Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 61 (2018) 402–412

ADVANCES IN SPACE RESEARCH (a COSPAR publication)

www.elsevier.com/locate/asr

A new weighted mean temperature model in China

Jinghong Liu, Yibin Yao *, Jizhang Sang

School of Geodesy and Geomatics, Wuhan University, Wuhan, China

Received 15 January 2017; received in revised form 15 September 2017; accepted 17 September 2017
Available online 23 September 2017

Abstract

The Global Positioning System (GPS) has been applied in meteorology to monitor the change of Precipitable Water Vapor (PWV) in atmosphere, transformed from Zenith Wet Delay (ZWD). A key factor in converting the ZWD into the PWV is the weighted mean temperature (T_m), which has a direct impact on the accuracy of the transformation. A number of Bevis-type models, like $T_m - T_s$ and $T_m - T_s, P_s$ type models, have been developed by statistics approaches, and are not able to clearly depict the relationship between T_m and the surface temperature, T_s. A new model for T_m, called weighted mean temperature norm model (abbreviated as norm model), is derived as a function of T_s, the lapse rate of temperature, δ, the tropopause height, h_{trop}, and the radiosonde station height, h_s. It is found that T_m is better related to T_s through an intermediate temperature. The small effects of lapse rate can be ignored and the tropopause height be obtained from an empirical model. Then the norm model is reduced to a simplified form, which causes fewer loss of accuracy and needs two inputs, T_s and h_s. In site-specific fittings, the norm model performs much better, with RMS values reduced averagely by 0.45 K and the Mean of Absolute Differences (MAD) values by 0.2 K. The norm model is also found more appropriate than the linear models to fit T_m in a large area, not only with the RMS value reduced from 4.3 K to 3.80 K, correlation coefficient R^2 increased from 0.84 to 0.88, and MAD decreased from 3.24 K to 2.90 K, but also with the distribution of simplified model values to be more reasonable. The RMS and MAD values of the differences between reference and computed PWVs are reduced by on average 16.3% and 14.27%, respectively, when using the new norm models instead of the linear model.

Keywords: Tropospheric zenith delay; Precipitable water vapor; Weighted mean temperature; Norm model

1. Introduction

Water vapor is one of the most active gases on the Earth surface, and is unevenly distributed in time and space (Yao et al., 2012). It affects the long-term and short-term changes in weather and climate. It has been a main research focus for meteorologists to monitor the change of water vapor. In the ground-based GPS meteorology, the GPS signal delay due to the atmospheric refraction is used to study water vapor contents. The delay in the zenith direction is called zenith total delay, and the Zenith Wet Delay (ZWD) can be separated from the zenith total delay by removing the zenith hydrostatic delay. Bevis et al. (1992) discussed the principle of applying GPS technology to detect water vapor contents, derived from the relationship between the ZWD and the precipitable water vapor (PWV), and obtained a linear regression model relating the weighted mean temperature to the surface temperature using 8718 radiosonde data distributed in the middle latitude region of the USA. That makes the transformation of ZWD to PWV possible, and the GPS technology since then has become an important technology to monitor the change of water vapor contents.

GPS can obtain high spatial and temporal resolution of the precipitation, and has a wide range of applications. The variation characteristics of precipitable water vapor can be used as an indicator to detect the fog (Lee et al., 2010), to

* Corresponding author.
E-mail address: ybyao@whu.edu.cn (Y. Yao).

https://doi.org/10.1016/j.asr.2017.09.023
0273-1177/

correct the InSAR data affected by the water vapor (Lindenbergh et al., 2009), and to estimate the arrival time of plum rain in China (Cao et al., 2007). Therefore, it has an important significance for numerical weather prediction and climate related research (Heise et al., 2015).

When the wet refractivity on a profile is available, the ZWD can be obtained from the following equation (Singh et al., 2014)

$$ZWD \approx 10^{-6}\sum_{i=1}^{n}\frac{N_w^i + N_w^{i+1}}{2}\Delta H_{i+1,i} \tag{1}$$

where N_w is the wet refractivity, i is a point on the profile, $H_{i+1,i}$ is the geopotential height difference between points $i+1$ and i, and n is the total number of points available along the profile. The wet refractivity N_w is given by Thayer (1974)

$$N_w = \left[k_2'\frac{e}{T} + k_3\frac{e}{T^2}\right]Z_w^{-1} \tag{2}$$

where e is the partial pressure of the water vapor in mbar, k_2' and k_3 are the atmospheric refractivity constants (Davis et al., 1985; Bevis et al., 1994), $k_2' = 16.5 \pm 10$ K/mbar, $k_3 = 377.600 \pm 3000$ K^2/mbar, and Z_w^{-1} is

$$Z_w^{-1} = 1 + 1650\left(\frac{e}{T^3}\right)[1 - 0.01317T_c + 1.75\times10^{-4}T_c^2 + 1.44T_c^3] \tag{3}$$

where T_c and T are the dew point temperatures in Celsius and Kelvin, respectively.

When the ZWD is available, the PWV can be computed by

$$PWV = \Pi \times ZWD \tag{4}$$

where $\prod$ can be expressed by equation

$$\Pi = \frac{10^6}{\rho_w R_v[k_3/T_m + k_2']} \tag{5}$$

where ρ_w is the density of water, R_v is the specific gas constant.

Therefore, the ZWD can be estimated from radiosonde data using Eqs. (1)–(3). The estimated PWV from radiosonde data using Eq. (4) will serve as reference in analyzing the performance of estimating the PWV using different T_m models.

The weighted mean temperature, T_m, is given by Bevis et al. (1994)

$$T_m = \frac{\int e/T dh}{\int e/T^2 dh} \tag{6}$$

where the integration is made from the station to the tropopause, their heights being h_s and h_{trop}, respectively.

Achieving accuracy of 1% and 2%, respectively, in PWV would require errors in T_m less than 2.74 K and 5.48 K on average, respectively (Wang et al., 2005). This means the accuracy of T_m has to be improved. For an area without radiosonde data, it is impossible to calculate T_m from Eq. (6). So many regional or global linear models have been developed, such as linear models between T_m and T_s (Bevis et al., 1992; Mendes et al., 2000; Solbrig, 2000), and T_m and (T_s, e) (Singh et al., 2014). In addition, a non-meteorological parameter model for computing T_m was suggested by Yao et al. (2012), only requiring the input parameters of time and 3D coordinates of the observing station. All models mentioned above have been developed by using mathematical fitting techniques, but they lack theoretical basis, and thus it is subjective in determining which parameters are included in the fitting.

