Reliability Analysis of European ERA5 Water Vapor Content Based on Ground-based GPS in China

Hongzhen Yang[*], Chen He, Zhiqiang Wang and Weiping Shao
National Grid Zhejiang Electric Power Corporation, China
[*]Corresponding author

Abstract—On the basis of GPS data at stations of different climatic types in China, precipitable water vapor (PWV) products in year 2016 were obtained through precision data processing and water vapor inversion. On the basis of these products, the reliability of accuracy, annual variation, and diurnal variation of water vapor content for the first hourly resolution global reanalysis data ERA5 was systematically evaluated in China and compared with the previous generation of ERA-Interim analysis data. Results show that the annual and diurnal variation characteristics of PWV in different regions are mainly affected by monsoon and sunlight, respectively. The PWV accuracy provided by ERA5 is better than that of ERA-Interim. The two kinds of reanalysis data adequately reflect the annual variation characteristics of PWV, and the diurnal variation deviation of ERA5 is smaller than that of ERA-Interim in general, especially in the Qinghai–Tibet Plateau area.

Keywords—GPS; ERA5; Precipitable Water Content (PWV); annual variation; diurnal variation

I. INTRODUCTION

Water vapor, an abundant greenhouse gas in the atmosphere, has a strong positive feedback mechanism for atmospheric warming, thereby amplifying the warming effect caused by other greenhouse gases (Held and Soden, 2000). According to the Clausius–Clapeyron equation, atmospheric temperature increases by approximately 7% with every one-degree increase in atmospheric temperature, and increase in water vapor as a greenhouse gas further enhances the warming effect of the atmosphere. The specific heat capacity of water vapor is a principal factor in the formation and evolution of many disastrous weather events and has an important influence on the Earth's atmospheric energy system (Trenberth et al., 2005). Moreover, atmospheric delay due to atmospheric water vapor is the main source of error in space geodetic methods, such as GNSS, VLBI, and InSAR. Therefore, the distribution of and variation in atmospheric water vapor content have remarkable influences on long-term climate research, short-term weather forecasting, and spatial geodetic data processing.

However, different from other greenhouse gases, such as carbon dioxide, atmospheric water vapor exhibits complex changes in time and space, and thus measuring and modeling it is difficult. At present, the primary observation methods used for determining atmospheric water vapor content are radio sounding, ground-based microwave radiometry, satellite-based radiometry, and ground-based GPS. Of these methods, ground-based GPS has the advantages of high precision, all-weather, low cost, good uniformity, and high time sampling rate (Bevis et al., 1992) and has thus been widely used to calibrate water vapor observation mean system errors since its introduction in the 1990s (Wang and Zhang, 2008). Furthermore, it has greatly contributed to the improvement of numerical weather forecast reliability (Poli et al., 2007) and climate analysis (Nilsson and Elgered, 2008).

Meteorological reanalysis data is a multivariate, spatially complete, and continuous global atmospheric circulation data obtained by assimilating meteorological observation data. This type of data provides an important source of information on atmospheric circulation laws and climate change mechanisms for the credible assessment of water cycle and energy balance. Meteorological reanalysis data have greatly contributed to the development of modern atmospheric science (Zhao et al., 2010). At present, many water vapor studies are available on meteorological reanalysis data (e.g. Cai et al., 2004; Dai and Yang, 2009). However, meteorological reanalysis data, as a product of numerical prediction and observation data fusion, contain errors due to changes in numerical models, assimilation schemes, and observation systems and by the systematic deviation of assimilation data sources. Therefore, the distribution of variables obtained by meteorological reanalysis data and the reliability of the variation characteristics of each time scale are a matter of great concern in the field of meteorology.

ERA5 is the first international hourly resolution of global reanalysis data, which was recently launched by the European Center for Medium-Range Weather Forecasts (ECMWF). The dataset, funded by the European Union and developed by the Copernicus Climate Change Service center, which is operated by ECMWF, uses updated observational data and improved physical processes compared to previous re-analysis data, including ERA-Interim. The dataset is the first to offer global atmospheric variables with a flat resolution of up to 30 km and a time resolution of up to 1 hour. Previous studies have assessed the reliability of multi-type reanalysis data products (Zhang et al., 2018). However, the accurate assessment of PWV in ERA5 and reliability at multiple time scales are unavailable at present. In this study, PWV reliability provided in ERA5 was systematically evaluated with ground-based GPS data of different climatic types in China, and then compared with that of the previous generation ERA-Interim.

