Iranian Journal of Geophysics, Vol 19 No 3, 2025, Page 119 – 136 (Research Article)
DOI: 10.30499/ijg.2024.461309.1607

Performance of pressure-temperature models and reanalysis products for estimating GPS precipitable water vapor

Ali Sam-Khaniani [1*]

[1] *Assistant Professor, Babol Noshirvani University of Technology, Babol, Iran*

(Received: 05 June 2024, Accepted: 20 August 2024)

Abstract

GPS meteorology has been regarded as one of the most advanced techniques for estimating Precipitable Water Vapor (PWV) over the past two decades. To calculate PWV from estimated tropospheric delay values, surface pressure and temperature data are needed, while many GPS stations lack meteorological sensors. This study presents a comprehensive evaluation of the performance of different sources of surface pressure and temperature data for estimating GPS PWV in Iran. To do this, real observations were compared with the ERA5L reanalysis data, and the global empirical models GPT2w and GPT3 at six GPS stations across Iran over one year. In addition to assessing the statistical quality of these data sources, their impact on the accuracy of GPS PWV estimation was investigated. While reanalysis data generally shows higher correlation with real observations, the experimental models (GPT2w and GPT3) often lead to comparable or better PWV estimates, particularly in stations where ERA5L pressure data shows significant bias. The results showed that by replacing the temperature and surface pressure data obtained from the GPT2w/GPT3 models with the actual observed values, the maximum increase in the RMSE of the GPSPWV values compared to the radiosonde PWV is less than 0.3 mm. The findings of this research show the potential of using global empirical models as alternatives to real observations for GPS stations lacking meteorological sensors in Iran, providing valuable insights for improving GPS meteorology techniques in the region.

Keywords: Surface pressure-temperature, GPS PWV, GPT2w, GPT3, ERA5

[*]Corresponding author: ali.sam@nit.ac.ir

Introduction

Water vapor is one of the most important components of the atmosphere, which, in addition to its effective role in global climate change and hydrological cycles, has a vital effect on regulating the temperature on Earth (Andrews, 2000; Book et al., 2007). This parameter has a great contribution to the phenomena affecting social and agricultural activities, such as rainfall and floods (Mazany et al., 2002).

There are many parameters to express the amount of water vapor in the Earth's atmosphere. Meanwhile, Integrated Water Vapor (IWV) and Precipitable Water Vapor (PWV) are two widely used terms to express the amount of water vapor. IWV is the total water vapor in a vertical column of the atmosphere, which is used in Kg/m2 units. On the other hand, if all the water vapor in the vertical column of the atmosphere with a cross section of one square meter is compressed and turns into a liquid, its height in millimeters is called PWV.

Several methods, such as radiosonde, satellite remote sensing observations, ground-based solar photometers, radar observations, and GPS meteorology techniques, have been used to measure water vapor in the atmosphere, each of which is different in terms of spatial and temporal resolution (Gao et al., 2003; Divakarla et al., 2006; Alexandrov et al., 2009; Whiteman et al., 2012).

In many studies, PWV values obtained from radiosondes are considered reliable values. However, factors such as non-uniform coverage across the land, low temporal resolution (twice a day), reduced sensor performance in cold and dry conditions, and high cost are the limitations of this technique (Pramualsakdikul et al., 2007). Remote sensing techniques have expanded the spatial coverage of water vapor products, while they suffer from low temporal resolution and cloudiness (Divakarla et al., 2006).

Bevis et al. (1992) introduced the GPS meteorology technique to measure atmospheric water vapor using ground-based GPS receiver observations Today, due to its unique features such as the ability to be used in any weather conditions, long-term stability, and continuous observations with a very high temporal resolution, GPS meteorology is a powerful tool in estimating PWV values, which has attracted the attention of the hydrometeorological community. (Van Baelen et al., 2005; Sibylle et al., 2010; Vaquero-Martínez et al., 2017).

Accurate values of surface temperature and pressure are crucial for improving the accuracy of evapotranspiration calculations, hydrological model outputs, and precipitation forecasts (Singh and Woolhiser, 2002; Pelosi et al., 2020; Sam-Khaniani et al., 2021). In addition, this data plays an important role in estimating PWV with the help of the GPS meteorology technique. In the precise processing of GPS observations, the tropospheric delay is usually considered equal to the sum of two wet and hydrostatic (or dry) terms. To calculate the hydrostatic delay of the signal, the surface pressure at the height of the GPS antenna is very important. Furthermore, in the direction of the zenith, there is a correlation between the tropospheric delay parameter and the correction made to the receiver clock and the estimated height of the stations during the processing of GPS measurements (Rothacher, 2002). The hydrostatic delay (ZHD) of GPS signals is caused by the presence of dry gases in the atmosphere, and the wet delay (ZWD) is due to the presence of water vapor and condensed water in the clouds. Using Saastamoinen's model, the hydrostatic part of the delay can be calculated by using accurate pressure observations at the station (Saastamoinen, 1972).

Additionally, the mean atmospheric temperature (Tm) is required to obtain PWV values after computing ZWD. Tm can be estimated using numerical weather models, radiosonde observations, or linear empirical relationships based on surface temperature observations at stations (Bevis et al., 1994). Utilizing a linear function of the surface temperature is a less expensive solution to obtain the Tm values (Bevis et al. 1992). Therefore, access to the surface temperature at the GPS station is essential for PWV calculation. In general, the higher the accuracy of ZHD and Tm, which can be calculated with the help of surface meteorological parameters, will lead to a more accurate estimation of ZWD and finally PWV.

For many GPS stations, the lack of temperature and pressure sensor makes it difficult to examine and track changes in water vapor with high spatial and temporal precision (Huang et al., 2023). As an alternative solution in such situations, it is advised to use experimental models like GPT, GPT2, GPT2w, and GPT3 (Bohm et al. 2007; Lagler et al., 2013; Li et al., 2020). Experimental models may be used anywhere and do not require a meteorological sensor. Then, these models have been extensively utilized in research to obtain the temperature and pressure needed for GPS PWV estimation. Another valuable source of meteorological data is the reanalysis product, which provides various parameters in a wide range of space and time around the world. The reanalysis data is freely available to the public and provides researchers with meteorological variables globally or locally, with various spatial and temporal resolutions in long-term time intervals (Sheffield et al., 2006 Therefore, several studies use reanalysis data as helpful and valuable resources in the study of water vapor with the help of GPS (Zhang et al., 2019). However, these products are always available with a delay. Therefore, experimental models with proper accuracy will be favored for real-time applications in any GPS station without a weather sensor.

Iran has a network with approximately 100 permanent GPS stations that have been established for navigation and geodynamic purposes. However, this network can also be utilized to estimate accurate PWV data. Unfortunately, the majority of these stations lack meteorological sensors. Consequently, it would be beneficial for researchers to find a suitable source for obtaining surface temperature and pressure data to use as input for the GPS meteorology process. On the other hand, experimental models are advantageous as they do not rely on meteorological sensors. However, before using them in the GPS meteorology technique, it is necessary to evaluate the accuracy of the temperature and pressure obtained from the implementation of these models in comparison with real observations of each region. The importance of this issue has caused that in this study, the main objective is to evaluate the surface temperature and pressure resulting from two experimental models GPT2w and GPT3 in the Iranian region and examine their efficiency in calculating PWV from GPS data processing. In addition, temperature and surface pressure data from ERA5L reanalysis will also be used in GPS PWV calculation. Finally, the PWV obtained with the help of surface variables obtained from GPT2w, GPT3 models, ERA5 reanalysis and real observations will be compared with the corresponding reference values estimated from the radiosonde measurements in six stations of the country.

