Atmospheric Controls on Water-Vapor-Weighted Column Mean Temperature (T_m)

N. Z. Wong[1], **L. Feng**[2] **and E. M. Hill**[1,2]

[1]Asian School of the Environment, Nanyang Technological University
[2]Earth Observatory of Singapore, Nanyang Technological University

Key Points:

- T_m is governed by dynamics in the mid-upper troposphere.
- Comparison with other models highlight deficiencies in their estimates of T_m variability
- A new T_m dataset based on the latest ERA5 reanalysis is now available.

Corresponding author: Nathanael Wong, `nathanaelwong@fas.harvard.edu`

Abstract

We used the recently released ERA5 reanalysis data to create a global, gridded dataset (RE5) for the constant of proportionality Π which converts Global Navigation Satellite Systems (GNSS) zenith wet delay signals into precipitable water vapour based on the work of Askne and Nordius (1987). A comparison of this dataset with ERA-Interim (REI) shows that the diurnal variability is more pronounced in the RE5 dataset, and that Π is slightly overestimated in regions of high topography in REI compared to RE5. Comparison with other datasets such as GGOS Atmosphere (RGA) or GPT2w (EG2) also highlight deficiencies in the variability of Π in these models are different scales. We also use our new dataset to show that the linear approximation of T_m based on T_s that was first elaborated upon by Bevis et al. (1992) (EBB) gives rise to significant bias in both the mean value and variability of Π. Lastly, we also perform the calculation of T_m using pressure coordinates instead of vertical coordinates, and find that the error in both the climatological mean and variance that results from this transformation is $< 0.01\%$. Since reanalysis data is given in pressure coordinates as opposed to vertical coordinates, using pressure coordinates in the calculation of T_m provides a much simpler method to calculate Π as opposed to integrating in vertical coordinates to find T_m.

Plain Language Summary

Text

1 Introduction

The high spatial and temporal variability of atmospheric water vapor makes it one of the more difficult parameters to monitor in the atmosphere. However, the use of Global Navigation Satellite System (GNSS) tropospheric-delay data can help to improve efforts in monitoring atmospheric water vapor. The conversion of GNSS signals to precipitable water depends on the water-vapor-weighted mean column temperature T_m, which can be estimated or calculated through various means. Previous studies of T_m largely focus on the validation of empirical T_m models (e.g., GPT2w, GTm-III) against radiosonde-derived values or the GGOS Atmosphere model, but there has been little investigation into the large-scale spatial pattern of T_m and its variability. In this study, we calculate the global-gridded T_m using ERA-Interim and ERA5 reanalysis data, investigate its spatial pattern and variability, and compare our results with T_m from the GGOS Atmosphere and GPT2w models. We also test the viability of using linear approximation methods such as those of Bevis et al. [1992].

2 Methodology

The relationship between zenith wet delay (ZWD) and precipitable water vapour (PWV) was first established by Askne and Nordius (1987) as follows:

$$\mathrm{PWV} = \Pi \cdot \mathrm{ZWD}; \Pi = \frac{10^6}{\rho R_v \left(k_2' + \frac{k_3}{T_m}\right)} \tag{1}$$

where Π is the dimensionless constant of proportionality, ρ is the density of water, R_v is the specific gas constant for water vapour k_2' and k_3 are refractivity constants, and T_m is the water-vapour-weighted column air temperature. All variables except for T_m are assumed to be constant, such that Π is dependent only on T_m.

T_m is the water-vapour-weighted mean column temperature. Therefore, it is calculated by integrating the column air temperature $T(z)$ from the top-of-atmosphere to the surface, with the vapour-pressure $e(z)$ as the weight as shown in the formula by (Davis et al., 1985):

$$T_m = \frac{\int_0^\infty \frac{e(z)}{T(z)}\,\mathrm{d}z}{\int_0^\infty \frac{e(z)}{T^2(z)}\,\mathrm{d}z} = \frac{\sum_i \dfrac{e_i}{T_i}\Delta h_i}{\sum_i \dfrac{e_i}{T_i^2}\Delta h_i} \tag{2}$$

To obtain an accurate estimation of T_m, a vertical profile of both temperature and humidity is needed. However, such vertical profiles have historically been made available only through use of radiosonde. Areas where radiosonde measurements have been made are sparse due to the expense, and even then are often conducted once or twice daily only, and therefore many different methods have been developed to estimate T_m. These methods are listed in Table 1 and elaborated upon in the rest of Section 2.

