



A new voxel-based model for the determination of atmospheric weighted mean temperature in GPS atmospheric sounding

Changyong He^{1,2}, Suqin Wu², Xiaoming Wang², Andong Hu², Qianxin Wang¹, and Kefei Zhang^{1,2}

¹School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, P.R. China

²Satellite Positioning for Atmosphere, Climate and Environment (SPACE) Research Centre, School of Science, Mathematical and Geospatial Sciences, RMIT University, Melbourne, VIC 3001, Australia

Correspondence to: Kefei Zhang (kefei.zhang@rmit.edu.au, profkzhang@cumt.edu.cn)

Received: 8 October 2016 – Discussion started: 1 December 2016

Revised: 16 March 2017 – Accepted: 5 April 2017 – Published: 7 June 2017

Abstract. The Global Positioning System (GPS) is a powerful atmospheric observing system for determining precipitable water vapour (PWV). In the detection of PWV using GPS, the atmospheric weighted mean temperature (T_m) is a crucial parameter for the conversion of zenith tropospheric delay (ZTD) to PWV since the quality of PWV is affected by the accuracy of T_m . In this study, an improved voxel-based T_m model, named GWMT-D, was developed using global reanalysis data over a 4-year period from 2010 to 2013 provided by the United States National Centers for Environmental Prediction (NCEP). The performance of GWMT-D was assessed against three existing empirical T_m models – GTm-III, GWMT-IV, and GTm_N – using different data sources in 2014 – the NCEP reanalysis data, surface T_m data provided by Global Geodetic Observing System and radiosonde measurements. The results show that the new GWMT-D model outperforms all the other three models with a root-mean-square error of less than 5.0 K at different altitudes over the globe. The new GWMT-D model can provide a practical alternative T_m determination method in real-time GPS-PWV remote sensing systems.

atmosphere (Duan et al., 1996). The main advantages of using GPS technique to derive PWV are its high quality, wide coverage, and all-time availability under all-weather conditions. These features are vital for meteorological applications of GPS such as the prediction of short-term rainstorms and rainy seasons (Song et al., 2003; Zhang et al., 2007) and the monitoring of severe weather events, including thunderstorms, hailstorms, strong winds, and hurricanes (Choy et al., 2001; Zhang et al., 2015).

PWV is defined as the equivalent height to a column of liquid water. GPS-derived PWV values above a given GPS station, i.e. GPS-PWV, are converted from the zenith tropospheric delay (ZTD) estimated from GPS measurements. The GPS-PWV can be used to compare different techniques of WV detection, such as radiosonde, WV radiometer, Moderate-Resolution Imaging Spectroradiometer (MODIS), and sun photometer (Yang et al., 1999; Li et al., 2003; Prasad and Singh, 2009; Kwon et al., 2010). It can also be used for evaluating improvements in numerical weather prediction (NWP) systems (Gutman and Benjamin, 2001; Song et al., 2004). Moreover, temporal and spatial variations in PWV can be precisely identified using GPS-PWV over GPS networks (Champollion et al., 2004; Jin and Luo, 2009; Van Baelen and Penide, 2009).

The GPS-derived ZTD (GPS-ZTD) generally consists of two components: the zenith hydrostatic delay (ZHD) and the zenith wet delay (ZWD). The ZWD is caused by WV in the atmosphere below ~ 10 km height (Saastamoinen, 1972). Furthermore the ZWD can be converted to PWV by multiplying it with a dimensionless conversion factor. This factor is a function of atmospheric weighted mean temperature (T_m),

1 Introduction

Water vapour (WV), a component of the Earth's atmosphere, plays a crucial role in global atmospheric radiation, energy equilibrium, and the hydrological cycle (Wang et al., 2007). Since the Global Positioning System (GPS) became fully operational in 1994, it has been possible to use GPS measurements to retrieve precipitable WV (PWV) information in the

as expressed below (Askne and Nordius, 1987; Davis et al., 1985; Jade et al., 2005).

$$\text{PWV} = \Pi \cdot (\text{ZTD} - \text{ZHD}) = \Pi \cdot \text{ZWD}, \quad (1)$$

$$T_m = \frac{10^6}{\rho_w R_v (k_3/T_m + k'_2)}, \quad (2)$$

$$T_m = \frac{\int_{h_T}^h \rho_v dz}{\int_{h_T}^h \rho_v/T dz}, \quad (3)$$

where Π is the conversion factor; ρ_w and ρ_v are the density of liquid water and WV, respectively; R_v is the specific gas constant for water vapour; k'_2 and k_3 are the atmospheric refractivity constants given in Bevis et al. (1994); and T is the absolute temperature of the atmosphere in kelvin (K). Using the ideal gas law for water vapour, ρ_v can be written as $\rho_v = eT/R_v$, where e is the WV pressure in hPa (Picard et al., 2008). The heights of the tropopause and the GPS station are h_T and h , respectively. Note that both PWV and ZWD are in millimetres in Eqs. (1)–(3).

The T_m over a GPS station or at any given point can be accurately determined using a ray tracing method that takes profiles of atmospheric temperature and WV pressure. However, atmospheric profiles are almost impossible to obtain in real-time/near-real-time (RT/NRT) (Wang et al., 2016). Traditionally the following two alternatives have been used in practical applications: the Bevis formula ($T_m = a + b \cdot T_s$ where T_s is the atmospheric temperature) and the empirical model (Bevis et al., 1994; Ross and Rosenfeld, 1997, 1999).

In the Bevis formula $T_m = a + b \cdot T_s$, the coefficients (a and b) are season and location dependent and can be estimated from meteorological measurements (Wang et al., 2011; Bevis et al., 1992; Schueler et al., 2001; Mendes et al., 2000; Ewardson and Derks, 2000). The root-mean-square error (RMSE) of T_m from the Bevis formula is 2–5 K. Nevertheless, the Bevis formula becomes invalid when in situ temperature measurements are unavailable at some GPS stations, which could adversely affect the continuous operation of a RT/NRT GPS-PWV remote sensing system. Therefore, an empirical T_m model, fed only by coordinates of the site and the time, is essential for ground-based GPS atmospheric sounding. Although the accuracy of empirical models has been shown to be lower than that of the ray tracing method and the Bevis formula, it is still used as a practical alternative to calculate the T_m in RT/NRT if meteorological measurements are unavailable.

Table 1 summarises existing empirical T_m models developed in the last decade. The data source column presents the type and time span of the data used to develop the models: NCEP-DOE Atmospheric Model Intercomparison 2 data (NCEP2) released by National Centers for Environmental Prediction (NCEP), ERA-Interim data from the European Centre for Medium-range Weather Forecasts (ECMWF), and

the Global Geodetic Observing System (GGOS) data generated from ECMWF reanalysis data.

Building upon the global pressure and temperature (GPT) model proposed by Böhm et al. (2007), Yao et al. (2012) developed the season-specific Global Weighted Mean Temperature (GWMT) model based on radiosonde data of 135 global stations in the period 2005–2009. The RMSE of T_m over the ground is around 4.6 K. However, due to its poor performance in the southern Pacific Ocean, the coefficients were recalculated for an updated model – GTm-II using T_m over the ocean calculated from the Bevis formula, where T_s is derived from the GPT model (Yao et al., 2013). This GTm-II model was further improved into GTm-III using GGOS surface T_m by taking semi-annual and diurnal variations of T_m into account (Yao et al., 2014a). In the later GWMT-IV model, the T_m lapse rate is considered as a function of geodetic coordinates instead of a constant value adopted in the former models (He et al., 2013). In a study by Chen et al. (2014), the non-linear model in GTm-III was expressed into a linear one and developed it further into GTm_N. Unlike the spherical harmonics applied in GTm_N, Chen and Yao (2015) developed the GTm-X model based on the semi-annual and diurnal variations in T_m with a global resolution of a $1^\circ \times 1^\circ$ geographical grid. More details for these three models can be found in Appendix B. Note that UNB3m and GPT2w are not specific T_m models even though they can output T_m values (Leandro et al., 2008; Böhm et al., 2015). In addition, no comprehensive intercomparison with the same reference T_m has been carried out among existing empirical T_m models.

However, given that the diurnal variation and the lapse rate of T_m are either ill-modelled or neglected in most of these empirical models, this study presents a recent development towards an improved T_m model (i.e. GWMT-D). This paper is structured as follows. Section 2 reviews the data sets used for determination of T_m and validation of empirical T_m models. Section 3 presents the new model GWMT-D and its procedure of calculation. The performance of the new model GWMT-D is assessed against three other selected models using reference T_m derived from 2014 NCEP2, radiosonde, and GGOS data in Sect. 4, followed by conclusions in Sect. 5.

2 The determination of T_m

Three data sets with various temporal and spatial resolutions are used to calculate T_m : NCEP2 reanalysis data, GGOS data, and radiosonde measurements. NCEP2 data in the period 2010–2013 are used to develop the new GWMT-D model, while all these three data sets in 2014 are used to evaluate GWMT-D as well as the other three selected empirical T_m models.

Table 1. A list of the latest global empirical T_m models^a.

Model name	Feature	Data source	Input variable	Surface T_m error (K)	Reference
UNB3m	Annual	US Standard Atmosphere Supplements	$\varphi, \theta, \text{DOY}, h$	–	Leandro et al. (2008)
GWMT	Spherical Harmonic	Radiosonde (2005–2009)	φ, DOY, h	4.6	Yao et al. (2012)
GTm-II	Spherical Harmonic	Radiosonde (2005–2009)	φ, DOY, h	4.0	Yao et al. (2013)
GTm-III	Spherical Harmonic	GGOS (2005–2011)	$\theta, \text{DOY}, \text{HOD}, h$	4.2	Yao et al. (2014a)
GWMT-IV	Spherical Harmonic	NCEP2 (2005–2009)	φ, DOY, h	~ 4.1	He et al. (2013)
GTm_N	Spherical Harmonic	NCEP (2006–2012)	φ, DOY, h	3.38	Chen et al. (2014)
GTm_X	Grid	ERA-Interim (2007–2010)	φ, DOY, h	~ 4.0	Chen and Yao (2015)
GPT2w	Grid	ERA-Interim (2001–2010)	φ, DOY, h	$< 4.0^*$	Böhm et al. (2015)

^a The inputs are day of year (DOY), hour of day (HOD), latitude (φ), longitude (θ), and surface height (h) of a site; the values in the surface error column are the RMSEs of the model on the surface given by the authors, except for the 4.0^* of GPT2w, which is a post-fit standard deviation according to the reference.