II. DATA AND PROCESSING METHODS

A. Data and Processing Methods

GPS data solution and water vapor inversion are conducted in five GPS stations (AHAQ, BJFS, LHAS, QION, and URU2), which have different climatic types, in the Crustal Movement Observation Network of China (CMONOC). Table 1 shows the climate type in the station locations.

TABLE I. LOCATION OF GPS STATION AND CORRESPONDING CLIMATE TYPE

GPS test station name	Location of test station	Climate type
QION	Hainan, Qiongzhou	Tropical monsoon climate
AHAQ	Anqing, Anhui	Subtropical monsoon climate
BJFS	Fangshan, Beijing	Temperate monsoon climate
URU2	Urumqi, Xinjiang	Temperate continental climate
LHAS	Lhasa, Tibet	Plateau alpine climate

The GPS data are processed throughout 2016 (some data missing in April) and at sampling intervals of 30 seconds with PANDA software, which was developed by Wuhan University (Shi et al., 2008). The GPS data are used for daily precise point positioning (PPP) processing. Table 2 shows the strategy for data processing.

The accurate valuation of zenith tropospheric total delay (*ZTD*) is obtained by processing the GPS data, then zenith wet delay (*ZWD*) is separated using the following equation:

$$ZWD = ZTD - ZHD \qquad ,(1)$$

where zenith hydrostatic delay (*ZHD*) can be calculated by using the Saastamoinen model based on the air pressure data that are accurately measured at the station, and ZWD can be converted to *PWV* by using the following formula (Bevis et al., 1992):

$$PWV = \prod \cdot ZWD \qquad ,(2)$$

where the conversion factor $\prod$ can be represented as

$$\prod = \frac{10^6}{\rho_w R_v [(k_3 / T_m) + k_2']} \qquad ,(3)$$

where ρ_w is the liquid water density, R_v indicates the specific heat capacity of water vapor, k_2' and k_3 are atmospheric refraction constants $(17 \pm 10) K\ mbar^{-1}$ $(3.776 \pm 0.004) \times 10^5 K^2 \cdot mbar^{-1}$, and T_m is the weighted average temperature of the atmosphere and defined by the following formula:

$$T_m = \frac{\int (P_v / T) dz}{\int (P_v / T^2) dz} \qquad ,(4)$$

where P_v is the atmospheric water vapor pressure (mbar), T is the atmospheric temperature (K), and z represents the elevation.

T_m can be estimated using Bevis linear equations, but the deviation of the equation in China is considerable (Zhang et al., 2017). To obtain enhanced accuracy, we use ERA5 pressure-level products. Then, T_m is calculated by using the Equation (4) integral.

TABLE II. STRATEGY FOR GPS DATA PROCESSING

Observation Model	Observational value	GPS satellite LC, PC combination observational value
	Height cutoff angle	7°
	Weighting strategy	$p = 1, e > 30°$; $p = sin^2 e, e \le 30°$
Error correction	Phase windup	Correction
	Phase center correction	IGS_08 model
	Atmospheric load	Not considered
	Tide correction	Solid and ocean tides
	Relativity correction	Correction
Parameter estimation	Satellite orbit	Fixed (IGS final product)
	Satellite clock	Fixed (IGS final product)
	ZTD Model	Piece-wise constant (5 minutes) Random walk between segments (20 mm/√h)
	Horizontal gradient model	Piece-wise constant (12 hours) Random walk between segments (5 mm/√h)
	Mapping function	GPT2
	Station coordinate	Daily constant
	Receiver clock	White noise

B. ERA5 Product and Processing Methods

In this study, the ERA5 pressure-level product, which contains temperature, specific humidity, and potential high variables on each air pressure layer, is used. The product level resolution is 0.25° and time resolution is 1 hour, and elevation is divided into 37 layers (to 0.1 hPa on the top layer). As the grid point of the reanalysis data is inconsistent with the GPS station data, interpolating or extrapolating the products at the grid point is necessary to the calculation of PWV at the GPS station, and the specific algorithm described by Zhang et al. (2017) is used. The air pressure layer variables in the GPS test station are obtained, then PWV is obtained by integrating the parameters of different air pressure layers with the following formula:

$$PWV = \frac{1}{\rho_W} \int \frac{q}{g} dP \qquad ,(5)$$

where q indicates specific humidity, g is the gravitational acceleration corresponding to the air pressure layer, and P is the air pressure.

