and, consequently, the PWV retrieval.

Water-vapour-weighted mean temperature can be obtained from three main methodologies. The most accurate one is using temperature and humidity profiles obtained from radiosonde or numerical weather models. This approach has the advantage of its high precision but the disadvantage of low temporal and spatial resolution (Wang et al. 2016). Moreover, it cannot be applied in real time. Another approach to obtain T_m is its estimation through its strong correlation with the surface temperature, T_s . This correlation can be used to derive a linear model between the two temperatures. One of the most well-known relations is the Bevis model, that was established using 8718 profiles of atmospheric parameters from 2 years of radiosonde data along North America, achieving a root mean square (rms) on the T_m calculation of 4.7 k (Bevis et al. 1992):

$$T_m = 70.2 + 0.72 \cdot T_s \quad (11)$$

Another well-known relation is the Mendes model (Mendes 1998). This linear model was found using 32,467 radiosonde profiles, achieving a rms of 3.07 k on T_m :

$$T_m = 50.4 + 0.789 \cdot T_s \quad (12)$$

In Spain, Ortiz de Galisteo Marín, (2011) made a study using 8 radiosonde sites, analysing 37,179 profiles from 2000 to 2008, developing the relationship:

$$T_m = 61.92 + 0.75 \cdot T_s \quad (13)$$

The rms of the water-vapour-weighted mean temperature obtained from this model was found in 3.19 K.

Apart from those relations, some have been developed, considering the spatial, seasonal, or even daily variation of the relation of the surface temperature and the water-vapour-weighted mean temperature. For example, Baldysz and Nykiel, (2019) established a set of

models to Europe, using 49 radiosonde sites and covering a period of 24 years, with a rms of 2.8k on the estimation of T_m . However, no matter if the relationship is simpler or more complex, it is required to know the surface temperature to be able to apply this approach.

A third approach to determine the needed water-vapour-weighted mean temperature is the use of empirical blind models. These empirical models can even use different approaches to determine the T_m , but are characterized by simple input data, like site position and the day of year, and most of them can be used to determine T_m globally. Using this approach no meteorological data in situ is needed so it could be useful for those stations without a collocated meteorological station or in a real time retrieval. Some global empirical models have been developed. UNB3m model is an improvement of UNB3 model related to the water vapour pressure estimations. It permits T_m to be derived from a set of atmospheric parameters that are calculated from a linear interpolation in latitudinal direction using a set of prefixed tabulated values (Leandro et al. 2006). Global empirical T_m model X, (GTm_X) was established from global reanalysis data provided by the ECMWF, based on the semiannual and diurnal variations in T_m , and achieves a global spatial resolution of $1^\circ \times 1^\circ$ (Chen and Yao 2015). Global Weight Mean Temperature D, GWMT_D considers the diurnal variation using the UTC hour as an input, as well as GTm_X model, and achieves a rms less than 5.0 k on T_m (He et al. 2016). Global Pressure and Temperature 2 wet model (GPT2w), as well as its successor, Global Pressure and Temperature 3 (GPT3) (Landskron and Böhm 2017) was derived using ten years mean monthly pressure level data from ERA-Interim and is provided in two resolution grids of $1^\circ \times 1^\circ$ and $5^\circ \times 5^\circ$ (Böhm et al. 2015). These two last blind models provide not only the T_m but also a set of atmospheric parameters, so can be useful in the whole process of the retrieval of PWV from GNSS observations. Therefore, the latest, GPT3 model, is the chosen model to provide the needed atmospheric parameters on this work.

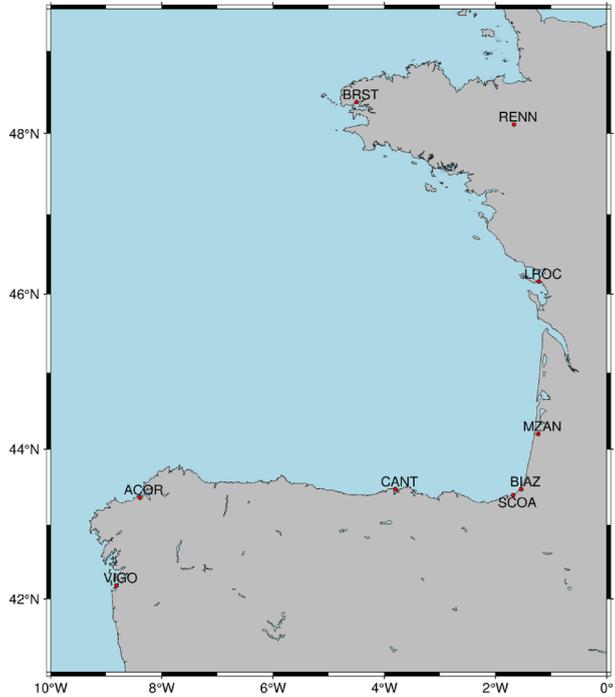
The uncertainty of the PWV retrieved from the GNSS observations can be evaluated from the comparison with other techniques or using a theoretical analysis where the uncertainty of the PWV is calculated from the uncertainties of each variable from which PWV is derived. The uncertainty of PWV comes mainly from the ZTD uncertainty, contributing to more than 75% uncertainty of PWV (Ning et al. 2015). Apart from that, the uncertainty of the atmospheric pressure, used on ZHD calculation coupled with the variables that are involved in the calculation of conversion factor, especially the uncertainty of the water-vapour-weighted mean temperature, affects the final quality of the retrieved PWV. Some studies found, using a similar approach, a less than 2 mm theoretical error for the GNSS-retrieved PWV (van Malderen et al. 2014; J. Wang et al. 2007).

3 Data sets and methodology

3.1 Study area

The area of this study is the Atlantic coast of Spain and France. This area has an oceanic climate, typified by low annual temperature variation, with a mild winter and warm summer, but without high temperatures. Regarding the precipitation, it is consistent the whole year, especially on coastal places. One remarkable characteristic of this area is its high humidity. Covering the whole area, nine GNSS sites were used (Fig. 1). Five of the sites contribute to the EUREF network. The stations of Spain are managed by the National Geographical Institute of Spain (IGE) while the French GNSS stations are managed by the

Fig. 1 Location of the GNSS stations used in the study in the Atlantic coast of Spain and France



National Institute of Geographic and Forest Information of France (IGN). Table 1 summarizes the locations of the GNSS sites.

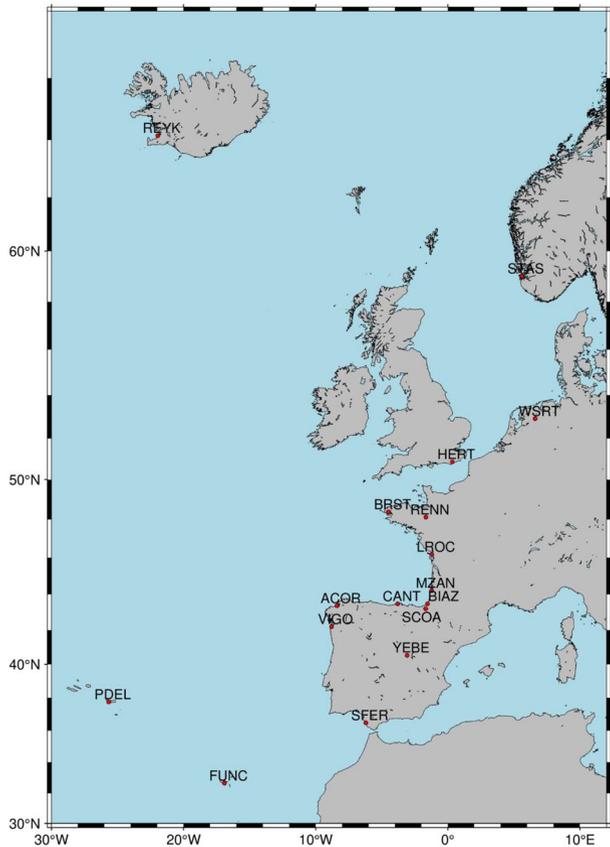
3.2 Processing of GNSS observations

The processing of GNSS observations, to obtain the ZTD, was done with the Bernese software, version 5.2, which was developed by the Astronomical Institute of the University of Bern (Dach et al. 2015). This software was used on a double difference phase approach. The cluster of the relative processing (Fig. 2) was designed considering that, on a double

Table 1 Locations of the GNSS sites used in this study. Latitudes and longitudes are given in decimal degrees. The height is given with respect to the WGS84 ellipsoid

GNSS site	Latitude (°N)	Longitude (°W)	Ellipsoidal height (m)
VIGO	42.184	8.813	87.780
ACOR	43.364	8.399	66.920
CANT	43.472	3.798	99.250
SCOA	43.395	1.682	59.470
BIAZ	43.472	1.537	121.420
MZAN	44.201	1.230	78.080
LROC	46.159	1.219	57.860
RENN	48.109	1.667	93.270
BRST	48.380	4.497	65.810

Fig. 2 Distribution of the GNSS stations used in the GNSS processing



difference processing approach, the quality of the result is sensitive to the network configuration used on the calculation and should include long baselines, with the aim of obtaining absolute ZTD estimates. Tregoning et al. (1998) showed that networks including baselines longer than 2000 kms provide more accurate PWV estimates from GNSS observations, increasing the recommended value of 500 kms found by the study of Rocken et al. (1993). The cluster consists of 17 stations. In this set, 13 GNSS stations contribute to the EUREF network, whereas 6 of them contribute also to the IGS network.