Name	Classification	Output:	Data Source / Input	Section	Reference
RE5	reanalysis	T_m	ERA5 $(T, T_s, T_d, q, z, \Phi)$	2.1	Hersbach and Dee (2016)
REP			ERA5 (T, T_s, T_d, q, p_s)	2.1.1	
REI			ERA-Interim (T, T_s, T_d, q, p_s)	2.1.1	Dee et al. (2011)
RGA			GGOS Atmosphere	2.1.2	Böhm and Schuh (2013)
EBB	empirical	T_m	ERA5 T_s; (a, b) Bevis et al. (1992)	2.2.1	Bevis et al. (1992)
EBM			ERA5 T_s; (a, b) Manandhar et al. (2017)		Manandhar et al. (2017)
EG2			day-of-year; (lon,lat)	2.2.2	Böhm et al. (2015)
EMN		Π	day-of-year; (lat,height)		Manandhar et al. (2017)

Table 1. A summary of the methods used to calculate T_m, and therefore Π.

We convert T_m to Π using Eqn 1, and this is relevant to all datasets except EMN, which returns Π directly instead. We take RE5 dataset, derived using ERA5 reanalysis output at each time of year and at each point, to be the reference dataset against which all other Π method estimates were compared and validated against.

In this paper, we also aim to investigate the diurnal $\Delta_d(\cdot)$, intraseasonal $\Delta_i(\cdot)$, seasonal $\Delta_s(\cdot)$ and interannual $\Delta_a(\cdot)$ mean-weighted variability (Eqn. 3) of Π. Based on Eqn. 4 (X. Wang et al., 2016), we see that $\Delta\Pi \approx \Delta T_m$. Therefore, the variability we see in Π is due to fluctuations either in air temperature T or water vapor content e. We also compare the temporal variability at different scales across different datasets to evaluate the strengths and weaknesses of each dataset compared to RE5.

$$\Delta(\cdot) = \frac{\delta(\cdot)}{\overline{(\cdot)}}, \text{ where } \overline{(\cdot)} \text{ denotes the time-averaged value of } (\cdot) \tag{3}$$

$$\Delta\mathrm{PWV} = \Delta\Pi + \Delta\mathrm{ZWD} \approx \Delta T_m + \Delta\mathrm{ZWD} \tag{4}$$

2.1 Calculating T_m using reanalysis data

Although the best way to derive T_m is to use in-situ measurements of water vapour pressure and temperature from vertical profiles, generally derived through radiosonde, the density of radiosonde measurements is sparse in many regions around the globe. Therefore, J. Wang et al. (2005) championed to use of reanalysis data (e.g. ERA-40) to estimate water vapour pressure and temperature at different pressure heights to obtain T_m. X. Wang et al. (2016) used ERA-Interim reanalysis data to calculate T_m and found that the difference between reanalysis T_m and the "ground-truth" calculated using in-situ radiosonde data was 0.5% on average. In this study, we use both ERA5 (Hersbach & Dee, 2016) and ERA-Interim (Dee et al., 2011) reanalysis data at 1.0° resolution in both longitude and latitude.

The following variables are required to calculate T_m in vertical coordinates:

- Pressure Level: geopotential Φ, air temperature T, specific humidity q
- Surface Level: surface geopotential z, surface temperature T_s, dewpoint temperature T_d

Vapor pressure e can be calculated from either specific humidity q (Eqn. 5) or dewpoint temperature T_d (Eqn. 6).

$$e = p \cdot \frac{q}{q + (1-q)\varepsilon} \tag{5}$$

$$e = e_0 \exp\left[\frac{L}{R_v}\left(\frac{1}{T_0} - \frac{1}{T_d}\right)\right] \tag{6}$$

where p is the pressure, $\varepsilon = R_d/R_v$ is the ratio of the specific gas constants of the dry atmosphere and water vapour, e_0 and L are the saturation vapor pressure and enthalpy of vaporization at a reference temperature T_0.

