Developing Iran's empirical zenith wet delay model (IR-ZWD)

Dehvari, Masood; Farzaneh, Saeed; Forootan, Ehsan

Published in:
Journal of Atmospheric and Solar-Terrestrial Physics

DOI (link to publication from Publisher):
10.1016/j.jastp.2023.106163

Publication date:
2023

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):
Dehvari, M., Farzaneh, S., & Forootan, E. (2023). Developing Iran's empirical zenith wet delay model (IR-ZWD). *Journal of Atmospheric and Solar-Terrestrial Physics*, *253*, Article 106163.
https://doi.org/10.1016/j.jastp.2023.106163

Developing Iran's empirical Zenith Wet Delay model (IR-ZWD)

Masoud Dehvari[1], Saeed Farzaneh[1,*], Ehsan Forootan[2],

[1]School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

[2]Geodesy Group, Department of Planning, Aalborg University, Rendburggade 14, 9000 Aalborg, Denmark.

[*]Corresponding author: Associate Professor, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran; * Farzaneh@ut.ac.ir

Abstract

The presence of water vapor in the lower atmosphere can introduce errors in satellite-based geodetic observations. Accurate modeling of this part of atmospheric delay is particularly challenging due to the considerable variations of water vapor. Therefore, constructing a reasonable model to predict Zenith Wet Delay (ZWD) can improve the accuracy of geodetic observations and positioning techniques. In this study, we aim at constructing a regional ZWD model for Iran and nearby regions (called the IR-ZWD model) using base functions with local support. The mode is based on the five-year outputs of the Empirical Reanalysis Fifth generation (ERA5) data with the spatial resolution of about 0.25 degree from 2017-2021. The B-spline base functions are used to effectively represent local spatial changes in the spectral domain and to decrease the number of unknown parameters. A B-spline model with the order and surface resolution of about 3 and 5 (scalar values) is found to be efficient, which has an equivalent spatial resolution of ~0.5 degree. Temporal variations are accounted for by applying a constant term, along with periodic components with annual, semi-annual, 3-, and 4-monthly periods. Our results demonstrate that the proposed model has a mean Root Mean Squared Error (RMSE) of about 0.035 m within Iran, which represents an improvement of approximately 12.5% compared to the commonly used global empirical models such as GPT3w, GTrop, and HGPT2. The squared correlation coefficient value of 0.55 is found between IR-ZWD and ERA5 data, which is about 10% higher than that of, e.g., GPT3w and GTrop. The IR-ZWD model is also evaluated against five radiosonde stations and ZWD from the Jason-3 satellite mission. In both cases, the results indicate that IR-ZWD can reduce the RMSE and MAE values of about 10%, and it improves the squared correlation coefficient value about 9%.

Commented [EF1]: Why squared?

Commented [EF2]: Why squared?

Keywords: Empirical model; Zenith Wet Delay (ZWD); ERA5; B-spline; radiosonde station; Jason-3 mission

1. Introduction

Satellite geodetic observations of the Global Navigation Satellite System (GNSS) have been frequently used for positioning, navigation, and remote sensing of atmospheric parameters (Bender et al., 2011; Bevis et al., 1992; Forootan et al., 2021). GNSS observations are of very high temporal resolution with the benefits of all-weather capabilities, and they induce relatively low cost (Sun et al., 2017b). However, these observations encounter errors caused by the movement of GNSS signal through the atmosphere, namely the ionospheric and tropospheric errors. The ionosphere delay can be mitigated via dual frequency observations in the form of ionosphere-free or other types of combinations (Subirana et al., 2013). Therefore, the remaining tropospheric delay, also known as the Zenith Total Delay (ZTD), is a considerable error source, which cannot be reduced by combining GNSS bands (Sun et al., 2017a). The ZTD consists of two parts, i.e., the Zenith Hydrostatic Delay (ZHD) and the Zenith Wet Delay (ZWD), where ZHD can be modeled quite precisely using surface pressure and temperature observations (Davis et al., 1985; Dogan and Erdogan, 2022). The spatial and temporal variations of the ZWD are, however, driven by weather and might contain many local features that make it difficult to be accurately modeled (Forootan et al., 2021; Tunalı and Özlüdemir, 2019). Presenting ZWD with high accuracy can be helpful for predicting seasonal weather, and increasing the accuracy of space-based geodetic observations, especially enhancing the accuracy of the Single Point Positioning (SPP) technique (Kalita and Rzepecka, 2017; Tregoning and Herring, 2006; Vedel and Huang, 2004; Yan et al., 2009). Therefore, it is beneficial to construct models that present ZWD as precisely as possible .

Many empirical models have been constructed to mitigate the effects of tropospheric delay, which can be separated into two different types based on the required parameters. The first one consists of empirical models such as the Hopfield model (Hopfield, 1969), the Saastamoinen model (Saastamoinen, 1972), and the Black model (Black, 1978), which requires surface meteorological parameters for the calculation of atmospheric delays. However, due to a lack of measurements about the vertical profile of water vapor, the reliability of these models is found to be low (Yao and Hu, 2018). Also, the real-time application of these models in positioning and navigation may be limited due to the high dependence of models on meteorological measurements (Sun et al., 2017b). The second category comprises empirical models that are constructed based on numerical analysis or reanalysis datasets such as ERA5 (Hersbach et al., 2020), ERA-Interim (Dee et al., 2011), or the Global Geodetic Observing System (GGOS) (Plag et al., 2009). From these models, the UNB series (Collins and Langley, 1997), EGNOS (Penna et al., 2001), IGGtrop (Li et al., 2012), GZTD Series (YAO et al., 2013; Yao et al., 2016), GZTDS (Sun et al., 2017a), GPT2w (Böhm et al., 2015), IGPT2W (Du et al., 2020), GPT3w (Landskron and Böhm, 2018), GEOFT (Sun et al., 2017b), GTrop (Sun et al., 2019), EGtrop (Ma et al., 2021), AGtrop (Ma et al., 2022), and the HGPT series (Mateus et al., 2020; Mateus et al., 2021) have been constructed with global coverage but with different spatial and temporal resolutions. The data source for the construction of the mentioned models and the corresponding spatial resolution of each model are listed in Table 1.

Table 1. An overview of the available empirical atmospheric delay models.