Troposphere modelling in the Bernese software is based on a priori model that is corrected by the site-specific troposphere parameters, that are represented in a piecewise linear, continuous way (Dach et al. 2015). Because of the fact that the a priori model absorbs virtually the entire hydrostatic component of the troposphere delay, the wet component is estimated by those site-specific parameters and its corresponding mapping wet function. Moreover, Bernese software allows to estimate the gradient parameters, as a time-dependent horizontal north and east parameters. Due to the use of low elevation data, estimation of these tropospheric gradient parameters is essential to decorrelate station heights and tropospheric corrections. Then, the total tropospheric delay is the sum of the a priori model, its correction by the site-specific troposphere parameters, and the horizontal gradients. The needed tasks to do the described troposphere modelling on Bernese processing can be summarized as follows. First step is the preparation of the pole and orbit

information, generating the standard orbits from IGS final orbits. Also, the observation Rinex data from GNSS sites are converted, pre-processed, and synchronized. After the formation of the baselines, essential in the Double Difference approach, the phase baselines files are pre-processed to detect and correct cycle-slip. Then, a residual screening of the observation data is done. Next step is the resolution of phase ambiguities. According to the characteristics of the network, the Quasi-Iono Free (QIF) strategy was chosen for this study. Then, the ambiguity-fixed network solution is computed and coordinates as well as troposphere parameters are obtained. As the final step, the Back Substitution procedure is applied, and the final TRP and TRO files, with 15 min rate, are written. The main parameters used on these described tasks are listed on Table 2. The elevation cutoff was set to 3° according to the Guidelines of EUREF EPN (https://www.epncb.oma.be/_documentation/guidelines/ accessed in November 2022).

3.3 Global pressure and temperature 3

Global Pressure and Temperature 3, GPT3, is an empirical model that can act as a full empirical troposphere model since it outputs all the information that may be required in troposphere modelling (Landskron 2017). As explained, the GPT3 model was developed using 10 years (2001–2010) of monthly mean pressure levels from ERA-Interim reanalysis and is provided in two resolution grids of 1° × 1° and 5° × 5°. Even though the GPT3 and its predecessor GPT2w are very similar, GPT3 achieves slightly better results using the 1° × 1° grid version (Landskron and Böhm 2017). As input parameters, GPT3 uses the modified Julian date and the coordinates of the site (latitude, longitude, and ellipsoidal height). Also, two different cases with or without time variation can be chosen on the input. If the time variation is chosen, the model will consider the characteristics of annual and semi-annual cycles on the atmospheric parameters using the equation:

$$y_i = A_{0i} + A_{1i} \cos\left(\frac{DOY}{365.25} 2\pi\right) + B_{1i} \sin\left(\frac{DOY}{365.25} 2\pi\right) + A_{2i} \cos\left(\frac{DOY}{365.25} 4\pi\right) + B_{2i} \sin\left(\frac{DOY}{365.25} 4\pi\right) \quad (14)$$

where y_i is any of the atmospheric parameters or coefficients that can be obtained by the model and DOY is the day of year. A_{0i} represents the mean value of y_i , A_{1i} and B_{1i} are the annual amplitudes and A_{2i} and B_{2i} are the semi-annual amplitudes of y_i , and all are obtained from the grid of the model. As a last step, a bilinear interpolation is done to calculate the coefficients and atmospheric parameters on the desired location.

Table 2 Main processing parameters used in Bernese GNSS Software 5.2

Parameter	Bernese processing
Elevation Cutoff	3°
Sampling Rate	30s
Satellite Orbit	Final IGS orbits
A priori troposphere model	GPT dry with GMF dry mapping
Mapping Function	GMF
Tropospheric Gradients	Estimated
Gradient Model	CHENHER
Ambiguity Strategy	Quasi Ionosphere-Free (QIF)

In this study, the GPT3 model was used on the $1^\circ \times 1^\circ$ grid version with time variation (VMF Data Server). The scripts and the corresponding needed files and grids can be downloaded from <http://vmf.geo.tuwien.ac.at/codes/> (accessed in November 2022). From the set of GPT3 output quantities, the water-vapour-weighted mean temperature and pressure were selected. Then, atmospheric pressure on the sites was applied on the Saastamoinen model to estimate the ZHD (Eq. 6). Once ZTD and ZHD were calculated, ZWD was obtained from the subtraction of these two quantities (Eq. 8). Hereafter, the estimation of the water-vapour-weighted mean temperature from GPT3 model was used to calculate the factor . Using this factor, ZWD was finally converted into PWV (Eq. 9).

The behaviour of GPT3 model is influenced by geographical location, with worse performance in higher latitudes or with drastic climate changes (S. Li et al. 2022a, b, c) and has a limited horizontal resolution of $1^\circ \times 1^\circ$. Moreover, even the annual and semi-annual cycles are considered by the model, the diurnal variations are not, so GPT3 model is unable to capture the fluctuations of the atmospheric parameters along the day. However, GPT3 model, as a full troposphere model, provides each needed atmospheric parameter so can be used in a consistent way in the retrieval of PWV from GNSS observations without in situ meteorological measurements and in a real-time retrieval requirement. GPT3 model has been considered a highly efficient model and has been adopted in different applications on PWV retrieval (Tunali, 2022; J. Li et al. 2022a, b, c).

3.4 Radiosonde data

Radiosonde technique consists of a balloon that carries sensors that can measure different parameters while ascending in the atmosphere. Radiosonde is then, a direct measure of the vertical profile of atmospheric parameters from which different atmospheric variables can be retrieved. Precipitable Water Vapour from radiosonde data can be obtained by:

$$PWV = \frac{10^3}{\rho} \int_{z=0}^{z=\infty} \rho_w dz \quad (15)$$

where ρ and ρ_w are the liquid water and vapour water densities respectively, and z is the height in meters (Castro-Almazán et al. 2016).

Radiosonde data can be downloaded from different data sets. In this work, radiosonde data was obtained from the Integrated Global Radiosonde Archive (IGRA) version 2 (Durre et al. 2018). Version 2 of the archive contained observations from almost 2800 radiosonde sites around the world and it has a growing volume of data and spatial coverage. IGRA archive has also a data set of derived quantities, including the Precipitable Water Vapour. One remarkable characteristic of this derived data set is that, for its calculation, pressure level is up to 500 hPa. This is justified by the poorer behaviour of the humidity sensors on small temperatures (Ross and Elliott 1996).

The IGRA dataset used in this work can be obtained from the NCDC FTP Web site (online at <ftp://ftp.ncdc.noaa.gov/pub/data/igra> accessed in November 2022).

The accuracy of radiosonde data is affected by the sensor design, calibration, data processing and contamination (Moradi et al. 2013). Moreover, the changes in humidity sensors on the radiosonde sites introduce a discontinuity in the radiosonde climatic records that can lead to an inaccurate climate trend estimate (Wang and Zhang 2008). Sapucci et al. (2005) focused on the quality of humidity measurement of the radiosonde on tropical areas and its impact over the IWV retrieved. The study concluded that different sensors on the

radiosonde had a similar behaviour especially on the lower atmosphere, where a major part of the water vapour is contained, and that the IWV could be retrieved with a rms of only $1.3 \text{ kg}\cdot\text{m}^{-1}$. J. Wang and Zhang (2008) studied the uncertainty on the radiosonde data focused on the identification of its systematic errors using GNSS observations. They used 130 matched GNSS and Radiosonde sites, analysing 14 different models of radiosonde. Regarding to the measurement errors of the radiosonde, the cited study concluded that Vaisala RS92 model, the one used at the matched sites in this study, had a dry bias to values of PWV less than 40 mm. Miloshevich et al. (2009) investigated the accuracy of Vaisala RS92 radiosonde model, comparing the humidity measurements of this sensor to simultaneous measurements from three reference instruments, and estimated a precision of 5% of the PWV on the retrieval of the water vapour.

3.5 EUREF REPRO-2

The second EUREF Permanent Network (EPN) reprocessing campaign, EPN REPRO-2, might be used as a reference data set for meteorological and climate research (Pacione et al. 2017). On that campaign, GNSS data of about 280 stations were reprocessed by 5 Analysis Centres using 3 GNSS software packages: Bernese (Dach et al. 2015), GAMIT (Herring et al. 2010) and GIPSY-OASIS II (Webb 1997). The resulting values of the ZTD for the network were evaluated to independent data sets as radiosonde and ERA-Interim, showing an improvement in the overall standard deviation compared with the first reprocessing campaign (EPN REPRO-1). The period covered by the second reprocessing campaign, for the tropospheric products, is from GPS week 834 to GPS week 1824. In order to compare with this latest reprocessing and reference data of ZTD from EUREF, the processing on this study started in July 2013 covering 87 common weeks and with information of the four-weather season. The values of troposphere parameters from EPN REPRO-2 can be downloaded at <ftp://igs.bkg.bund.de/EPNrepro2/products/> (accessed in November 2022).