In the absence of a continuous profile, we show in Eqn. 2 that the integral can be rewritten into the form of a discretized summation. In order to calculate T_m, we therefore calculate the column temperature from the top of the atmosphere down to the different individual pressure levels of ERA-Interim for each grid-point. From there, at each grid-point, we convert the coordinates from pressure to the respective geopotential level, and thus obtain T_m for 37 different geopotential heights. It therefore remains to linearly interpolate T_m to the surface geopotential for each grid-point.

We note that coastal regions and the ocean surface are often found beneath the 1000 hPa pressure-level height, which is the lowest pressure level in the reanalysis models. Therefore, a method of extrapolation to beneath this pressure-level height is needed, while providing constraints for reasonable values of T_m. Direct extrapolation can produce unreasonable oscillations, so we therefore add another pressure level at 1012.35 hPa, which is assumed to be the average pressure at sea-level, and calculate the geopotential height of this level using the hydrostatic balance. At this level, we use Eqn. 6 to calculate the surface vapor pressure e_s using the surface dewpoint temperature T_d, and the surface temperature. Again, we find T_m by integrating to the surface geopotential.

2.1.1 Integration using pressure coordinates

As mentioned above, ECMWF reanalysis model output is given at 37 different pressure levels. Due to hydrostatic balance in the atmosphere and the relatively slow nature of vertical transport as opposed to horizontal flow it can be assumed that pressure is monotonically increasing downwards. Therefore, instead of retrieving the geopotential height of each pressure level at every different space-time coordinates and integrating in vertical coordinates, we integrate in pressure coordinates that are fixed in time and space. This reduces the amount of data needed to be downloaded and extracted in order to calculate T_m and therefore saves time and computational cost.

Assuming that R and g are constant throughout the atmosphere, the transformation of coordinates from vertical height to pressure can be performed using Eqn. 7-8:

$$\begin{aligned}
&\frac{\mathrm{d}p}{\mathrm{d}z} = -\rho g \quad \text{(By hydrostatic balance)} \\
&\therefore \mathrm{d}p = -\rho g \,\mathrm{d}z = -\frac{p}{RT} g \,\mathrm{d}z \\
&\therefore \mathrm{d}z = -\frac{T}{p}\frac{R}{g}\,\mathrm{d}p
\end{aligned} \tag{7}$$

$$\therefore T_m = \frac{\sum_i \frac{e_i}{T_i}\Delta z_i}{\sum_i \frac{e_i}{T_i^2}\Delta z_i} = \frac{-\sum_i \frac{e_i}{T_i}\frac{T_i}{p_i}\frac{R}{g}\Delta p_i}{-\sum_i \frac{e_i}{T_i^2}\frac{T_i}{p_i}\frac{R}{g}\Delta p_i} = \frac{\sum_i \frac{e_i}{p_i}\Delta p_i}{\sum_i \frac{e_i}{p_i T_i}\Delta p_i} \tag{8}$$

2.1.2 The GGOS Atmospheres Model

The GGOS Atmosphere model (Böhm & Schuh, 2013) is also similarly based off reanalysis data, specifically ERA-Interim data. The T_m output from the GGOS Atmosphere model has been used to create empirical models such as GTm-III (Yao et al., 2014) and has itself been used as ground truth in various studies (e.g. Lan et al. (2016); Liu et al. (2015)). We evaluate our method above against the GGOS Atmosphere T_m values from 1979-2018 by comparing the RGA dataset against the REI dataset, as REI also uses ERA-Interim data to calculate T_m. However, we are unable to do a direct comparison of our methodology with GGOS Atmosphere, because the method that Böhm and Schuh (2013) used to calculate T_m was not explicitly given.