2.1 NCEP2 data

A state-of-the-art analysis and forecast system has been used to assimilate multi-source data since 1948 and the NCEP2 data set is an updated version from its former reanalysis data (available on www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.pressure.html) (Kanamitsu et al., 2002). The NCEP2 global reanalysis data set has a latitude and longitude grid resolution of $2.5^\circ \times 2.5^\circ$ at 17 different pressure layer from 1000 to 10 hPa and also has a temporal resolution of 6 h (namely, at 00:00, 06:00, 12:00, 18:00 UTC). In this study, temperature, geopotential height, pressure, and humidity from NCEP2 data over the period of 2010–2014 are used for the development and validation of the new GWMT-D model.

2.2 Radiosonde data

Radiosondes released from ground-based stations can directly measure the atmospheric profiles. Radiosonde records from 585 global Integrated Global Radiosonde Archive (IGRA) stations (Fig. 1) in 2014 are utilised to validate the new GWMT-D model. They are retrieved from the upper-air archive at the website of University of Wyoming (available on <http://weather.uwyo.edu/upperair/sounding.html>). The daily observations at a site usually consist of 1–4 radiosonde observations, containing pressure, temperature, geopotential height, dew point temperature, relative humidity (RH), and mixing ratio at the surface, tropopause, and standard pressure levels (i.e. 1000, 925,

850, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, and 10 hPa) (Wang et al., 2005). T_m values are obtained through numerical integration (see Appendix A) under the assumption that the collected pressure, temperature, and humidity measurements are along the zenith direction, even though radiosonde balloons often drift away from the vertical direction, especially in windy conditions.

In addition, raw radiosonde measurements are rejected as outliers in the data preprocessing under the following conditions:

1. the height of the first data record in the profile is greater than 20 m above the ground;
2. the difference in height between two successive pressure levels is greater than 10 km;
3. the gap between two successive atmospheric pressure levels is greater than 200 hPa;
4. the total number of valid radiosonde levels is less than 20;
5. the highest humidity level (at the pressure level of 200–350 hPa) is less than $\mu - 4\sigma$ (where μ and σ are the mean height of the tropopause and its standard deviation) obtained from an empirical model (Liu, 2015);
6. the height of the last data record in the profile is lower than 20 km.

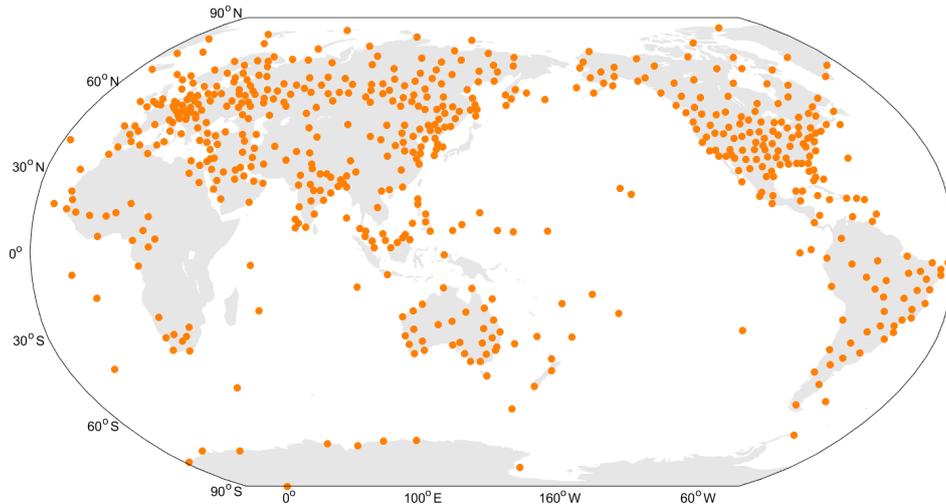


Figure 1. Distribution of the 585 radiosonde stations selected to validate the new GWMT-D model (only those data that pass a quality check are used).

2.3 Surface T_m from GGOS Atmosphere

In this study, global surface T_m values are used for the validation of the new GWMT-D model. GGOS Atmosphere provides the daily global surface T_m with a horizontal resolution of $2^\circ \times 2.5^\circ$ (latitude and longitude) at 00:00, 06:00, 12:00, and 18:00 UTC (available on <http://ggosatm.hg.tuwien.ac.at/DELAY/ETC/TMEAN/>). This data set has been applied in the development of GTm-III and will be also used in the performance assessment of this study. Nevertheless, the discrepancies between these different data sets are noticeable and may affect the validation results, which will be shown in Sect. 4.

3 GWMT-D model

The NCEP2 data from the 4-year period 2010–2013 are employed to develop the new GWMT-D (D stands for diurnal variation) model. All global T_m values with a latitudinal and longitudinal grid resolution of 2.5° at 17 pressure levels are first calculated (see Eq. 3). For more details of the calculation, refer to Appendix A. Note that the geopotential height in the radiosonde and NCEP2 data needs to be converted to ellipsoidal height (refer to Appendix A), which is simplified as “height” hereafter.

3.1 Improvements in GWMT-D

Compared with other empirical T_m models, the improvement achieved by the new GWMT-D model are the modelling of diurnal variation and lapse rate in T_m . The T_m lapse rate in this paper is the decreasing rate of T_m (Bevis et al., 1994; Yao et al., 2012). The heights of 17 pressure levels in NCEP2 data are time-variant. In order to investigate a time series of

T_m for a specific location, NCEP2-derived T_m at the pressure level for each grid points are first interpolated at four selected reference heights: 0, 2, 5, and 9 km. The spline interpolation is carried out in this procedure to avoid the Runge’s phenomenon (Fornberg and Zuev, 2007). Also the T_m time series at each of the reference times (00:00, 06:00, 12:00, 18:00 UTC) of day for a specific reference location are assumed to follow a seasonal cycle:

$$T_m(\text{DOY}) = \alpha_1 + \alpha_2 \cos\left(2\pi \frac{\text{DOY}}{365.25}\right) + \alpha_3 \sin\left(2\pi \frac{\text{DOY}}{365.25}\right) + \alpha_4 \cos\left(4\pi \frac{\text{DOY}}{365.25}\right) + \alpha_5 \sin\left(4\pi \frac{\text{DOY}}{365.25}\right) \quad (4)$$

where DOY is the day of year, α_1 is the annual mean value; α_2 and α_3 are the coefficients of the annual variation, and α_4 and α_5 are that of the semi-annual variation. Here the reference locations mean the geodetic coordinates of the grid points on the four reference height levels.

These coefficients are estimated from the time series of T_m values at the specific reference time using the least-squares method. The new model is developed upon a four-dimensional (4-D) global T_m field with a resampled horizontal resolution of $5^\circ \times 5^\circ$ at the four reference height levels and the four reference times. The voxel-based feature of GWMT-D’s coefficients is its primary difference from the other models.

3.1.1 Diurnal variation

Annual and semi-annual variations in a NCEP2-derived T_m time series for a reference location can be detected using the spectrum analysis (Chen et al., 2014). Although simple sine and cosine functions have been widely used to model the diurnal variations of T_m , few studies have been conducted to

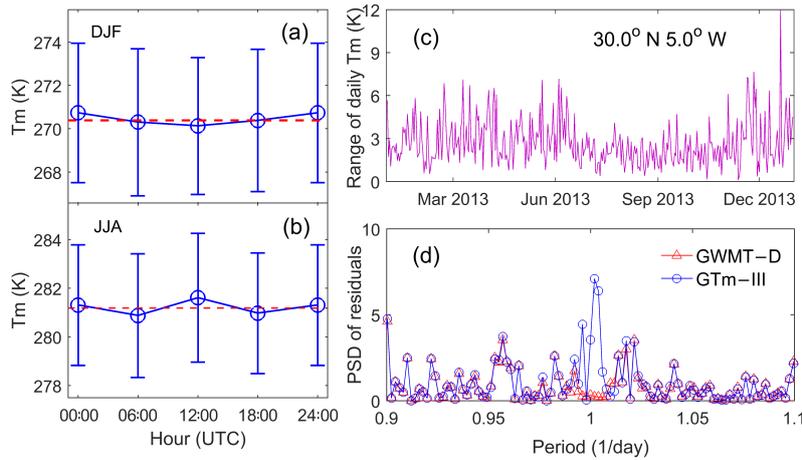


Figure 2. Statistical results of diurnal T_m (mean \pm standard deviation) at 2 km height and four reference times during (a) December–January–February (DJF) and (b) June–July–August (JJA) in 2013, (c) the range of daily T_m (max – min), and (d) power spectrum density (PSD) of T_m residuals.

analyse the periodic nature of the diurnal variation in T_m . Diurnal variations in different seasons and locations are first investigated to study the voxel-based modelling process in GWMT-D.

Figure 2 shows an example of the diurnal variation at 2 km above the ground for 30° N, 5° W. It clearly shows that, on the one hand, the diurnal variation in T_m is season and location dependent and is difficult to be modelled through simple trigonometric functions as used in the GTm-III model (see Appendix B). On the other hand, the plot of power spectrum density (PSD) in Fig. 2d shows that GWMT-D efficiently captures the diurnal variations in T_m but GTm-III does not. This study takes into account this feature by modelling T_m at each of the four reference times so that T_m values at any other times can be obtained by the spline interpolation.

3.1.2 Vertical lapse rate of T_m

The T_m lapse rate along the vertical direction can be affected by several factors, e.g. the moisture content of air, atmospheric pressure and the surface height. Figure 3 illustrates the global distribution of annual mean T_m lapse rate in the height layer from the ground up to 2 km in 2013. It shows that global annual mean T_m lapse rate varies with latitude and land–sea distribution. Therefore, it is essential to consider season- and location-dependent T_m lapse rate in the development of empirical T_m models. More results can be found in Sect. 4.3.