Model	Spatial resolution(deg)	Source
UNB3	15	Multi source observations
EGNOS	15	European Centre for Medium-Range Weather Forecast (ECMWF) reanalysis dataset
IGGtrop	2.5	National Centers for Environmental Prediction (NCEP) reanalysis data
GZTD	18	GGOS ZTD data

GZTD2	10	GGOS ZTD data
GZTDS	2 × 2.5	GGOS ZTD data
GPT2w	1	ERA-Interim
IGPT2W	1	ERA-Interim
GPT3w	1	ERA-Interim
GEOFT	2.5	GGOS
GTrop	1	ERA-Interim
EGtrop	12	ERA5
AGtrop	1	ERA5
HGPT2	0.25	ERA5

The empirical models in Table 1 present the global average of tropospheric delay at the best spatial resolution of about 1° (except HGPT2 model). Considering the high spatial changes of water vapor, it can be concluded that these models have limitations for representing the regional spatio-temporal changes of ZWD values. Also, these models calculate atmospheric delay in a grid-based approach or using a functional model (e.g., in terms of spherical harmonics coefficients) with global support. Compared with grid-based models, functional models might be more numerically stable, because a smaller number of parameters needs to be computed to numerical build the model. For example, for 1-degree resolution global ZWD model one needs to estimate 64800 parameters in the grid domain, which is equivalent with 32761 coefficients in terms of spherical harmonics up to degree and order 180. Among the empirical models listed in Table 1, the HGPT2 model is a grid-based model for atmospheric parameters, featuring a spatial resolution of approximately 0.25 degrees. This model is Constructed using 20 years of ERA5 data and incorporating a time segmentation concept, which boasts an hourly temporal resolution.

The aim of this study is to produce an empirical ZWD model for Iran and nearby regions, which is called here IR-ZWD. This model is constructed based on five years of the high-resolution ERA5 dataset from January 1, 2017 to 31 December 2021. IR-ZWD considers a constant component along with the annual, semi-annual, 3-, and 4-monthly harmonics to account for the temporal variations of the ZWD values. To represent the spatial anomaly maps, the B-spline base functions (Limberger, 2015) are used. Compared with the functional model with global coverage (e.g. spherical harmonics in (Dehvari et al., 2023; YAO et al., 2013; Yao et al., 2016), the B-splines can retrieve the local features of ZWD with a lower number of required unknown coefficients. Therefore, our motivation for selecting this configuration is the usage of a high-resolution dataset (ERA5) for the reconstruction of the empirical ZWD model using base functions with local support that benefits from a lower number of unknown parameters compared to the grid-based model. The proposed model provides ZWD values with spatial resolution of about 0.5 degrees (B-spline functions with order 3 and surface resolution of about 5). For constructing this model in the grid domain, one must use 36936 unknown points, 25921 spherical cap harmonics coefficients (Al-Fanek, 2013; Forootan et al., 2021), where the cap extension is required to solve the local orthogonality.

To evaluate the performance of IR-ZWD, the outputs are compared with those of GPT3w (Landskron and Böhm, 2018), GTrop (Sun et al., 2019), and HGPT2 (Mateus et al., 2021) models, as well as the outputs of ERA5 during 1 January 2022 to 31 December 2022, which are not used within the IR-ZWD model. Additionally, the outputs of IR-ZWD are compared with 5 located radiosonde stations in Iran and the ZWD estimated from the Jason-3 satellite altimetry mission (Dumont et al., 2016) over the Caspian Sea and the Persian Gulf.

In what follows, in section 2, the datasets and the study region are introduced. In Section 3, the construction of the proposed empirical model is explained. In Section 4, the numerical results of the study are presented, and finally in section 5, this study is concluded.

2. Data and region of study

2.1. Data

The European Centre for Medium-Range Weather Forecasts (ECMWF) has introduced the Empirical Reanalysis Fifth Generation (ERA5) global meteorological model, which has a spatial resolution of about 0.25 degrees and an hourly temporal resolution for 37 different pressure levels ranging from 1 hPa to 1000 hPa (baba shaeb Kannemadugu et al., 2022; Forootan et al., 2021). This dataset is freely accessible from https://cds.climate.copernicus.eu/. The numerical model provides a wide range of meteorological parameters, including temperature, geopotential, and relative humidity, which can be used to calculate wet refractivity indices at different pressure levels. The geopotential data is employed to determine the corresponding ellipsoidal height of the dataset (Ma et al., 2021). By incorporating the remaining meteorological parameters, the wet refractivity indices can be calculated using :

Commented [EF3]: Check the reference!

$$e = \frac{RH.\alpha_1.e^{(\frac{\alpha_2.t}{t+\alpha_3})}}{100}, \tag{1}$$

$$N_w = 3.732 * 10^5 * \frac{e}{T^2}, \tag{2}$$

(Forootan et al., 2023), where e is the water vapor pressure in hPa, RH is relative humidity, t and T are temperature in Celsius and Kelvin, respectively, and N_w is the wet refractivity index. The constant coefficients of α_1, α_2, and α_3 are selected to be 6.1121, 17.502, and 240.97 respectively. Using the wet refractivity indices of Eq. (?) at different altitudes, the Zenith Wet Delay (ZWD) can be calculated by integrating them over the zenith direction as in (Bevis et al., 1992):

$$ZWD = 10^{-6} \int_L N_W \, dl \ . \tag{3}$$

The aim of this study was to develop a regional model for estimating ZWD values. To achieve this goal, we used ERA5 data that include temperature, geopotential, and relative humidity from January 1, 2017 to December 31, 2021. Since the wet refractivity indices tend to zero for altitudes above 10 km, we selected only 16 levels (out of 37) from 1000 hPa to 250 hPa of ERA5 pressure levels. Additionally, we used the data with a time resolution of approximately 6 hours (0, 6, 12, and 18 UTC times) to establish the regional ZWD model. This is also justified because the finest harmonics base functions used in the regional model has a period of three months (see section 4.1). The study region is shown in Fig. 1 and contains 4104 unique grid points, for which we calculated ZWD values using the 6-hourly inputs of 2017-2021.

2.2. The region of study

The study area has been selected to include Iran and the nearby regions, covering the latitudes between 25.5 to 39.5° and the longitudes 44.5 to 62.5°, with its altitudes ranging from approximately -37 to 2524 m (Fig. 1). This region is located in the mid-latitude zone and is influenced by its proximity to the Caspian Sea, the Persian Gulf, as well as Zagros, and Alborz Mountains, thus, resulting in diverse climates with local variations (Heydarizadeh Shali et al., 2020). Some regions experience arid and hot climates with low water vapor content, while others exhibit a more moderate climate with high water vapor content and rapid changes. To evaluate the proposed model's derived ZWD values, the study uses observations from 5 existing radiosonde stations and the valid microwave radiometer of Jason-3. Fig. 1 shows the location of the radiosonde stations and Jason-3 ZWD observations in the study area, and Table 2 lists the geographical coordinates of the radiosonde stations.