The resolution of the RGA dataset is 2.5° longitude by 2.0° latitude. In order to

2.2 Approximating T_m using empirical relationships

2.2.1 Empirical models based on surface temperature

Because vertical profiles of the atmosphere are sparse compared to surface measurements, Bevis et al. (1992) used a linear method to estimate T_m from surface temperature T_s. From over 8000 radiosonde profiles in the United States, he derived a linear relationship between T_m and T_s:

$$T_m = a + b \cdot T_s \tag{9}$$

where $a = 70.2$ and $b = 0.72$. However, it has long been recognised that these coefficients are highly dependant on location and season, and therefore there have been studies done to estimate a and b on regional scales for better estimates of T_m. Manandhar et al. (2017) derived estimates for a and b based on three different latitude categories: tropical, subtropical and temperature, which we also used to see if there was any significant difference when adjusting the linear relationship based on region. As it has been previously found that reanalysis models are able to model surface temperature with a high degree of accuracy, we use ERA5 surface temperature data here to estimate T_m via the linear method of Eqn. 9 using the coefficients of both Bevis et al. (1992) and Manandhar et al. (2017).

2.2.2 Empirical models based on location and time

Unlike the methods described in Section 2.1 and 2.2.1 that are used to find T_m, empirical models such as GPT2w (Böhm et al., 2015), GTm-III and Manandhar et al. (2017) (hereafter referred to as MN2017) do not require any meteorological or climatological variables as input to calculate T_m. Instead, these models require as input only time/day-of-year and (x, y, z) positions. These empirical models are of course based on T_m values either directly calculated from reanalysis data (e.g. GPT2w), or from other models and approximations (e.g. GTm-III, MN2017). In our study, we look at the GPT2w and MN2017 models and compare the mean values and variability of T_m in these models compared to T_m values calculated directly from reanalysis. As GPT2w is ultimately derived from ERA-Interim reanalysis data, it would be more appropriate to compare Π_{EG2} to Π_{REI} instead of Π_{RE5} in order to determine the veracity of the GPT2w model creation.