C. ERA-Interim Product and Processing Methods

To compare with the latest generation of ERA5, the processing method used for ERA5 is used for the calculation of the PWV of the GPS test station in the previous generation of meteorological reanalysis data ERA-Interim. The ERA-Interim pneumatic layer product also contains 37 layers with a horizontal resolution of 0.75° and a time resolution of 6 h.

The reliability of ERA5 product is evaluated by using PWV based on ERA-Interim product. A total of 6 h of resolution for time linear interpolation is obtained, along with ERA-Interim PWV products with 1 h resolution.

III. RESULTS AND DISCUSSION

A. PWV Comparison

Figures 1–5 show the comparison results of PWV at station QION, AHAQ, BJFS, URU2, and LHAS. The left subdiagram in each figure presents the PWV time series diagram of GPS, ERA5, or ERAI. The intermediate subdiagram illustrates PWV correlation analysis between GPS and ERA5. The correlation fitting formula, error in fitting of residual unit weight (sig.), and correlation coefficient (corr.) are obtained. The right subdiagram shows PWV correlation analysis between GPS and ERAI.
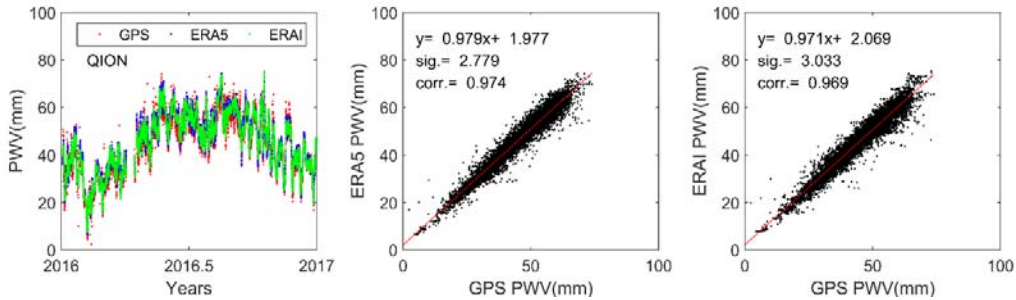