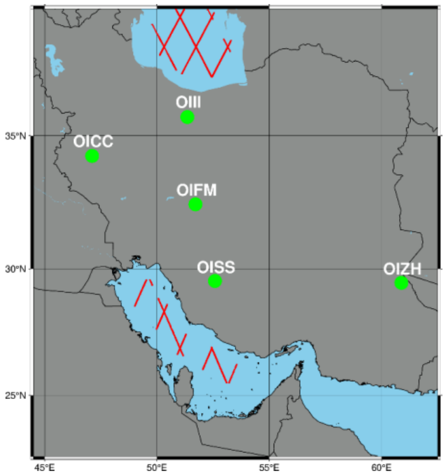


Fig. 1. The presentation of the region of study. The green dots show the location of 5 radiosonde stations (location are reported in Table 2) in the region. Also, the red dots display the Jason-3 ZWD observations (cycles 217-325), including the path numbers ?.

Table 2. Geographical coordinates of radiosonde stations.

Station name	Latitude(deg)	Longitude(deg)	Altitude(m)
OICC	34.26	47.11	1322
OIFM	32.46	51.71	1550
OIII	35.68	51.35	1191
OISS	29.53	52.58	1491
OIZH	29.46	60.88	1370

3. Methodology

3.1. Quadratic B-spline Functions

The B-spline functions are often used for local and global signal localization due to their compact support (Nohutcu et al., 2010). These functions are commonly implemented in Euclidean space and can be used to decompose target parameters into a series of detailed signals via consecutive low-pass filters (Limberger, 2015). A one-dimensional signal can be expanded in terms of B-spline functions using:

$$\mathrm{g}(\mathrm{x}) = \sum_{\mathrm{h}=0}^{h_s} \alpha_h^s \varphi_{\mathrm{h}}^{\mathrm{m,s}}(\mathrm{x}), \tag{4}$$

(Schmidt et al., 2008), where $\mathrm{g}(.)$ is the considered signal, α_k^s are the coefficients of the base function, $\varphi_{\mathrm{k}}^{\mathrm{s}}$ are the kernel of the normalized polynomial B-splines of order m, s is the surface resolution for corresponding kernel, and h_s is set of polynomial B-splines for selected surface resolution and is equal to:

$$h_s = 2^s + m - 1. \tag{5}$$

The number of coefficients for retrieving g(x) is also equal to h_s. The base function with order m and the corresponding surface resolution can be calculated in a recursive procedure as:

$$\varphi_{\mathrm{h}}^{\mathrm{m.s}}(\mathrm{x}) = \frac{x - t_h^s}{t_{h+m}^s - t_h^s}\varphi_{\mathrm{h}}^{\mathrm{m-1.s}}(\mathrm{x}) + \frac{t_{h+m+1}^s - x}{t_{h+m+1}^s - t_{h+1}^s}\varphi_{\mathrm{h+1}}^{\mathrm{m-1.s}}(\mathrm{x}) \tag{6}$$

$$\varphi_{\mathrm{h}}^{0.\mathrm{s}}(\mathrm{x}) = \begin{cases} 1 & t_h^s \le x < t_{h+1}^s \\ 0 & else \end{cases}. \tag{7}$$

(Limberger, 2015). In these expressions, t_h^s are knots and the number of them are equals to $2^s + 2$ and control the spatial resolution of B-spline functions (Schmidt et al., 2008). For the 2D modeling of a parameter, Eq. (1) can be expanded as:

$$\mathrm{g}(\mathrm{x}_1.\mathrm{x}_2) = \sum_{\mathrm{h}_1=0}^{h_{s_1}} \sum_{\mathrm{h}_2=0}^{h_{s_2}} \alpha_{h_1.h_2}^{s_1.s_2} \varphi_{h_1}^{\mathrm{m}.s_1}(\mathrm{x}_1) \varphi_{h_2}^{\mathrm{m}.s_2}(\mathrm{x}_2)\,. \tag{8}$$

In this case, the number of coefficients ($\alpha_{h_1.h_2}^{s_1.s_2}$) are about $h_{s_1} \times h_{s_2}$. In order to reconstruct the $\mathrm{g}(\mathrm{x}_1.\mathrm{x}_2)$ function, the mentioned coefficients need to be estimated, where the computation procedure follows a least squares optimization, see section 3.2.

3.2. The IR-ZWD model

The values of ZWD exhibit a periodical change, which is usually modeled using sinusoidal and cosine functions (YAO et al., 2013; Yao et al., 2016). Based on the characteristics of ZWD values, the considered periodical model can be written as follows:

$$ZWD(\varphi.\lambda.h) = \left[a_0(\varphi.\lambda) + \sum_{i=1}^{4}\left(a_i(\varphi.\lambda)\cos\left(\frac{i.2\pi.doy}{365.25}\right) + b_i(\varphi.\lambda)\cos\left(\frac{i.2\pi.doy}{365.25}\right)\right)\right]e^{-\vartheta h} \tag{9}$$

$$a_i(\varphi,\lambda) = \sum_{\mathrm{h}_1=0}^{h_{s_1}}\sum_{\mathrm{h}_2=0}^{h_{s_2}} \alpha_{h_1.h_2}^{s_1.s_2}\,\varphi_{h_1}^{\mathrm{m}.s_1}(\varphi)\varphi_{h_2}^{\mathrm{m}.s_2}(\lambda). \tag{10}$$

where φ and λ are the geographical latitude and longitude, h is the height, doy is the day of the year, ϑ is the height scale, and is about -0.00013137 (Yao et al., 2016). The considered periodical variation of ZWD values are the mean value (a_0), annual variations (a_1 and b_1), semi-annual variations (a_2 and b_2), 4-monthly variations (a_3 and b_3), and 3-monthly variations (a_4 and b_4). In this model, the coefficients a_i & b_i , $i = 1,2,3,4$, are determined using B-spline functions. Therefore, for implementing the IR_ZWD models, the B-spline coefficients ($\alpha_{h_1,h_2}^{s_1,s_2}$) for each amplitude must be determined. Here the ZWD values are computed using the ERA5 data of 2017-2021. The corresponding B-spline coefficients are computed using the least squares method (Koch, 2007). After the estimation of B-spline coefficients, the amplitude (*Amp*, except for the mean value (a_0)) can be calculated as:

$$Amp(\varphi,\lambda) = \sqrt{a_i(\varphi,\lambda)^2 + b_i(\varphi,\lambda)^2}. \tag{11}$$

To build the B-spline expansion of the ZWD fields, one needs to select the order and resolution of B-splines. The higher number of parameters, results in a higher resolution. However, considering too high number for these parameters results in huge number of model coefficients which are not desired (reference?). Therefore, the accuracy of the results is correlated with these parameters and needs to be examined. Our numerical experiment is described in section 4.1.