3.6 Statistical parameters

Mean Bias Error (MBE), Standard Deviation (SD) and Root Mean Square (RMS), the statistical parameters used in this study, were calculated using the following formulas:

$$MBE = \frac{\sum_{i=1}^N (PWV_{GNSS} - PWV_{RS})}{N} \quad (16)$$

$$SD = \sqrt{\frac{\sum_{i=1}^N (PWV_{GNSS} - PWV_{RS})^2}{N - 1}} \quad (17)$$

$$RMS = \sqrt{\frac{\sum_{i=1}^N (PWV_{GNSS} - PWV_{RS})^2}{N}} \quad (18)$$

where PWV_{GNSS} is the Precipitable Water Vapour obtained as explained on Sect. 2 and PWV_{RS} corresponds to the value extracted from the IGRA radiosonde archive.

Table 3 Overall statistical results of the comparison of ZTD estimated versus EPN Repro 2 from GPS Week 1737 to GPS Week 1824

GNSS Site	N	Correlation coefficient	MBE (mm)	SD (mm)	RMS (mm)
VIGO	12,201	0.99	-0.65	2.37	2.46
ACOR	11,863	0.99	0.14	2.54	2.55
CANT	12,289	0.99	0.11	2.43	2.43
SCOA	12,088	0.99	-0.13	2.48	2.48
BRST	12,526	0.99	0.26	2.36	2.38

Table 4 Relation of GNSS and RS matched stations

GNSS Site	RS IGRA Code	Horizontal Distance	Height Difference (GNSS-RS)
ACOR	SPM00008001	~ 1.8 km	~ +9 m
CANT	SPM00008023	~ 2.1 km	~ +47 m

4 Results

4.1 Validation of ZTD estimates using EUREF REPRO-2

The common weeks between EPN REPRO-2 and this work, from GPS Week 1737 to GPS Week 1824, were analysed on the five GNSS sites of the network that contribute to the EUREF network. Statistical results of the comparison are summarized on Table 3.

Results showed a very high correlation factor, with values greater than 0.996 on all sites. Mean Bias Error are low and clearly under one millimetre. Standard deviation and root mean square have no remarkable differences between the sites. Considering all the stations, mean values for correlation factor, MBE, SD and RMS are 0.99, -0.05, 2.44 and 2.46 respectively. The achieved statistical parameters show that the ZTD calculated by the processing are consistent with the official products and are then validated to proceed to the next step.

4.2 Validation of PWV using radiosonde data

Radiosonde, as a traditional technique to retrieve the PWV, it is often used as a reference to validate the results of other techniques, like GNSS. The GNSS and radiosonde site, if not collocated, must be in a maximum range of horizontal and vertical separation to be able to compare their PWV values. In fact, minimizing the vertical difference between the sites minimizes the PWV differences between the two techniques (J. Wang and Zhang 2009). Ohtani and Naito (2000) suggests a maximal horizontal distance of 40 kms with a difference of height of 100 m. Gui et al. (2017) increases the horizontal distance to 50 kms, with the same limit on the height difference. According to those limitations, in this work, two radiosonde sites, managed by the State Meteorological Agency of Spain (AEMET) were used. The radiosonde stations are located near to the ACOR

and CANT GNSS sites. Table 4 shows these matched stations and their differences on position and height.

Apart from the spatial difference between the matched sites, the difference in temporal resolution of the PWV obtained from the two techniques must also be considered. While ZTD from the processing of the GNSS observations were obtained with a sampling rate of 15 min, temporal resolution of the two radiosonde sites involved in this study was 12 h, with 2 launches per day, at 0UTC and 12UTC. In addition, the radiosonde is often released before this time to reach the tropopause around the synoptic time. The release time of the sounding is provided on the radiosonde data obtained from the IGRA archive. A review of this release time shows that, in the case of ACOR station, more than 92% of the launches were done more than 15 min before the synoptic time. On the CANT site, around 85% of the radiosondes were released more than 15 min before the synoptic time. This is important because most of the water vapour retrieved by the radiosonde is measured in the first 15 min of its ascending way because it is contained in the lower part of the troposphere. Some studies, taking all of this into account, do not match the GNSS data with radiosonde data at the exact time of 0UTC and 12 UTC. For example, Ning et al. (2012) suggests a shift of 30 min on values from the GNSS data set. Ohtani and Naito (2000) with a temporal resolution of the PWV from the GNSS observations of two hours, used a mean value from their two values calculated before and after the synoptic time. In this study, the mean of the 15 min rate PWV retrieved from GNSS observations from the release time to the synoptic time is calculated and then compared to the PWV retrieved from the radiosonde.

The validation was done between GPS Week 1737 to 1824, with information of each one of the four climate seasons. Within this period, on the ACOR site, values of PWV from the IGRA archive varied from a minimum of 4.83 mm, in December of 2013, on the local season of winter, to a peak maximum of 43.19 mm in October of 2013. Apart from that peak, the maximum values for ACOR site are found in the local summer season. In the case of CANT, the range of PWV varies from a minimum of 3.59 mm to a maximum of 41.08 mm in winter and summer local seasons respectively. According to the values found in the IGRA archive for PWV at the stations, on ACOR, 70% of the PWV values are from 10 to 25 mm while on CANT, same percentage is obtained for values between 15 and 32 mm.

The statistical parameters are shown in Table 5. The two data sets agree well with a high correlation and MBE values are low and under 1.5 mm. The linear regression (Fig. 3) yields a high R^2 value for both sites, with values of 0.86 and 0.88 to ACOR and CANT respectively. The slopes of the linear regression lines show a direct relation between the two datasets and the slope of the linear regression, less than 1, indicating the dependence on the amount of PWV since when PWV increases the regression line separates from the diagonal. Besides, considering the intercept values, when PWV are low, the GNSS technique tends to underestimate the value of PWV but, when PWV are high, the technique has the opposite behaviour. So, the difference between the two

Table 5 Statistics summary of the comparison between PWV-retrieved from GNSS and Radiosonde

GNSS Site	N	Correlation coefficient	MBE (mm)	SD (mm)	RMS (mm)
ACOR	980	0.93	1.03	2.70	2.89
CANT	982	0.94	1.48	2.62	3.00

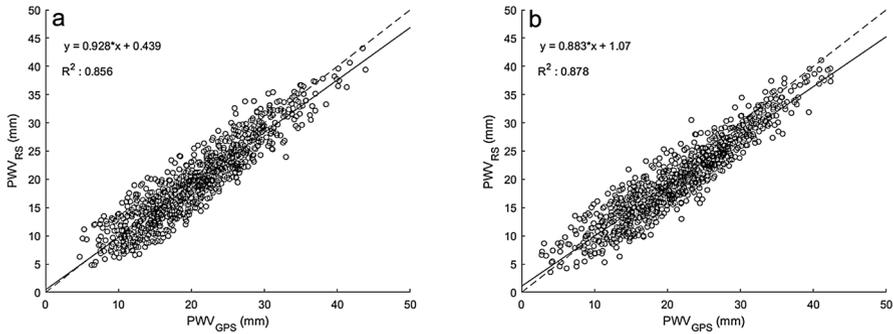


Fig. 3 Scatter diagrams of PWV derived from GNSS and radiosonde at **a** ACOR and **b** CANT sites

techniques is greater as the PWV increases. This performance is clearer seen on the CANT site. A similar behaviour can be found in other previous studies (Deblonde et al. 2005; Namaoui et al. 2017).

Since the matched radiosonde sites had two launches, at 0 UTC and 12 UTC, the performance at day and night could be studied. Statistical comparisons calculated with Eqs. 10, 11, 12 are shown in Table 6. Correlation coefficients are again high, all above 0.93 and with no significant differences in this parameter between daytime and night-time measurements. Related to the MBE, SD and RMS statistical parameters, on the ACOR site the differences are also not significant showing on CANT site a higher difference on the three parameters associated with the day or night retrieval. This can also be seen in the linear regression over the data (Fig. 4). The analysis shows a very small difference on the ACOR site both on slopes and R^2 parameter while on the CANT site differences are more significant. The different behaviour at daytime and night-time of the radiosonde Vaisala RS92-SGP, the one used on ACOR and CANT sites, have been observed to a greater or lesser extent in different studies and it is mainly related to the different performance of the humid sensor because of the solar heating during the day (Cady-Pereira et al. 2008; van Malderen et al. 2014).