In the following, Eq. (6) is re-formulated into a new function form by using the functional inner product and functional norm. The resulting equation is called the weighted mean temperature norm equation. The norm equation gives theoretical descriptions of how the influencing parameters affect T_m. In site-specific situations, the effects of temperature lapse rate, δ, and $(h_{trop} + h_s)$, on the fitting performance are assessed to propose a simplified model. The norm model and its simplified model are then determined using the radiosonde data of 84 stations in China between 2008 and 2010, and compared with linear models of T_m. Finally, four radiosonde stations at different latitudes are chosen to study the performance of the simplified model in retrieving the PWV. In the performance evaluation, the ZWD, T_m and PWV, computed respectively from Eq. (1), (6) and (4), are used as reference.

2. Derivation of the weighted mean temperature norm model

By using the functional norm, Eq. (6) can be expressed as (Yao et al., 2015)

$$T_m = \frac{\|T\|\cdot\varphi}{\sqrt{h_{trop} - h_s}} \tag{7}$$

where T is the norm of the temperature along the profile above the observing station. The derivation process from Eq. (6) to Eq. (7) is given in Appendix A. The models deduced from Eq. (7) are collectively called the weighted mean temperature norm models, norm models for short.

The temperature at a height h over an observing station can be expressed by $T = T_s + \delta(h - h_s)$, assuming that the temperature decreases linearly with the height. And the expression of $T = T_s + \delta(h - h_s)$ can be simplified as follow

$$T = T_k + \delta h \tag{8}$$

where $T_k = T_s - \delta h_s$, the expression of T is then obtained as

$$\|T\| = \sqrt{\int_{h_s}^{h_{trop}}(T_s + \delta(h - h_s))^2 dh} = \sqrt{\int_{h_s}^{h_{trop}}(T_k + \delta h)^2 dh} = \sqrt{\frac{1}{3\delta}(T_k + \delta h)^3\Big|_{h_s}^{h_{trop}}}$$

$$= T_k\sqrt{h_{trop}-h_s} \times\sqrt{1+\frac{\delta(h_{trop}+h_s)}{T_k}+\frac{\delta^2(h_{trop}^2+h_{trop}h_s+h_s^2)}{3T_k^2}} \quad (9)$$

The approximate expressions of φ is given by Yao et al. (2015)

$$\varphi=\frac{b}{T_k}-\frac{b\delta(h_{trop}+h_s)}{2T_k^2}+bw_1 \quad (10)$$

where b is related to the temperature and water vapor pressure.

Making the Taylor series expansion to the first order, the specific expression of T is given as follow

$$\|T\|=T_k \times\sqrt{h_{trop}-h_s}\left(1+\frac{\delta(h_{trop}+h_s)}{2T_k}+\frac{\delta^2(h_{trop}^2+h_{trop}h_s+h_s^2)}{6T_k^2}+w_2\right) \quad (11)$$

where w_1 and w_2 have relationships with δ, h_{trop}, T_k, and h_s below

$$w_1=\frac{1}{T_k}\left\{-\frac{\delta^2(h_{trop}^2+h_{trop}h_s+h_s^2)}{6T_k^2}+\sum_{i=1}^{\infty}\frac{-\frac{1}{2}\left(-\frac{1}{2}-i\right)}{(i+1)!}\left(\frac{\delta(h_{trop}+h_s)}{T_k}+\frac{\delta^2\left(h_{trop}^2+h_{trop}h_s+h_s^2\right)}{3T_k^2}\right)^{i+1}\right\} \quad (12)$$

$$w_2=\frac{\delta^2(h_{trop}^2+h_{trop}h_s+h_s^2)}{6T_k^2}+\sum_{i=1}^{\infty}\frac{\frac{1}{2}(\frac{1}{2}-i)}{(i+1)!}\left\{\frac{\delta\left(h_{trop}+h_s\right)}{T_k}+\frac{\delta^2\left(h_{trop}^2+h_{trop}h_s+h_s^2\right)}{3T_k^2}\right\}^{i+1} \quad (13)$$

Substituting Eq. (10) and Eq. (11) into Eq. (7), T_m can be expressed as follow

$$\begin{aligned}T_m=&(1+w_2)bw_1T_k-\frac{b\delta w_2(h_{trop}+h_s)}{2T_k}\\&+\frac{bw_1\delta^2(h_{trop}^2+h_s^2+h_{trop}h_s)}{6T_k}-\frac{b\delta^2(h_{trop}+h_s)^2}{4T_k^2}\\&+\frac{b\delta^2(h_{trop}^2+h_s^2+h_{trop}h_s)}{6T_k^2}\\&-\frac{b\delta^3(h_{trop}+h_s)(h_{trop}^2+h_s^2+h_{trop}h_s)}{12T_k^3}\\&+\frac{bw_1\delta(h_{trop}+h_s)}{2}+(1+w_2)b\end{aligned} \quad (14)$$

Introducing coefficients $a_i, i=1,2,\ldots,7$, Eq. (14) is simplified to

$$\begin{aligned}T_m=&a_0+a_1T_k+a_2\frac{\delta(h_{trop}+h_s)}{T_k}+a_3\frac{\delta^2(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_k}\\&+a_4\frac{\delta^2(h_{trop}+h_s)^2}{T_k^2}+a_5\frac{\delta^2(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_k^2}\\&+a_6\frac{\delta^3(h_{trop}+h_s)(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_k^3}+a_7\delta(h_{trop}+h_s)\end{aligned} \quad (15)$$

where coefficients $a_i, i=1,2,\ldots,7$ are

$$a_0=(1+w_2)b, a_1=(1+w_2)bw_1, a_2=-\frac{bw_2}{2}, a_3=\frac{bw_1}{6}$$
$$a_4=-\frac{b}{4}, a_5=\frac{b}{6}, a_6=-\frac{b}{12}, a_7=\frac{bw_1}{2}$$

Consider the facts that δ is about 6.5 K/km, the average station height h_s is about 1 km, and the maximum value of h_s is no more than 10 km. If the surface temperature is 273 K, then $T_s^2\approx 74529$ and $\delta h_s\approx 65$. Therefore $T_s^2\gg\delta h_s$, it is appropriate to make the following approximation for $\frac{1}{T_k}$

$$\frac{1}{T_k}=\frac{1}{T_s-\delta h_s}=\frac{1}{T_s(1-\delta h_s/T_s)}\approx\frac{1}{T_s}(1+\frac{\delta h_s}{T_s})\approx\frac{1}{T_s} \quad (16)$$

Substituting Eq. (16) into Eq. (15) and rearranging, we can get the expression

$$\begin{aligned}T_m\approx&a_0+a_1T_s+a_2\frac{\delta(h_{trop}+h_s)}{T_s}+a_3\frac{\delta^2(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_s}\\&+a_4\frac{\delta^2(h_{trop}+h_s)^2}{T_s^2}+a_5\frac{\delta^2(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_s^2}\\&+a_6\frac{\delta^3(h_{trop}+h_s)(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_s^3}+a_7\delta(h_{trop}+h_s)\end{aligned} \quad (17)$$

It can be found, from Eq. (14), that the coefficients $a_i, i=1,2,\ldots,7$, are functions of δ, h_{trop}, T_s, and h_s. Eq. (14) reveals that T_m is not a simple linear function of T_s, but a complicated function of T_k.