For examining the accuracy and reliability of the proposed model, the Root Mean Squared Error (RMSE), the squared Correlation Coefficients (CC), and the Mean Absolute Error (MAE) are calculated as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} ZWD_m^i - ZWD_o^i. \tag{12}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\,ZWD_m^i - ZWD_o^i\,)^2}{n}}. \tag{13}$$

$$CC = 1 - \left(\frac{\sum_{i=1}^{n} \left(ZWD_m^i - ZWD_o^i \right)^2}{\sum_{i=1}^{n} \left(ZWD_o^i - \overline{ZWD_o} \right)^2} \right). \quad (14)$$

In the mentioned expressions, ZWD_m and ZWD_o correspond to the modeled and observed values, respectively, and $\overline{ZWD_o}$ represents the mean observed value.

4. Results

4.1. Determination of the B-spline parameters

The spatial resolution of the functional models depends on the properties of their base functions, which is in this case, consists of the defined B-spline parameters in Eq. (10) . Besides, the temporal resolution of the model is controlled by the number and frequency of the (temporal) harmonics in Eq. (9). Therefore, for constructing the IR-ZWD model, the order of temporal resolution (i in Eq. (9)), the surface resolution of B-splines (s), and the order of B-splines (m) in Eq. (10) needs to be fixed. These parameters have been determined empirically by developing the IR-ZWD model for the considered groups of parameters. For this, in an empirical approach, the model has been developed for different groups of parameters and the corresponding performance has been examined (Al-Fanek, 2013; Forootan et al., 2021; Razin and Voosoghi, 2017). For implementing the B-spline, the order and surface resolutions are considered to be between 2 and 3 and 2 to 5, respectively. The temporal resolution is selected to be up to semi-annual variations ($i = 1$ and 2), up to 4-month variations ($i = 1.2.$ and 3), and up to 3-month variations ($i = 1.2.3$, and 4). For each case, a set of model parameters (e.g. B-spline order, B-spline surface resolution, and temporal resolution) has been used for developing the proposed model using 20% of the ERA5 data from years 2017-2021 that have been randomly used. Afterwards, ZWD values are computed using the developed model, for each case, and the RMSE value (Eq. (13)) was calculated with respect to the original ERA5 values. Accordingly, the group of parameters that provided the minimum mean RMSE is selected as the “optimal” functional model parameters. Table 3 reports the results of this comparison, where the 3 last columns list the RMSE values for different temporal resolutions of the proposed model.

Table 3. RMSE values for various group of functional model parameters.

		Semi-annual (i=2)	**4-monthly (i=3)**	**3-monthly (i=4)**
B-spline order(m)	**B-spline resolution(s)**	**RMSE(m)**	**RMSE(m)**	**RMSE(m)**
2	**2**	0.040	0.0393	0.0382
2	**3**	0.0388	0.0380	0.0369
2	**4**	0.0380	0.0372	0.0361
2	**5**	0.0374	0.0367	0.0355
3	**2**	0.0395	0.0388	0.0376
3	**3**	0.0386	0.0379	0.0367
3	**4**	0.0379	0.0371	0.0360

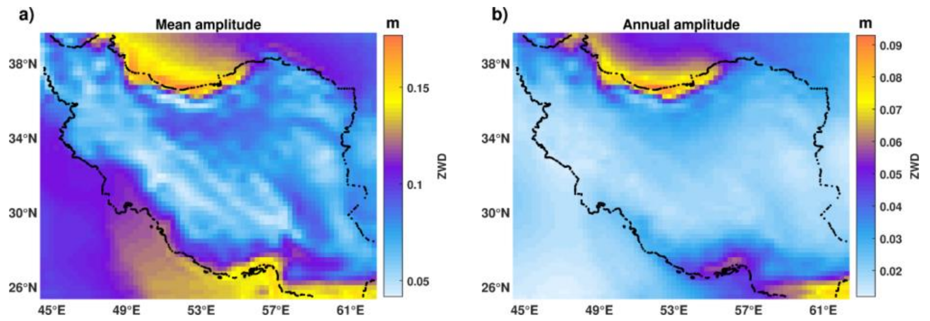
3	**5**	0.0374	0.0367	0.0355

Given the RMSE values in Table 1, it can be concluded that by increasing the number of model parameters, the RMSE value drops, and the minimum RMSE derived by considering model parameters as $m = 3. s = 5. i = 4$. As mentioned, the motivation for using the B-spline function as the functional model was to decrease the model parameters compared to the grid-based approach. Considering the grid-based formulation with a temporal resolution up to 3-months results in 9 model parameters for each point, and overall 36936 parameters for this model. However, our proposed model, that contains B-spline functions, contains10404 parameters, which are less than the grid-based approach making the computation more stable. The region of study covers about $18^{\circ} \times 14^{\circ}$ in longitude and latitude directions. According to the considered surface resolution of the B-spline, the number of knots is about 32. Therefore, the spatial resolution of the considered B-spline functions is $\sim 0.5^{\circ}$. As a result, given the geographical coordinates and the day of the year, using the IR-ZWD model, the ZWD value can be calculated for the entire study region.

Commented [EF4]: Table 3?

4.2. The amplitude of temporal variations

The regional ZWD model was constructed using ERA5 data from 2017 to 2021, and the corresponding B-spline model with $m = 3. s = 5. i = 4$ as model parameters. Using Eq. (11), the amplitude of each considered periodic variation for the corresponding regional model can be calculated. Fig. *2* displays the amplitudes of the mean, annual, semi-annual, 4-month, and 3-month variations over the study area.