Four specific reference height levels (0, 2, 5, and 9 km) are selected covering most of the troposphere in the new GWMT-D model. All global T_m values on these reference heights are first calculated (see Eq. 3). Then T_m value for any other heights can be obtained by piecewise linear interpolating its two nearest height levels. This improvement is a distinguished feature of the new model in comparison with

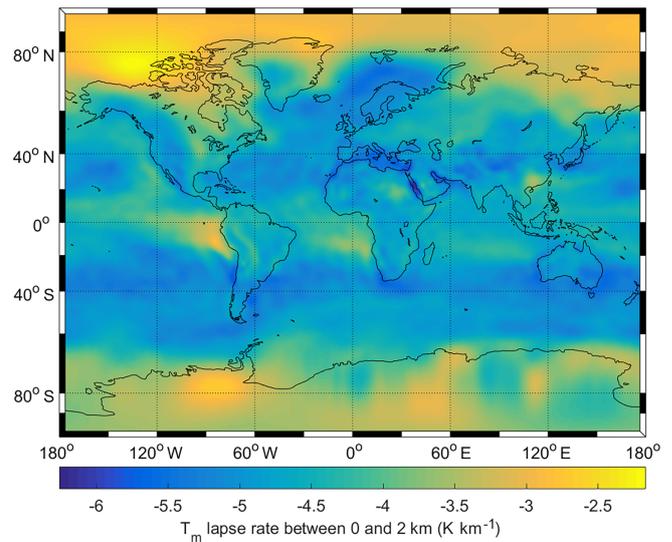


Figure 3. Global annual mean T_m lapse rate in the height interval 0–2 km from NCEP2 in 2013.

the aforementioned empirical models where a constant T_m lapse rate is adopted for the different heights over the globe. This step has been taken for each reference time.

3.1.3 Data span used in T_m modelling

Another important task is to determine the optimal length of reanalysis data required for the development of empirical T_m models. Long-term T_m data (> 10 years) over the globe can be used for climatological analysis, but the temporal correlation of T_m time series may be too weak to be considered in the T_m modelling process. However, a set of short-term T_m data (< 1 year) may be insufficiently for reliable results.

Consequently, an optimal length of period needs to be investigated.

Different sets of coefficients of the GWMT-D are calculated using the NCEP2-derived T_m data for a period of 1 (2013) to 9 years (2005–2013). The GWMT-D with different sets of coefficients are compared with 1-year T_m time series at five pressure levels (1000, 925, 850, 700, and 600 hPa) derived from NCEP2 data in 2014. Table 1 lists the statistical results of the comparison. In this research, the NCEP2 T_m time series from the 4-year period are adopted to develop the GWMT-D model for its best-fitting results (shown in bold fonts).

3.2 The procedure to determine T_m using GWMT-D

Assuming T_m is a function of the target location (φ, λ, h) , day of year (DOY), and UTC hour (HOD), the key steps of determining T_m in GWMT-D can be described as follows:

1. Two nearest reference height levels close to h , (see Fig. 4 in grey) and the other four vertical surfaces containing the eight voxels closest to (φ, λ) are determined. Then T_m values for the reference times on the eight voxels are given by

$$T_m(\varphi_i, \lambda_j, h_l, t_k) = \alpha_1(\varphi_i, \lambda_j, h_l, t_k) + \alpha_2(\varphi_i, \lambda_j, h_l, t_k) \cos\left(2\pi \frac{\text{DOY}}{365.25}\right) + \alpha_3(\varphi_i, \lambda_j, h_l, t_k) \sin\left(2\pi \frac{\text{DOY}}{365.25}\right) + \alpha_4(\varphi_i, \lambda_j, h_l, t_k) \cos\left(4\pi \frac{\text{DOY}}{365.25}\right) + \alpha_5(\varphi_i, \lambda_j, h_l, t_k) \sin\left(4\pi \frac{\text{DOY}}{365.25}\right), \quad (5)$$

where φ_i and λ_j are the latitude and longitude of the vertex (at a $5^\circ \times 5^\circ$ resolution); l ($l = 1, 2, 3, 4$) is the index of the reference height h_l corresponding to 0, 2, 5 or 9 km, respectively; and t_k ($k = 1, 2, \dots, 5$) is the index of the reference time corresponding to 00:00, 06:00, 12:00, 18:00, and 24:00 UTC, respectively.

2. For each of the four reference time, a vertical interpolation is performed for the four grid points at the height of h (four corners of the dashed rectangular in Fig. 4). At each grid point, T_m is linearly interpolated from the T_m values on the two nearest reference height levels h_l and h_{l+1} :

$$T_m(\varphi_i, \lambda_j, h) = T_m(\varphi_i, \lambda_j, h_l) + \frac{T_m(\varphi_i, \lambda_j, h_{l+1}) - T_m(\varphi_i, \lambda_j, h_l)}{h_{l+1} - h_l} \cdot (h - h_l). \quad (6)$$

3. The T_m values on the four corners are horizontally interpolated to obtain the T_m at the target point (φ, λ, h) using

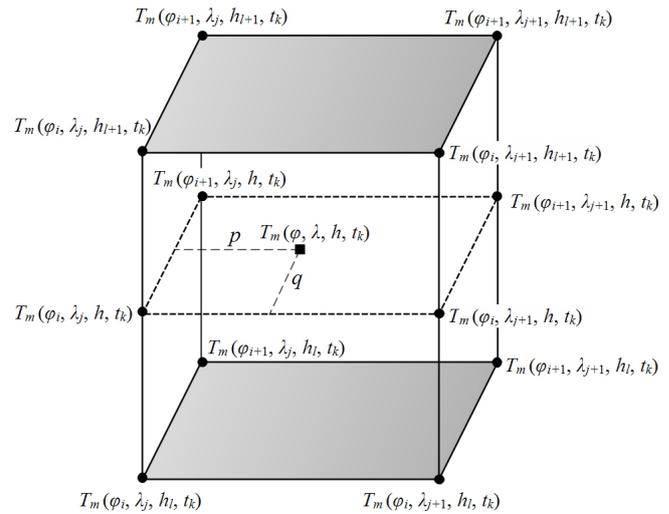


Figure 4. Spatial interpolation for the target point located at (φ, λ, h) using the T_m values at the eight closest voxels determined by the GWMT-D model. The first interpolation is for each of the four vertical edges of the box, and the second interpolation is on the 2-D plane at the height of the target point (the dashed rectangular).

$$T_m(\varphi_i, \lambda_j, h) = T_m(\varphi_i, \lambda_j, h_l) + \frac{T_m(\varphi_i, \lambda_j, h_{l+1}) - T_m(\varphi_i, \lambda_j, h_l)}{h_{l+1} - h_l} \cdot (h - h_l), \quad (7)$$

$$p = (\lambda - \lambda_j)/5, \quad q = (\varphi - \varphi_i)/5, \quad (8)$$

$$T_m(\varphi, \lambda, h) = (1 - p)(1 - q) T_m(\varphi_i, \lambda_j, h) + p(1 - q) T_m(\varphi_i, \lambda_{j+1}, h) + (1 - p)q T_m(\varphi_{i+1}, \lambda_j, h) + pq T_m(\varphi_{i+1}, \lambda_{j+1}, h). \quad (9)$$

Now T_m values of the target point are determined on the four reference times of the day (i.e. 00:00, 06:00, 12:00, and 18:00 UTC). All of notations in Eqs. (7) and (8) can be found in Fig. 4. The number “5” in Eq. (7) is the horizontal resolution of the new model.

4. After the aforementioned spatial interpolations, a spline interpolation in the time domain is carried out to find the T_m of the target location (φ, λ, h) for the specific time of the day t_k using the T_m values from the previous step.

4 Validation of T_m models

Different empirical T_m models (Table 1) are developed based on different data sets. The accuracies of these models claimed in relevant literature are referenced to different reference values (e.g. He et al., 2013; Chen et al., 2014; Yao

Table 2. The global mean RMSE of various GWMT–D models built with different lengths of time periods (2005–2013) of NCEP2 data at five pressure levels (in K). The reference T_m are derived from NCEP2 data in 2014. The values inside square brackets are the minimum and maximum, and the fourth row (in bold) shows the best fitting results.

Period length (year)	1000 hPa	925 hPa	850 hPa	700 hPa	600 hPa
1	3.31 [1.16, 12.47]	3.40 [1.17, 11.66]	3.50 [1.18, 10.84]	4.19 [1.16, 9.81]	4.74 [1.31, 15.38]
2	3.24 [1.19, 12.07]	3.32 [1.17, 11.24]	3.42 [1.18, 10.40]	4.13 [1.14, 9.31]	4.68 [1.30, 14.61]
3	3.23 [1.18, 12.33]	3.32 [1.19, 11.48]	3.43 [1.21, 10.61]	4.13 [1.14, 9.47]	4.67 [1.28, 13.84]
4	3.22 [1.15, 11.96]	3.32 [1.14, 11.14]	3.42 [1.18, 10.34]	4.13 [1.14, 9.28]	4.67 [1.27, 11.54]
5	3.22 [1.18, 12.13]	3.31 [1.18, 11.29]	3.42 [1.18, 10.45]	4.12 [1.13, 9.37]	4.66 [2.10, 11.75]
6	3.22 [1.17, 12.02]	3.31 [1.19, 11.20]	3.42 [1.21, 10.37]	4.12 [1.14, 9.29]	4.66 [2.10, 11.75]
7	3.22 [1.16, 12.13]	3.31 [1.19, 11.30]	3.41 [1.21, 10.47]	4.11 [1.14, 9.41]	4.66 [2.10, 13.80]
8	3.22 [1.20, 12.16]	3.31 [1.19, 11.33]	3.41 [1.21, 10.50]	4.11 [1.15, 9.43]	4.66 [1.79, 12.43]
9	3.22 [1.20, 12.28]	3.31 [1.19, 11.42]	3.41 [1.21, 10.56]	4.11 [1.15, 9.48]	4.66 [1.27, 11.55]