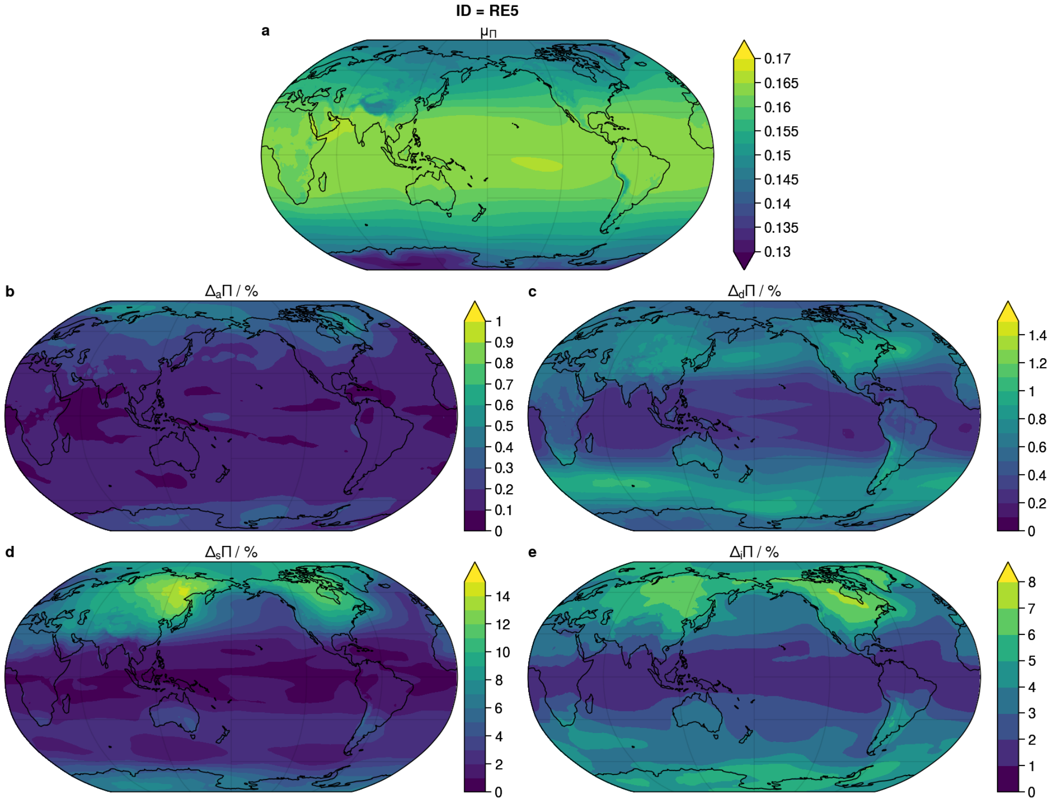

Figure 1. The spatial distribution of the (a) mean value, (b) mean-weighted trend, (c) diurnal variability, (d) seasonal variability and (e) interannual variability of Π_{AD}

3 Results

3.1 Validation of our calculations

From comparing our results with Π_{RGA}, we find that the T_m values derived from GGOS Atmosphere, which forms the basis upon which the GPT2w and GTm-III models are built upon, are very similar to that of Π_{REI}. However, we also find that Π_{RGA} shows relatively high $\Delta_a\Pi$ over areas of high topography, as compared to $\Delta_a\Pi$ from Π_{AD}. This is indicative of differences in the method extrapolation of T_m values to the surface, though the exact cause has yet to be determined. Nonetheless, the overall similarity of the spatial pattern of Π_{REI} and Π_{RGA} indicate the robustness of using reanalysis data to calculate T_m as different methodologies are able to provide similar results.

Π_{RE5} is taken to be the most accurate estimation of Π out of all the other methodologies, and therefore unless otherwise specified, is taken to be the reference against which all other Π estimates were validated against. Therefore, we use Π_{RE5} in our analysis to estimate the spatial and temporal variability of Π (Fig. 1).

3.2 Variability of Π

From Fig. 1, we see that the variability of Π_{AD} is relatively small at the diurnal and interannual scale at $< 1.5\%$ and $< 0.5\%$ in most area respectively, with $\Delta_a\Pi$ increasing polewards. $\Delta_s\Pi$ is also relatively small in the tropics, hovering at $\sim 10\%$ at around 30° latitude over land and $\sim 6\%$ over oceans, but also increases polewards. However, $\Delta_s\Pi$ in the northern hemisphere has a maximum not at the pole, but over eastern Siberia and Hudson Bay, likely reflecting the variability in the northern Polar vortex, while in the southern hemisphere $\Delta_s\Pi$ increases roughly monotonically with latitude, indicating a stronger and more consistent southern polar vortex. The zone of extremely high $\Delta_s\Pi$ along the coast of Antartica is likely due to interpolation errors as a result of complex orography. We also see that $\Delta_s\Pi$ is generally higher over land than over the ocean at equivalent latitudes.

Meanwhile, we see that the magnitude of the diurnal variability of Π ($\Delta_d\Pi$) seems to be dominated by processes near the upper troposphere such as the mid-latitude jet-stream, with maximum diurnal variability coinciding with regions where the mid-latitude jet-stream is strongest. The jet-stream is known to modulate diurnal weather, but the exact mechanisms of this signal remain as of yet unexplored. The amplitude of the interannual variability $\Delta_a\Pi$ also seems to correlate well with the amplitude of the trend of Π_{AD} over the years. The trend is increasing over most of the globe, while decreasing over the south pole and central Africa and the Andes, but as it is $< 0.05\%$ of μ_Π, the overall change due to the trend can be considered to be negligible around the globe. We also see that

3.3 Comparing Π_{AD} against Linear Estimation Methods (Π_{ABB} and Π_{ABM})

As mentioned previously in Section 2.2, T_m is often estimated via a linear model relating T_m to T_s (see Eqn 7), first proposed by Bevis et al. (1992). Therefore, we compare our results (Π_{AD}) with two different linear models, Π_{ABB} and Π_{ABM}. Π_{ABB} uses the original coefficients found by Bevis et al. (1992), while Π_{ABM} uses updated coefficients based on Manandhar et al. (2017) for three different latitudinal boundaries for more accurate estimation.