In the second part, meteorological data, radiosonde, GPS observations, as well as ERA5 reanalysis data and GPT2w/GPT3 models will be introduced. How to evaluate models and reanalysis data is given in the third section. The comparison

of the temperature and pressure of different sources with the actual values measured at the stations and also their effect on GPS PWV estimation is stated in section 4. Finally, section 5 will depict the most important results of the present research.

2 Data and study area

As mentioned in the previous section, the focus of this study is to evaluate the efficiency of different temperature and surface pressure sources in converting the GPS tropospheric delay values to PWV. To achieve this goal, different data sources, including real temperature and surface pressure observations at synoptic stations, ERA5 reanalysis surface data, radiosonde measurements, GPS observations at ground stations and temperature and pressure obtained from two experimental models GPT2w and GPT3 have been used. In this section, the details of these data will be given.

The studied stations were selected from the Iranian Permanent GPS Network (IPGN) stations in Iran in order to estimate PWV. Among the IPGN network, stations were selected that have a radiosonde station near them for statistical evaluation. In Figure 1, the positions of six GPS stations (stars) along with the nearest radiosonde (red dots) and synoptic (triangles) stations are displayed. Additionally, the longitude and latitude of the meteorological, radiosonde, and GPS stations, as well as the distance between the corresponding GPS and radiosonde stations, are given in Table 1.

2.1 Radiosonde and synoptic observation

One of the most established methods for measuring atmospheric conditions is the use of radiosondes, which have been used to measure temperature, relative humidity, wind direction and speed, geopotential height, and other characteristics at various elevations above the earth's surface. These observations are typically accessible twice a day (at 00 UTC and 12 UTC) due to financial constraints. PWV values can be computed using vertical profiles of several atmospheric parameters. Even if there is a 5–8% uncertainty in the PWV values acquired using radiosonde data when accounting for different types of particles, this method is still regarded as reliable for validating other PWV measurement techniques (Turner et al., 2003).

PWV can be computed using the following equation through numerical integration using profiles of meteorological parameters obtained by radiosonde (Bai, 2005).

$$\mathrm{PWV}=\frac{1}{\rho_w}\int \rho_v .\mathrm{dh}=\frac{1}{\rho_w}\sum\left(\mathrm{h}_{j+1}-\mathrm{h}_j\right).\left(\rho_v^{i+1}+\rho_v^{i}\right)/2 \quad (1)$$

where ρ_w is the density of liquid water and ρ_v is the density of water vapor which could be obtained by ideal gas equation:

$$\rho_v=\frac{e}{R_v.T} \quad (2)$$

where e and R_v are the water vapor's gas constant and partial pressure, respectively. With the following equation, one can determine the value of e given relative humidity observations (rh) obtained with a radiosonde (Godson, 1995).

$$e=\mathrm{rh}.\exp\left(-37.2465+0.213166.T-2.56908.10^{-4}.T^2\right) \quad (3)$$

where T the absolute temperature in Kelvin is expressed in the above equation. Data for a year from six radiosonde stations in Iran close to permanent GPS stations were used in this study (Table 1). The data were retrieved from the University of Wyoming's Department of Atmospheric Sciences website (http://weather.uwyo.edu/upperair/sounding.html). One year of precipitable water vapor values obtained from radiosonde observations (RSPWV) at OIII, AMND, SFHN, TEHN, AHVZ, BIJD and MSHN stations are used as reliable values for evaluating PWV estimates obtained from

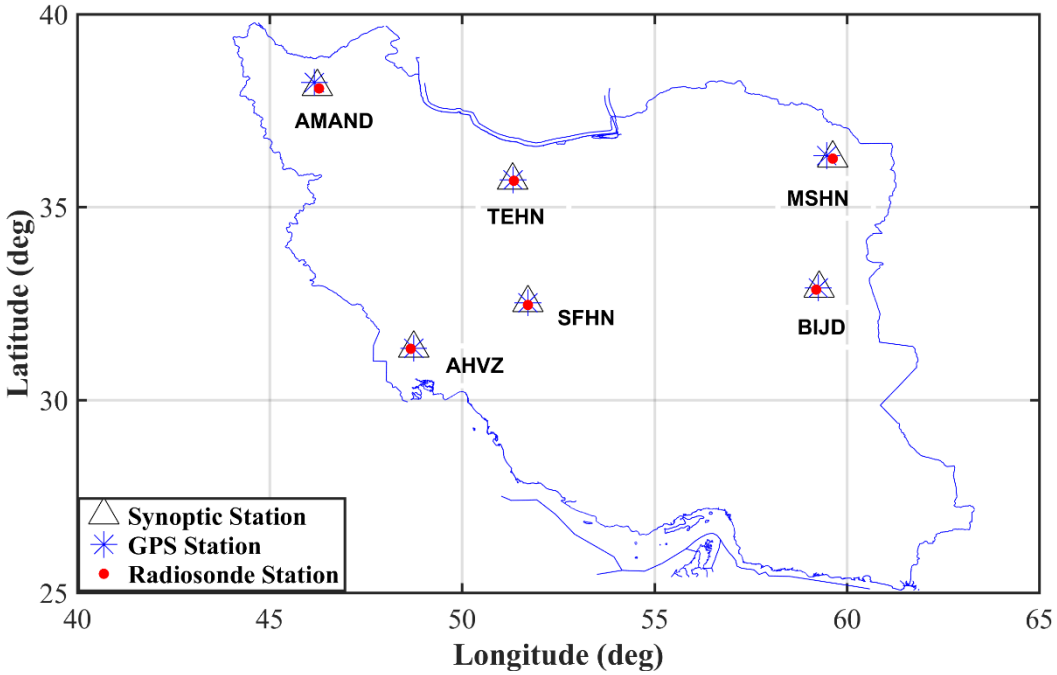

Figure 1. Spatial distribution of GPS stations used in this study along with the closest radiosonde and synoptic stations to them.

Table 1. Geodetic coordinates of GPS stations together with the nearest synoptic and radiosondes location.

	Radiosonde		GPS		Synoptic		GPS-RS
Station	Latitude	Longitude	Latitude	Longitude	Latitude	Longitude	Dist (Km)
amnd	38.08	46.28	38.23	46.16	38.12	46.24	21.13
tehn	35.68	51.35	35.7	51.33	35.69	51.31	3.11
ahvz	31.33	48.67	31.34	48.74	31.34	48.74	7.77
mshn	36.26	59.63	36.33	59.48	36.27	59.63	18.20
bijd	32.86	59.2	32.9	59.26	32.89	59.28	7.93
sfhn	32.46	51.71	32.52	51.71	32.52	51.71	6.6

different sources of surface data in the GPS meteorology process.

In addition, surface pressure and temperature observations from the nearest synoptic station to GPS receivers are needed to calculate the mean atmospheric temperature and zenith hydrostatic delay. For this purpose, three-hourly pressure and temperature data were extracted from the synoptic stations introduced in Table 1 during the study period. It should be noted that the precipitable water vapor resulting from the use of real temperature and pressure observations in the GPS meteorology process is called PWV in the text from now on.