When Eq. (17) is truncated to $T_m=a_0+a_1T_s$, it is then similar to the Bevis linear model $T_m=b_0+b_1T_s$. Eq. (15) demonstrates that a_1 is affected by the four influencing factors. These factors have different features. The temperature lapse rate is closely related to the environment and latitude. The tropopause height near the equator is maximal, and gradually reduces toward the poles (Añel et al., 2007). Some studies concluded that regionally optimized models do not

have superior performance compared to the global models (Singh et al., 2014; Ross and Rosenfeld, 1997), and the site-specific model has a better accuracy. This can be explained by examining Eq. (17), where the coefficients are affected by the water vapor pressure (P_s), δ, h_{trop}, T_s, and h_s. The relationships between the weighted mean temperature and these influencing factors are given for the first time.

3. Assessment of weighted mean temperature norm model using radiosonde data in China

Considering the complexity of the expression, the data from China's 84 radiosonde stations over a 3-yr interval is used to analyze the relevant influencing factors in Eq. (17). Fig. 1 shows the locations of the 84 radiosonde stations in China, which are generally well distributed except in the Tibet region. This section is to analyze the effects of δ and $(h_{trop} + h_s)$ on the fitting performance, which will result in a simplified model for convenient application without much loss of accuracy. The assessment results, obtained from using site-specific data and regional data, demonstrate the validity of the new models.

The experiments are conducted following the workflow illustrated in Fig. 2. Three model variants, originated from Eq. (17), are obtained by removing some parts of Eq. (17). Each model variant will then have a RMS value, and a RMS ratio is defined as a metric to assess the effects of relevant influencing factors on the performances of data fitting. In Section 4, four stations are selected to illustrate how much a simple norm model can improve the accuracy of PWV.

3.1. The effects of δ and $(\boldsymbol{h_{trop}} + \boldsymbol{h_s})$ on the fitting performance

The radiosonde balloons are launched twice a day. The data contains the temperature profiles and station information, such as the height of station (h_s). It's easy to obtain the lapse rate, based on the assumption that the temperature decreases linearly with the height. The tropopause is defined by the change rate of temperature (Nagurny, 2003; Seidel et al., 2001) and can be obtained by the analysis of COSMIC occultation data (Schmidt et al., 2004). Liu (2015) proposed a global tropopause model, using global COSMIC data of 2008–2010. This model is based on the hypothesis that the tropopause height is a function of time and latitude of observing station given by Eq. (18). Table 1 lists the coefficients of the model.

$$h_{trop}(t,\varphi) = a_0(t) + a_1(t)\cos(w(t)\varphi) + b_1(t) \times \sin(w(t)\varphi) \tag{18}$$

where t is the time of the observation and φ is the latitude of observe station.

Now, the focus is turned to the performance of the norm model. Clearly, the T_m in Eq. (17) can be denoted as $f(h_s, h_{trop}, T_s, \delta)$. Three variants are proposed by removing some parts in the equation with an intention to understand the effects of δ and $(h_{trop} + h_s)$ on the fitting performances. The expressions of three model variants of Eq. (17) are given as follow:

$$f(T_s) : T_m \approx a_0 + a_1 T_s + \frac{a_2}{T_s} + \frac{a_3}{T_s^2} + \frac{a_4}{T_s^3} \tag{19}$$

$$f(T_s,\delta) : T_m \approx a_0 + a_1 T_s + a_2\frac{\delta}{T_s} + a_3\frac{\delta^2}{T_s} + a_4\frac{\delta^2}{T_s^2} + a_5\frac{\delta^3}{T_s^3} + a_6\delta \tag{20}$$

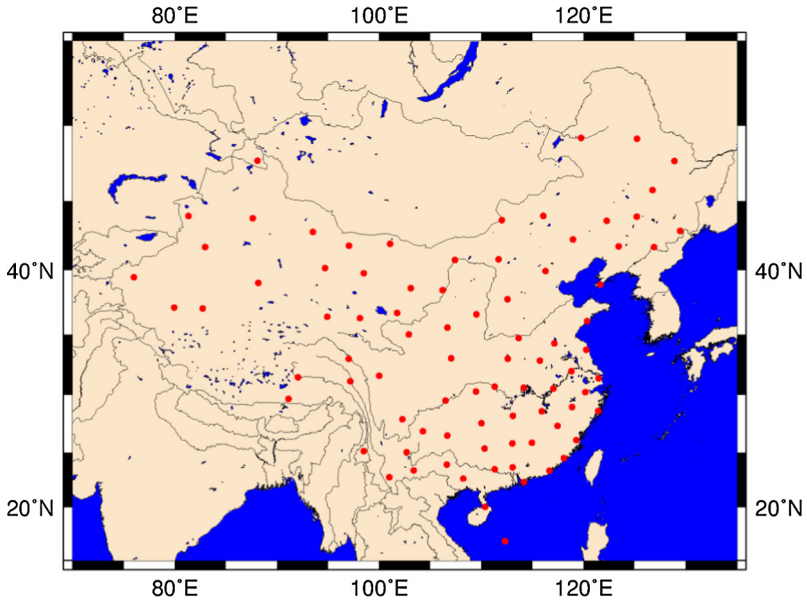

Fig. 1. Distribution of radiosonde stations in China.