Wang and Zhang (2008) explained that the differences between the two techniques, that can possibly make the expected theoretical error unachievable, come from three main sources: the distance, both horizontal and vertical, between the matched sites, the differences in the measurement technique and finally the errors in the data sets, both radiosonde and GNSS. Regarding the effect of not being collocated, the cited study claimed that there could arise an important difference in the water vapour retrieved from the two techniques, especially with strong precipitable water vapour gradients. About

Table 6 Statistical parameters of the comparison of Radiosonde and GNSS PWV regarding to the launching time

GNSS Site	N	Correlation coefficient	MBE (mm)	SD (mm)	RMS (mm)	Synoptic time
ACOR	488	0.93	1.02	2.66	2.85	0 UTC
CANT	483	0.94	1.38	2.62	2.96	0 UTC
ACOR	492	0.93	1.04	2.75	2.94	12 UTC
CANT	499	0.94	1.60	2.60	3.05	12 UTC

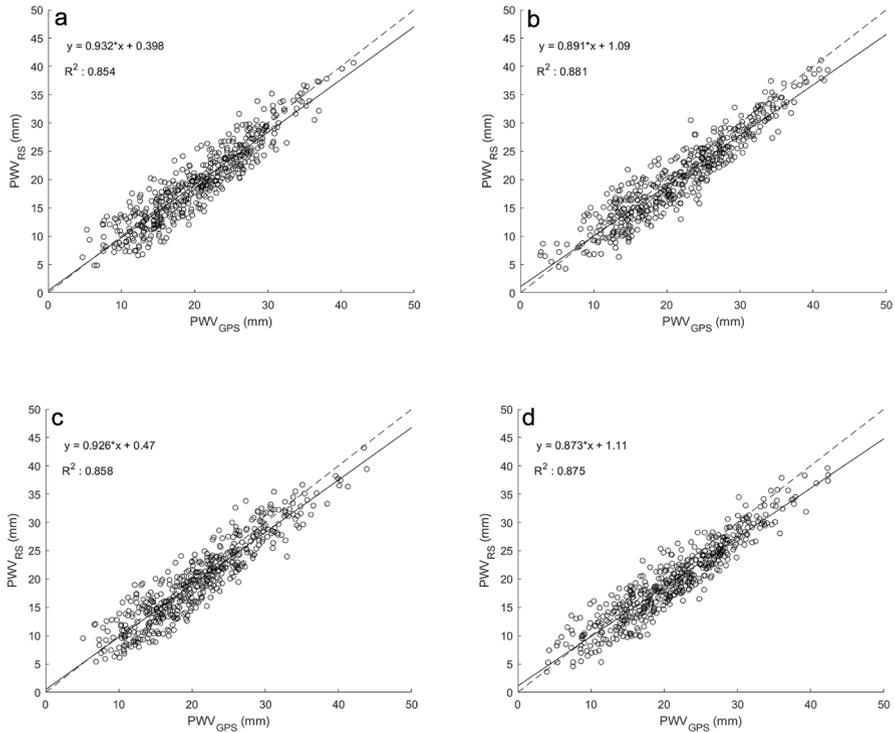


Fig. 4 Scatter plots of the comparison of Radiosonde and GNSS PWV regarding to the launching time at **a** ACOR 0 UTC, **b** CANT 0 UTC, **c** ACOR 12 UTC and **(d)** CANT 12 UTC

the measurement techniques, the PWV-retrieved from the two techniques comes from a different volume of air. In the case of the GNSS technique, the ZTD corresponds to the vertical mapped value, using mapping functions, from all the paths between the antenna and the satellites, with different azimuths and elevation angles. In the case of the radiosonde, the profile of atmospheric parameters, from which PWV is obtained, is measured in the radiosonde ascending path. This path could be non-vertical, suffering a displacement by the wind. Although the displacement that the radiosonde suffers on its complete path is difficult to quantify, the cited study explained that it could be as much as 20 kms in the horizontal direction in 10 kms of ascent. This displacement could affect more on sites situated near the coast because some part of the PWV profile retrieved by the radiosonde could be measured over the sea, increasing the differences between the GNSS and radiosonde technique (Vaquero-Martínez et al. 2019). Since the ACOR and CANT stations are located on the coast, the above-mentioned effect could have occurred. Related to the errors of the two techniques, even though radiosonde is considered as a reference, it has its own errors. The use of a third technique could further clarify if the differences could come from a mix of errors on the two techniques (Ning et al. 2015). However, the results achieved in this comparison between the two techniques and shown on Table 5 and Table 6, are consistent with other comparisons. For example, similar values of MBE can be found in other studies (Gui et al. 2017; Rezaei

and Khazaei 2022). Related to the parameters SD and RMS, similar values can be found in previous studies (Gui et al. 2017; Ohtani and Naito 2000; Torres et al., 2010).

4.3 Characterization of PWV over the area of study

Once the processing was validated, the PWV of each GNSS site were calculated using the TRP and TRO files derived from our Bernese processing, with GPT3 as a model for the pressure and T_m . The calculated values of PWV covered four years, from July 2013 to the end of June 2017, between GPS weeks 1747 to 1955, with a sampling rate of 15 min. Table 7 shows the number of observations (N) and some statistical parameters of the PWV retrieved from GNSS for all the stations, using the daily mean. Table 7 also shows the mean value of PWV considering four seasons: June to August (JJA), September to November (SON), December to February (DJF) and March to May (MMA).

Multiannual mean, using all the stations, is 19.70 mm. LROC, RENN and BRST stations have values below this mean. Considering each station, the multiannual maximum mean is found in SCOA, with a value of 21.30 mm while the multiannual minimum mean is found in RENN with 17.85 mm, the only station situated inland. As noticed in Table 7, PWV retrieved from the GNSS observations capture the seasonal PWV variation, with minimum values observed in the local winter season and maximum values in the local summer season. The multiannual daily mean in the local summer season is almost double the corresponding value in the winter season. The minimum mean value is found in winter in RENN of 12.88 and the maximum mean value is 29.61 in the summer season at SCOA site. The mean of autumn and spring seasons are between the values of the other extreme seasons, although the smallest differences are found between winter and spring season. As noticed in Table 7, PWV retrieved from the GNSS observations using the blind empirical model shows the seasonal PWV variation, with minimum values observed in the local winter season and maximum values in the local summer season. This behaviour can also be observed using the multiannual monthly mean of the PWV (Fig. 5). Clearly, in all the sites, the PWV rises from the minimum in February to its maximum in July. This evident seasonal variation is related with the rise in temperature associates with the changes on seasons in this area. SCOA, BIAZ and CANT, with higher values according to Table 7, have the higher changes between months, while on RENN or BRST the variation are small.

The temporal variation of PWV series were described from its mean and annual and semi-annual terms to the retrieved PWV in each GNSS site. The fit model was constructed as:

$$PWV = PWV_0 + a_1 \sin(2\pi t) + a_2 \cos(2\pi t) + a_3 \sin(4\pi t) + a_4 \cos(4\pi t) \quad (19)$$

where PWV_0 is the estimated mean, a_1 is the linear trend and t is the time in years. The annual amplitudes are the pair (a_2, a_3) while (a_4, a_5) corresponds to the semi-annual amplitudes. Parameters on Eq. 19 were estimated using the least square method and setting the first day of the series, 1st of July 2013, as the time epoch $t=0$. Table 8 shows the results on each site.

Annual amplitudes are clearly more important than the semi-annual in all the sites. Sites that are located closer, tend to have similar amplitudes. The annual amplitudes are stronger in SCOA, BIAZ and CANT, sites with the biggest variability in their PWV over the 4-year period. This annual cycle can also be observed in Fig. 6, which shows the daily mean for the whole four-year period on each site. Regarding each site, minimum values were found on BRST and RENN with values less than 1 mm, while the maximum, higher than 44 mm

Table 7 Statistical parameters of the 4-year PWV-retrieved. The row ALL corresponds with the mean values in the network of study

GNSS Site	N	Mean (mm)	Maximum (mm)	Minimum (mm)	SD (mm)	Mean JJA (mm)	Mean SON (mm)	Mean DJF (mm)	Mean MAM (mm)
VIGO	1433	20.00	43.40	4.36	6.41	23.90	22.45	16.36	17.31
ACOR	1430	19.93	43.65	4.98	6.37	24.39	22.01	16.14	17.19
CANT	1438	20.83	44.66	4.17	7.22	28.00	22.49	14.96	17.88
SCOA	1434	21.30	44.60	3.87	7.81	29.61	22.90	14.57	18.14
BLAZ	1397	20.13	41.55	3.48	7.66	28.18	21.72	13.64	17.00
MZAN	1233	20.08	43.07	3.45	7.40	27.45	21.68	14.43	16.77
LROC	1312	18.70	42.49	1.15	6.93	25.21	19.84	13.60	16.14
RENN	1421	17.85	40.29	0.87	6.92	24.19	19.02	12.88	15.31
BRST	1452	18.51	42.23	0.38	6.79	24.08	20.02	14.02	15.93
ALL		19.70	42.88	2.97	7.06	26.11	21.35	14.51	16.85