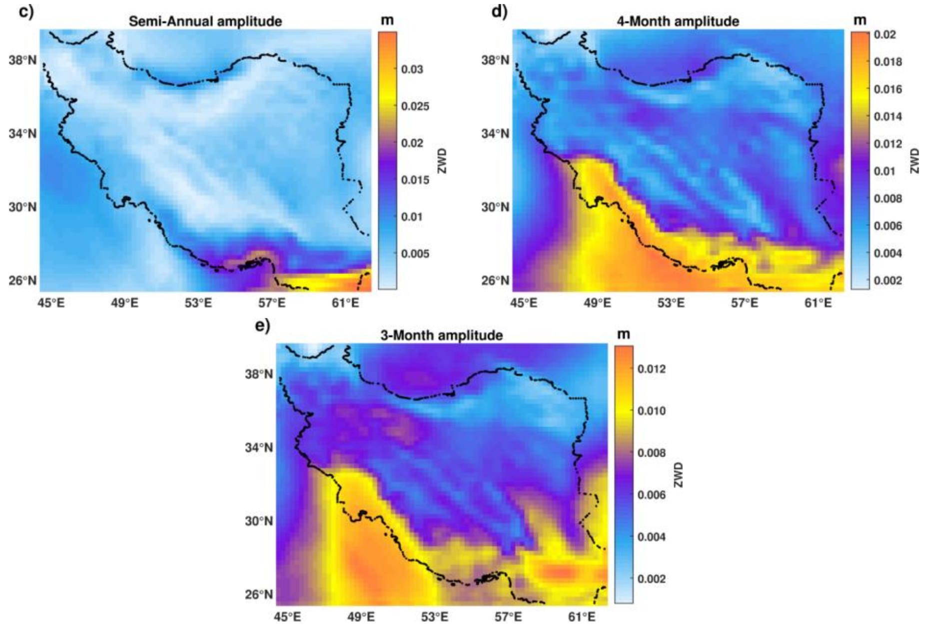


Fig. 2. An overview of the amplitude of the mean (a), annual (b), semi-annual (c), 4-monthly (d), and 3-monthly (e) variations of IR-ZWD over the considered region.

From Fig. 2, it can be observed that the reduction of the period of variations is accompanied by a decrease in the amplitude of the corresponding component. As depicted in Fig. 2, the maximum mean amplitude of ZWD corresponds to the regions over or near the oceans in the study area, which is expected due to the higher amount of water vapor in these regions. Similarly, the maximum amplitude of the annual variations also corresponds to these regions due to the high contribution of water vapor content. Moreover, the strongest semi-annual amplitude is observed in the southeast regions, reaching up to 3 centimeters, and can be related to the monsoon effect (Vuille et al., 2005). We found the near ocean region such as those in the southern part of the study area to be associated with the maximum amplitude of the 4- and 3-monthly variations, which reach up to 2 and 1 cm, respectively. This is due to seasonal changes in the ZWD values which show a higher amplitude in coastal areas.

Using the estimated B-spline coefficients, geographical longitude and latitude, and the day of the year (DOY), ZWD can be calculated using the proposed regional model. In the following, the constructed model has been evaluated using ERA5 ZWD data for the entire year 2022. For this case, the ZWD values derived from the proposed model, as well as those calculated from the GPT3w, GTrop, and HGPT2 models, will be compared with the corresponding ERA5 values. Afterwards, to evaluate the proposed model using a different dataset, the obtained ZWD values from the IR_ZWD model will be compared with the corresponding values from six radiosonde stations and Jason-3 radiometer measurements in the study region.

4.3. Comparison with the ERA5 ZWD values

In this section, the performance of models is compared to the ERA5 ZWD values for the time period of January 1, 2022 to December 31, 2022. For this, we calculated ZWD values from IR-ZWD, GPT3w, GTrop, and HGPT2 at the same grid points as the ERA5 dataset and with a time resolution of approximately 6 hours (0, 6, 12, and 18 UTC times). The calculated ZWD time series from these models are then compared with the corresponding ERA5 values at each grid point, and the statistical values are computed for each grid point. For example, Fig. 3 shows the comparison for a grid point with the latitude=39°, the longitude=45.75°, and the altitude=833m. This location is selected because that of IRI-ZWD was different from other models.

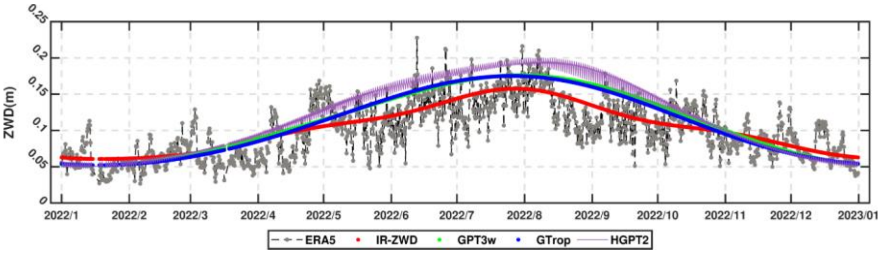
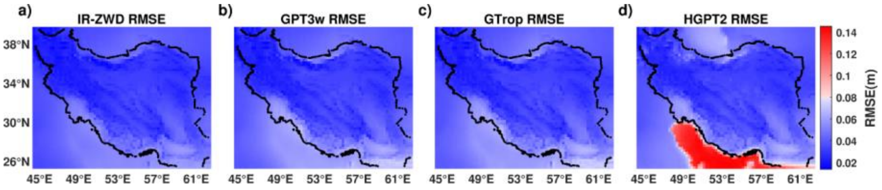


Fig. 3. The ZWD values of a grid point with latitude=39°, longitude=45.75°, and altitude=833m from ERA5 data, IR-ZWD, GPT3w, GTrop, and HGPT2 models with gray, red, green, blue, and purple lines, respectively.

In Fig. 3, the RMSE of the IR-ZWD, GPT3w, GTrop , and HGPT2 models are found to be about 0.025, 0.032, 0.031, and 0.039 m, respectively. This shows IR-ZWD provides a slightly better regional fit to ERA5. Fig. ***4*** displays the calculated statistical values for each model at each grid point across the study region. Additionally, Fig. 5 depicts the Taylor diagram for the mean RMSE and CC values obtained from the evaluation of these models.



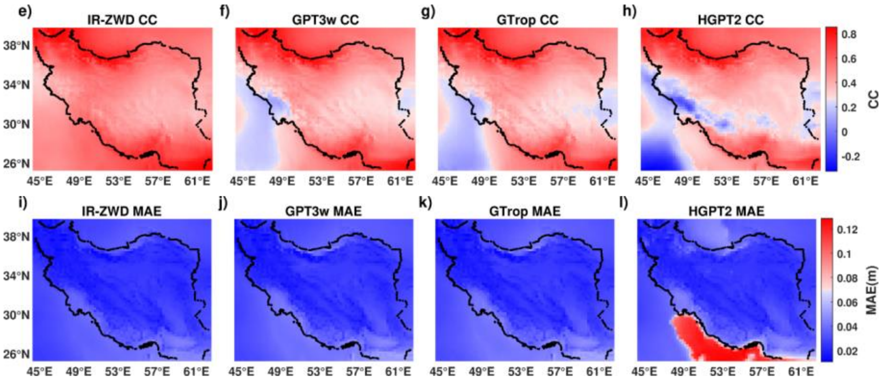


Fig. 4. An overview of the RMSE values for the IR-ZWD, GPT3w, GTrop, and HGPT2 models over the study region, which are in the subplot a, b, c, and, d, respectively. Figures e, f, g, and h show the corresponding CC values of the considered models. Also, the distribution of the calculated MAE values for considered models has been depicted in i, j, k, and l figures. These parameters are calculated through comparison with the ERA5 ZWD values from 1 January 2020–31 December 2022.