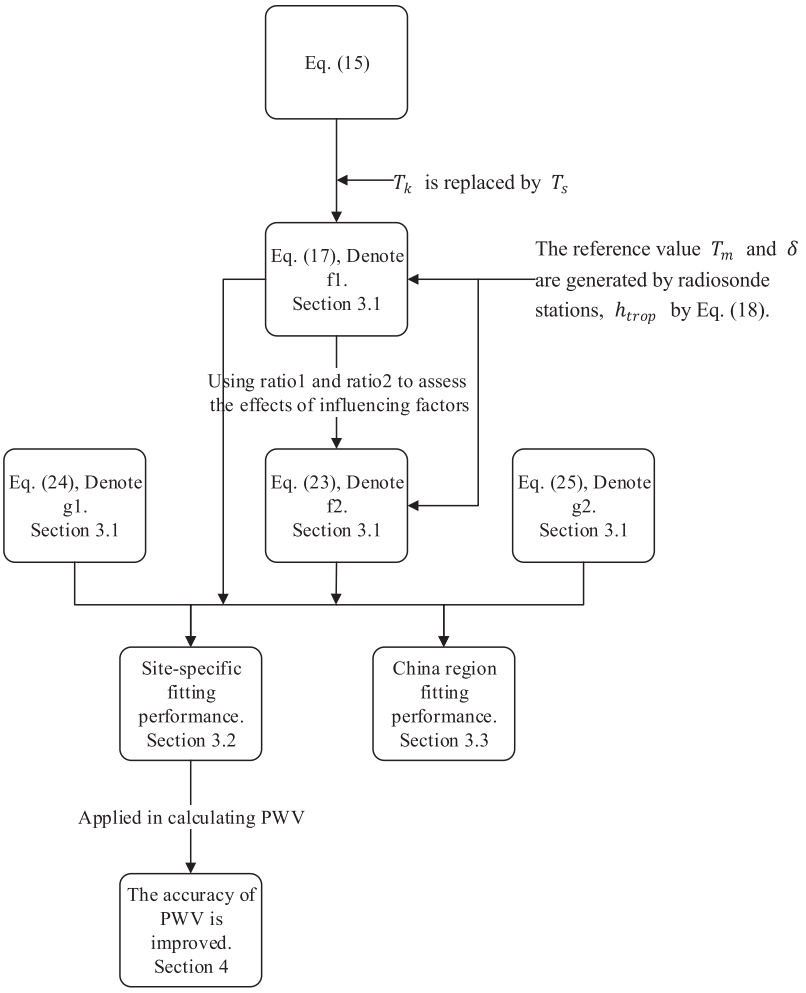

Fig. 2. Experiment workflow.

Table 1
the coefficients of the tropopause model.

Coefficients	Values
a_0	$13.503 - 0.1089\cos(2\pi t/365) - 0.2525\sin(2\pi t/365)$
a_1	$4.01015 + 0.08538\cos(2\pi(t+38)/365) + 0.60328\sin(2\pi(t+38)/365)$
b_1	$0.2887 - 0.1836\cos(2\pi(t+38)/365) - 0.9238\sin(2\pi(t+38)/365)$
w	$0.046345 + 0.001109\cos(2\pi t/365)\text{-}0.001885\sin(2\pi t/365)$

$$f(T_s,h_s,h_{trop}):T_m \approx a_0 + a_1T_s + a_2\frac{(h_{trop}+h_s)}{T_s} + a_3\frac{(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_s} + a_4\frac{(h_{trop}+h_s)^2}{T_s^2} + a_5\frac{(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_s^2} + a_6\frac{(h_{trop}+h_s)(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_s^3} + a_7(h_{trop}+h_s) \quad (21)$$

Denote RMS0 as the root mean square of $f(T_s)$, RMS1 for $f(T_s,\delta)$, and RMS2 for $f(T_s,h_s,h_{trop})$. RMS0 can be used as a reference value because of the strong correlation between T_m and T_s, demonstrated by many scholars. The following ratios are introduced:

$$ratio_1 = \frac{RMS1}{RMS0}, \quad ratio_2 = \frac{RMS2}{RMS0} \quad (22)$$

The ratio values represent the proportion of the RMS1 and RMS2 in the RMS0. The higher the ratio value is, the smaller the effects of including more influencing factors will be. Site-specific data is used to analyze the properties of three variants, with the corresponding results shown in Fig. 3.

The inclusion of δ and $(h_{trop}+h_s)$ in the norm equation is intended to reduce the root mean square values of data fitting. The range of $ratio_1$ is between 0.86 and 1.02. The range of $ratio_2$ is between 0.65 and 0.99. It appears that δ

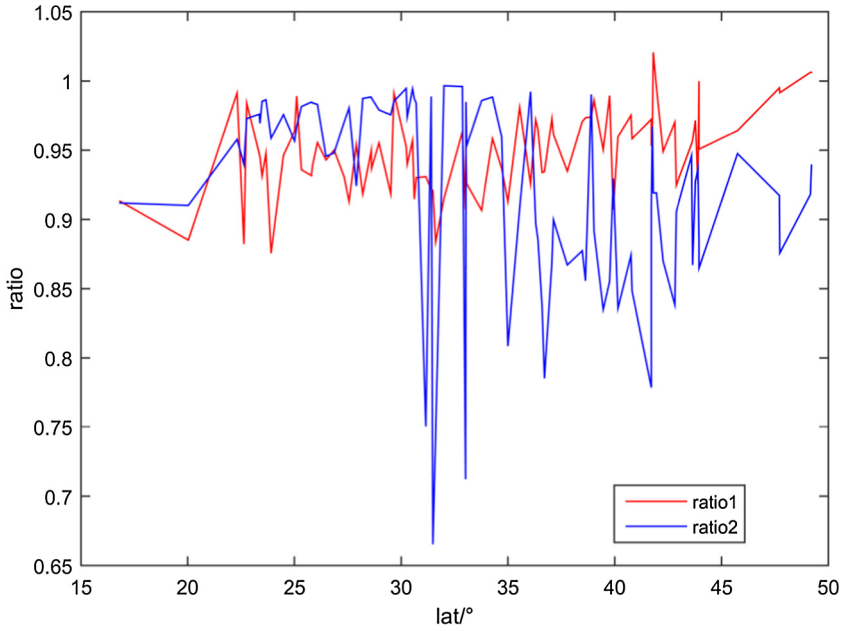

Fig. 3. The ratios of corresponding results.

has a strong impact on the model fitting than that of $(h_{trop}+h_s)$ in the region of latitudes between 15° and 32°, while the differences of the ratios in these areas are very small, in fact, no more than 0.05. Outside this region, the $ratio_2$ of $(h_{trop}+h_s)$ has lower values which decrease with the increasing latitude, and the differences between two ratios exceed 0.1, which implies that $(h_{trop}+h_s)$ has dominant effects on the fitting performance. The values of $ratio_2$ in the region of latitude from 32° to 36° drop sharply due to drastic tropopause change. This behavior is consistent with the study of tropopause height of Xu et al. (2013), and it demonstrates that the tropopause height indeed has a major impact on the weighted mean temperature. In the high latitude regions, tropopause plays a more important role in affecting the weighted mean temperature, and Eq. (17) gives a clear description of the relationship between tropopause and the weighted mean temperature for the first time.