FIGURE I. PWV COMPARISON AT QION STATION (TROPICAL MONSOON CLIMATE TYPE)
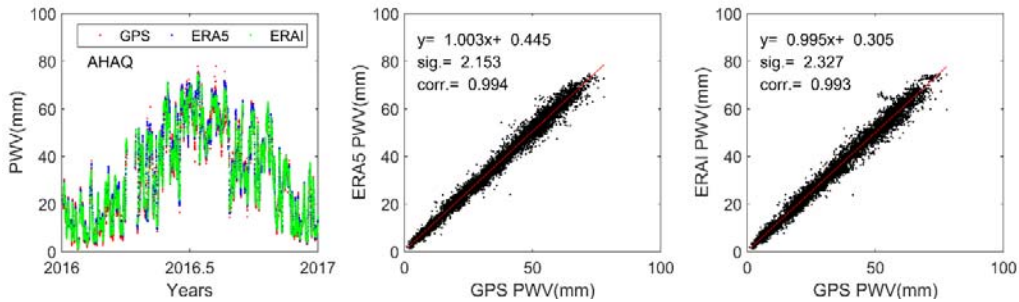

FIGURE II. PWV COMPARISON AT AHAQ STATION (SUBTROPICAL MONSOON CLIMATE TYPE)
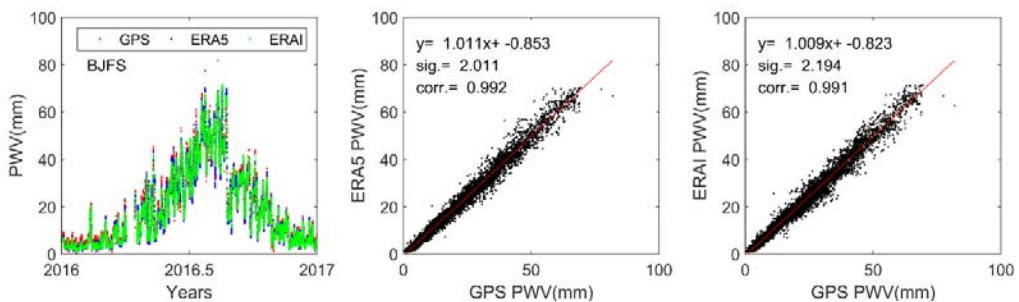

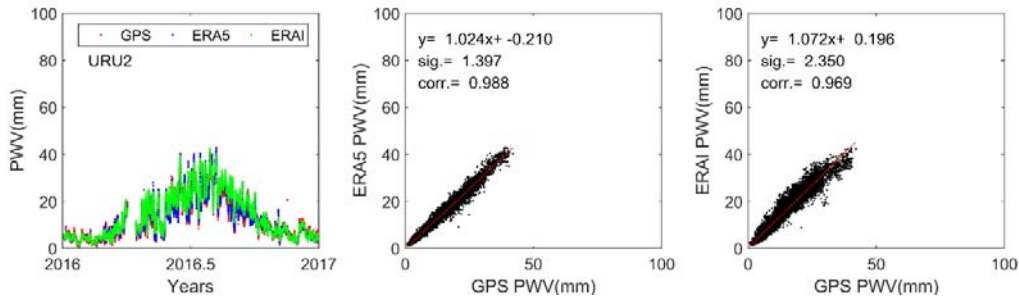FIGURE III. PWV COMPARISON AT BJFS STATION (TEMPERATE MONSOON CLIMATE TYPE)

FIGURE IV. PWV COMPARISON AT URU2 STATION (TEMPERATE CONTINENTAL CLIMATE TYPE)
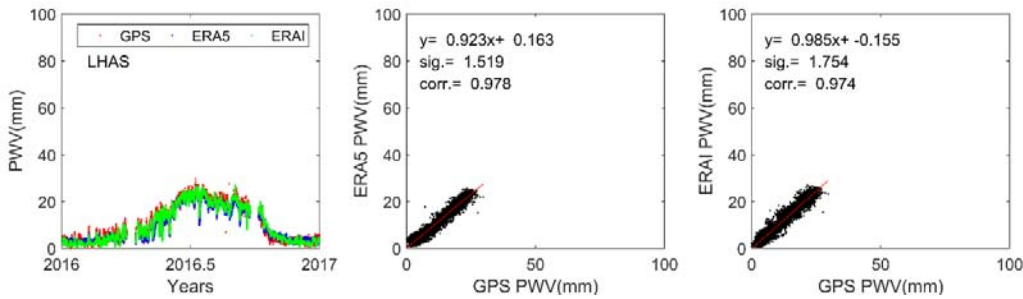

FIGURE V. PWV COMPARISON AT LHAS STATION (ALPINE PLATEAU CLIMATE TYPE)

As shown in the PWV time series diagrams in Figures 1–5, PWV in the regional stations exhibits annual variations. PWV at QION station in the tropical monsoon climate area maintains a relatively high value throughout the year and exceeds 20 mm even in winter, except for a short period of time during winter. The PWV peak at AHAQ station is comparable to that of QION station at around 70 mm, but the decline in PWV in winter is large (less than 10 mm) mainly because of the summer monsoon, which brings abundant water vapor from the ocean to the area. Meanwhile, the area is dominated by cold and dry winter winds from Siberia, thus showing a remarkable drop in water vapor content during winter. The station BJFS and AHAQ have the same annual variation amplitude, but the time period of PWV exceeding 50 mm in the BJFS station is remarkably shorter than that in the AHAQ station, which is mainly because the water vapor needs a strong monsoon to bring to the area. The location of URU2 is far from the ocean, and thus the water vapor brought by the monsoons to the area are relatively low compared with the other areas. Hence, the PWV peak in this area is less than 50 mm. The LHAS station, which is in the Qinghai–Tibet Plateau area, has a PWV content of less than 30 mm for the whole year owing to its high altitude, low temperature, and small monsoon impact. The PWV time series comparison of GPS, ERA5, and ERAI show that ERA5 and ERAI are more accurate in reflecting the mean and annual variation characteristics of PWV.