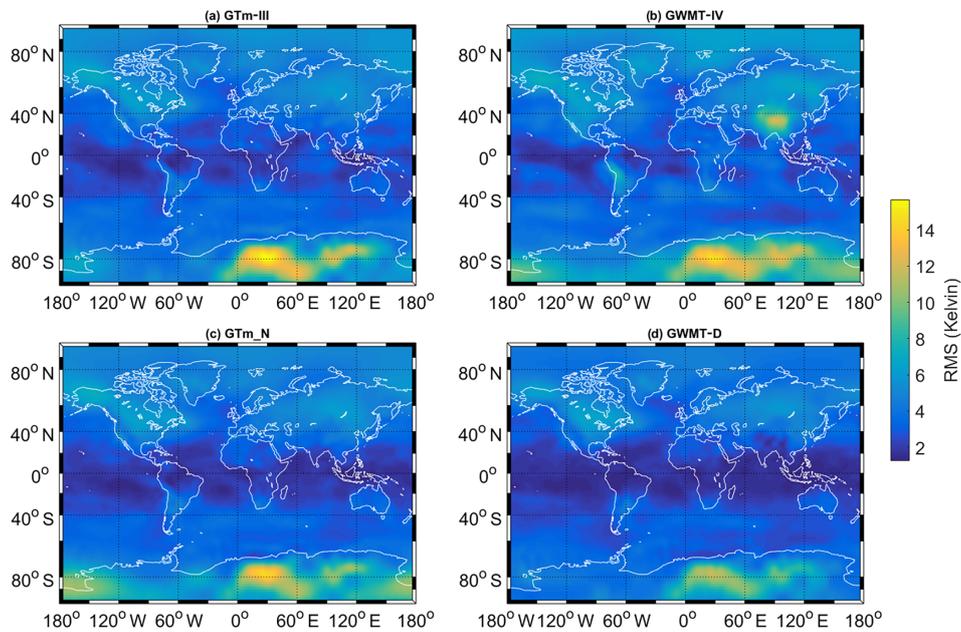


Figure 5. The global RMSE distribution of the differences between the T_m derived from each model and the NCEP2 data on pressure level of 925 hPa in 2014.

et al., 2014a). Consequently, cross comparisons of these accuracy values for their performance may not be appropriate. In this study, the performance of three selected empirical T_m models and the new GWMT-D model is assessed using the same reference T_m values derived from NCEP2, GGOS, and radiosonde data sets.

Due to the fact that GTm_X is unavailable to the public and GWMT and GTm-II have been proven inferior to GTm-III, GWMT-IV, and GTm_N, only GTm-III, GWMT-IV, GTm_N, and the new GWMT-D model are assessed. The methodologies for obtaining T_m from NCEP2 radiosonde data sets are given in Appendix A. The two statistical quantities used to measure the performance of these models are bias and RMSE, which are calculated by

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^N (T_m^{C_i} - T_m^i), \tag{10}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (T_m^{C_i} - T_m^i)^2}, \tag{11}$$

where $T_m^{C_i}$ and T_m^i are the T_m values from the models and the reference, respectively, and N is the number of the samples.

4.1 Comparison with NCEP2 data

The globally mean biases and RMSEs of the differences between NCEP2-derived and model-derived T_m at all global grid points are summarised in Table 3. Only the results of pressure levels of 925 hPa (~0.6 km) and 600 hPa (~5 km)

Table 3. The globally mean biases and RMSEs of the differences between the T_m (in K) derived from four empirical models and 2014 NCEP2 data on pressure levels of 925 and 600 hPa. Values within square brackets are the minimum and maximum, and the % column is the percentage of those global grids with a value ≤ 5 K.

Pressure level	Model	Bias	%	RMSE	%
925 hPa	GTm-III	-1.31 [-5.19, 9.63]	98.0	3.91 [1.26, 15.38]	77.1
	GWMT-IV	-1.89 [-11.40, 4.77]	96.2	4.36 [1.36, 14.61]	70.3
	GTm_N	-1.25 [-8.53, 9.18]	97.1	3.84 [1.16, 13.84]	77.4
	GWMT-D	-0.03 [-2.50, 4.62]	100	3.32 [1.14, 11.14]	91.1
600 hPa	GTm-III	-1.25 [-9.30, 4.92]	89.4	5.63 [2.10, 11.75]	33.2
	GWMT-IV	-5.83 [1.69, 12.35]	38.4	7.28 [2.10, 13.80]	13.7
	GTm_N	2.65 [-9.10, 8.81]	72.9	6.38 [1.79, 12.43]	26.0
	GWMT-D	0.03 [-2.48, 3.28]	100	4.67 [1.27, 11.54]	58.3

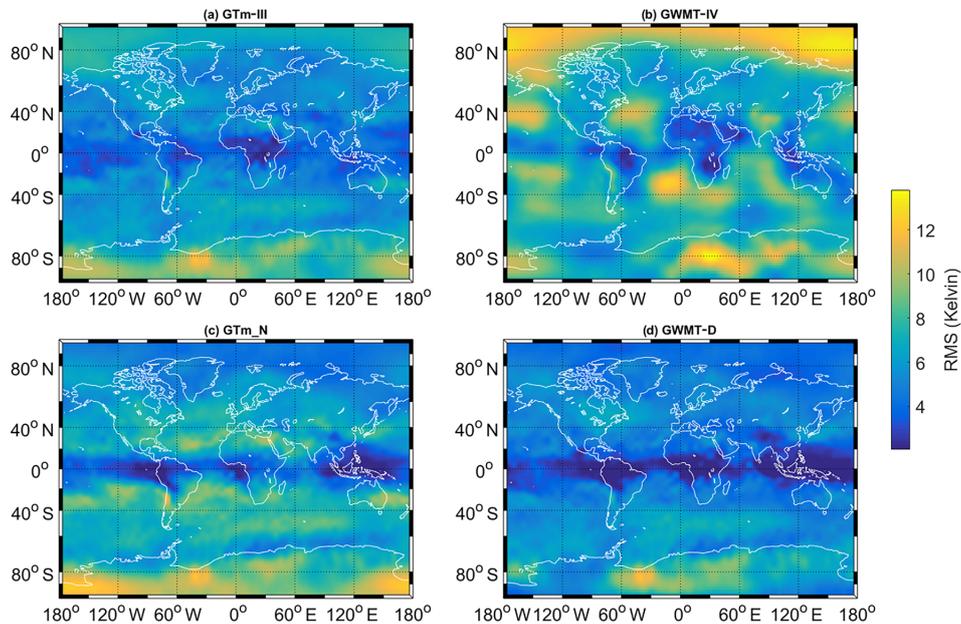


Figure 6. The global RMSE distribution of the differences between the T_m derived from each model and the NCEP2 data on pressure level of 600 hPa in 2014.

are given since similar results can be obtained from the new GWMT-D model on the other pressure levels less than 600 hPa (refer to Table 2). One can find from Table 3 that GWMT-D significantly outperforms all the other three empirical models.

Figures 5–6 illustrate the distribution of the RMSE (not the mean RMSE of all grid points listed in Table 3) of the differences between the T_m derived from the models and the NCEP2 data in 2014 on two pressure levels. Figures 5d and 6d present the best agreement for GWMT-D over the globe. On the pressure level of 925 hPa, more than 91 % of the grid points had RMSE less than 5 K, compared with 77 % from GTm_N and GTm-III, and less than 71 % from GWMT-IV. Whilst on the 600 hPa level, GWMT-IV is worse, especially in the Arctic Circle. The RMSE values of GWMT-D ranged from 1.27 K to 11.55 K, outperforming the other three mod-

els with a global average RMSE of only 4.73 K and an approximately 25 % improvement over the other models.

It is worth pointing out that all these four models have relatively low RMSE values near the tropical areas, and all have a similar performance globally except for the Antarctic. This finding is consistent with recent studies, (e.g. He et al., 2013; Chen et al., 2014; Yao et al., 2014a). It may be explained by the fact that the T_m on the pressure levels under the terrain of Antarctica (~ 6 km) may contain large systematic biases caused by the extrapolation of actual meteorological records.

4.2 Comparison with GGOS data

The GGOS surface grid T_m data in 2014 are used as the reference in this section to evaluate the performance of the four models. The statistical results of the four selected models are

Table 4. Global statistics of the differences between the surface T_m derived from four models and GGOS data in 2014 (in K).

Model	Bias	%	RMSE	%
GTm-III	-0.02 [-4.44, 4.93]	100	3.29 [0.98, 6.62]	92.3
GWMT-IV	-0.88 [-20.05, 13.61]	92.5	3.95 [0.91, 20.37]	76.4
GTm_N	-0.27 [-7.07, 10.09]	98.2	3.70 [1.08, 10.66]	83.8
GWMT-D	1.20 [-1.48, 6.23]	99.5	3.54 [0.83, 7.51]	86.2

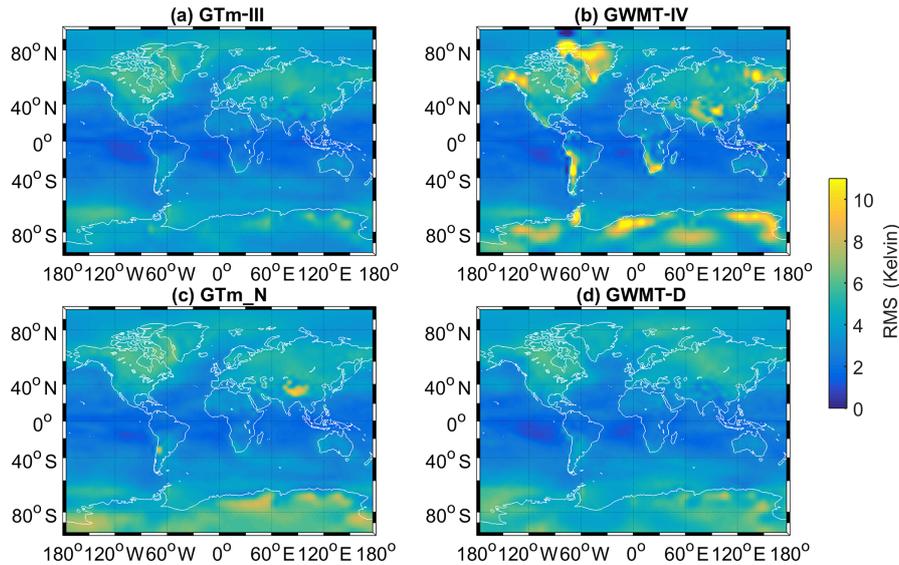


Figure 7. The global RMSE distribution of the differences between surface T_m derived from each model and GGOS data in 2014.

shown in Table 4 and Fig. 7. The GTm-III performed the best this time because the GGOS Atmosphere data derived from ECMWF reanalysis data are used in the development of the GTm-III. From Table 3, GWMT-D is almost unbiased while the GTm-III showed a bias of -1.25 to -1.31 K in comparison with the NCEP2-based T_m . In contrast, a bias of about $+1.2$ K (warmer) compared to the GGOS-based T_m is found with GWMT-D (see Table 4). This discrepancy of 1.2 K between the NCEP2-derived and GGOS-derived T_m may result from differences of NWP systems, e.g. different observations, physical models, data assimilation processes, and boundary conditions (Buizza et al., 2005).