From the above analysis, the relative small effects of δ on the fitting performance can be ignored. Therefore, Eq. (17) can be further simplified:

$$T_m \approx a_0 + a_1 T_s + a_2 \frac{(h_{trop}+h_s)}{T_s} + a_3 \frac{(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_s} + a_4 \frac{(h_{trop}+h_s)^2}{T_s^2} + a_5 \frac{(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_s^2} + a_6 \frac{(h_{trop}+h_s)(h_{trop}^2+h_s^2+h_{trop}h_s)}{T_s^3} + a_7(h_{trop}+h_s) \tag{23}$$

Obviously, Eq. (23) is a concise form. Eq. (17) and Eq. (23), denoted as f_1 and f_2, respectively, are assessed using the site-specific radiosonde data of 84 stations in 2008–2010 all over China. As comparisons, two linear models from Sapucci (2014), denoted as g_1 and g_2, are also assessed:

$$g_1 : T_m = b_0 + b_1 T_s \tag{24}$$

$$g_2 : T_m = c_0 + c_1 T_s + c_2 P_s \tag{25}$$

3.2. Modelling performance in site-specific station

In this section, the site-specific modelling performance, using the four models mentioned above, is assessed to investigate the differences between these models. This paper will use RMS and MAD as the measures to evaluate qualities of the data modelling, and R^2 to judge the degree of correlation between reference and fitted T_m. The larger the R^2 value is, the stronger the correlation will be.

The results are illustrated in Fig. 4, generated from the data of 84 radiosonde stations. The model f_1 is colored by red,[1] f_2 by blue, g_1 by cyan, and g_2 by green. The blue and green lines are surrounded by the red and cyan lines in the RMS picture, and the blue line is below the green line. The blue and cyan lines are surrounded by the red and green lines in MAD picture, the blue line is overlapped with the cyan line in the low latitudes, while the blue line gradually falls and is under the cyan line with increasing latitude. The red line shows strong correlations of the weighted mean temperature, and the blue and cyan lines share similar trends in the R^2 picture.

Therefore, the conclusions can be drawn from Fig. 4 that, in the site-specific situations, f_1 has the best performance, f_2 is inferior to f_1, g_2 is inferior to f_2, and the g_1

[1] For interpretation of color in Figs. 4–6, the reader is referred to the web version of this article.

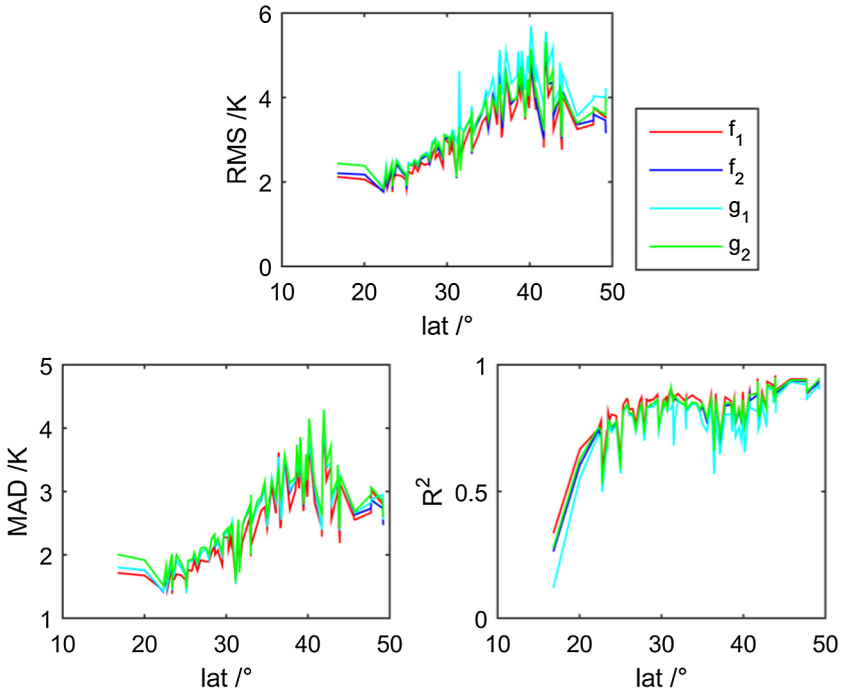

Fig. 4. Site-specific fitting of RMS, MAD, and R^2 at China radiosonde stations.

is worst among the four models. In the three figures, although f_2 and g_2 have similar trends, f_2 performs better than g_2 with lower values of RMS and MAD on the whole.

3.3. Modelling performance in China region

The performances of f_1, f_2, g_1, and g_2 over the whole China region are analyzed using all the radiosonde data. The RMS values are 3.80 K, 3.79 K, 4.30 K and 3.92 K for f_1, f_2, g_1, and g_2, respectively; the MAD values are 2.91 K, 2.90 K, 3.24 K, 3.06 K; and the R^2 values are: 0.88, 0.88, 0.84, 0.87. The differences in the RMS, MAD and R^2 values of f_1 and f_2 are relatively small. It indicates that the simplified Eq. (23) doesn't cause much loss of accuracy.

Fig. 5 illustrates the performance of the four models. The first picture reflects the distribution of f_1 (red colored) and g_1 (cyan colored), the second one reflects that of f_2 (red colored) and g_2 (cyan colored). The distribution of f_1 has obvious stripping patterns. This phenomenon reflects the combination effects of the temperature lapse rate and tropopause height on the weighted mean temperature, which have strong relationships with latitudes. The distribution f_2 becomes more concentrated because of ignoring the effects of temperature lapse rate. The distribution of g_2 seems unreasonable. The amount of water vapor can be very scarce in winter and plentiful in summer. The seasonal changes in the water vapor content perhaps make the distribution of g_2 unreasonable and the g_2 model gradually degenerates into g_1 model when temperature gets cold. Therefore, it will be much better to use the tropopause instead of water vapor as argument when fitting the models and the simplified Eq. (23) can achieve better accuracy on the whole.