As shown in Figures 1–5, the PWV of ERA5 has slightly stronger correlation with that of GPS is slightly stronger than the PWV of ERAI for each test station. The PWV correlation coefficients of ERA5 and GPS at LHAS and GPS ERA5 at QION, AHAQ, BJFS, URU2, and LHAS stations were 0.974, 0.994, 0.992, 0.988, and 0.978, whereas those of ERAI were 0.969, 0.993, 0.991, 0.969, and 0.974, respectively. The error in

the unit weight obtained using linear fitting also shows that ERA5 is consistent with GPS, thereby showing that the accuracy of PWV extracted from the latest generation of ERA5 products is slightly better than that of ERAI when GPS data are used as references.

B. PWV Diurnal Variation Comparison

ERA5 is the world's first global reanalysis data with hourly resolution. In this section, we compare the PWV diurnal variation characteristics of the five GPS test stations on the basis of GPS, ERA5, and ERA-Interim data to assess the reliability of the water content information of the two reanalysis data on the daily scale.

First, the PWV distance flat value sequence is obtained by removing the daily average value in the PWV sequence, and then the average monthly value of each hour distance is obtained. As shown in Figures 6–10, the horizontal axis is the UTC hour, longitudinal axis is the month, and the figures show the diurnal variation characteristics of PWV of different months. The diurnal variation of PWV in different climatic regions has significant differences, wherein the largest diurnal variation of PWV is in QION station with tropical monsoon climate; the minimum diurnal variation is at LHAS station, and the difference of the diurnal variation phase of PWV in different regions is also remarkable and is mainly affected by sunrise and sunset. The peak moments of QION, AHAQ, and BJFS in China's eastern region occur earlier than China's URU2 and LHAS stations in the western region, and the peak moments vary with the seasons of sunrise and sunset, that is, the peaks usually occur earlier in summer than in winter.
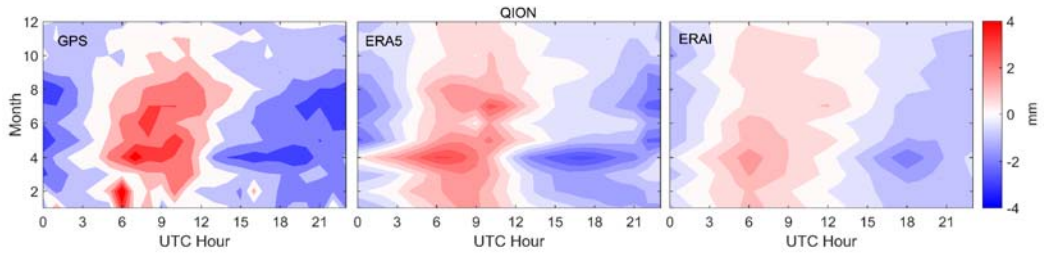
FIGURE VI. PWV DIURNAL VARIATION COMPARISON AT QION STATION (TROPICAL MONSOON CLIMATE TYPE)
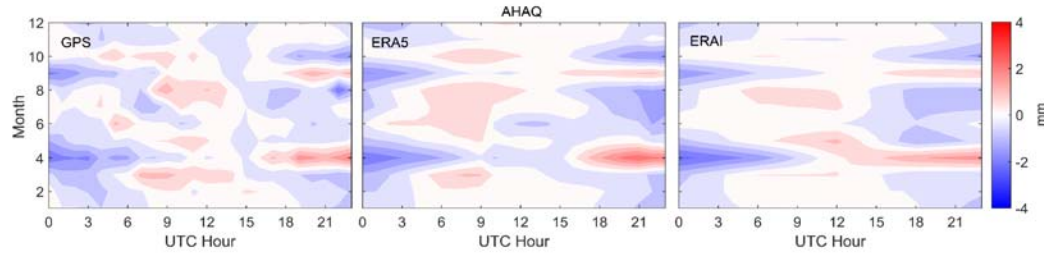