Comparison of μ_Π for Π_{ABB} and Π_{ABM} against Π_{AD} indicates that the linear approximation method slightly underestimates μ_Π in the tropics ($< 30°$), and overestimates it everywhere else. Regions of high topography tend to slightly overestimate μ_Π as well (i.e. the Tibetian Plateau and the Andes), more so for Π_{ABM} than for Π_{ABB}. Furthermore, we see that both linear methods overestimate $\Delta_d\Pi$ over land and underestimate it over the oceans, likely due to the fact that surface temperature variability over land is much greater than that of the atmospheric column above it. This overestimation of diurnal variability ($\Delta_d\Pi$) is also perhaps a factor that accounts for the global underestimation of seasonal variability ($\Delta_s\Pi$), except over Antartica.

Interestingly, Π_{ABB} and Π_{ABM} shows a strong ENSO pattern in $\Delta_a\Pi$ that reflects the corresponding interannual variability in the sea surface temperature. This signal is not as evident in $\Delta_a\Pi$ for Π_{AD}. This is further highlights the influence of the upper troposphere and its importance in modulating the column air temperature.

3.4 Comparison against Empirical Models (e.g. GPT2w)

Lastly, we also compare our results against values of Π derived from blind empirical models such as GPT2w (Böhm et al., 2015) or the empirical formula of Manandhar et al. (2017). We note that these empirical models do not display interannual or diurnal variability, and therefore we only compare μ_Π and $\Delta_s\Pi$.

The GPT2w model is based on ERA-Interim data from 2001 to 2010, and therefore it is no surprise that μ_Π for Π_{AG2} shows very close agreement to Π_{AD}. However, what is surprising is that though the spatial pattern of $\Delta_s\Pi$ for the GPT2w model is very similar to our results, with $\Delta_s\Pi$ reaching maximum at the locations near the polar vortex, the magnitude of $\Delta_s\Pi$ is also increasingly underestimated at high latitudes and especially in North America by over 5%. Comparison of Π_{AD} values of mean and variability calculated from 2001-2010, which are the values used to estimate the GPT2w model respectively, show very similar values of seasonality to the whole 40 years worth of ERA-