4. Applications of norm models in retrieving PWV

From the discussion on the performance of the four models, we recommend model f_2. The new weighted mean temperature model can improve the accuracy of PWV. It has been proved by Bevis et al. (1994). The uncertainties in $\prod$ can be derived from the uncertainties in the weighted mean temperature, T_m, and in the physical constants k_2', and k_3. Let the errors in these quantities be σ_T, σ, and σ_3, respectively. We can derive the relative error of the important parameter $\prod$ in Eq. (5):

$$\frac{\sigma_\Pi}{\Pi} = \frac{\Pi \rho_w R_v}{10^6}\left(\frac{\sigma_3^2}{T_m^2} + \sigma^2 + k_3^2\frac{\sigma_T^2}{T_m^4}\right)^{1/2} \tag{26}$$

By neglecting the small contribution of k_2' to the value of the leading term $\prod$ on the right hand side of Eq. (26), we find

$$\frac{\sigma_\Pi}{\Pi} \approx \left(\frac{\sigma_3^2}{k_3^2} + \frac{\sigma_T^2}{T_m^2}\right)^{1/2} \approx \frac{\sigma_T}{T_m} \tag{27}$$

Eq. (27) approximates Eq. (26) in an accuracy of better than 2%. Although we use Eq. (26) to evaluate the error in $\prod$, we can gain interesting insight into the evolution of this error by examining Eq. (27). In particular we can see that if σ_T is sufficiently large and the term σ_T^2/T_m^2 dominates σ_3^2/k_3^2, then the relative error in $\prod$ will closely approximate the relative error in T_m.

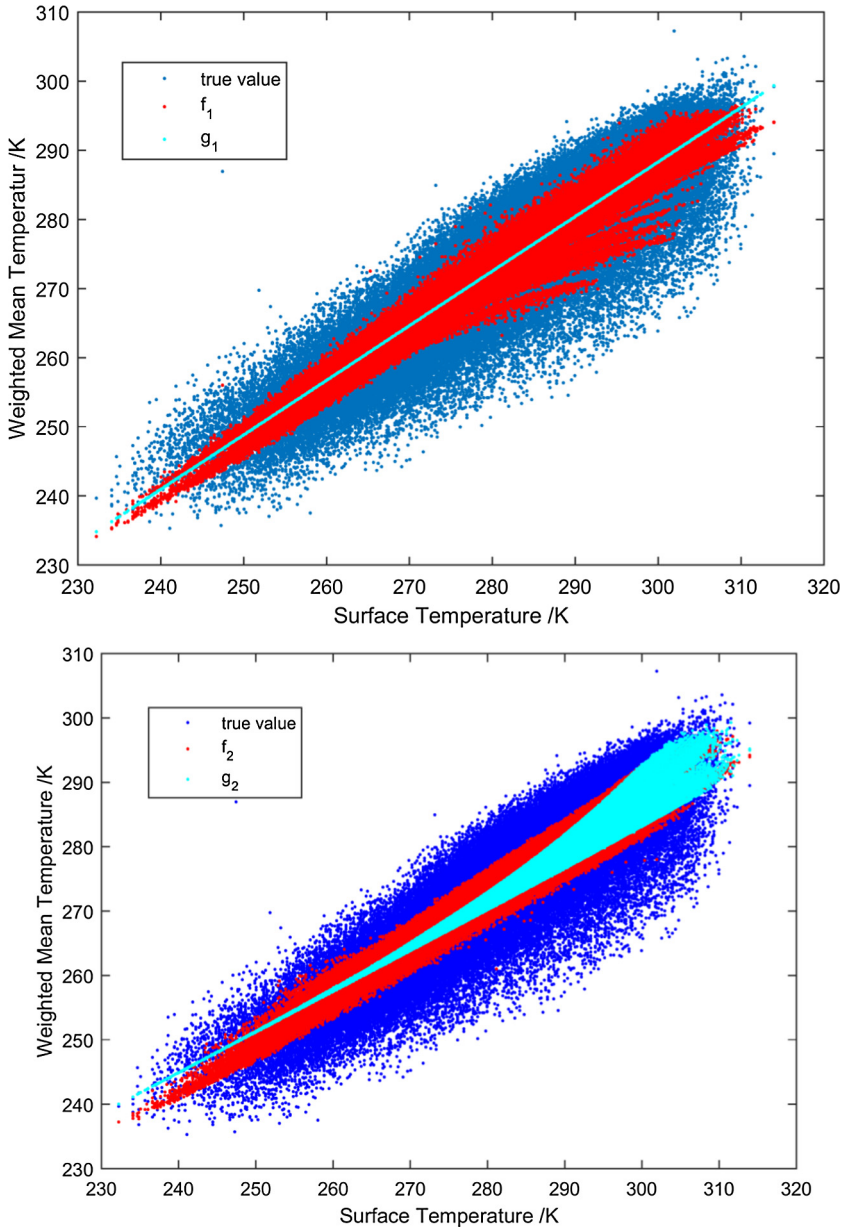

Fig. 5. Fitting performance using f_1, f_2, g_1 and g_2 in China region.

Assuming that the values of ZWD are given without errors, if T_m has less error, $\prod$ will have less error too. This will improve the converting accuracy of ZWD to PWV. Therefore, our new model of the weighted mean temperature will improve the accuracy of PWV indirectly.

In order to further illustrate how much the new model improves the accuracy of PWV, four stations are chosen. Table 2 lists the information of the four stations. First, Eq. (1) is used to calculate the reference ZWD values of the four stations, and then Eq. (4) is used to calculate the reference values of PWV. Finally, the reference PWV values are compared to those computed using f_2 and g_1. The reference PWV and ZWD values and their corresponding computed PWV values are shown in Fig. 6. In this figure, the red colored values using f_2 are closely following the reference values, which suggests f_2 performs better than g_1 when retrieving PWV.

Table 2
The information of the four stations.

Station number	Latitude/°	Longitude/°	Altitude/m
59981	16.83	112.33	5.0
58457	30.23	120.16	43.0
52418	40.15	94.68	1140.0
50557	49.16	125.23	243.0

Table 3 shows that $f_2 and g_1$ perform almost the same at station 58,457 located in the middle latitude. At other three stations, the RMS values of f_2 are better off by 16.3% on average than that of g_1, and the MAD values are reduced by 14.27% on average. This proves the superiority of f_2 over g_1 when retrieving the PWV.