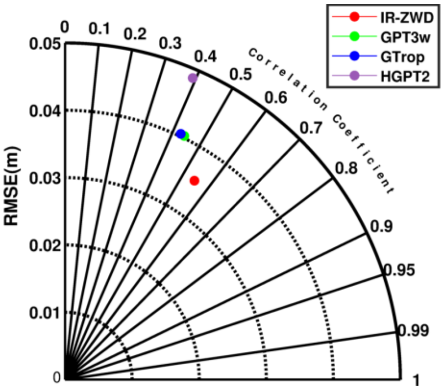


Fig. 5. An overview of the calculated mean RMSE and CC values over the region of study based on the Taylor diagram.

Considering Fig. 4a, b, c, and d, we found that the RMSE values of the IR-ZWD model to be lower than the corresponding values of GPT3w ,GTrop, and HGPT2 models in regions over or nearby oceans or seas. These areas have the highest water vapor content and its variations. Therefore, developing the IR-ZWD model with a higher spatial and temporal resolution has improved the accuracy in these regions. Additionally, Fig. 5 shows that the mean RMSE of IR-ZWD is about 0.035 m, indicating an improvement of approximately 12.5% compared to the GPT3w and GTrop models, and improvement of about 27.5% compared to HGPT2 model. Comparing the CC values in Fig. 4e, f, g, and h, it can be seen that the proposed model greatly improved the CC values in regions where the corresponding values of GPT3w, GTrop, and HGPT2 were low (the region with the blue colors in Fig. 4f, g, and h). Moreover, as illustrated in Fig. 5, the mean CC of IR-ZWD, i.e., about 0.55, is approximately 10, 11, and 15% higher than the corresponding CC values of GPT3w, GTrop, and HGPT2. The mean MAE values for IR-ZWD, GPT3w, GTrop, and HGPT2 models are found to be about 0.028, 0.030, 0.030, and 0.037 m, respectively. These values indicate an improvement of approximately 7, 7, and 29% in the MAE value of IR-ZWD compared to the GPT3w, GTrop, and HGPT2 models. By comparing the statistical parameters in Fig. 5, it can be seen that GPT3w slightly outperforms the GTrop model over the study region. It should be noted that the spatial resolution of the two models is the same and GTrop is constructed with a higher temporal resolution. Referring Fig. 4d and l, it becomes evident that, for the southern part of the study region (including the Persian Gulf and the Oman Sea), the RMSE and MAE values of the HGPT2 model are higher than those of other models.

To compare the performance of IR-ZWD at different altitudes, we examined the mean RMSE calculated at each grid point of ERA5 data (Fig. 4a, b, c, and d) and considered the corresponding altitudes of the points. Fig. ***6*** shows the mean RMSE of each ERA5 grid point for the IR-ZWD, GPT3w, GTrop, and HGPT2 models, plotted against the altitude of each point. The vertical profile of the mean RMSE for each model is also shown in this figure.

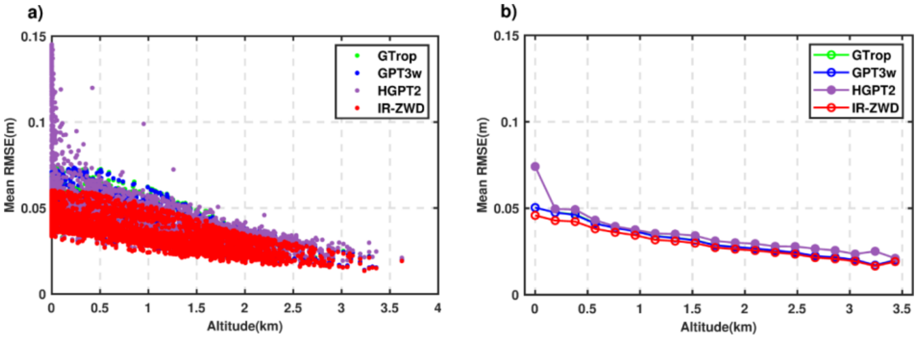


Fig. 6. An overview of a) the comparison of the mean RMSE values for different grid points with respect to the altitude of each point. the mean RMSE of GTrop, GPT3w, HGPT2, and IR-ZWD models are displayed with green, blue, purple, and red lines respectively. In b) the vertical profile of the mean RMSE value for each model is provided.

From Fig. 6, it can be seen that the performance of the IR-ZWD model for all altitudes is relatively closer to observations compared to other models. Fig. ***6***b compares the mean RMSE of three models for points with a low altitude (between 0 and 1 km), which that of IR-ZWD is 0.045 m compared with 0.05, 0.05, and 0.074 m from GPT3w, GTrop, and HGPT2 models. while the performance of all three models is found almost similar for points with high altitudes. According to the RMSE values depicted in Fig. 6a, for the HGPT2 model, one can infer that the model's accuracy is comparatively lower for points with low altitudes compared to the GPT3w and GTrop models.

4.4. Comparison with the radiosonde measurements

A radiosonde is an instrument attached to a balloon that rises up to different atmospheric layers and measures several atmospheric parameters that can be used to calculate wet refractivity indices and thus ZWD values. These measurements are provided twice per day (at 12 and 24 UTC time) and, due to their high accuracy, are always considered as a reference for evaluation in atmospheric-related research (Adavi and Mashhadi-Hossainali, 2014; Bender et al., 2011; Forootan et al., 2023; Forootan et al., 2021). In this section, to further evaluate the estimated ZWD values from models, they are compared with the data from five existing radiosonde stations in the study region (Fig. 1). This evaluation is done over the period of January 1, 2022 to December 31, 2022, where data were not used in the modelling but were available for validation. The radiosonde data have been obtained from http://weather.uwyo.edu/upperair/sounding.html. Fig. 7 shows the results of IR-ZWD, GPT3w, GTrop, and HGPT2 ZWD. The calculated statistical parameters are depicted in Fig. 8.