Nevertheless, the good performance of GWMT-D indicates that the modelling method of T_m can significantly improve the model’s accuracy. Figure 7d illustrates that GWMT-D has RMSE values of less than 6 K at most grids, except the areas in the Antarctic, northeast North America, and Middle East (6–10 K).

4.3 Comparison with radiosonde data

These four empirical T_m models of interest are also evaluated using independent measurements (i.e. radiosonde). A number of comparisons are carried out in this section:

1. Surface T_m values calculated from radiosonde measurements are used as the reference to compare with various model-derived surface T_m .
2. T_m derived from the GWMT-D and three other selected empirical models is compared with radiosonde-derived T_m to investigate models’ performance in different heights.
3. The accuracy of the T_m models in different seasons is also investigated.

Figure 8 illustrates the RMSE of model-derived surface T_m in 2014 at the 585 selected radiosonde stations. The spatial (horizontal) variation in the accuracy of these models can be seen from this figure. An accuracy of better than 8 K has been achieved at most stations for the GWMT-D (Fig. 8d) and a similar accuracy can be achieved by the GTm_N as well (Fig. 8c). These two models outperformed the other two models, GTm-III and GWMT-IV, especially in the Middle East, Siberia, and South Africa regions.

Figure 9 shows the histogram of the difference (i.e. model-derived T_m minus radiosonde-derived T_m) at all heights from 0 to 9 km in terms of the mean, standard deviation, median, and mode values. As one of the new trials in this study, we use mode and median values to estimate the sample bias. The

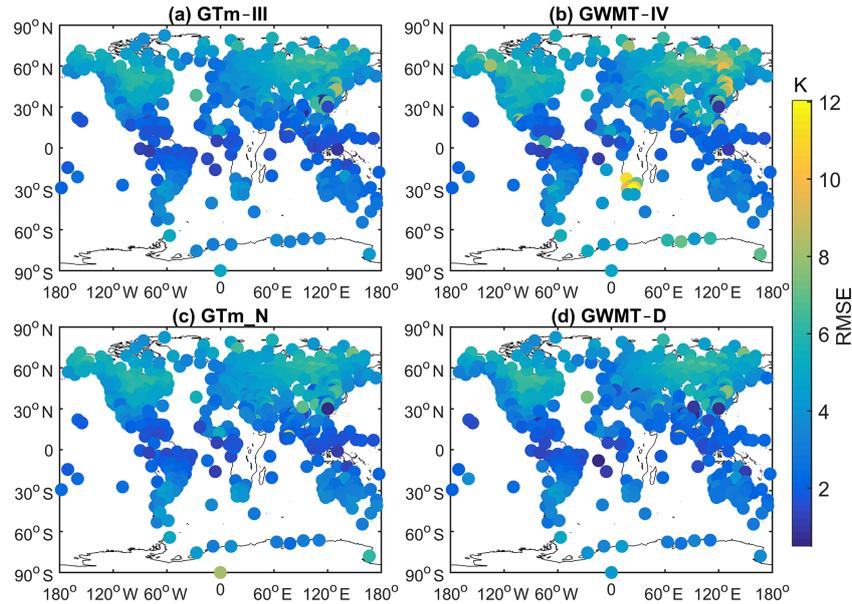


Figure 8. RMSE of model-derived surface T_m in 2014 at 585 global radiosonde sites.

Table 5. Statistics of the differences between model-derived and radiosonde-derived T_m in various height intervals for the year 2014 at 585 global radiosonde sites (in K). The % column in this table is the percentage of radiosonde records within the height interval in that of all height intervals.

Height (km)	%	GWMT-D bias (RMSE)	GTm_N bias (RMSE)	GWMT-IV bias (RMSE)	GTm-III bias (RMSE)
< 2	30.1	0.52 (4.42)	-0.39 (4.50)	-3.21 (5.20)	-0.73 (4.48)
2–5	34.1	0.94 (4.67)	3.23 (6.00)	-8.18 (10.11)	3.23 (4.82)
5–9	35.8	0.51 (4.50)	9.83 (11.55)	-14.53 (18.33)	9.83 (6.50)

main advantage of using mode and median values is that they are more robust than the arithmetic mean value, especially in skewed distributions (see Fig. 9b and c). As a result, a warm (cold) bias of 3.8 K (−4.4 K) can be found in the GTm_N (GWMT-IV). The histograms of both GTm-III and GWMT-D (Fig. 9a and d) are normally distributed and the GWMT-D is slightly better than the GTm-III.

The entire radiosonde-derived T_m is grouped into three height intervals 0–2, 2–5, and 5–9 km according to their station heights. The accuracy comparisons between the GWMT-D and other models in different height intervals are listed in Table 5. It can be concluded that the accuracies of all the models except for the GWMT-D are significantly degraded with the increase of the height of the site. In contrast, it shows that the accuracy of GWMT-D is nearly stable in three different height ranges. Compared with the GTm_N model, the better performance of GTm-III may result from the fact that GGOS T_m , which was derived from ECMWF reanalysis data, is more consistent with the radiosonde data than the NCEP-derived T_m . The RMSE values of GWMT-D, GTm-III, and GTm_N are plotted in the Fig. 10 as a function of height relative to the ground surface in order to reveal the represen-

tative effect of terrain on the models. The GWMT-D model’s RMSEs are in the range of 4–5 K, while the other two models have large RMSE values at high altitudes. It is noted that the GWMT-IV model is excluded due to its poorer performance shown in Table 5. It shows that the accuracy of GWMT-D is better than 5 K, even at the top of the troposphere.

Figure 11 shows the monthly or seasonal performance of these four selected models. The monthly-mean RMSEs of all the models vary with month (or season) and only the GTm_N shows a variation pattern opposite to that of the other three models. The GWMT-D and GWMT-III present very similar results in both pattern of variation and monthly-mean RMSEs. The GWMT-IV performs the worst and GWMT-D performs the best among all these four models.

4.4 Impact of T_m on GPS-derived PWV

The purpose of determining T_m is to convert the ZWD of GPS signals to PWV. From Eq. (1), the relationship of the RMSEs between T_m and PWV can be obtained as

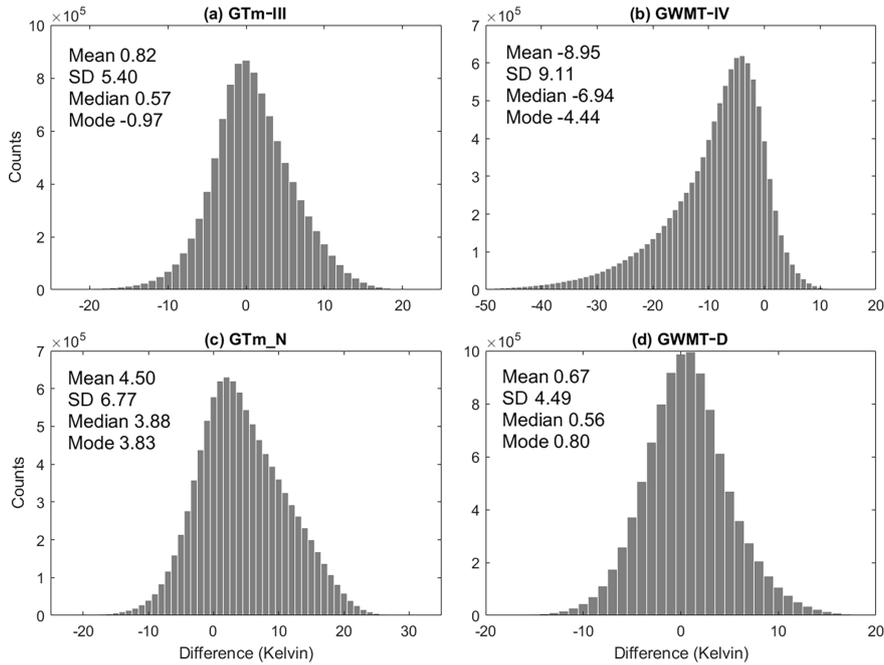


Figure 9. Histogram of model-derived T_m minus radiosonde-derived T_m in 2014 at different heights.

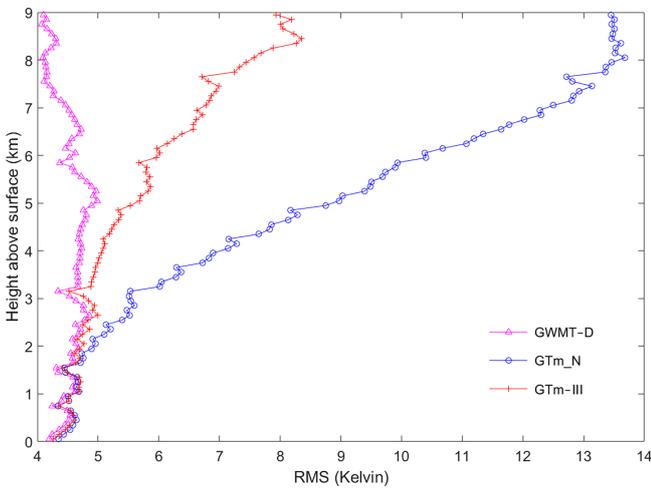


Figure 10. RMSE profile of the T_m from GTm-III, GTm_N, and GWMT-D models. The reference T_m values are derived from records of 585 radiosonde sites in 2014.

$$\begin{aligned} \frac{\text{RMSE}_{\text{PWV}}}{\text{PWV}} &= \frac{\text{RMSE}_{\Pi}}{\Pi} = \frac{k_3 \text{RMSE}_{T_m}}{(k_3/T_m + k'_2) T_m^2} \\ &= \frac{k_3}{(k_3/T_m + k'_2) T_m} \cdot \frac{\text{RMSE}_{T_m}}{T_m}, \end{aligned} \quad (12)$$

where the three RMSEs are defined for the differences between observed and true values (more details see Appendix C) and the relative error of PWV can be defined as $\text{RMSE}_{\text{PWV}}/\text{PWV}$ here.