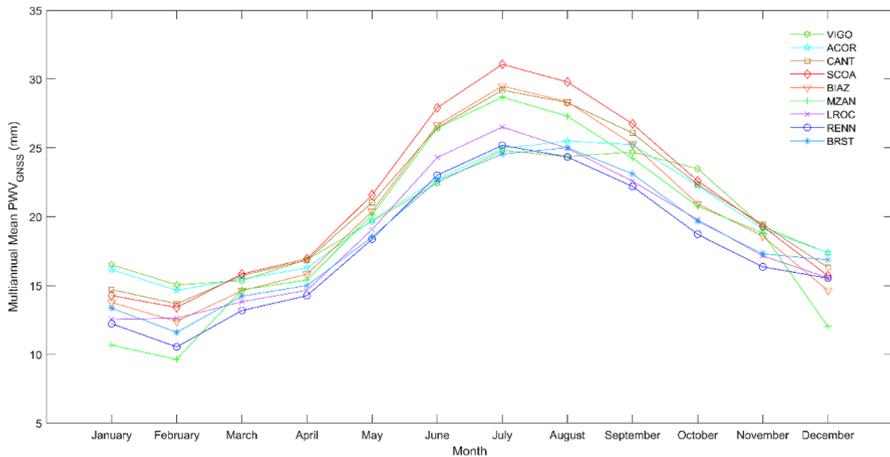


Fig. 5 Distribution of multiannual monthly mean of PWV in each GNSS site

Table 8 Estimated mean and annual and semi-annual terms of GNSS-retrieved PWV series

GNSS Site	PWV ₀ (mm)	a ₁ (mm)	a ₂ (mm)	a ₃ (mm)	a ₄ (mm)
VIGO	20.0	3.8	3.4	0.1	0.0
ACOR	19.9	3.9	3.7	0.5	0.2
CANT	20.9	4.3	6.1	0.6	0.6
SCOA	21.3	4.5	7.1	0.9	0.9
BIAZ	20.1	4.4	6.9	0.8	1.0
MZAN	19.9	4.0	6.0	0.8	1.4
LROC	18.6	3.8	5.4	0.3	1.1
RENN	17.9	3.8	5.3	0.2	1.1
BRST	18.5	3.9	4.6	0.3	0.7

were found on CANT and SCOA, both in 2015. These two stations presented a high variability behaviour for the whole period. Maximum values of 43.4 and 43.7 were also seen in 2015 in VIGO and ACOR respectively. In general, minimum values are observed in most of the sites on the end of 2013 or the beginning of 2014. The distribution of mean daily values shows a bigger amplitude between maximum and minimum in 2015 and 2014 in VIGO, ACOR, CANT, SCOA, BIAZ and MZAN and a clear lower amplitude in 2016. However, this tendency is the opposite in LROC, RENN and BRST where the differences between maximum and minimum are bigger in 2016.

Related to the spatial distribution of water vapour in the area of study, the PWV multiannual mean tends to decrease with increasing latitude, according to the distribution of the multiannual mean with the latitude (Fig. 7). In fact, the line on Fig. 7, that represents the best linear fit between the multiannual PWV mean and latitude, has a negative slope.

The diurnal cycle of PWV is also one of the most important of its variations (J. Wang and Zhang 2009). The high temporal resolution of the PWV retrieved from the GNSS observations, 15 min, facilitates the study of the short-term variations of the PWV on every site. The daily precipitable water vapour anomaly is calculated as the difference between the mean daily and each PWV value calculated from the processing. Figure 8 shows the

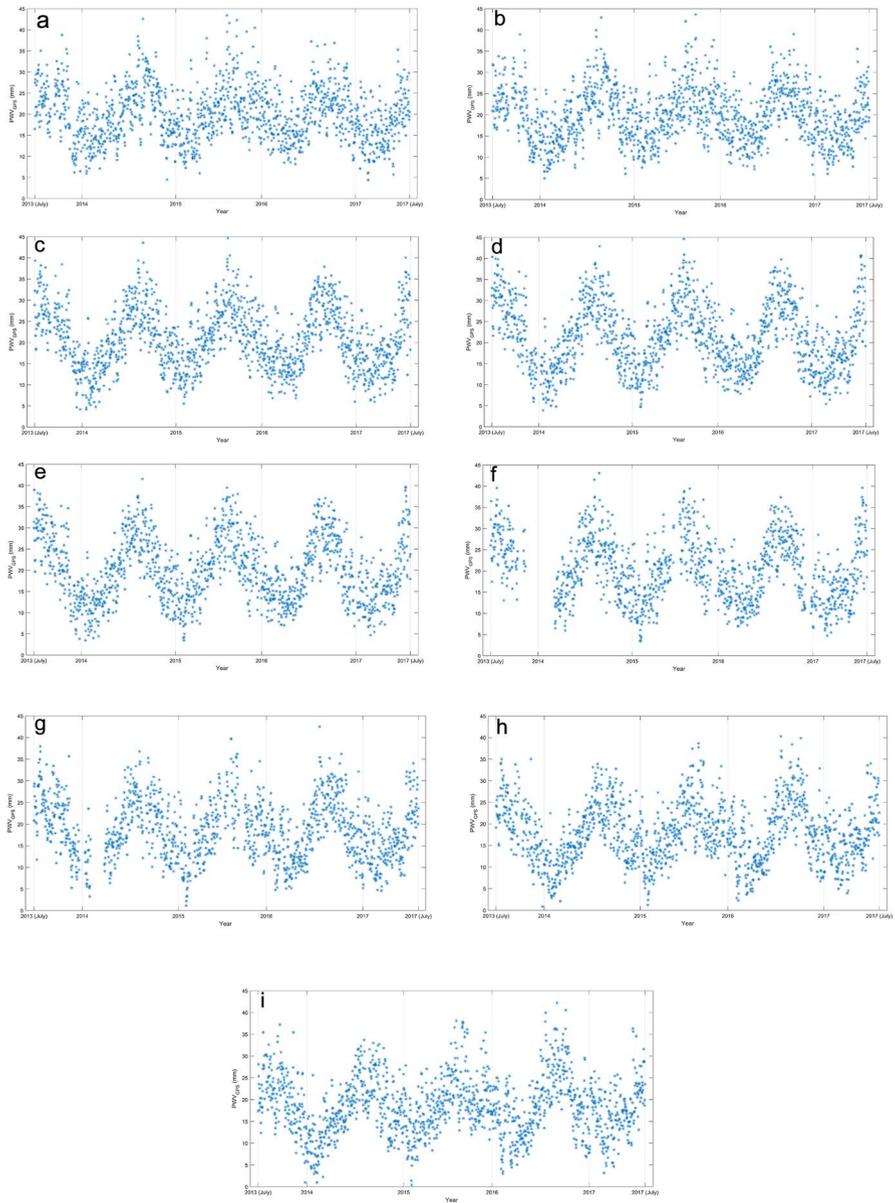


Fig. 6 Time series of daily mean PWV on the GNSS sites from July 2013 to June 2017 at **a** VIGO, **b** ACOR, **c** CANT, **d** SCOA, **e** BIAZ, **f** MZAN, **g** LROC, **h** RENN and **i** BRST

multiannual diurnal anomalies on each site, calculated through the entire period, and considering All seasons, and each of the four seasons: DJF, MAM, JJA and SON.

Using the whole data set of the stations, the maximum multiannual amplitude considering all seasons is found in CANT with a value of 1.31 while LROC has the smallest: 0.54 mm. Diurnal variation is clearly smaller at night than during the daytime at most of

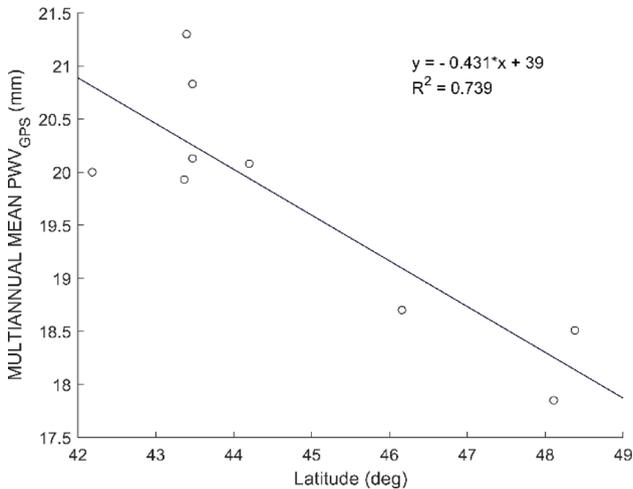


Fig. 7 Distribution of multiannual mean of PWV with latitude

the stations. Related to the extreme values, most of the sites have their minimum value between 5 and 7 UTC, while the maximum values are observed between 20 and 22 UTC except at the RENN site where the maximum is at 19 UTC but with a close value to the one observed at 20 UTC. Temperature mainly affects the diurnal cycle (Ortiz de Galisteo et al. 2011). This explains the minimum at night and the morning increase of PWV according to the rise of temperature that was also found in a study of the characterization of PWV in the Gulf of Cadiz, in Spain (Torres et al., 2010). Some stations exhibit double peaks on their multiannual diurnal anomalies. For example, MZAN has a minimum at 7 UTC, but another minimum is observed on this site at 16 UTC, to reach the maximum at 22 UTC. A similar behaviour is found in LROC, with two minimums, one at 6 UTC and another at 16 UTC to reach the maximum at 21 UTC, but with lower values. At the ACOR site, a change is observed on the diurnal anomaly around 18 UTC.