FIGURE VII. PWV DIURNAL VARIATION COMPARISON AT AHAQ STATION (SUBTROPICAL MONSOON CLIMATE TYPE)
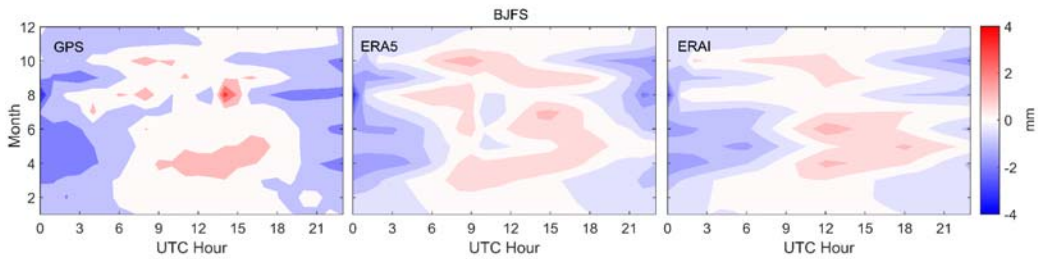

FIGURE VIII. PWV DIURNAL VARIATION COMPARISON AT BJFS STATION (TEMPERATE MONSOON CLIMATE TYPE)
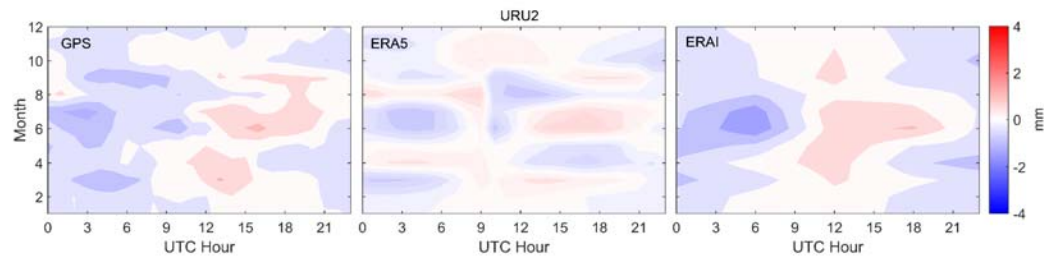

FIGURE IX. PWV DIURNAL VARIATION COMPARISON AT URU2 STATION (TEMPERATE CONTINENTAL CLIMATE TYPE)
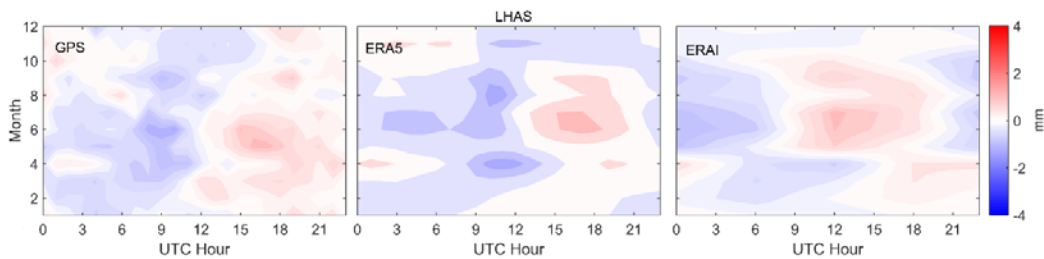

FIGURE X. PWV DIURNAL VARIATION COMPARISON AT LHAS STATION (ALPINE PLATEAU CLIMATE TYPE)

The diurnal variation amplitude of ERA5 and ERAI at the QION station is remarkably smaller than that of the GPS. This result indicates that the analysis data near the equatorial region are underestimated. However, ERA5 is more precise than ERAI regardless of amplitude or phase. In the AHAQ and BJFS stations, ERA5 and ERAI can reflect the amplitude and phase of diurnal variation to a certain extent. By contrast, in the URU2 station, large deviations exist in the diurnal variation phase of the two reanalysis data. In the LHAS station, the change trend of peak time in ERAI is opposite to that of GPS. This deviation is improved remarkably in ERA5.