Figure 12 illustrates the global distribution of both RMSE_{PWV} and $\text{RMSE}_{\text{PWV}}/\text{PWV}$ obtained from Eq. (11) and radiosonde data in 2014. The value of RMSE_{T_m} is obtained from Sect. 4.3 and PWV and T_m are set to annual mean values. Some radiosonde stations have been removed with insufficient observations or near the polar areas. As we can see the global mean values of RMSE_{PWV} and $\text{RMSE}_{\text{PWV}}/\text{PWV}$ are around 0.25 mm and 1.3 %, respectively.

5 Conclusion and discussion

T_m is a critical parameter in PWV detection using the GPS atmospheric sounding technique. Robust empirical T_m models are required as a practical alternative in RT/NRT PWV remote sensing systems, especially when in situ meteorological measurements cannot be obtained. The accuracies of the existing empirical T_m models are limited by the ill-modelling or neglect of global lapse rate and diurnal fluctuations in T_m . Therefore, a new voxel-based T_m model, namely GWMT-D, is developed in this study using global NCEP2 data from 2010 to 2013. The voxel-based modelling method effectively captures diurnal variations and lapse rates in T_m .

Moreover, comprehensive comparisons with GTm-III, GWMT-IV, and GTm_N show that GWMT-D is unbiased and can achieve a RMSE accuracy of 4–5 K for different seasons and locations. The improvement of the new model is around 25 % over the other three models when using NCEP2- and radiosonde-derived T_m as reference. Comparisons with GGOS surface T_m show that GWMT-D’s accuracy is slightly worse than that of GTm-III with a bias of ~ 1.2 K, possi-

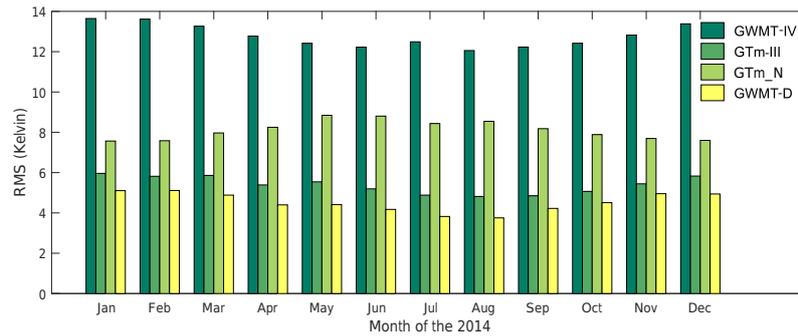


Figure 11. Monthly-mean RMSE of the T_m from the four models and reference values is global radiosonde-derived T_m in 2014.

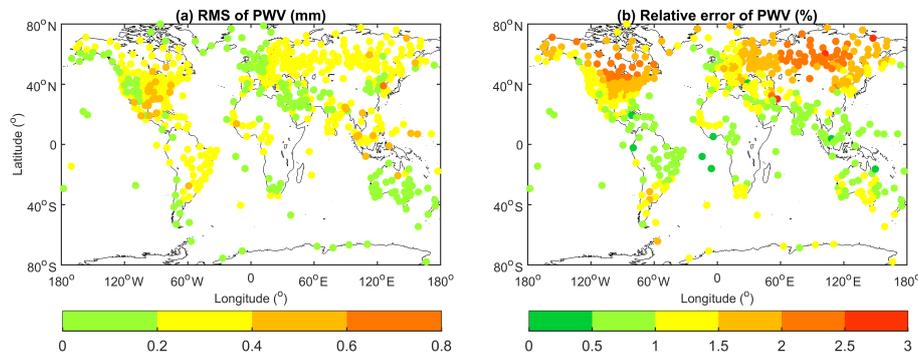


Figure 12. The theoretical RMSE (a) and relative error (b) of PWV resulting from the GWMT-D model using radiosonde observations in 2014.

ably due to the systematic difference between NCEP2 and ECMWF reanalysis data. This difference is significant, especially for the Antarctic. These comparisons also confirm that the piecewise linear interpolation of T_m used in GWMT-D is better than the direct modelling of T_m lapse rate in GWMT-IV and the constant-value method in GTm-III and GTm_N. An approximate 0.3 mm global mean RMSE in PWV read as 1.3 % relative error will be brought about by GWMT-D for ground stations.

It is suggested that the sets of coefficient for empirical T_m models (e.g. GWMT-D) need to be redetermined regularly using state-of-the-art data sets with appropriate length of time periods. The new GWMT-D model is an essential alternative T_m determination method to RT/NRT GPS PWV remote sensing system. The continuous operation of this system can be maintained even if in situ meteorological measurements on the GPS station are unavailable. Further inter-

comparisons are recommended between empirical T_m models and other measurements over the ocean (for instance with the Constellation Observation System of Meteorology, Ionosphere, and Climate, COSMIC), since radiosonde measurements used in this study are mainly collected over the land. An optimal integral approach for the determination of T_m and quantification of the stochastic characteristics of T_m time series are the main focus of our future work.

Data availability. NCEP2: <http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.pressure.html>.
 Radiosonde: <http://weather.uwyo.edu/upperair/sounding.html>.
 GGOS: <http://ggosatm.hg.tuwien.ac.at/DELAY/ETC/TMEAN/>.

Appendix A: Determination of T_m and water vapour pressure

This appendix presents the calculation of T_m and water vapour pressure from layered meteorological data (e.g. re-analysis and radiosonde). The numerical integration in the Eq. (3) can be approximated as

$$T_m \approx \frac{\sum_{i=1}^N \left(\frac{e_i}{T_i} + \frac{e_{i+1}}{T_{i+1}} \right) \frac{\Delta z_i}{2}}{\sum_{i=1}^N \left(\frac{e_i}{T_i^2} + \frac{e_{i+1}}{T_{i+1}^2} \right) \frac{\Delta z_i}{2}}, \quad (\text{A1})$$

where e_i and e_{i+1} are the water vapour pressure and T_i and T_{i+1} are the temperature on the lower and upper boundary of the i th layer of the atmosphere, Δz_i is the thickness of the i th layer, and N is the total number of layers.

Note that the height used in NCEP2 and radiosonde data is the geopotential height, which is widely used in meteorology, whilst the height used in the Eq. (A1) is a geometric height. The equations for the conversion of a geopotential height to a geometric height (ellipsoidal height) are (Ge, 2006)

$$h = \frac{R_e(\varphi) \cdot H}{\frac{g(\varphi)}{g_0} R_e(\varphi) - H}, \quad (\text{A2})$$

$$g(\varphi) = 9.80620(1 - 2.6442 \times 10^{-3} \cos 2\varphi + 5.8 \times 10^{-6} \cos^2 2\varphi), \quad (\text{A3})$$

$$R_e(\varphi) = \frac{a}{1 + f + m - 2f \sin^2 \varphi}, \quad (\text{A4})$$

where φ is the latitude, h is the ellipsoidal height (in km), and H is the geopotential height (in km); the constant g_0 is assigned to 9.80665 m s^{-2} ; $g(\varphi)$ is the gravity on the geoid; $R_e(\varphi)$ is the radius of curvature of the Earth at the latitude of φ ; and the parameters $a = 6378.137 \text{ km}$, $f = 1/298.257223563$, $m = 0.00344978650684$.

Since the humidity in layered, meteorological data are recorded as dew point temperature (T_d) or RH or specific humidity (q) instead of partial pressure of water vapour (e). The water vapour pressure needs to be computed first in the determination of T_m with T_d , RH, and q , i.e.

$$e = f(P) \cdot 6.112 \exp\left(\frac{17.62t}{243.12 + t}\right), \quad (\text{A5})$$

$$e = \frac{qP}{q + \varepsilon(1 - q)}, \quad (\text{A6})$$

$$\begin{aligned} \log_{10}(e)_{\text{liquid}} &= \log_{10}\left(\frac{\text{RH}}{100}\right) + \log_{10}(f(P)) \\ &+ 10.79574 \left(1 - \frac{273.16}{T}\right) - 5.028 \log_{10}\left(\frac{T}{273.16}\right) \\ &+ 1.50475 \times 10^{-4} \left(1 - 10^{-8.2969\left(\frac{T}{273.16} - 1\right)}\right) \\ &+ 0.42873 \times 10^{-3} \left(10^{-4.76955\left(1 - \frac{273.16}{T}\right)} - 1\right) \end{aligned}$$

$$+ 0.78614, \quad (\text{A7})$$

$$\begin{aligned} \log_{10}(e)_{\text{ice}} &= \log_{10}\left(\frac{\text{RH}}{100}\right) + \log_{10}(f(P)) \\ &- 9.096853 \left(\frac{273.16}{T} - 1\right) - 3.566506 \log_{10}\left(\frac{273.16}{T}\right) \\ &+ 0.876812 \left(1 - \frac{T}{273.16}\right) + 0.78614, \quad (\text{A8}) \end{aligned}$$

where t is the temperature in Celsius degree, and $t = T - 273.15$; $\varepsilon = M_w/M_d$ is the ratio of the molar masses of vapour and dry air, respectively; $f(P)$ is an enhancement factor defined as the ratio of the saturation vapour pressure of moist air to that of pure water vapour (WMO, 2000); Eqs. (A7) and (A8) are deduced from the Goff's formulation and the unit of water vapour pressure e is Pa (Goff, 1957). T_m in this study is computed with RH data. Note that interpolations of meteorological measurements are not applied in Eq. (A1).

Appendix B: Empirical T_m models

B1 UNB3m

Strictly, the UNB3m model is not a specific T_m model, but it can be used to calculate T_m from the following equation (Leandro et al., 2008):

$$T_m^{\text{UNB3m}} = (T_0 - \beta_T \cdot h) \left(1 - \frac{\beta_T R}{g_m(\lambda + 1)}\right), \quad (\text{B1})$$

where T_0 is the temperature at the mean sea level, λ is the dimensionless water vapour pressure height factor, β_T is temperature lapse rate, g_m is the acceleration of gravity at the atmospheric column centroid, R is the gas constant for dry air, and h is the height of unknown position.

The UNB3m model neglects the longitudinal variations in T_m . The meteorological variables in Eq. (B1), i.e. T_0 , λ , and β_T , are linearly interpolated in latitudinal direction based on a simple look-up table.