2.2 ERA5

Reanalysis data are in common forms and publicly accessible. On a global or local scale, these weather data sources offer meteorological variables with a variety of spatiotemporal resolutions over extended time periods (Sheffield et al., 2006). Based on a combination of atmospheric and ground-based observations combined

with physical model integration, this kind of weather data source provides an accurate account of historical weather conditions (Soci et al., 2016). Without taking into account the atmospheric module of the integrated ECMWF forecasting system, ERA5 Land (ERA5L) is generated using a single simulation. From the surface to 0.01 hPa, it offers hourly meteorological variables at 137 pressure levels. The enhanced spatial resolution of ERA5L, which can reach up to 9 kilometers, is its main benefit over ERA5. Surface temperature and pressure data from the ERAL reanalysis throughout the research period were extracted and downloaded in grib format from the website: (https://cds.climate.copernicus.eu/cdsapp#!/dataset/). Then, hourly temperature and surface pressure values were derived using bilinear interpolation at the chosen station locations.

ERA5L reanalysis data has been processed using the geopotential height system. The distinction between the geopotential height system and the geodetic height system may have a substantial influence on the vertical corrections of temperature and pressure data from grid points to the desired station height. Therefore, it is vital to standardize the height system between the two datasets: the reanalysis data and ground measurements such as synoptic and GPS observations (Wang et al., 2016; Huang et al., 2023).

Vedel (2000) proposed a two-step elevation transfer from the geopotential height to the WGS84 ellipsoidal height. The ellipsoidal height must first be transformed into orthometric height, or the height above the geoid surface.

$$H = h - N \quad (4)$$

where N is the geoid height above the ellipsoid at the station's position, H is the orthometric height of the station, and h is its geodetic height. Using the EGM2008 model, one may obtain the geoid height N at a given location. This value can then be subtracted from the geodetic height in Equation 4 to calculate the orthometric height. The following formulas are utilized to convert the orthometric height to the geopotential height (H_g),), which is the subsequent step (Vedel, 2000).

$$H_g = \frac{\gamma_s(\varphi)}{\gamma_{45}} . \left[\frac{R(\varphi).H}{R(\varphi)+H}\right] \quad (5)$$

$$\gamma_s(\varphi)=9.780325. \left[\frac{1+0.00193185.\sin(\varphi)^2}{1-0.00669435.\sin(\varphi)^2}\right]^{0.5} \quad (6)$$

$$R(\varphi) = \frac{6378.137}{1.006803-0.006706.\sin(\varphi)^2} \quad (7)$$

where $R(\varphi)$ is the Earth's effective radius at the geodetic latitude under consideration, γ_{45} is the normal gravity on the ellipsoidal surface at a 45-degree latitude, and $\gamma_s(\varphi)$ is the normal gravity on the rotating ellipsoidal surface at latitude (φ). It is assumed that γ_{45}=9.80665 $^{m}/_{s^2}$. Equations 6 and 7 can be used to derive the values of $R(\varphi)$ and $\gamma_s(\varphi)$.

It is crucial to carry out vertical corrections between the reanalysis grid points and the ground station elevation once the height systems of the ground-based meteorological data and the reanalysis grid points have been unified. Vertical corrections are required for the pressure and temperature data if the geopotential height of the reanalysis grid points exceeds that of the ground station (Wang et al., 2016).

$$T_s = T_{grid} - \Delta(H_s - H_{grid}) \quad (8)$$

$$P_s = P_{grid}\left(\frac{T_{grid}-\Delta(H_s-H_{grid})}{T_{grid}}\right)^{\frac{g.M}{R.\Delta}} \quad (9)$$

where H_s is the station height, $M = 0.0289644\ ^{Kg}/_{mol}$ is the molar mass of dry air, $\Delta = 0.0065\ ^{Kelvin}/_{m}$ is the standard lapse rate of temperature , $R = 8.31432 \left[\frac{N.m}{mol.Kelvin}\right]$ is the universal gas constant, and g is the gravitational

parameter, which can be calculated using the following Equation:

$$g = 9.8063\left\{1 - 10^{-7}\left(\frac{H_s+H_{grid}}{2}\right)[1 - 0.0026373\cos(2\varphi) + 5.9 * 10^{-6}\cos(2\varphi)^2]\right\} \quad (10)$$

The reanalysis temperature and pressure are compared separately with the actual observations from the synoptic stations. Then, ERA5L temperature and pressure are utilized to convert the tropospheric zenith delay values estimated from GPS observations into PWV. We refer to this result as ERA5PWV throughout this article. It is important to mention that ERA5PWV is derived solely from reanalysis surface data in conjunction with tropospheric estimates from GPS processing.

2.3 GPT2w/GPT3 models

In 2013, based on 10 years of global data from the monthly average profile of pressure, temperature, relative humidity, and geopotential obtained from ERA-Interim reanalysis, the GPT2 model was introduced. It has a 5-degree spatial resolution and provides outputs such as pressure, temperature, water vapor pressure, lapse rate, and coefficients of the VMF1 image function (Lagler et al., 2013). The GPT2 model can consider the annual and semi-annual changes of atmospheric parameters using the following relationship.

$$y_i = A_{0i} + A_{1i}\cos\left(2\pi\frac{doy}{365.25}\right) + B_{1i}\sin\left(2\pi\frac{doy}{365.25}\right) + A_{2i}\cos\left(2\pi\frac{doy}{365.25}\right) + B_{2i}\sin\left(2\pi\frac{doy}{365.25}\right) \quad (11)$$

where y_i is one of the model output parameters such as temperature and surface pressure. Also, $A_{0i}$is the average of that parameter, A_{1i} and B_{1i} are annual domains and A_{2i} and B_{2i} are semi-annual domains that can be extracted from the grid of the model. The inputs of this model are the geodetic coordinates of the desired point and the modified Julian date. Then the GPT2 model was developed and the GPT2w version was presented with a spatial resolution of one degree, which added the water vapor reduction factor and the average atmospheric temperature to its outputs (Schindelegger et al., 2015).

The GPT3 model is the third version of the global temperature and pressure models. It is an experimental model that uses station position (geodetic longitude, latitude, and altitude) and time (modified Julian date) as inputs. The model provides the necessary information to calculate the tropospheric delay (Landskron, 2017). It has been developed based on the monthly average data of ERA-Interim reanalysis from 2001 to 2010. The model is available to users with two spatial resolutions of 1 and 5 degrees. In fact, GPT3 is an updated model of GPT2w, which is developed based on the VMF3 mapping function, and east and north gradients are added to the output parameters. In this study, by using both GPT2w and GPT3 models with a spatial resolution of 1°×1°, the pressure and temperature data are generated and will be used to calculate ZHD and Tm. In the following, the values of PWV calculated from the GPS tropospheric estimates and surface temperature/pressure obtained from GPT2w and GPT3 models are called GPT2WPWV and GPT3PWV, respectively.

2.4 GPS PWV

Deflection and delay occur in the GPS signal when it passes through the atmosphere and reaches the ground-based receivers. This error is partially due to the ionosphere layer, which can be determined by using observations from dual frequency GPS receivers because the delay in this layer is associated with the frequency of the satellites (Bohm et al., 2013). The delay resulting from the neutral region of the atmosphere, known

as the troposphere, is independent of frequency and cannot be eliminated. There are numerous weather applications for the valuable information found in the signal's total Zenith Tropospheric Delay (ZTD). Zenith Hydrostatic Delay (ZHD), also known as the dry element, and Zenith Wet Delay (ZWD), which represents the wet part, are the two components of ZTD; the latter is related to atmospheric water vapor (Iwabuchi et al., 2000).

$$ZTD = ZHD + ZWD \quad (12)$$

ZHD is calculated with high precision using the Sasstamoinen model and surface pressure (P_0) measured at the GPS station (Davis et al., 1985).