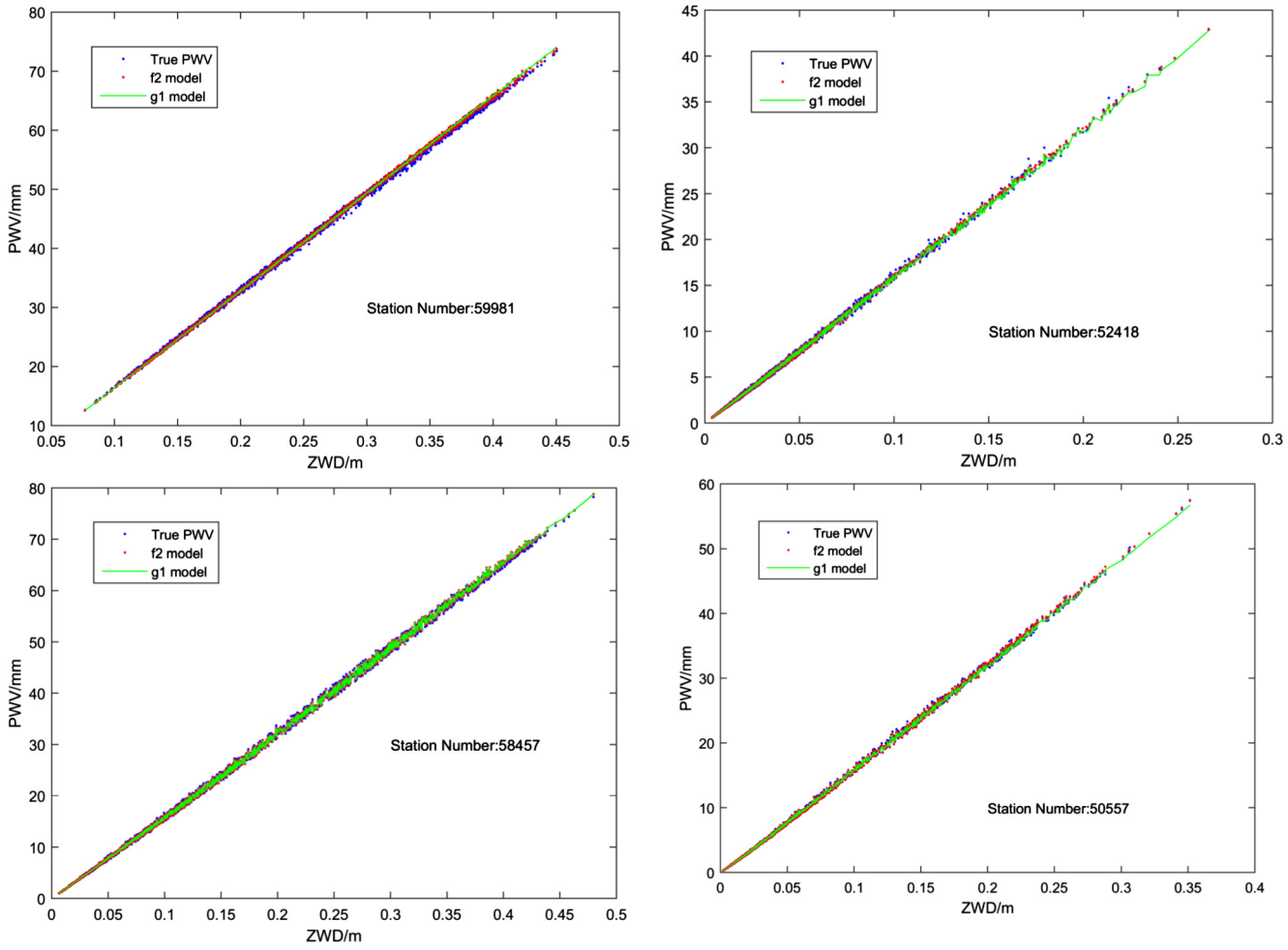

Fig. 6. Application of f_2 and g_1 in calculating PWV at four radiosonde stations.

Table 3
The RMS, MAD of the differences between the computed and the reference PWV values.

Station number	f_2		g_1		Change in %	
	RMS/mm	MAD/mm	RMS/mm	MAD/mm	RMS	MAD
59,981	0.337	0.264	0.372	0.296	10.39	12.12
58,457	0.272	0.204	0.278	0.205	2.21	0.49
52,418	0.140	0.094	0.170	0.110	21.43	17.2
50,557	0.140	0.089	0.164	0.101	17.14	13.48

5. Conclusions

The weighted mean temperature is a key parameter for calculating the precipitable water vapor. Many linear models have been developed with the regression method, which lack of solid theoretical reasoning. Starting from its definition, a new reliable model of the weighted mean temperature considering four influencing factors, δ, h_{trop}, T_s, and h_s, is presented in this paper, and theoretical explanations are provided to reveal the relationship between the weighted mean temperature and the influencing factors. A more rigorous relationship between the weighted mean temperature and the surface temperature is obtained. The new model appears the first of its kind to quantitatively analyze the influencing factors of weighted mean temperature.

Ignoring the effects of temperature lapse rate results in a simpler norm model. The tropopause height can be obtained from the empirical model, and the numbers of the simpler model inputs are further reduced to two, the surface temperature and station height.

When fitting site-specific data, the RMS and MAD values of f_1 are smallest and the values of f_2 are smaller than those of the other two models. Over larger areas the performance of f_2 is even better than that of f_1. Model g_2 is not recommended. Although it performs better in summer time, seasonal changes of water vapor will make it degenerate into model g_1 in other seasons. The RMS and

MAD values of the normal model are 3.8 K and 2.9 K respectively in China region.

Finally, the use of the model f_2 is clearly better than model g_1 in retrieving the PWVs at the four testing radiosonde stations, with the RMS and MAD values reduced by about 16.3% and 14.3%, respectively.

The effect of the water vapor pressure on the weighted mean temperature norm model has not been discussed in this paper, which will be part of the future work. The long term objective of this work is to take more careful consideration of all influencing factors, such as the appropriateness of the temperature being linearly related to the height and the tropopause model, in the construction of more accurate empirical models.

Acknowledgements

The authors would like to express gratitude towards China Meteorological Bureau for providing radiosonde data. This research was supported by the National Key Research and Development Program of China (No. 2016YFB0501803) and the National Natural Science Foundation of China (41574028; 41474024).

Appendix A

A.1. The functional inner product (Zhang, 2004)

If $x(t)$ and $y(t)$ are two function vectors of variable t, their inner product is defined as:

$$\langle x(t), y(t)\rangle = \int_a^b x(t)^T y(t)dt \tag{A.1}$$

where $t \in [a, b]$ and $a < b$.