B2 GPT2w

The GPT2w, an improved GPT model, was developed by Böhm et al. (2015). This empirical model can provide pressure, temperature, tropospheric delay, and T_m with the annual and semi-annual amplitudes. The updated model was established on a regular resolution of 5° with monthly meteorological data of 10-year (2001–2010) ERA-Interim. The T_m is calculated by Eq. (B5), but the coefficients in this equation are determined based on a regular grid of 5 or 1° . Note that the GPT2w is not specifically designed for T_m computation.

B3 GWMT series models

The GWMT series models are global models developed and consistently improved by Yao et al. using the state-of-the-art

data sources and improved methodologies (Yao et al., 2015, 2014a, b, 2013, 2012).

The GWMT model was based on spherical harmonics of degree nine and order nine and is a function of the geodetic coordinates of the site, as expressed below:

$$T_m^{\text{GWMT}} = \alpha_1 + \alpha_2 h + \alpha_3 \cos\left(2\pi \frac{\text{DOY} - 28}{365.25}\right), \quad (\text{B2})$$

$$\alpha_i = \sum_{n=0}^9 \sum_{m=0}^n P_{nm}(\sin \varphi) \cdot \left[A_{nm}^i \cos(m\lambda) + B_{nm}^i \sin(m\lambda) \right] \quad (i = 1 \text{ or } 3), \quad (\text{B3})$$

where the globally mean lapse rate of T_m and α_2 is -4.1 K km^{-1} ; φ , λ , and h are the latitude, longitude, and height of the site, respectively; DOY is the day of year; P_{nm} is the Legendre function; A_{nm}^i and B_{nm}^i in Eq. (6) are two coefficients estimated from the least-squares estimation.

The GTm-II and GWMT models are developed using the same methodology but with different data.

Considering the semi-annual and diurnal variations in T_m , the GTm-III model can be expressed as

$$T_m^{\text{GTm-III}} = \alpha_1 + \alpha_2 h + \alpha_3 \cos\left(2\pi \frac{\text{DOY} - C_1}{365.25}\right) + \alpha_4 \cos\left(4\pi \frac{\text{DOY} - C_2}{365.25}\right) + \alpha_5 \sin\left(2\pi \frac{\text{HOD} - C_3}{24}\right), \quad (\text{B4})$$

where HOD is the hour of the day. The coefficients α_i ($i = 1, 2, \dots, 3$) are expanded to spherical harmonics similar with the case in GWMT and GTm-II.

Since the adjustment model in Eq. (B4) for the GTm-III is non-linear, the coefficients determined may be unstable or biased. Chen et al. (2014) established the GTm_N model with a global grid of $2.5^\circ \times 2.5^\circ$ NCEP reanalysis data neglecting the diurnal variation in T_m . The GTm_N model linearises the Eq. (B4) as (Chen et al., 2014):

$$T_m^{\text{GTm-N}} = \alpha_1 + \alpha_2 h + \alpha'_3 \cos\left(2\pi \frac{\text{DOY}}{365.25}\right) + \beta'_3 \sin\left(2\pi \frac{\text{DOY}}{365.25}\right) + \alpha'_4 \cos\left(4\pi \frac{\text{DOY}}{365.25}\right) + \beta'_4 \sin\left(4\pi \frac{\text{DOY}}{365.25}\right). \quad (\text{B5})$$

All the aforementioned models are based on such an assumption that the vertical lapse rate of T_m is the same over the globe, i.e. the α_2 in these equations are constants. In fact, this assumption is not always true (He et al., 2013). Therefore, the horizontal variation of T_m lapse rate (β) is considered in the GWMT-IV model. It is a function of the horizontal location. Thus, the global T_m at the height of h can be expressed as a function of the mean sea level T_m (T_m^0) and β in Eq. (9), both of which can be further separated into annual and semi-annual components. Parameters of both amplitude and initial

phase for annual and semi-annual variations are similarly expanded into a spherical harmonics form.

$$T_m^{\text{GWMT-IV}} = T_m^0 + \beta \cdot h \quad (\text{B6})$$

Appendix C: Approximated propagation of RMSE

Given a series of observations V collected at the same time (Ning et al., 2016):

$$V_i = \tilde{V} + M + \varepsilon_i, \quad (\text{C1})$$

where M is a time-independent bias (systematic error); \tilde{V} is the true value of observations, V_i ; and ε_i is the zero-mean stationary Gaussian random error. Hence, the RMSE of the difference between estimates and true values is given by

$$\text{RMSE}_V = \sqrt{\frac{1}{N} \sum_{i=1}^N (V_i - \tilde{V})^2} = \sqrt{\frac{1}{N} \sum_{i=1}^N (M + \varepsilon_i)^2}, \quad (\text{C2})$$

where N is the total number of observations. Since the mean value of ε_i will be close to zero for massive repeated observations, Eq. (C2) can be approximately reduced to

$$\text{RMSE}_V = \sqrt{\frac{1}{N} \sum_{i=1}^N (M^2 + \varepsilon_i^2)} = \sqrt{M^2 + \sigma_\varepsilon^2}, \quad (\text{C3})$$

where σ_ε is the standard deviation of ε . As can be seen from this equation, the RMSE will be identical to standard deviation when observations are free of systematic bias. Consider a linear or non-linear function $W = f(V) : R \rightarrow R$ whose RMSE can be expressed by

$$\text{RMSE}_W = \sqrt{\frac{1}{N} \sum_{i=1}^N [f(V_i) - f(\tilde{V})]^2}. \quad (\text{C4})$$

Using first-order Taylor expansion, we have

$$f(V) - f(\tilde{V}) \approx (V - \tilde{V}) \cdot \left. \frac{\partial f(V)}{\partial V} \right|_{V=\tilde{V}}. \quad (\text{C5})$$

Substituting Eq. (C5) into Eq. (C4),

$$\begin{aligned} \text{RMSE}_W &\approx \left. \frac{\partial f(V)}{\partial V} \right|_{V=\tilde{V}} \cdot \sqrt{\frac{1}{N} \sum_{i=1}^N (V_i - \tilde{V})^2} \\ &= \left. \frac{\partial f(V)}{\partial V} \right|_{V=\tilde{V}} \cdot \text{RMSE}_V. \end{aligned} \quad (\text{C6})$$

As a result, the RMSE of $f(V)$ can be approximately propagated from that of V .

The Supplement related to this article is available online at <https://doi.org/10.5194/amt-10-2045-2017-supplement>.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We would like to thank the NOAA/OAR/ESRL PSD (Boulder, Colorado, USA) for providing NCEP-DOE Reanalysis 2 data (2010–2014), the Department of Atmospheric Science in the University of Wyoming for providing access to radiosonde data in 2014, the GGOS (Global Geodetic Observing System) Atmosphere for providing 2014 surface T_m product. Jennifer Anderson and Timothy Kodikara at RMIT University and the three anonymous reviewers are thanked for their assistance in evaluating this paper. This research is partially supported by the Australian Research Council grant (ARC-LP130100243) and the National Science Foundation of China (No. 41404033).

Edited by: J. Jones

Reviewed by: three anonymous referees

References

- Askne, J. and Nordius, H.: Estimation of Tropospheric Delay for Microwaves from Surface Weather Data, *Radio Sci.*, 22, 379–386, <https://doi.org/10.1029/RS022i003p00379>, 1987.
- Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A., and Ware, R. H.: Gps Meteorology – Remote-Sensing of Atmospheric Water-Vapor Using the Global Positioning System, *J. Geophys. Res.-Atmos.*, 97, 15787–15801, 1992.
- Bevis, M., Businger, S., Chiswell, S., Herring, T. A., Anthes, R. A., Rocken, C., and Ware, R. H.: Gps Meteorology – Mapping Zenith Wet Delays onto Precipitable Water, *J. Appl. Meteorol.*, 33, 379–386, [https://doi.org/10.1175/1520-0450\(1994\)033<0379:Gmmzwd>2.0.Co;2](https://doi.org/10.1175/1520-0450(1994)033<0379:Gmmzwd>2.0.Co;2), 1994.
- Böhm, J., Heinkelmann, R., and Schuh, H.: Short note: a global model of pressure and temperature for geodetic applications, *J. Geodesy*, 81, 679–683, <https://doi.org/10.1007/s00190-007-0135-3>, 2007.
- Böhm, J., Moller, G., Schindelegger, M., Pain, G., and Weber, R.: Development of an improved empirical model for slant delays in the troposphere (GPT2w), *GPS Solut.*, 19, 433–441, <https://doi.org/10.1007/s10291-014-0403-7>, 2015.
- Buizza, R., Houtekamer, P. L., Toth, Z., Pellerin, G., Wei, M. Z., and Zhu, Y. J.: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems, *Mon. Weather Rev.*, 133, 1076–1097, <https://doi.org/10.1175/Mwr2905.1>, 2005.
- Champollion, C., Masson, F., Van Baelen, J., Walpersdorf, A., Chéry, J., and Doerflinger, E.: GPS monitoring of the tropospheric water vapor distribution and variation during the 9 September 2002 torrential precipitation episode in the Cévennes (southern France), *J. Geophys. Res.-Atmos.*, 109, D24102, <https://doi.org/10.1029/2004JD004897>, 2004.
- Chen, P. and Yao, W.: GTm_X: A New Version Global Weighted Mean Temperature Model, China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume II, 605–611, 2015.
- Chen, P., Yao, W. Q., and Zhu, X. J.: Realization of global empirical model for mapping zenith wet delays onto precipitable water using NCEP re-analysis data, *Geophys. J. Int.*, 198, 1748–1757, <https://doi.org/10.1093/gji/ggu223>, 2014.
- Choy, S., Zhang, K., Wang, C.-S., Li, Y., and Kuleshov, Y.: Remote Sensing of the Earth's Lower Atmosphere during Severe Weather Events using GPS Technology: a Study in Victoria, Australia, Proceedings of the 24th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2011), 559–571, 2001.
- Davis, J. L., Herring, T. A., Shapiro, I. I., Rogers, A. E. E., and Elgered, G.: Geodesy by Radio Interferometry – Effects of Atmospheric Modeling Errors on Estimates of Baseline Length, *Radio Sci.*, 20, 1593–1607, <https://doi.org/10.1029/RS020i006p01593>, 1985.
- Duan, J. P., Bevis, M., Fang, P., Bock, Y., Chiswell, S., Businger, S., Rocken, C., Solheim, F., vanHove, T., Ware, R., McClusky, S., Herring, T. A., and King, R. W.: GPS meteorology: Direct estimation of the absolute value of precipitable water, *J. Appl. Meteorol.*, 35, 830–838, [https://doi.org/10.1175/1520-0450\(1996\)035<0830:Gmdeot>2.0.Co;2](https://doi.org/10.1175/1520-0450(1996)035<0830:Gmdeot>2.0.Co;2), 1996.
- Emardson, T. R. and Derks, H. J. P.: On the relation between the wet delay and the integrated precipitable water vapour in the European atmosphere, *Meteorol. Appl.*, 7, 61–68, <https://doi.org/10.1017/S1350482700001377>, 2000.
- Fornberg, B. and Zuev, J.: The Runge phenomenon and spatially variable shape parameters in RBF interpolation, *Comput. Math. Appl.*, 54, 379–398, <https://doi.org/10.1016/j.camwa.2007.01.028>, 2007.
- Ge, S.: GPS radio occultation and the role of atmospheric pressure on spaceborne gravity estimation over Antarctica, The Ohio State University, 2006.
- Goff, J. A.: Saturation pressure of water on the new Kelvin temperature scale, *Transactions of the American Society of Heating and Ventilating Engineers*, 63, 347–353, 1957.
- Gutman, S. I. and Benjamin, S. G.: The role of ground-based GPS meteorological observations in numerical weather prediction, *GPS Solut.*, 4, 16–24, 2001.
- He, C., Yao, Y., Zhao, D., Li, K., and Qian, C.: GWMT Global Atmospheric Weighted Mean Temperature Models: Development and Refinement, China Satellite Navigation Conference (CSNC) 2013 Proceedings, 487–500, 2013.
- Jade, S., Vijayan, M. S. M., Gaur, V. K., Prabhu, T. P., and Sahu, S. C.: Estimates of precipitable water vapour from GPS data over the Indian subcontinent, *J. Atmos. Sol.-Terr. Phys.*, 67, 623–635, <https://doi.org/10.1016/j.jastp.2004.12.010>, 2005.
- Jin, S. and Luo, O.: Variability and climatology of PWV from global 13-year GPS observations, *IEEE T. Geosci. Remote Sens.*, 47, 1918–1924, 2009.
- Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S. K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP-DOE AMIP-II reanalysis (R-2), *B. Am. Meteorol. Soc.*, 83, 1631–1643, <https://doi.org/10.1175/Bams-83-11-1631>, 2002.
- Kwon, H. T., Jung, E. H., and Lim, G. H.: A comparison of GPS- and NWP-derived PW data over the Korean