IV. Conclusions

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

PANDA software is used to conduct PPP processing on the data of five GPS reference stations located in different climatic types in CMONOC in 2016, and the PWV products are obtained by inversion. On the basis of these products, PWV accuracy and characteristics of the changes on different scales are evaluated and analyzed. Data are obtained from the latest generation of meteorological reanalysis data ERA5 with the world's first hourly resolution and the previous generation of meteorological reanalysis data ERA-Interim. The following observations are obtained:

(1) The annual variation characteristics of PWV in different regions are mainly affected by monsoon, and the two reanalysis data are in good agreement with the PWV mean and annual variation characteristics of GPS;

(2) Considerable differences are observed among the diurnal variation amplitudes of PWV of different regions, whereas the diurnal variation phase is mainly dominated by sunrise and sunset times. Remarkable deviations are obtained in the PWV diurnal variation of the two reanalysis data. However, the diurnal variation deviation of ERA5 is smaller than that of ERA-Interim products, especially in the Qinghai–Tibet Plateau area. The reverse phenomenon of the diurnal variation phase trend in ERA-Interim is also improved remarkably in ERA5.

The relevant results of this study can provide an important reference for studies using reanalysis data to determine water vapor distribution and variation.

Acknowledgment

References

[1] Bevis M., S. Businger, T. Herring, C. Rocken, R. Anthes, and R. Ware (1992): GPS meteorology: Remote sensing of the atmospheric water vapor using the Global Positioning System. J. Geophys. Res., 97 (D14): 15787–15801.

[2] Held I. M. and B. J. Soden (2000): Water vapor feedback and global warming. Annu. Rev. Energy Env., 25, 445–475.

[3] Nilsson T. and G. Elgered (2008): Long-term trends in the atmospheric water vapor content estimated from ground-based GPS data. J. Geophys. Res., 113, D19101.

[4] Poli P., et al. (2007): Forecast impact studies of zenith total delay data from European near real-time GPS stations in Météo France 4DVAR. J. Geophys. Res., 112, D06114.

[5] Shi, C., Q. Zhao, J. Geng, Y. Lou, M. Ge, and J. Liu, 2008: Recent development of PANDA software in GNSS data processing. International Conference on Earth Observation Data Processing and Analysis (ICEODPA), D. Li, J. Gong, and H. Wu, Eds., International Society for Optical Engineering (SPIE Proceedings, Vol. 7285).

[6] Trenberth K. E., J. Fasullo, and L. Smith (2005): Trends and variability in column-integrated atmospheric water vapor. Climate Dyn., 24.

[7] Wang J. and L. Zhang (2008): Systematic errors in global radiosonde precipitable water data from comparisons with groundbased GPS measurements. J. Climate, 21: 2218–2238.

[8] Zhang, W., Y. Lou, J. Haase, R. Zhang, G. Zheng, J. Huang, C. Shi, and J. Liu, 2017: The use of ground-based GPS precipitable water measurements over China to assess radiosonde and ERA-Interim moisture trends and errors from 1999 to 2015. J. Climate, 30, 7643–7667.

[9] Zhang, W., Y. Lou, J. Huang, F. Zheng, Y. Cao, H. Liang, C. Shi and J. Liu, 2018: Multiscale Variations of Precipitable Water over China Based on 1999–2015 Ground-Based GPS Observations and Evaluations of Reanalysis Products. J. Climate, 31, 945–962.

[10] Cai Ying, Qian zhengan, Wu Tongwen, et al. Distribution,Changes of Atmospheric Precipitable Water over Qinghai-Xizang Plateau and Its Surroundings and Their Changeable Precipitation Climate [J]. PLATEAU METEOROLOGY, 23: 1–10, 2004.

[11] Dai Ying, Yang Xiuqun. Spatial-temporal variations of precipitable water over China [J], SCIENTIA METEOROLOGICA SINICA, 29(2): 143–149, 2009.

[12] Zhao Tianbao, Tu Kai, Yan Zhongwei. Advances of Atmospheric Water Vapor Change and Its Feedback Effect [J], PROGRESSUS INQUISITIONES DE MUTATIONE CLIMATIS, 9(2): 079–088, 2013