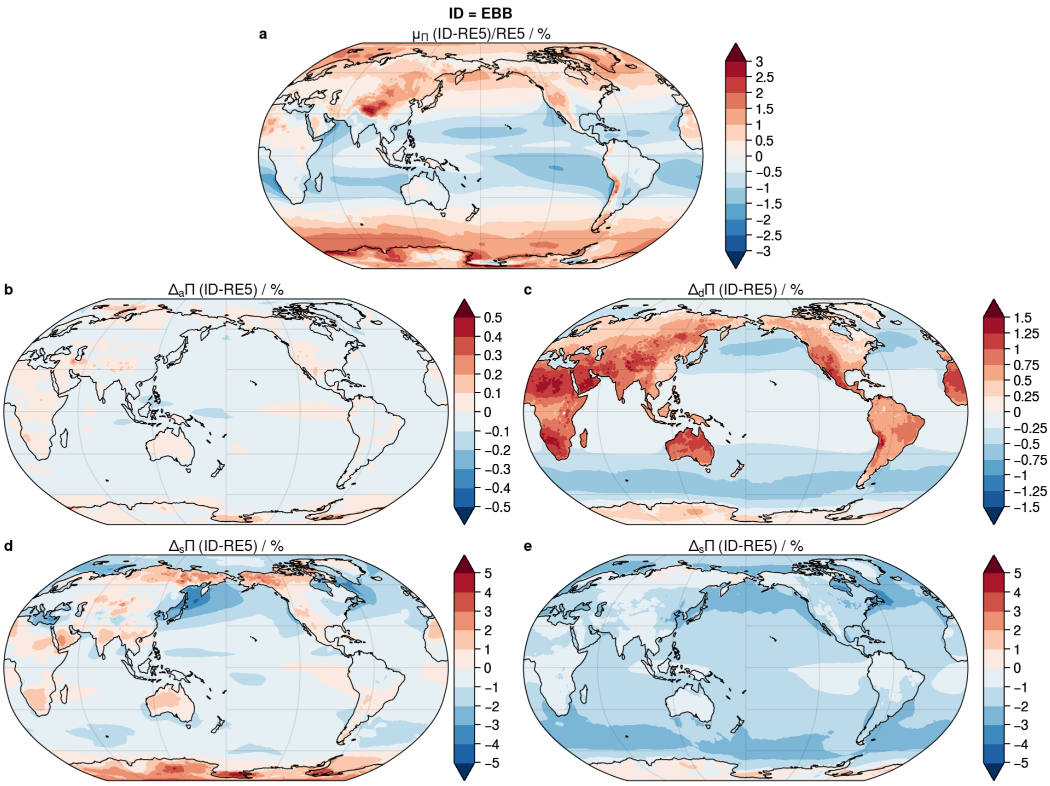

Figure 2. The spatial distribution of the difference in (a) mean value, (b) mean-weighted trend, (c) diurnal variability, (d) seasonal variability and (e) interannual variability between Π_{ABB} and Π_{AD}.

Interim data available. Therefore, the inability of the GPT2w model to model the seasonal variability is not due to sampling error in the years chosen to adjust the model.s

Acknowledgments

Enter acknowledgments, including your data availability statement, here.

References

Askne, J., & Nordius, H. (1987, 5). Estimation of tropospheric delay for microwaves from surface weather data. *Radio Science*, *22*(3), 379–386. Retrieved from `http://doi.wiley.com/10.1029/RS022i003p00379` doi: 10.1029/RS022i003p00379

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. *Journal of Geophysical Research*, *97*(D14), 15787–15801. Retrieved from `http://doi.wiley.com/10.1029/92JD01517` doi: 10.1029/92JD01517

Böhm, J., Möller, G., Schindelegger, M., Pain, G., & Weber, R. (2015, 7). Development of an improved empirical model for slant delays in the troposphere (GPT2w). *GPS Solutions*, *19*(3), 433–441. Retrieved from `http://link.springer.com/10.1007/s10291-014-0403-7` doi: 10.1007/