- Peninsula, *Advances in Atmospheric Sciences*, 27, 871–882, <https://doi.org/10.1007/s00376-009-9069-4>, 2010.
- Leandro, R. F., Langley, R. B., and Santos, M. C.: UNB3m_pack: a neutral atmosphere delay package for radiometric space techniques, *GPS Solut.*, 12, 65–70, <https://doi.org/10.1007/s10291-007-0077-5>, 2008.
- Li, Z. H., Muller, J. P., and Cross, P.: Comparison of precipitable water vapor derived from radiosonde, GPS, and Moderate-Resolution Imaging Spectroradiometer measurements, *J. Geophys. Res.-Atmos.*, 108, <https://doi.org/10.1029/2003jd003372>, 2003.
- Liu, J.: Tropopause heights with its applications in the key parameters of the troposphere models Master, School of Geodesy and Geomatics, Wuhan University, Wuhan, 2015.
- Mendes, V., Prates, G., Santos, L., and Langley, R.: An evaluation of the accuracy of models for the determination of the weighted mean temperature of the atmosphere, *Proceedings of ION*, 433–438, 2000.
- Ning, T., Wang, J., Elgered, G., Dick, G., Wickert, J., Bradke, M., Sommer, M., Querel, R., and Smale, D.: The uncertainty of the atmospheric integrated water vapour estimated from GNSS observations, *Atmos. Meas. Tech.*, 9, 79–92, <https://doi.org/10.5194/amt-9-79-2016>, 2016.
- Picard, A., Davis, R. S., Glaser, M., and Fujii, K.: Revised formula for the density of moist air (CIPM-2007), *Metrologia*, 45, 149–155, <https://doi.org/10.1088/0026-1394/45/2/004>, 2008.
- Prasad, A. K. and Singh, R. P.: Validation of MODIS Terra, AIRS, NCEP/DOE AMIP-II Reanalysis-2, and AERONET Sun photometer derived integrated precipitable water vapor using ground-based GPS receivers over India, *J. Geophys. Res.-Atmos.*, 114, D05107, <https://doi.org/10.1029/2008jd011230>, 2009.
- Ross, R. J. and Rosenfeld, S.: Estimating mean weighted temperature of the atmosphere for Global Positioning System applications, *J. Geophys. Res.-Atmos.*, 102, 21719–21730, <https://doi.org/10.1029/97jd01808>, 1997.
- Ross, R. J. and Rosenfeld, S.: Correction to “Estimating mean weighted temperature of the atmosphere for Global Positioning System applications” by Rebecca J. Ross and Simon Rosenfeld, *J. Geophys. Res.-Atmos.*, 104, 27625–27625, 1999.
- Saastamoinen, J.: Contributions to the theory of atmospheric refraction, *B. Geod.*, 105, 279–298, 1972.
- Schueler, T., Pósfay, A., Hein, G. W., and Biberger, R.: A global analysis of the mean atmospheric temperature for GPS water vapor estimation, *Proceedings of ION-GPS*, 11–14, 2001.
- Song, S., Zhu, W., Ding, J., and Cheng, Z.: Coming process of Meiyu season in 2002 over Yangtze river delta monitored by Shanghai GPS network, *Progress in Astronomy*, 21, 180–184, 2003.
- Song, S. L., Zhu, W. Y., Ding, J. C., Liao, X. H., Cheng, Z. Y., and Ye, Q. X.: Near Real-Time Sensing of PWV from SGCAN and the Application Test in Numerical Weather Forecast, *Chinese J. Geophys.-Ch.*, 47, 719–727, 2004.
- Van Baelen, J. and Penide, G.: Study of water vapor vertical variability and possible cloud formation with a small network of GPS stations, *Geophys. Res. Lett.*, 36, L02804, <https://doi.org/10.1029/2008GL036148>, 2009.
- Wang, J., Zhang, L., and Dai, A.: Global estimates of water-vapor-weighted mean temperature of the atmosphere for GPS applications, *J. Geophys. Res.-Atmos.*, 110, D21101, <https://doi.org/10.1029/2005JD006215>, 2005.
- Wang, J., Zhang, L., Dai, A., Van Hove, T., and Van Baelen, J.: A near-global, 2-hourly data set of atmospheric precipitable water from ground-based GPS measurements, *J. Geophys. Res.-Atmos.*, 112, D11107, <https://doi.org/10.1029/2006JD007529>, 2007.
- Wang, X., Dai, Z., Cao, Y., and Song, L.: Weighted mean temperature T (m) statistical analysis in ground-based GPS in China, *Geomatics and Information Science of Wuhan University*, 36, 412–416, 2011.
- Wang, X. M., Zhang, K. F., Wu, S. Q., Fan, S. J., and Cheng, Y. Y.: Water vapor-weighted mean temperature and its impact on the determination of precipitable water vapor and its linear trend, *J. Geophys. Res.-Atmos.*, 121, 833–852, <https://doi.org/10.1002/2015jd024181>, 2016.
- WMO: General Meteorological Standards and Recommended Practices, Appendix A, WMO Technical Regulations, edited by: WMO, Geneva, Switzerland, 2000.
- Yang, X. H., Sass, B. H., Elgered, G., Johansson, J. M., and Emardson, T. R.: A comparison of precipitable water vapor estimates by an NWP simulation and GPS observations, *J. Appl. Meteorol.*, 38, 941–956, [https://doi.org/10.1175/1520-0450\(1999\)038<0941:Acopwv>2.0.Co;2](https://doi.org/10.1175/1520-0450(1999)038<0941:Acopwv>2.0.Co;2), 1999.
- Yao, Y. B., Zhu, S., and Yue, S. Q.: A globally applicable, season-specific model for estimating the weighted mean temperature of the atmosphere, *J. Geodesy*, 86, 1125–1135, <https://doi.org/10.1007/s00190-012-0568-1>, 2012.
- Yao, Y. B., Zhang, B., Yue, S. Q., Xu, C. Q., and Peng, W. F.: Global empirical model for mapping zenith wet delays onto precipitable water, *J. Geodesy*, 87, 439–448, <https://doi.org/10.1007/s00190-013-0617-4>, 2013.
- Yao, Y. B., Xu, C. Q., Zhang, B., and Cao, N.: GTm-III: a new global empirical model for mapping zenith wet delays onto precipitable water vapour, *Geophys. J. Int.*, 197, 202–212, <https://doi.org/10.1093/gji/ggu008>, 2014a.
- Yao, Y. B., Zhang, B., Xu, C. Q., and Yan, F.: Improved one/multi-parameter models that consider seasonal and geographic variations for estimating weighted mean temperature in ground-based GPS meteorology, *J. Geodesy*, 88, 273–282, <https://doi.org/10.1007/s00190-013-0684-6>, 2014b.
- Yao, Y. B., Xu, C. Q., Zhang, B., and Cao, N.: A global empirical model for mapping zenith wet delays onto precipitable water vapor using GGOS Atmosphere data, *Sci. China Earth Sci.*, 58, 1361–1369, <https://doi.org/10.1007/s11430-014-5025-y>, 2015.
- Zhang, K. F., Manning, T., Wu, S. Q., Rohm, W., Silcock, D., and Choy, S.: Capturing the Signature of Severe Weather Events in Australia Using GPS Measurements, *IEEE J-Stars*, 8, 1839–1847, <https://doi.org/10.1109/Jstars.2015.2406313>, 2015.
- Zhang, M., Ni, Y. Q., and Zhang, F. Q.: Variational assimilation of GPS precipitable water vapor and hourly rainfall observations for a meso-beta scale heavy precipitation event during the 2002 Mei-Yu season, *Adv. Atmos. Sci.*, 24, 509–526, <https://doi.org/10.1007/s00376-007-0509-8>, 2007.