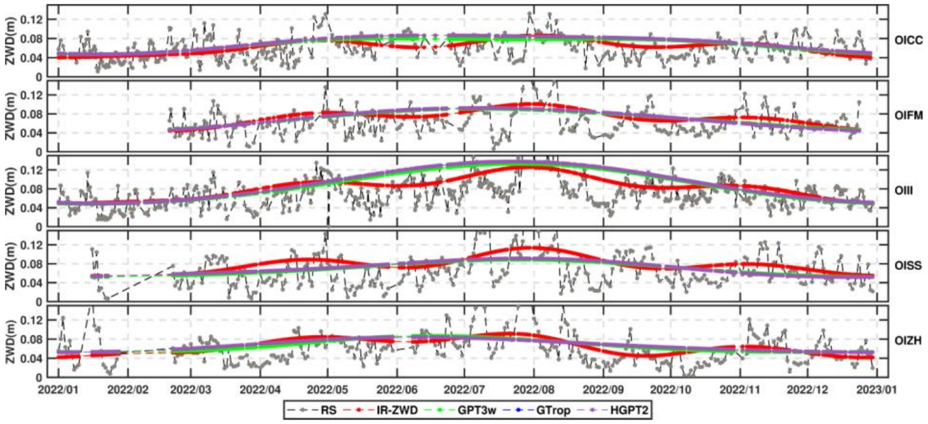


Fig. 7. An overview of the validation of ZWD estimated from the four considered models with corresponding measurements of 5 existing radiosonde stations in the study region. The gray, red, green, blue, and purple dots correspond to the radiosonde, IR-ZWD, GPT3w, GTrop, and HGPT2 ZWD values, respectively.

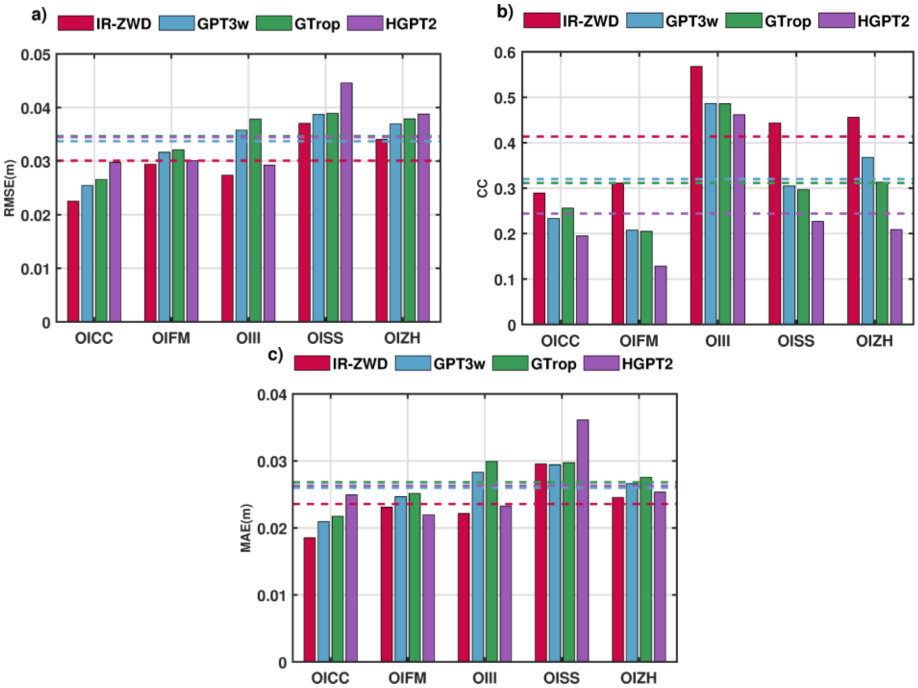


Fig. 8. A summary of the calculated RMSE, CC, and MAE statistic parameters for each radiosonde station in a, b, and c, respectively. The bars with the red, cyan, green, and purple colors are for the IR-ZWD, GPT3w, GTrop, and HGPT2 models, respectively. Also, the mean value of each parameter has been shown with the corresponding dashed lines.

Fig. ***8***a shows that the mean RMSE of IR-ZWD model is about 0.03 m, indicating an improvement of about 11%, 13%, and 13% compared to GPT3w, GTrop, and HGPT2, respectively. In Fig. 8b, the mean CC value of IR-ZWD is found to be about 0.41, which is approximately 10, 10, and 17% higher than the corresponding values for the GPT3w, GTrop, and HGPT2 models, respectively. Additionally, Fig. 8c displays the IR-ZWD mean MAE value of 0.023 m, indicating an improvement of about 9, 12, and 12% compared to the GPT3w, GTrop, and HGPT2 models, respectively. These statistical measures are found to be consistent with those reported in the comparison with ERA5 data except for statistic parameters from the HGPT2 model. It is worth noting that the higher improvement in RMSE for IR-ZWD is observed at the OIII station, as shown in Fig. 8a. Referring to Table 2, we can see that this station has the lowest altitude among the considered radiosonde stations. Thus, consistent with the results of the comparison with ERA5 data, the IR-ZWD model outperforms the GPT3w and GTrop models for locations of lower altitudes with higher vapor fluctuations. Comparing RMSE values of the HGPT2 model in Fig. ***8*** and Fig. 5, it can be seen that the RMSE value of the HGPT2 model is lower when compared to radiosonde stations (about 28% lower). Referring to Fig. 4d, it can be seen that the RMSE values of the HGPT model for land areas are almost the same as the GPT3w and GTrop models. Thus, considering the fact that the considered radiosonde stations are located inland areas, it can be concluded that the performance of the HGPT2 model only in the Persian Gulf and Oman Sea is weaker than the GPT3w and GTrop models.

4.5. Comparison with the Jason-3 radiometer measurements

One of the valuable sources of water vapor-related observations over oceans is measurements from equipped radiometers in altimetry missions. The Jason-3 mission was launched in 2016 with the Advanced Microwave Radiometer-2 (AMR-2) to correct the effect of ZWD value in satellite range observations (Gong and Liu, 2020). The presence of the Caspian Sea and Persian Gulf in the study region provides an opportunity to further evaluate the IR_ZWD model with respect to estimated ZWD values from Jason-3 radiometer observations. For this purpose, IR-ZWD model was compared with the Jason-3 ZWD values over the time interval of 1 January 2022 to 31 December 2022 (Jason-3 cycles 217-325). The Jason-3 data can be downloaded from https://www.aviso.altimetry.fr. Radiometer observations are unreliable and biased near coastal regions (Desportes et al., 2007). Therefore, only observations with a distance of more than 50 km from the nearest coastal region were considered (the points with valid ZWD observations have been displayed in Fig. 1). Hence, the valid Jason-3 radiometer measurements were about 6251 in the study region. Fig. ***9*** shows the joint Kernel Density Estimate plots of IR-ZWD, GPT3w, GTrop, and HGPT models in comparison with the Jason-3 estimated ZWD value that have been plotted using the Seaborn library in the Python (Waskom et al., 2014). These plots show the comparison of the modeled ZWD values versus the objective one and the contributed blue colorized points, indicate the probability of each value. The corresponding Taylor diagram for calculated statistical parameters has been displayed in Fig. 10.