The diurnal anomalies also capture the seasonal variability (Fig. 8). Thus, as expected, the highest values of amplitude occur in summertime, with SCOA and its neighbour site BIAZ, with the highest amplitudes in this season. In fact, SCOA has the biggest value of amplitude with 2.04 mm in the summer season. The differences between summer and winter are more significant in VIGO, RENN, SCOA and ACOR where amplitude values are doubled in summer than in winter. This performance is consistent with the study in Europe from Wang and Zhang (2009) where the diurnal cycle detected at summertime was clearly higher than in winter. In the case of CANT, the site with the highest amplitude considering the entire 4-years period, shows a similar behaviour among the four seasons, although it maintains a greater variability in summer and spring than winter and autumn. In most sites, spring and autumn have intermediate amplitudes compared to those found for summer and winter; spring however tends to be more like winter. LROC, RENN and BRST, the sites with higher latitude, have less amplitude in every season than the rest of the sites.

There are many different causes for the diurnal variations of PWV. Dai et al. (2002) in their study over North America, explained some principal causes, for example, the surface evapotranspiration, the atmospheric large-scale vertical motion, the atmospheric low-level moisture convergence, the precipitation and, finally, the vertical mixing in the lowest

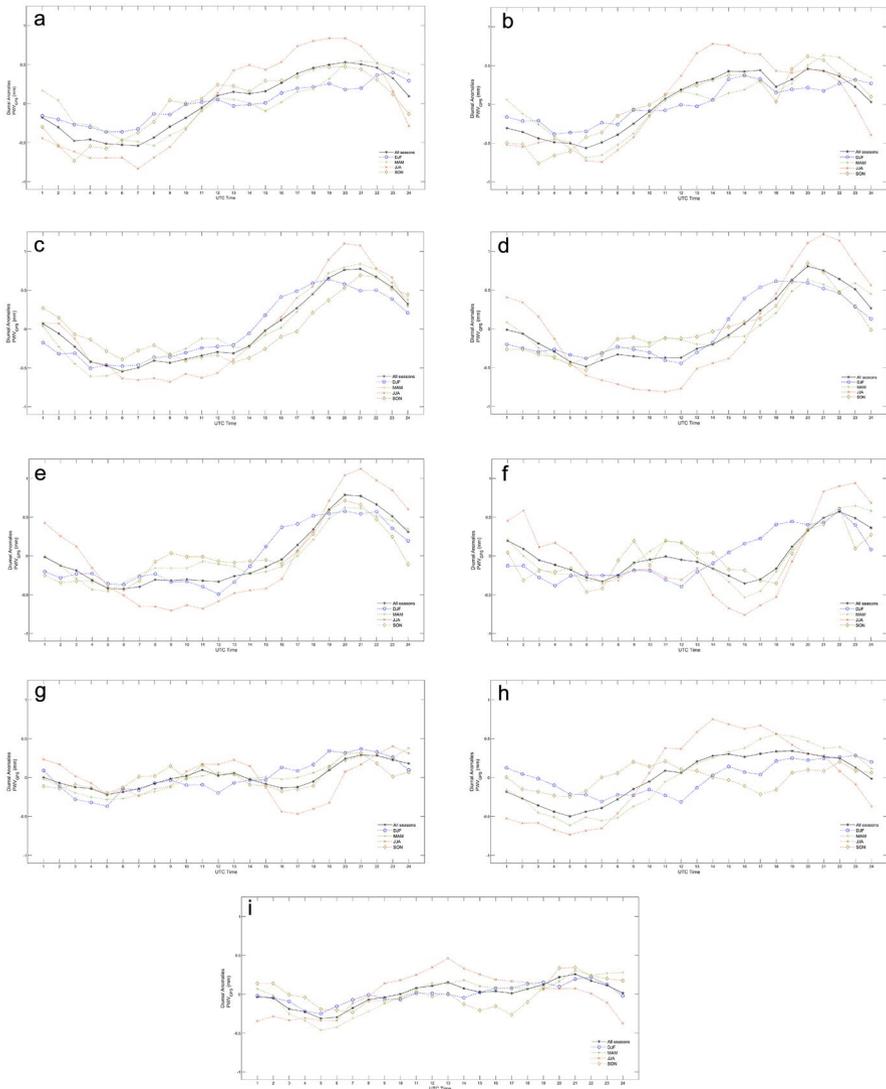


Fig. 8 Multiannual diurnal PWV anomalies considering All Seasons, and each DJF, MAM, JJA and SON season as indicated in the legend, at **a** VIGO, **b** ACOR, **c** CANT, **d** SCOA, **e** BIAZ, **f** MZAN, **g** LROC, **h** RENN and **i** BRST

atmosphere. Wind also plays an important role. On their study over a network in Japan, G. Li et al. (2008) concluded that, the phase of the diurnal cycle of stations situated on the coast, like the sites of this study, is clearly different from the one found on other areas because of the sea breeze that introduces humidity on the atmosphere of coastal sites and affects the PWV variation during the day. Related to the seasonal variation, Kannemadugu et al. (2022) in their study over East Africa concluded that the small variations of temperature during the winter season accounts for the small amplitude found on this season as in

the occurrence in this study. In summary, the diurnal variations of the PWV on every site are caused by the complex mix of the aforementioned causes and other factors such as the orography that plays an important role. Therefore, these short-term variations, as well as the seasonal variability found with GNSS-retrieved PWV could be applied on further studies on climate over the area of study and can be useful to model the complex variability of the PWV. Biondi (2015) indicates that, the short-term variations captured by the GNSS technique, mixed with a correct distribution of the GNSS sites, help to characterize the rainy events. Consequently, GNSS-retrieved PWV could be used to improve real time short-term rainfall forecasting (Cao et al. 2015). In fact, the characterization of GNSS-retrieved PWV before a rainfall event has been used to develop rainfall forecasting models (H. Li et al., 2020, L. Li et al. 2022a, b, c). Besides that, Ortiz de Galisteo et al. (2011) claimed that the knowledge of the diurnal cycle of PWV could improve the atmospheric correction of the satellite images.

5 Summary and conclusions

This study presents the characterization of PWV on the Atlantic coast of Spain and France using GNSS observations without any meteorological in situ measurements, using the empirical model GPT3. The starting point of the PWV retrieval from the GNSS observations is the calculation of the ZTD that is estimated with other parameters in the double difference phase processing of the observations. As given its importance as the starting point, the ZTD obtained from the processing were validated with the official values from EUREF REPRO-2 obtaining a mean correlation factor, MBE, SD and RMS of 0.99, -0.05, 2.44 and 2.46 respectively. Then, the rest of the parameters needed for the calculation of the PWV, as the pressure and the weighted mean temperature were extracted from the global blind model GPT3, using its grid of $1^\circ \times 1^\circ$ resolution. Thereafter, 15-min rate PWV values were calculated through a four-year period at the nine sites distributed on the area of study. The PWV calculated was validated using two matched Radiosonde stations. The statistical parameters found showed a good agreement between the two data sets with values of MBE less than 1.5 mm and mean SD and RMS of 2.66 and 2.95 mm respectively. Even if the values agree with previous studies, it will be desirable to include a third independent PWV retrieval technique in order to verify the discovered differences.

The characterization of PWV over the area of study using GPT3 model clearly captures the seasonal variability with the maximum value on local summer season and minimums on winter season in all sites. In fact, the multiannual mean in JJA season is almost double than of DJF season. Maximum mean values are retrieved on SCOA, CANT and BIAZ with 21.30, 20.83, 20.13 mm respectively, while the minimum corresponds to the only inland station: RENN with 17.85 mm. The annual cycle is clearly more important, in all the sites, than the semi-annual cycle. Related to the spatial distribution, PWV multiannual mean tends to decrease as the latitude increases. The multiannual diurnal anomalies also reflect the seasonal variability, with higher amplitudes in summer local season and smallest on winter local season. The highest amplitudes found in the CANT site that exhibit a great variability, followed by SCOA and BIAZ. Even though the sites have different behaviour on their diurnal variation, a similar pattern is detected with a minimum in the late-night hours to reach a peak in the afternoon.

The PWV characterization on the area using GNSS observations demonstrates the potential of the use of the GNSS observations as an excellent source of information of the spatial and temporal behaviour of the water vapour present in the atmosphere, even without meteorological in situ measurements. Future research on the long-term PWV variations will be desirable, enlarging the time series of GNSS-retrieved PWV over the area in order to be able to estimate the trends. Moreover, future investigation of the behaviour of the PWV before the severe rainfall events in the area will be done to develop a real time short-term rainfall forecasting model with the use of the empirical model GPT3.