$$ZHD = \frac{0.0022768P_0}{1 - 0.00265\cos(2\varphi) - 0.000285H} \quad (13)$$

where P_0 is expressed in hPa, H is the height from the geoid in kilometers. During the GPS observation processing, ZTD values are estimated in conjunction with the station's location parameters. ZWD values might be derived by subtracting ZTD values from ZHD values, which could be calculated using the Sasstamoien model. ZWD depends on air temperature and water vapor. Bevis et al. (1994) state that a dimensionless factor is used to convert ZWD to PWV.

$$WV = \Pi \times ZWD \quad (14)$$

$$\Pi = \frac{10^6}{\rho_w R_v \left[\left({}^{k_3}/_{T_m}\right) + k_2 + k_1\left({}^{M_w}/_{M_d}\right)\right]} \quad (15)$$

where k_i coefficients are physical constants related to atmospheric refraction (Rueger, 2002), M_d and M_w are the molecular masses of dry air and water vapor, respectively. Additionally, T_m represents the average atmospheric temperature, which can be determined using water vapor partial pressure and temperature profiles (Davis et al., 1985).

$$T_m = \frac{\int \frac{e}{T}}{\int \frac{e}{T^2}} dh \quad (16)$$

T_m is derived by numerically solving equation 16 using numerical weather prediction models or atmospheric profiles measured by radiosondes. However, vertical atmospheric profiles are usually not available at the GPS station. Bevis et al. (1992) introduced an experimental linear model between surface temperature T_0 (in Kelvin) and average air temperature T_m..

$$T_m = 70.2 + 0.72\,T_0 \quad (17)$$

The geographical and seasonal conditions of the study area are related to the coefficients of this linear model. Based on 54330 radiosonde profile data that were dispersed throughout Iran, Sadeghi et al. (2014) computed the linear method (18) coefficients in accordance with the following for the study area:

$$T_m = 75.39 + 0.7103\,T_0 \quad (18)$$

Raw data from GPS receivers in the RINEX standard format were gathered to estimate the ZTD values. The next step involves estimating tropospheric delays and coordinate components for each station. Double-differenced observations were processed in base-line mode using GAMIT software, with ITRF2008 as the earth referencing framework. To estimate ZTD values at one-hour intervals, data processing techniques included using a 7-degree elevation cutoff angle, the Global Mapping Function (GMF), and the final products of clock and orbit from the International GNSS Service (IGS).

3 Evaluation process

At first, three-hourly raw pressure and temperature data recorded at the studied synoptic stations were collected. Then, the hourly ERA5L product at the position of these stations was prepared using bilinear interpolation. In the next step, reanalysis data were extracted for each station at times when there were real observations. In other words, real observations and ERA5L data were synchronized. On the other hand, global experimental models obtain temperature and surface pressure as

outputs by entering the longitude, latitude, and geodetic height of the desired station and time. Therefore, by entering the position of the stations and the times when the actual observations and the reanalysis data were adjusted with each other, the GPT2w and GPT3 models produced the temperature and pressure. These procedures allowed for the preparation of all the data in three hours. After providing different data sets of surface temperature and pressure and using equations 12 to 15, the estimated tropospheric delay at GPS stations will be converted to PWV. The computed values will then be compared with reliable radiosonde data. All of these steps were conducted using programming in the MATLAB software environment.

To assess the performance of different pressure/temperature sources compared to the real observations, the values of Mean Error (ME), RMSE, and correlation coefficient (R) were calculated using the GPT2w/GPT3 and ERA5L data for comparison with corresponding synoptic station data.

$$ME = \frac{1}{N_y}\sum_{i=1}^{N_y}\left(y_i^S - y_i^{obs}\right) \quad (19)$$

$$RMSE = \sqrt{\frac{1}{N_y}\sum_{i=1}^{N_y}\left(y_i^S - y_i^{obs}\right)^2} \quad (20)$$

$$R = \frac{\sum_{i=1}^{N_y}\left(y_i^S - y_m^S\right)\sum_{i=1}^{N_y}\left(y_i^{obs} - y_m^{obs}\right)}{\sqrt{\sum_{i=1}^{N_y}\left(y_i^S - y_m^S\right)^2}\sqrt{\sum_{i=1}^{N_y}\left(y_i^{obs} - y_m^{obs}\right)^2}} \quad (21)$$

where N_y represents the total number of 3-hourly data points in the time series of variable y for each station. Additionally, y_i^S and y_i^{obs} denote the ith values of the time series of variable y, extracted respectively from GPT2w/GPT3/ ERA5L products and surface real measurements. Also, the last three equations are used to evaluate the amount of GPS-based PWV compared to the reliable values obtained from radiosonde. In equation 21, the average of the time series values of the meteorological variable y at the specified station is denoted as y_m^S and y_m^{obs}, corresponding to GPT2w/GPT3/ ERA5L products and surface measurements, respectively.

4 Results and discussions

One year of temperature and surface pressure data was collected from ground sensor measurements or extracted from ERA5L reanalysis products at six studied stations. Additionally, another source of temperature and surface pressure data in this research is the output of experimental models GPT2w/GPT3 in the location of the stations and the time period under study. In the next step, the temperature and pressure values obtained from different methods were extracted at the same times so that they can be statistically compared with real observations before entering the GPS meteorology process. In this section, the statistical quality of each of the temperature and pressure data sets compared to real observations is discussed. Subsequently, the impact of each approach on the precision of GPS PWV calculation is examined based on how well it provides surface meteorological data.

4.1 Pressure/temperature comparison

After synchronizing one year of surface pressure data at six stations (OIAW, OIMM, OIII, OIMB, OITT, and OIFM), the surface pressure values obtained from ERA5L reanalysis and the global empirical models GPT2w/GPT3 are compared with actual observations at the considered stations. In Figure 2, the black, red, blue, and green graphs represent real observations, reanalysis values, the GPT2w model, and the GPT3 model, respectively. Based on the graphical results in Figure 2, it is evident that the GPT2w and GPT3 models are unable to

display short-term changes in surface pressure at any station. However, due to their nature, they can present the annual and semi-annual variations of real surface pressure measurements.

Also, since the efficiency of both GPT2w and GPT3 models is very close to each other, green graphs are placed on blue graphs in all stations. On the other hand, in all the stations, the short-term changes in the actual observations are well observed in the corresponding pressure data extracted from the ERA5L reanalysis (red graphs in Figure 2). However, in some stations such as OIMM, OIII and OITT, the pressure time series obtained from the reanalysis product have a significant bias compared to the actual observations.

Similar to Figure 2, a one-year time series of surface temperature data at six stations OIAW, OIMM, OIII, OIMB, OITT and OIFM, obtained from ERA5L reanalysis and GPT2w/GPT3 models, along with actual observations were compared in Figure 3. As shown in Figure 3, like the pressure data, the temperature values obtained from the reanalysis (red graphs) at all stations follow the behavior of short-term changes of actual observations (black graphs). However, based on the graphical results in Figure 3, the bias in the reanalysis temperature values is lower compared to the pressure data from this data source.

Another point that can be found from figure 3 is the significant bias of surface temperature data obtained from GPT2w (blue color) and GPT3 (green color) models compared to real observations. In addition, by comparing Figures 2 and 3, it can be seen that in most stations the temperature bias is higher than the pressure bias obtained from experimental models.