The intersection angle between the two function vectors is defined as

$$\cos(\varphi) = \frac{\langle x, y\rangle}{\sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}} = \frac{\int_a^b x(t)^T y(t)dt}{\|x(t)\| \cdot \|y(t)\|} \tag{A.2}$$

where $x(t)$ is the norm of $x(t)$, and is given by

$$\|x(t)\| = \sqrt{\int_a^b x(t)^T x(t)dt} \tag{A.3}$$

A.2. Derivation of norm equation of T_m

Let h_s be the height of a ground site, h_{trop} the tropopause height, and $[h_s, h_{trop}]$ the integral interval for the weighted mean temperature. Eq. (A.4) is then the definition of the weighted mean temperature, T:

$$T_m = \frac{\int_{h_s}^{h_{trop}} \frac{e}{T} dz}{\int_{h_s}^{h_{trop}} \frac{e}{T^2} dz} \tag{A.4}$$

The numerator in Eq. (A.4) can be expressed as the inner product of $\frac{e}{T^2}$ and T, and the denominator as the inner product of $\frac{e}{T^2}$ and 1. By use of Eqs. (A.1) and (A.2), the following two equations containing the numerator and denominator are obtained:

$$\cos(\theta_1) = \frac{\langle e/T^2, T\rangle}{\sqrt{\langle e/T^2, e/T^2\rangle}\sqrt{\langle T, T\rangle}} = \frac{\int_{h_s}^{h_{trop}} (e/T^2) \cdot T dz}{\|e/T^2\| \cdot \|T\|} \tag{A.5}$$

$$\cos(\theta_2) = \frac{\langle e/T^2, 1\rangle}{\sqrt{\langle e/T^2, e/T^2\rangle}\sqrt{\langle 1, 1\rangle}} = \frac{\int_{h_s}^{h_{trop}} (e/T^2) \cdot 1 dz}{\|e/T^2\| \cdot \|1\|} \tag{A.6}$$

Denote φ as $\frac{\cos(\theta_1)}{\cos(\theta_2)}$, where φ is determined by temperature and water vapor pressure. Taking Eq. (A.4) into consideration, we can obtain the norm equation of T_m:

$$T_m = \frac{\|T\| \cdot \varphi}{\sqrt{h_{trop} - h_s}} \tag{A.7}$$

References

Añel, J.A., Antuña, J.C., Torre, L.D.L., Nieto, R., Gimeno, L., 2007. Global statistics of multiple tropopauses from the IGRA database. Geophys. Res. Lett. 34 (6), 226–230.

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

Bevis, M., Businger, S., Chiswell, S., Herring, T.A., Anthes, R.A., Rocken, C., 1994. GPS meteorology: mapping zenith wet delays onto precipitable water. J. Appl. Meteorol. 33 (3), 379–386.

Cao, X.G., Ding, J.C., Ye, Q.X., Wang, J., Qiu, L.H., 2007. Method of diagnose the onset of Meiyu GPS/PWV data. J. Appl. Meteorol. Sci. 18 (6), 791–801.

Davis, J.L., Herring, T.A., Shapiro, I.I., Rogers, A.E.E., Elgered, G., 1985. Geodesy by radio interferometry: effects of atmospheric modeling errors on estimates of baseline length. Radiol. Sci. 20 (6), 1593–1607.

Seidel, D.J., Ross, R.J., Angell, J.K., Reid, G.C., 2001. Climatological characteristics of the tropical tropopause as revealed by radiosondes. J. Geophys. Res. Atmos. 106 (D8), 7857–7878.

Heise, S., Shangguan, M., Deng, Z., Dick, G., Wickert, J., Zus, F., 2015. Validation of GPS derived integrated water vapor with microwave radiometer measurements in zenith and satellite tracking mode. EGU Gen. Assembly Conf. 17, 1–9.

Lee, J., Park, J.U., Cho, J., Baek, J., Kim, H.W., 2010. A characteristic analysis of fog using GPS-derived integrated water vapor. Meteorol. Appl. 17 (4), 463–473.

Liu, J.H., 2015. Tropopause heights with its applications in the key parameters of the troposphere models. Wuhan University Press.

Lindenbergh, R., Hanssen, R., Haan, S.D., Marel, H.V.D., Liu, S., 2009. InSAR based validation of MERIS IWV cloud gap filling using GPS IWV. Aerospace Eng.

Mendes, V.B., Prates, G., Santos, L., Langley, R.B., 2000. An evaluation of the accuracy of models for the determination of the weighted mean temperature of the atmosphere. National Technical Meeting of the Institute of Navigation.

Nagurny, A.P., 2003. Climatic characteristics of the tropopause over the arctic basin. Ann. Geophys. 16 (1), 110–115.

Ross, R.J., Rosenfeld, S., 1997. Estimating mean weighted temperature of the atmosphere for global positioning system applications. J. Geophys. Res. Atmos. 102 (D18), 21719–21730.

Sapucci, L.F., 2014. Evaluation of modeling water-vapor-weighted mean tropospheric temperature for GNSS-integrated water vapor estimates in Brazil. J. Appl. Meteorol. Climatol. 53 (3), 715–730.

Schmidt, T., Wickert, J., Beyerle, G., Reigber, C., 2004. Tropical tropopause parameters derived from GPS radio occultation measurements with CHAMP. J. Geophys. Res. 109 (D13), 1097–1111.

Singh, D., Ghosh, J.K., Kashyap, D., 2014. Weighted mean temperature model for extra tropical region of India. J. Atmos. Solar Terr. Phys. 107 (1), 48–53.

Solbrig, P., 2000. Untersuchungen uber die Nutzung numerischer Wettermodellezur Wasserdampfbestimmung mit Hilfe des Global Positioning Systems. Diploma Thesis. Institute of Geodesy and Navigation, University FAF Munich, Germany.

Thayer, G.D., 1974. An improved equation for the radio refractive index of air. Radiol Sci. 9 (10), 803–807.

Wang, J., Zhang, L., Dai, A., 2005. Global estimates of water-vapor-weighted mean temperature of the atmosphere for GPS applications. J. Geophys. Res. Atmos. 110 (D21), 2927–2934.

Xu, X.H., Gao, P., Zhang, X.H., 2013. Structure and variation of the tropopause over China with COSMIC radio occultation bending angles. Chin. J. Geophys. 56 (5), 2531–2541.

Yao, Y.B., Zhu, S., Yue, S.Q., 2012. A globally applicable, season-specific model for estimating the weighted mean temperature of the atmosphere. J. Geodesy 86 (12), 1125–1135.

Yao, Y.B., Liu, J.H., Zhang, B., He, C.Y., 2015. Nonlinear relationships between the surface temperature and the weighted mean temperature. Geomatics & Information Science of Wuhan University 40 (1), 112–116.

Zhang, X.D., 2004. Matrix Analysis and Application. Tsinghua University Press, Beijing.