Acknowledgements We would like to thank the IGN, IGE and EUREF for providing access to GNSS observations, the NOAA Integrated Global Radiosonde Archive for providing radiosonde data and Vienna University of Technology for publishing VMF Data Server.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Raquel Perdiguer-Lopez. The first draft of the manuscript was written by Raquel Perdiguer-Lopez and Jose Luis Berne Valero and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interest to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Baldysz Z, Nykiel G (2019) Improved empirical coefficients for estimating water vapor weighted mean temperature over Europe for GNSS applications. In Remote Sens. <https://doi.org/10.3390/rs11171995>
- Bevis M, Businger F, Herring T, Rocken C, Anthes R, Ware R (1992) GPS meteorology: remote sensing of atmospheric water vapor using the global positioning system. *J Geophys Res.* <https://doi.org/10.1029/92JD01517>
- Bevis M, Businger S, Chiswell S, Herring TA, Anthes RA, Rocken C, Ware RH (1994) GPS meteorology: mapping zenith wet delays onto precipitable water. *J Appl Meteorol Climatol* 33(3):379–386. [https://doi.org/10.1175/1520-0450\(1994\)033%3c0379:GMMZWD%3e2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033%3c0379:GMMZWD%3e2.0.CO;2)
- Biondi R (2015) The usefulness of the global navigation satellite systems (GNSS) in the analysis of precipitation events. *Atmospheric Res* 167:15–23. <https://doi.org/10.1016/j.atmosres.2015.07.011>
- Bock O, Guichard F, Janicot S, Lafore JP, Bouin M-N, Sultan B (2007) Multiscale analysis of precipitable water vapor over Africa from GPS data and ECMWF analyses. *Geophys Res Lett.* <https://doi.org/10.1029/2006GL028039>
- Böhm J, Möller G, Schindelegger M, Pain G, Weber R (2015) Development of an improved empirical model for slant delays in the troposphere (GPT2w). *GPS Solut* 19(3):433–441. <https://doi.org/10.1007/s10291-014-0403-7>
- Cady-Pereira KE, Shephard MW, Turner DD, Mlawer EJ, Clough SA, Wagner TJ (2008) Improved daytime column-integrated precipitable water vapor from vaisala radiosonde humidity sensors. *J Atmos Oceanic Tech* 25(6):873–883. <https://doi.org/10.1175/2007JTECHA1027.1>

- Cao Y, Guo H, Liao R, Uradzinski M (2015) Analysis of water vapor characteristics of regional rainfall around Poyang Lake using ground-based GPS observations. *Acta Geod Geoph* 51:467–479. <https://doi.org/10.1007/s40328-015-0137-1>
- Castro-Almazán J, Perez Jordan G, Munoz-Tunon C (2016) A semiempirical error estimation technique for PWV derived from atmospheric radiosonde data. *Atmospheric Meas Tech* 9:4759–4781. <https://doi.org/10.5194/amt-9-4759-2016>
- Chen P, Yao W (2015) GTm_X: a new version global weighted mean temperature model. *Lecture Notes in Electr Eng* 341:605–611. https://doi.org/10.1007/978-3-662-46635-3_51
- Dach, R., Lutz, S., Walser, P., and Fridez, P. (2015). *2015: Bernese GNSS Software Version 5.2. User manual*. Astronomical Institute, University of Bern, Bern Open Publishing. <https://doi.org/10.7892/boris.72297>
- Dai A, Wang J, Ware RH, van Hove T (2002) Diurnal variation in water vapor over North America and its implications for sampling errors in radiosonde humidity. *J Geophys Res Atmos*. <https://doi.org/10.1029/2001JD000642>
- Deblonde G, MacPherson S, Mireault Y, Héroux P (2005) Evaluation of GPS precipitable water over Canada and the IGS network. *J Appl Meteorol - J APPL METEOROL* 44:153–166. <https://doi.org/10.1175/JAM-2201.1>
- Duan J, Bevis M, Fang P, Bock Y, Chiswell S, Businger S, Rocken C, Solheim F, van Hove T, Ware R, McClusky S, Herring TA, King RW (1996) GPS meteorology: direct estimation of the absolute value of precipitable water. *J Appl Meteorol* 35(6):830–838
- Durre I, Yin X, Vose RS, Applequist S, Arnfield J (2018) Enhancing the data coverage in the integrated global radiosonde archive. *J Atmos Oceanic Tech* 35(9):1753–1770. <https://doi.org/10.1175/JTECH-D-17-0223.1>
- Essen L, Froome KD (1951) The refractive indices and dielectric constants of air and its principal constituents at 24,000 Mc/s. *Proc Phys Soc Sect B* 64(10):862. <https://doi.org/10.1088/0370-1301/64/10/303>
- Firsov KM, Chesnokova TYu, Bobrov E, v, and Klitochenko, I. I. (2013) Total water vapor content retrieval from sun photometer data. *Atmospheric and Oceanic Optics* 26(4):281–284. <https://doi.org/10.1134/S1024856013040040>
- Foelsche U, Kirchengast G (2001) Tropospheric water vapor imaging by combination of ground-based and spaceborne GNSS sounding data. *J Geophys Res* 106(D21):27221–27231. <https://doi.org/10.1029/2001JD900230>
- Ortiz de Galisteo Marín (2011) *Análisis del contenido total en columna de vapor de agua atmosférico sobre la Península Ibérica medido con distintas técnicas: radiosondas, fotómetros solares y sistema GPS* [Universidad de Valladolid]. <https://doi.org/10.35376/10324/888>
- Gao B-C, Kaufman YJ (2003) Water vapor retrievals using moderate resolution imaging spectroradiometer (MODIS) near-infrared channels. *J Geophys Res*. <https://doi.org/10.1029/2002JD003023>
- Guerova G, Jones J, Douša J, Dick G, de Haan S, Pottiaux E, Bock O, Pacione R, Elgered G, Vedel H, Bender M (2016) Review of the state of the art and future prospects of the ground-based GNSS meteorology in Europe. *Atmos Measur Tech* 9(11):5385–5406. <https://doi.org/10.5194/amt-9-5385-2016>
- Gui K, Che H, Chen Q, Zeng Z, Liu H, Wang Y, Zhang X (2017) Evaluation of radiosonde, MODIS-NIR-Clear, and AERONET precipitable water vapor using IGS ground-based GPS measurements over China. *Atmos Res*. <https://doi.org/10.1016/j.atmosres.2017.07.021>
- He C, Wu S, Wang X, Hu A, Zhang K (2016) A new voxel-based model for the determination of atmospheric-weighted-mean temperature in GPS atmospheric sounding. *Atmos Measur Tech Disc*. <https://doi.org/10.5194/amt-2016-338>
- Heise S, Dick G, Gendt G, Schmidt T, Wickert J (2009) Integrated water vapor from IGS ground-based GPS observations: initial results from a global 5-min data set. *Ann Geophys* 27(7):2851–2859. <https://doi.org/10.5194/angeo-27-2851-2009>
- Herring, T. A., King, R. W., and McClusky, S. C. (2010). *Introduction to GAMIT/GLOBK, release 10.4. Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology*.
- Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, Thépaut J-N (2020) The ERA5 global reanalysis. *Q J Royal Meteorol Soc* 146(730):1999–2049. <https://doi.org/10.1002/qj.3803>
- Hicks-Jalali S, Sica RJ, Martucci G, MaillardBarras E, Voirin J, Haeefe A (2020) A Raman lidar tropospheric water vapour climatology and height-resolved trend analysis over Payerne. *Switzerland Atmos Chem Phys* 20(16):9619–9640. <https://doi.org/10.5194/acp-20-9619-2020>
- Hofmann-Wellenhof, B., Lichtenegger, H., and Wasle, E. (2007). *GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and more*. Springer Vienna. https://books.google.es/books?id=Np7y43HU_m8C