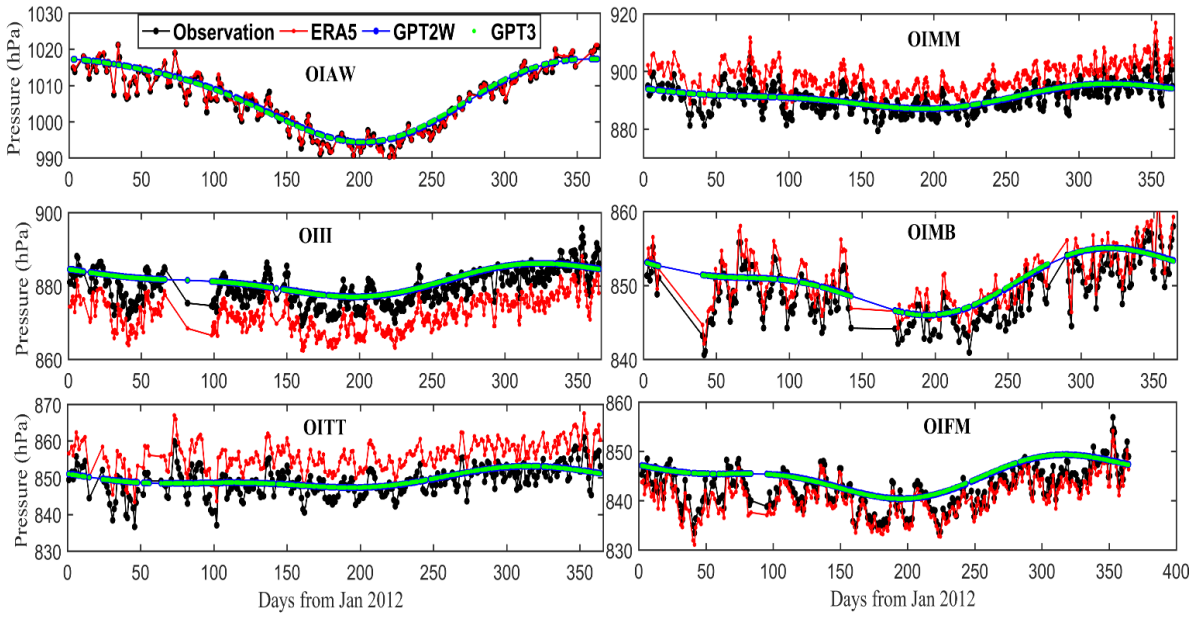

Figure 2. Comparison of the real surface pressure time series (black color) with the values obtained from ERA5L reanalysis (red color), GPT2w (blue color) and GPT3 (green color) models in six studied stations.

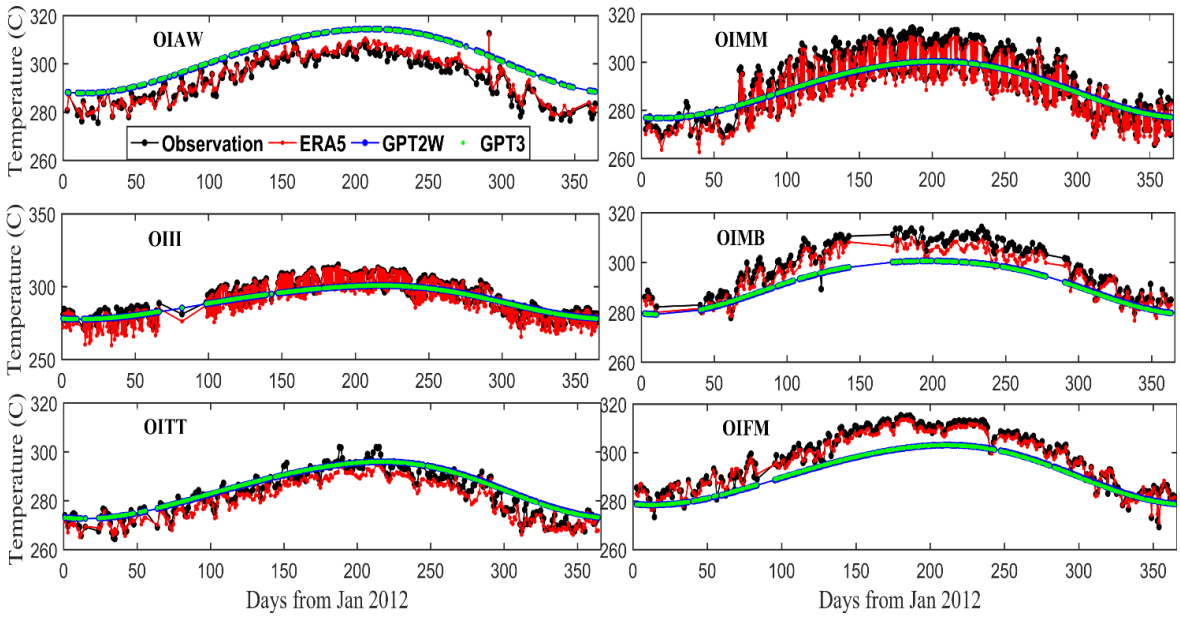

Figure 3. Comparison of the real surface temperature time series (black color) with the values obtained from ERA5L reanalysis (red color), GPT2w (blue color) and GPT3 (green color) models in six studied stations.

In addition to the graphical comparison between the time series of pressure and temperature observed in the stations compared to the corresponding values extracted from the ERA5L reanalysis and the GPT2w and GPT3 models, it is necessary to evaluate the quality of these data quantitatively in terms of statistics. For this purpose, the temperature and pressure data obtained from the reanalysis products and the output of the models were compared with the actual measurements at each station, and the results in terms of average bias, RMSE, and correlation coefficient are given in Table 2. The last column in Table 2 shows the number of common data pairs between actual observations and other temperature and pressure sources.

Based on the statistical values in Table 2, the absolute value of the average temperature bias obtained from experimental models in some stations such as AHVZ, BIJD and SFHN is much higher than the values obtained for ERA5L. Also, for the other three stations, although the absolute value of the temperature bias is lower, it is not significantly different from the corresponding values for ERA5L. Comparing the real pressure data with the values obtained from ERA5L and experimental models shows that in most stations the size of the bias values of the reanalysis data is greater than the GPT2w and GPT3 models. For example, the absolute value of the average pressure bias obtained from ERA5L in AMND, TEHN and MSHN stations has reached about 7hPa.

As seen in Figure 3, the RMSE values of the temperature data of GPT2w and GPT3 models were estimated to be higher in all stations compared to the reanalysis data. On average, in six study stations, the reanalysis pressure RMSE was higher than similar values for GPT2w and GPT3 models (Table 2). Additionally, the correlation coefficient between the surface data obtained from ERA5L reanalysis and measured observations was higher than the corresponding values estimated for experimental models.

Table 2. Comparison of pressure and temperature data extracted from GPT2W, GPT3 and ERA5 relative to real observation in six stations.