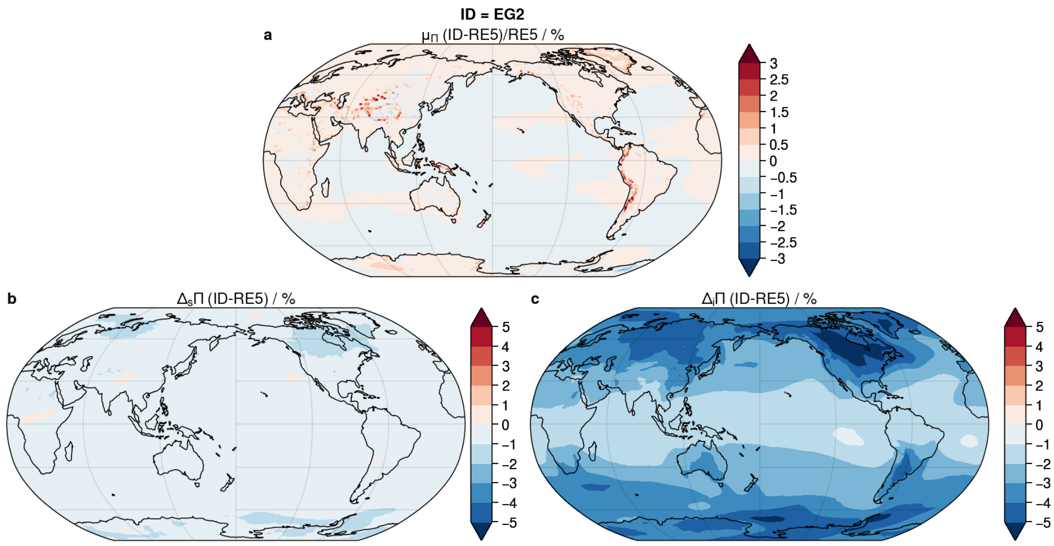

Figure 3. The spatial distribution of the difference in (a) mean value, (b) mean-weighted trend, (c) diurnal variability, (d) seasonal variability and (e) interannual variability between Π_{AG2} and Π_{AD}. (b),(c) and (e) are left blank because GPT2w is a blind model that only accounts for the seasonal variability and therefore there is no diurnal component or interannual variability.

s10291-014-0403-7

Böhm, J., & Schuh, H. (Eds.). (2013). *Atmospheric Effects in Space Geodesy.* Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from `http://link.springer.com/10.1007/978-3-642-36932-2` doi: 10.1007/978-3-642-36932-2

Davis, J. L., Herring, T. A., Shapiro, I. I., Rogers, A. E. E., & Elgered, G. (1985, 11). Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length. *Radio Science*, *20*(6), 1593–1607. Retrieved from `http://doi.wiley.com/10.1029/RS020i006p01593` doi: 10.1029/RS020i006p01593

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., ... Vitart, F. (2011, 4). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. Retrieved from `http://doi.wiley.com/10.1002/qj.828` doi: 10.1002/qj.828

Hersbach, H., & Dee, D. (2016). ERA5 reanalysis is in production. *ECMWF Newsletter*(147), 7. Retrieved from `https://confluence.ecmwf.int/pages/viewpage.action?pageId=74764925`

Lan, Z., Zhang, B., & Geng, Y. (2016, 3). Establishment and analysis of global gridded Tm Ts relationship model. *Geodesy and Geodynamics*, *7*(2), 101–107. Retrieved from `https://linkinghub.elsevier.com/retrieve/pii/S1674984716300015` doi: 10.1016/j.geog.2016.02.001

Liu, L., Li, J., Chen, X., & Cai, C. (2015, 12). Precision analysis on the weighted mean temperature of the atmosphere grid data offered by GGOS Atmosphere in Xinjiang. In G. Zhou & C. Kang (Eds.), (p. 98083I). Retrieved from `http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2207379` doi: 10.1117/12.2207379

Manandhar, S., Lee, Y. H., Meng, Y. S., & Ong, J. T. (2017, 11). A Simplified Model for the Retrieval of Precipitable Water Vapor From GPS Signal. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(11), 6245–6253. Retrieved from `http://ieeexplore.ieee.org/document/7994650/` doi: 10.1109/TGRS.2017.2723625

Wang, J., Zhang, L., & Dai, A. (2005). Global estimates of water-vapor-weighted mean temperature of the atmosphere for GPS applications. *Journal of Geophysical Research*, *110*(D21), D21101. Retrieved from `http://doi.wiley.com/10.1029/2005JD006215` doi: 10.1029/2005JD006215

Wang, X., Zhang, K., Wu, S., Fan, S., & Cheng, Y. (2016, 1). Water vapor-weighted mean temperature and its impact on the determination of precipitable water vapor and its linear trend. *Journal of Geophysical Research: Atmospheres*, *121*(2), 833–852. Retrieved from `http://doi.wiley.com/10.1002/2015JD024181` doi: 10.1002/2015JD024181

Yao, Y., Xu, C., Zhang, B., & Cao, N. (2014, 4). GTm-III: a new global empirical model for mapping zenith wet delays onto precipitable water vapour. *Geophysical Journal International*, *197*(1), 202–212. Retrieved from `http://academic.oup.com/gji/article/197/1/202/690874/GTmIII-a-new-global-empirical-model-for-mapping` doi: 10.1093/gji/ggu008