- Kannemadugu HB, Shaeb R, K., Gharai, B., (2022) GNSS-GPS derived integrated water vapor and performance assessment of ERA-5 data over India. *J Atmos Solar-Terrestrial Phys* 227:105807. <https://doi.org/10.1016/j.jastp.2021.105807>
- Köpken C (2001) Validation of Integrated Water Vapor from Numerical Models Using Ground-Based GPS, SSM/I, and Water Vapor Radiometer Measurements. *J Appl Meteorol* 40(6):1105–1117. [https://doi.org/10.1175/1520-0450\(2001\)040%3c1105:VOIWFV%3e2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040%3c1105:VOIWFV%3e2.0.CO;2)
- Landskron D, Böhm J (2017) VMF3/GPT3: refined discrete and empirical troposphere mapping functions. *J Geodesy*. <https://doi.org/10.1007/s00190-017-1066-2>
- Landskron, D. (2017). Modeling tropospheric delays for space geodetic techniques. *Vienna University of Technology*.
- Leandro, R., Santos, M., and Langley, R. (2006). *UNB Neutral Atmosphere Models: Development and Performance*. 2.
- Li Z, Muller J-P, Cross P (2003) Comparison of precipitable water vapor derived from radiosonde, GPS, and Moderate-Resolution Imaging Spectroradiometer measurements. *J Geophys Res*. <https://doi.org/10.1029/2003JD003372>
- Li G, Kimura F, Sato T, Huang D (2008) A composite analysis of diurnal cycle of GPS precipitable water vapor in central Japan during Calm Summer Days. *Theoret Appl Climatol* 92(1):15–29. <https://doi.org/10.1007/s00704-006-0293-x>
- Li H, Wang X, Wu S, Zhang K, Chen X, Qiu C, Zhang S, Zhang J, Xie M, Li L (2020) Development of an improved model for prediction of short-term heavy precipitation based on GNSS-derived PWV. *Remote Sens* 12(24):4101. <https://doi.org/10.3390/rs12244101>
- Li J, Li F, Liu L, Huang L, Zhou L, He H (2022a) A calibrated GPT3 (CGPT3) model for the site-specific zenith hydrostatic delay estimation in the Chinese mainland and its surrounding areas. *Remote Sens*. <https://doi.org/10.3390/rs14246357>
- Li L, Zhang K, Wu S, Li H, Wang X, Hu A, Li W, Fu E, Zhang M, Shen Z (2022b) An improved method for rainfall forecast based on GNSS-PWV. *Remote Sens*. <https://doi.org/10.3390/rs14174280>
- Li S, Xu T, Xu Y, Jiang N, Bastos L (2022c) Forecasting GNSS Zenith troposphere delay by improving GPT3 model with machine learning in antarctica. *Atmosphere*. <https://doi.org/10.3390/atmos13010078>
- Liu J, Sun Z, Liang H, Xu X, Wu P (2005) Precipitable water vapor on the Tibetan Plateau estimated by GPS, water vapor radiometer, radiosonde, and numerical weather prediction analysis and its impact on the radiation budget. In *J Geophys Res D Atmos*. <https://doi.org/10.1029/2004JD005715>
- Mendes, V. (1998). *Modeling the Neutral Atmosphere Propagation Delay in Radiometric Space Techniques*.
- Miloshevich LM, Vömel H, Whiteman DN, Leblanc T (2009) Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements. *J Geophys Res Atmos*. <https://doi.org/10.1029/2008JD011565>
- Moradi I, Soden B, Ferraro R, Arkin P, Vömel H (2013) Assessing the quality of humidity measurements from global operational radiosonde sensors. *J Geophys Res: Atmos* 118(14):8040–8053. <https://doi.org/10.1002/jgrd.50589>
- Namaoui H, Kahlouche S, Belbachir A, Van Malderen R, Brenot H, Pottiaux E (2017) GPS water vapor and its comparison with radiosonde and ERA-Interim data in Algeria. *Adv Atmos Sci* 34:623–634. <https://doi.org/10.1007/s00376-016-6111-1>
- Niell AE, Coster AJ, Solheim FS, Mendes VB, Toor PC, Langley RB, Upham CA (2001) Comparison of measurements of atmospheric wet delay by radiosonde, water vapor radiometer, GPS, and VLBI. *J Atmos Oceanic Tech* 18(6):830–850. [https://doi.org/10.1175/1520-0426\(2001\)018%3c0830:COMOAW%3e2.0.CO;2](https://doi.org/10.1175/1520-0426(2001)018%3c0830:COMOAW%3e2.0.CO;2)
- Ning T, Haas R, Elgered G, Willén U (2012) Multi-technique comparisons of 10 years of wet delay estimates on the west coast of Sweden. *J Geodesy* 86(7):565–575. <https://doi.org/10.1007/s00190-011-0527-2>
- Ning T, Wang J, Elgered G, Dick V, Wickert J, Bradke M, Sommer M (2015) The uncertainty of the atmospheric integrated water vapour estimated from GNSS observations. *Atmos Measur Tech Disc* 8:8817–8857. <https://doi.org/10.5194/amt-d-8-8817-2015>
- Ohtani R, Naito I (2000) Comparisons of GPS-derived precipitable water vapors with radiosonde observations in Japan. *J Geophys Res* 105(D22):26917–26929. <https://doi.org/10.1029/2000JD900362>
- Ortiz de Galisteo JP, Cachorro V, Toledano C, Torres B, Laulainen N, Bennouna Y, de Frutos A (2011) Diurnal cycle of precipitable water vapor over Spain. *Q J Royal Meteorol Soc* 137(657):948–958. <https://doi.org/10.1002/qj.811>
- Pacione R, Araszkiwicz A, Brockmann E, Dousa J (2017) EPN-Repro2: A reference GNSS tropospheric data set over Europe. *Atmospheric Measurement Techniques* 10:1689–1705. <https://doi.org/10.5194/amt-10-1689-2017>

- Rezaei M, Khazaei M (2022) Atmospheric precipitable water vapor over Iran using MODIS products: climatology and intercomparison. *Meteorol Atmos Phys* 134:15. <https://doi.org/10.1007/s00703-021-00854-6>
- Rocken C, Ware R, van Hove T, Solheim F, Alber C, Johnson J, Bevis M, Businger S (1993) Sensing atmospheric water vapor with the global positioning system. *Geophys Res Lett* 20(23):2631–2634. <https://doi.org/10.1029/93GL02935>
- Ross, R. J., and Elliott, W. P. (1996). Tropospheric Water Vapor Climatology and Trends over North America: 1973–93. *Journal of Climate*, 9(12), 3561–3574. <http://www.jstor.org/stable/26201470>
- Ross R, Elliott W (2001) Radiosonde-based northern hemisphere tropospheric water vapor trends. *J Clim* 14:1602–1612. [https://doi.org/10.1175/1520-0442\(2001\)014%3c1602:RBNHTW%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014%3c1602:RBNHTW%3e2.0.CO;2)
- Saastamoinen J (1972) Contributions to the theory of atmospheric refraction. *Bulletin Géodésique* 46(3):279–298. <https://doi.org/10.1007/BF02521844>
- Sapucci LF, Machado LAT, da Silveira RB, Fisch G, Monico JFG (2005) Analysis of Relative Humidity Sensors at the WMO Radiosonde Intercomparison Experiment in Brazil. *J Atmos Oceanic Tech* 22(6):664–678. <https://doi.org/10.1175/JTECH1754.1>
- Thayer GD (1974) An improved equation for the radio refractive index of air. *Radio Sci* 9(10):803–807. <https://doi.org/10.1029/RS009i010p00803>
- Tregoning P, Boers R, O'Brien D, Hendy M (1998) Accuracy of absolute precipitable water vapor estimates from GPS observations. *J Geophys Res* 103(D22):28701–28710. <https://doi.org/10.1029/98JD02516>
- Tunalı E (2022) Water vapor monitoring with IGS RTS and GPT3/VMF3 functions over Turkey. *Adv Space Res* 69(6):2376–2390. <https://doi.org/10.1016/j.asr.2021.12.036>
- van Malderen R, Brenot H, Pottiaux E, Beirle S, Hermans C, de Mazière M, Wagner T, de Backer H, Bruy-ninx C (2014) A multi-site intercomparison of integrated water vapour observations for climate change analysis. *Atmosph Measur Tech* 7(8):2487–2512. <https://doi.org/10.5194/amt-7-2487-2014>
- Vaquero-Martínez J, Antón M, Ortiz de Galisteo JP, Cachorro VE, Álvarez-Zapatero P, Román R, Loyola D, Costa MJ, Wang H, Abad GG, Noël S (2018) Inter-comparison of integrated water vapor from satellite instruments using reference GPS data at the Iberian Peninsula. *Remote Sens Environ* 204:729–740. <https://doi.org/10.1016/j.rse.2017.09.028>
- Vaquero-Martínez J, Antón M, Ortiz de Galisteo JP, Román R, Cachorro VE, Mateos D (2019) Comparison of integrated water vapor from GNSS and radiosounding at four GRUAN stations. *Sci Total Environ* 648:1639–1648. <https://doi.org/10.1016/j.scitotenv.2018.08.192>
- VMF Data Server: re3data.org: VMF Data Server; editing status 2021-08-24; re3data.org - Registry of Research Data Repositories. <https://doi.org/10.17616/R3RD2H>
- Wang J, Zhang L (2008) Systematic errors in global radiosonde precipitable water data from comparisons with ground-based GPS measurements. *J Clim* 21(10):2218–2238. <https://doi.org/10.1175/2007JCLI1944.1>
- Wang J, Zhang L (2009) Climate applications of a global, 2-hourly atmospheric precipitable water dataset derived from IGS tropospheric products. *J Geodesy* 83(3):209–217. <https://doi.org/10.1007/s00190-008-0238-5>
- Wang J, Zhang L, Dai A, van Hove T, van Baelen J (2007) A near-global, 2-hourly data set of atmospheric precipitable water from ground-based GPS measurements. *J Geophys Res Atmos*. <https://doi.org/10.1029/2006JD007529>
- Wang X, Zhang K, Wu S, Fan S, Cheng Y (2016) Water vapor-weighted mean temperature and its impact on the determination of precipitable water vapor and its linear trend. *J Geophys Res* 121(2):833–852. <https://doi.org/10.1002/2015JD024181>
- Webb, F. H. (1997). An Introduction to GIPSY/oasIs-II. *JPL D-11088*.
- Yao Y, Hu Y (2018) An empirical zenith wet delay correction model using piecewise height functions. *Ann Geophys* 36(6):1507–1519. <https://doi.org/10.5194/angeo-36-1507-2018>