		MBE		RMSE		R		Data number
Station		Temperature	Pressure	Temperature	Pressure	Temperature	Pressure	N
amnd	ERA5	-2.833	7.056	3.710	7.069	0.969	0.994	289
	GPT2W	2.435	0.654	4.291	3.584	0.929	0.485	
	GPT3	2.432	0.656	4.238	3.589	0.932	0.483	
tehn	ERA5	-2.645	-7.134	3.743	7.155	0.982	0.993	536
	GPT2W	-1.669	2.000	5.835	3.865	0.889	0.703	
	GPT3	-1.686	2.002	5.789	3.873	0.891	0.702	
ahvz	ERA5	1.124	0.155	2.218	0.504	0.980	0.998	228
	GPT2W	8.614	0.996	9.339	2.771	0.923	0.951	
	GPT3	8.627	0.992	9.320	2.776	0.926	0.951	
mshn	ERA5	-2.490	7.076	3.070	7.093	0.989	0.994	495
	GPT2W	-1.062	0.434	7.380	3.701	0.831	0.634	
	GPT3	-1.074	0.439	7.350	3.711	0.833	0.632	
bijd	ERA5	-3.030	2.025	3.4831	2.098	0.987	0.991	235
	GPT2W	-7.221	1.915	8.247	3.402	0.950	0.693	
	GPT3	-7.230	1.924	8.248	3.422	0.951	0.688	
sfhn	ERA5	-1.037	-1.693	1.881	1.826	0.992	0.987	292
	GPT2W	-6.5537	2.589	7.839	4.086	0.954	0.649	
	GPT3	-6.5627	2.588	7.809	4.098	0.956	0.644	

4.2 Impact on GPS PWV estimation

After processing the GPS observations at the selected stations, the time series of hourly ZTD values was estimated. First, the hourly ZTD values were resampled at the times when the pressure and temperature data from all three methods ERA5L/GPT2w/GPT3 were available. Then, using equation 13 and the pressure time series data, ZHD values were estimated at each station. The ZWD time series was obtained by subtracting ZHD from the total delay. The latter parameter is related to the water vapor in the atmosphere for each station (equation 12). Finally, by applying equations 18, 15, and 14, ZWD was converted to PWV. The estimated water vapor values obtained from real temperature and pressure data, ERA5L, GPT2w, and GPT3 were named PWV, ERA5PWV, GPT2WPWV, and GPT3PWV, respectively.

To evaluate these four categories of water vapor estimates at every GPS station, all these values were compared with the precipitable water vapor data obtained from the nearest radiosonde (RSPWV) observations. The time series of PWV differences (DPWV) of each of these estimates compared to RSPW for different stations were shown in Figure 4. According to the results, the DPWV time series estimated with the help of real meteorological observations (black color) at all stations are closer to the line y=0 than the other water vapor estimates (ERA5PWV, GPT2WPWV, and GPT3PWV). Additionally, the results show that the time series of ERA5PWV difference values from RSPWV data in three stations MSHN (corresponding to the OIMM radiosonde station), TEHN (corresponding to the OIII radiosonde station), and AMND (corresponding to the OITT radiosonde station) compared to the other water vapor estimates are further away from the y=0 line.

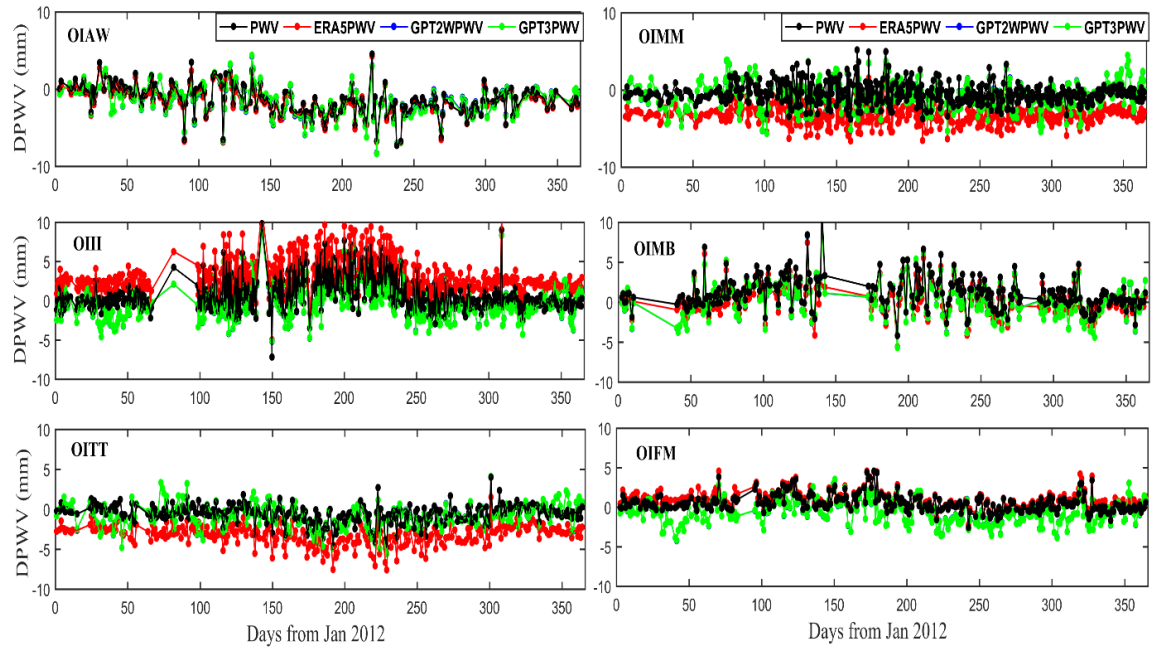

Figure 4. Time series comparison of differences in PWV (black color), ERA5PWV (red color), GPTWPWV (blue color) and GPT3PWV (green color) in six studied stations from the corresponding reference values obtained from the radiosonde observations (RSPWV).

Furthermore, average bias, RMSE and correlation coefficient of different estimates of water vapor named PWV, ERA5PWV, GPT2WPWV and GPT3PWV at each station compared to RSPWV values are calculated and the results are given in Table 3 and Figure 5. Statistical analysis has shown that the correlation coefficient of PWV values

Table 3. Evaluation of GPSPWV computed based on GPT2W (GPT2WPWV), GPT3 (GPT3PWV), ERA5 (ERA5PWV) and real observation (PWV) with respect to the radiosonde derived (RSPWV) values.

Station		MBE	RMSE	R
amnd	ERA5PWV	-3.233	3.461	0.980
	GPT2WPWV	-0.857	1.751	0.965
	GPT3PWV	-0.858	1.755	0.965
	PWV	-0.590	1.287	0.981
tehn	ERA5PWV	3.048	3.725	0.931
	GPT2WPWV	-0.051	2.133	0.932
	GPT3PWV	-0.052	2.134	0.931
	PWV	**0.223**	**2.122**	0.936
ahvz	ERA5PWV	-1.638	2.429	0.960
	GPT2WPWV	-1.552	2.460	0.954
	GPT3PWV	-1.551	2.474	0.954
	PWV	**-1.365**	**2.301**	0.957
mshn	ERA5PWV	-3.256	3.474	0.981
	GPT2WPWV	-0.832	1.871	0.963
	GPT3PWV	-0.834	1.877	0.963
	PWV	**-0.608**	**1.445**	0.978
bijd	ERA5PWV	0.227	1.975	0.887
	GPT2WPWV	0.248	2.218	0.856
	GPT3PWV	0.245	2.222	0.856
	PWV	**0.194**	**2.202**	0.886
sfhn	ERA5PWV	0.911	1.400	0.971
	GPT2WPWV	-0.654	1.568	0.951
	GPT3PWV	-0.654	1.580	0.951
	PWV	**0.508**	**1.187**	0.973

estimated by using all four categories of temperature and surface pressure data in almost all GPS stations is more than 90%. According to Table 4, as expected, in all GPS stations, the PWV estimates achieved by using real observations in the GPS meteorology process have the best quality in terms of average bias and RMSE. According to the results, it can be seen that by using surface meteorological observations, GPS PWV can be produced with an accuracy of about 2 mm compared to the water vapor values obtained from the radiosonde. Also, in general, using temperature and pressure data obtained from GPT2w and GPT3 models has almost the same and better statistical results than ERA5L data in estimating PWV with the help of GPS observations.

According to Figure 5, in most stations, the average bias of GPT2WPW and GPT3PWV is lower than ERA5PWV. Also, from the comparison of RMSE values obtained for GPT2WPW and GPT3PWV compared to ERA5PWV, it can be seen that the use of temperature and pressure data obtained from experimental models can be effective close to or better than ERA5L surface data in GPS PWV estimation. For example, in AHVZ, BIJD and SFHN stations, experimental models and reanalysis products have almost the same effect on the statistical quality of estimated PWV. However, in three stations AMND, TEHN and MSHN, the efficiency of reanalysis data compared to experimental models in estimating GPS PWV was much weaker (Figure 5).

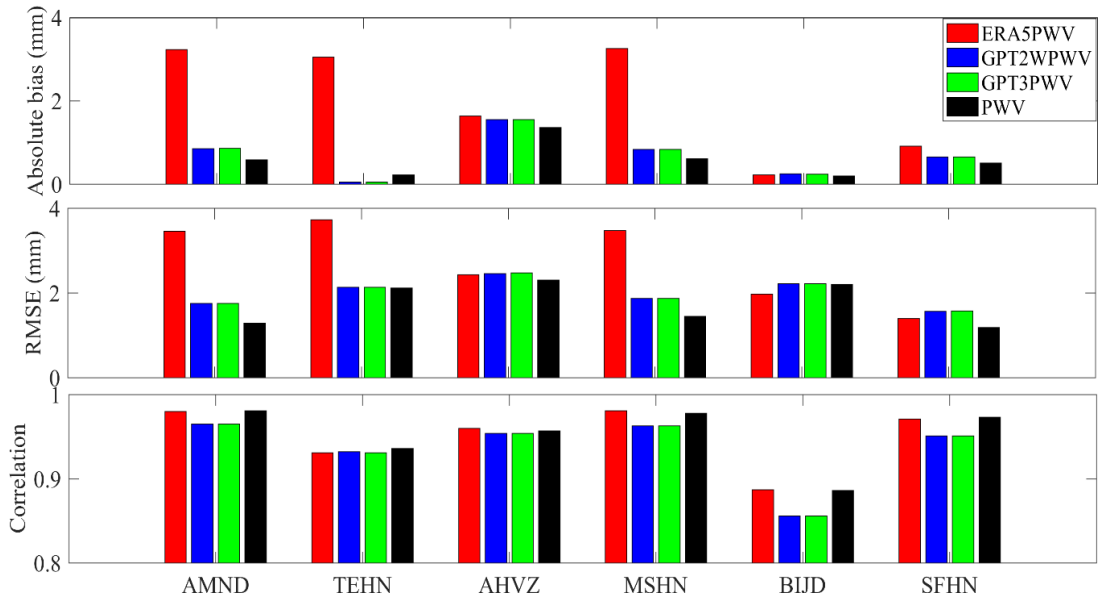

Figure 5. Comparison of mean bias, RMSE and correlation coefficient of GPSPWV computed based on GPT2W (GPT2WPWV), GPT3 (GPT3PWV), ERA5 (ERA5PWV) and real observation (PWV) with respect to the radiosonde derived (RSPWV) values.

As shown in Table 2, the average bias and RMSE of surface pressure obtained from ERA5L in three stations AMND, TEHN and MSHN were much higher than other stations. Surface pressure is a more influential parameter than surface temperature in the accuracy of the final GPSPWV. The RMSE values and the absolute value of the mean bias of ERA5PWV in these three stations are higher than other stations and the corresponding values obtained from experimental models. Again, according to Table 2, it can be seen that in three other stations, the average surface temperature obtained from experimental models in three other stations (AHVZ, BIJD and SFHN) was higher than the corresponding

values obtained for ERA5L reanalysis. However, the effect of surface temperature error in GPSPWV estimation is less than that of surface pressure error. Then, the accuracy of precipitable water vapor estimation in AHVZ, BIJD and SFHN stations using ERA5L temperature and pressure was not better than experimental models.

5 Conclusion

Atmospheric water vapor is one of the valuable parameters that affect weather forecasting, the hydrological cycle, and climate change. GPS meteorology has been regarded as one of the most advanced techniques for estimating PWV over the past two decades. Surface pressure and temperature data can have a significant impact on the accuracy of GPS PWV estimation. If these surface meteorological data are incorrect or not well-calibrated, it can lead to inaccuracies in the estimation of PWV. Therefore, it is important to have reliable surface pressure and temperature data to ensure the accuracy of GPS PWV estimation.

Unfortunately, meteorological sensors are absent from a significant portion of Iran's permanent GPS stations. Therefore, the use of other sources of surface temperature and pressure in this area is of great importance. Reanalysis data and empirical global models are proposed as alternatives to real observations. However, before using these data sources, it is necessary to check their statistical quality. In this study, the statistical quality of one year of ERA5L reanalysis temperature and surface pressure data and the GPT2w and GPT3 models were evaluated in six IPGN GPS stations near which radiosonde observations are available. After synchronizing the data sets of temperature and surface pressure obtained from different sources, they were compared to the actual recorded measurements.

Statistical analysis showed that surface temperature and pressure data extracted from both reanalysis products and global experimental models have a high correlation compared to synoptic station observations. The results showed that, in general, the average bias of the reanalysis pressure data was higher than the corresponding values obtained from the studied experimental models. Meanwhile, the surface temperature data obtained from the experimental models had a greater mean bias than the corresponding values extracted from the reanalysis products.

After the statistical evaluation of surface meteorological data obtained from different methods, the accuracy of its application in GPS PWV retrieval at the studied stations was analyzed. Based on the findings, the study finds that while reanalysis data generally shows a higher correlation with real observations, the experimental models (GPT2w and GPT3) often lead to comparable or better PWV estimates compared to the radiosonde-derived PWV. These results occur particularly at stations where ERA5L pressure data shows a significant bias. Among the six studied stations, the maximum difference in GPSPWV precision obtained from the experimental models and real observations was estimated to be about 0.3 mm in terms of RMSE. In general, the use of surface temperature and pressure data obtained from the GPT2w and GPT3 models had almost the same efficiency in the accuracy of GPSPWV calculation.

It is noteworthy that throughout Iran, there are only six GPS stations that are accessible for radiosonde observations within a 20km radius. Therefore, the results of this study were only examined at these points. Certainly, evaluating surface meteorological data of experimental models and reanalysis products at more points can provide a better comprehensive understanding for those interested in the

field of GPS meteorology. On the other hand, although the results of this research were conducted at a limited number of stations, the studied locations were selected from different regions of the country and had an almost appropriate distribution.

This research highlights the potential of using global empirical models as alternatives to real observations for GPS stations lacking meteorological sensors in Iran, providing valuable insights for improving GPS meteorology techniques in the region. The findings of this study have important implications for enhancing weather forecasting, understanding hydrological cycles, and monitoring climate change in Iran and potentially in other areas with similar challenges in meteorological data availability. For example, GPS PWV estimates can be assimilated into numerical weather prediction models. The use of temperature and pressure obtained from experimental models can help produce near real-time GPS PWV and ultimately predict weather conditions in a timely manner. Additionally, access to accurate water vapor values will be effective in correcting remote sensing images. The results of this study showed that in GPS stations without meteorological sensors, the use of GPT2w and GPT3 models as suitable alternatives can lead to accurate estimates of GPS PWV. Therefore, global experimental models can be used in the correction chain of remote sensing images in Iran.

Those interested in GPS meteorology can statistically evaluate the potential of experimental models in areas with more GPS and radiosonde stations in future studies to achieve more comprehensive and reliable results. They can also use water vapor obtained from methods other than radiosonde in their analyses to assess the accuracy of experimental models in GPS PWV estimation and compare their results with this research.

Acknowledgement

The author of the article acknowledged from the respected referees for their comments and their valuable suggestions to increase the quality of this research. Also, we would like to acknowledge the funding support of Babol Noshirvani University of Technology through Grant program No. 1132/M/P.

References

Alexandrov, M. D., Schmid, B., Turner, D. D., Cairns, B., Oinas, V., Lacis, A. A., ... & Eilers, J. 2009, Columnar water vapor retrievals from multifilter rotating shadowband radiometer data. Journal of Geophysical Research: Atmospheres, 114(D2).

Andrews, D. G. 2010, An introduction to atmospheric physics. Cambridge University Press.

Bai, Z., 2005, Near-Real-Time GPS Sensing of Atmospheric Water Vapour. Queensland University of Technology.

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-ALL SERIES-, 97, 15-787.

Bevis, M., Businger, S., Chiswell, S., Herring, T.A., Anthes, R.A., Rocken, C., Ware, R.H., 1994, GPS Meteorology: Mapping Zenith Wet Delays onto Precipitable Water. Journal of Applied Meteorology (1988-2005), 379-386.

Böhm, J., Heinkelmann, R., & Schuh, H. 2007, Short note: a global model of pressure and temperature for geodetic applications. Journal of Geodesy, 81, 679-683.

Böhm, J., Lagler, K., Schindelegger, M., Krásná, H., Weber, R., & Möller, G. 2013, GPT2: An improved model for tropospheric slant delays in VLBI and GNSS analysis. In *The European*

Navigation Conference Proceedings (p. 4). The European Navigation Conference.

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. Radio Science. 20, 1593-1607.

Divakarla, M. G., Barnet, C. D., Goldberg, M. D., McMillin, L. M., Maddy, E., Wolf, W., ... & Liu, X. 2006, Validation of Atmospheric Infrared Sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts. Journal of Geophysical Research: Atmospheres, 111(D9).

Gao, B.C., Yang, P., Guo, G., Park, S.K., Wiscombe, W.J., Chen, B., 2003, Measurements of water vapor and high clouds over the Tibetan Plateau with the Terra MODIS instrument. IEEE Transactions on geoscience and remote sensing, 41(4), 895-900.

Godson, W. L., 1995, Atmospheric radiation (current investigations and problems). WMO no. 38, p 32

Huang, L., Fang, X., Zhang, T., Wang, H., Cui, L., & Liu, L., 2023, Evaluation of surface temperature and pressure derived from MERRA-2 and ERA5 reanalysis datasets and their applications in hourly GNSS precipitable water vapor retrieval over China. Geodesy and Geodynamics, 14(2), 111-120.

Iwabuchi, T., Naito, I., & Mannoji, N. 2000, A comparison of Global Positioning System retrieved precipitable water vapor with the numerical weather prediction analysis data over the Japanese Islands. Journal of Geophysical Research: Atmospheres, 105(D4), 4573-4585.

Lagler, K., Schindelegger, M., Böhm, J., Krásná, H., & Nilsson, T., 2013, GPT2: Empirical slant delay model for radio space geodetic techniques. Geophysical research letters, 40(6), 1069-1073.

Landskron, D. 2017, Modeling tropospheric delays for space geodetic techniques. Vienna University of Technology.

Li, J., Zhang, B., Yao, Y., Liu, L., Sun, Z., & Yan, X., 2020, A refined regional model for estimating pressure, temperature, and water vapor pressure for geodetic applications in China. Remote Sensing, 12(11), 1713.

Mazany A., Businger S, Gutman S.I., Roeder W. 2002, A lightning prediction index that utilizes GPS integrated precipitable water vapor. Weather Forecast 17(5):1034–1047

Pelosi, A., Terribile, F., D'Urso, G., & Chirico, G. B. 2020, Comparison of ERA5L and UERRA MESCAN-SURFEX Reanalysis Data with Spatially Interpolated Weather Observations for the Regional Assessment of Reference Evapotranspiration. Water, 12(6), 1669.

Pramualsakdikul, S., Haas, R., Elgered, G., Scherneck, H.G., 2007, Sensing of diurnal and semi-diurnal variability in the water vapour content in the tropics using GPS measurements. Meteorolgical Applications. 14, 403-412.

Rothacher, M. 2002, Estimation of station heights with GPS. In Vertical Reference Systems: IAG Symposium Cartagena, Colombia, February 20–23, 2001 (pp. 81-90). Springer Berlin Heidelberg.

Saastamoinen, J. 1972, Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites, In: The Use of Artificial Satellites for Geodesy, Vol. 15, edited by S. W. Henriksen et al., pages 247–251, AGU, Washington, D.C.

Khaniani, A. S., Motieyan, H., & Mohammadi, A. 2021, Rainfall forecast based on GPS PWV together

with meteorological parameters using neural network models. Journal of Atmospheric and Solar-Terrestrial Physics, 214, 105533.

Schindelegger, M.; Pain, G.; Weber, R. 2015, Development of an improved empirical model for slant delays in the troposphere (GPT2w). GPS Solut. 19, 433–441.

Sheffield, J., Goteti, G., & Wood, E. F., 2006, Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. Journal of climate, 19(13), 3088-3111.

Sibylle, V., Dietrich, R., Ru¨lke, A., Fritsche, M., Steigenberger, P., Rothacher, M. 2010, Validation of precipitable water vapor within the NCEP/DOE reanalysis using global GPS observations from one decade. Journal of Climate 23:1675–1695. doi:10.1175/ 2009JCLI2787.1

Singh, V. P., & Woolhiser, D. A. 2002, Mathematical modeling of watershed hydrology. Journal of hydrologic engineering, 7(4), 270-292.

Turner, D. D., Lesht, B. M., Clough, S. A., Liljegren, J. C., Revercomb, H. E., & Tobin, D. C. 2003, Dry bias and variability in Vaisala RS80-H radiosondes: The ARM experience. Journal of Atmospheric and Oceanic Technology, 20(1), 117-132.

Van Baelen, J., Aubagnac, J.P., Dabas, A., 2005, Comparison of near-real time estimates of integrated water vapor derived with GPS, radiosondes, and microwave radiometer. Journal of Atmospheric and Oceanic Technology, 22(2), 201-210

Vaquero-Martínez, J., Antón, M., de Galisteo, J. P. O., Cachorro, V. E., Wang, H., Abad, G. G., and Costa, M. J. 2017, Validation of integrated water vapor from OMI satellite instrument against reference GPS data at the Iberian Peninsula. Science of the total environment, 580, 857-864.

Vedel, H. 2000, Conversion of WGS84 Geometric Heights to NWP Model HIRLAM Geopotential Heights, DMI Scientific Rep. 00-04, Danish Meteorolog. Inst., Copenhagen.

Wang, X.M., Zhang, K.F., Wu, S.Q., Fan, S.J., Cheng, Y.Y. 2016, Water vapor-weighted mean temperature and its impact on the determination of precipitable water vapor and its linear trend, Journal of Geophysical. Research. 121 833e852, https:// doi.org/10.1002/2015JD024181.

Whiteman, D.N., Cadirola, M., Venable, D., Calhoun, M., Miloshevich, L., Vermeesch, K., Twigg, L., Dirisu, A., Hurst, D., Hall, E., Jordan, A., Vömel, H. 2012, Correction technique for Raman water vapor lidar signal-dependent bias and suitability for water vapor trend monitoring in the upper troposphere. Atmospheric Measurement Techniques, 5(11), 2893-2916

Zhang, W., Zhang, H., Liang, H., Lou, Y., Cai, Y., Cao, Y., ... & Liu, W., 2019, On the suitability of ERA5 in hourly GPS precipitable water vapor retrieval over China. Journal of Geodesy, 93, 1897-1909.