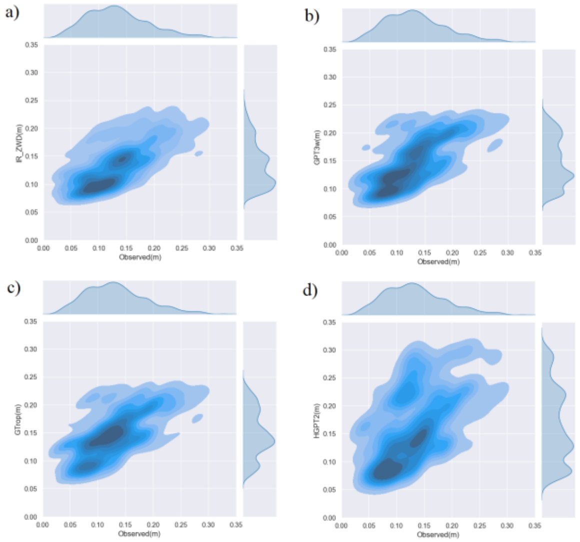


Fig. 9. An overview of the joint Kernel Density Estimate (KDE) plots of Jason-3 ZWD values versus the IR-ZWD (a), GPT3w (b), GTrop (c), and HGPT (d) models, respectively.

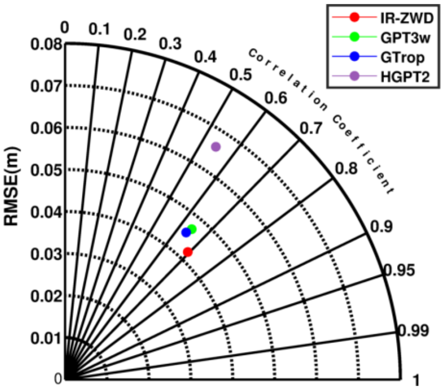


Fig. 10. An overview of the Taylor Diagram for comparing the ZWD estimated from Jason-3 with those of IR-ZWD, GPT3w, GTrop, and HGPT2. Radial and angular directions stand for different RMSE and CC values. The red, green, blue, and purple dots represent the IR-ZWD, GPT3w, GTrop, and HGPT2 models, respectively.

According to Fig. 10, the RMSE of IR-ZWD model is about 0.042 m, which indicates an improvement of approximately 10%, 7%, and 36% compared to the GPT3w, GTrop, and HGPT2 models, respectively. The CC of IR-ZWD model is found to be approximately 0.69, which is about 5%, 6%, and 15% higher than the corresponding CC values for GPT3w, GTrop, and HGPT2 models, respectively. Additionally, the calculated MAE value for IR-ZWD is found to be 0.031 m, which represents improvements of approximately 11%, 7%, and 36% compared to GPT3w, GTrop, and HGPT2 models, respectively. Referring to Fig. 10, in comparison with Jason-3 radiometer measurements, the HGPT2 model exhibits an RMSE of approximately 0.066 m, signifying the weakest performance among the considered models in the Persian Gulf and Oman Sea regions.

5. Conclusion

Atmospheric delay, specifically the Zenith Wet Delay (ZWD), is a challenging parameter to empirically parameterize due to high spatiotemporal variations. As it is the main source of error in space-based geodetic observations, constructing a reliable empirical model for real-time applications in positioning and weather prediction is crucial. However, in data sparse regions like Iran that exhibits considerable spatial and temporal vapor variations, the application of global atmospheric models might represent limited skills for positioning applications. In this study, we developed a regional ZWD model for a region including Iran, thus called the IR-ZWD model, which uses local base functions for an effective representation of the spatial distribution. This model incorporates ERA5 data from 2017-2022 to construct the model over five different time scales: mean, annual, semi-annual, 4-monthly, and 3-monthly variations. We optimized the structure of the model empirically and compared its result with the ZWD values from ERA5 data over the time interval outside of the fitting period (i.e., January 1, 2022, to December 31, 2022), and found that IR-ZWD has the least mean RMSE of about 0.035 m over the region of study. This is about 12.5% better than the global GPT3w and GTrop models, and 27.5% better than the HGPT2 model. Additionally, the squared correlation coefficient of IR-ZWD is found to be 0.55, i.e., about 10% higher than the other models. To further evaluate the accuracy of our proposed model, we compared the ZWD values from IR-ZWD with five located radiosonde stations in the region of study. We found that the mean RMSE of IR-ZWD is about 0.03 m, 11%, 13%, and 13% lower than the GPT3w, GTrop, and HGPT2 models, respectively. Furthermore, the derived CC value of the IR_ZWD model was about 0.41, which was about 10% higher than the corresponding GPT3w and GTrop models, and 17% higher than the CC value from the HGPT2 model. Finally, we compared the ZWD values from our model with corresponding Jason-3 radiometer measurements. The IR-ZWD model showed an improvement of about 10%, 6%, and 11% in RMSE, CC, and MAE values, respectively, compared to the GPT3w and GTrop models. Overall, our results suggest that the IR-ZWD model has higher accuracy than global empirical models and that our proposed method increases the reliability of the estimated ZWD values. In future, the impact of IR-ZWD for Standard Point Positioning (SPP) applications will be evaluated.

6. References

Adavi, Z., Mashhadi-Hossainali, M., 2014. 4D tomographic reconstruction of the tropospheric wet refractivity using the concept of virtual reference station, case study: northwest of Iran. Meteorology and Atmospheric Physics 126, 193-205.

Al-Fanek, O.J.S., 2013. Ionospheric imaging for Canadian polar regions. University of Calgary.

baba shaeb Kannemadugu, H., Ranganathan, K., Gharai, B., et al., 2022. GNSS-GPS derived integrated water vapor and performance assessment of ERA-5 data over India. Journal of Atmospheric and Solar-Terrestrial Physics 227, 105807.

Bender, M., Dick, G., Ge, M., et al., 2011. Development of a GNSS water vapour tomography system using algebraic reconstruction techniques. Advances in Space Research 47, 1704-20.

Bevis, M., Businger, S., Herring, T.A., et al., 1992. GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System. Journal of Geophysical Research: Atmospheres 97, 15787-801.

Black, H.D., 1978. An easily implemented algorithm for the tropospheric range correction. Journal of Geophysical Research: Solid Earth 83, 1825-8.

Böhm, J., Möller, G., Schindelegger, M., et al., 2015. Development of an improved empirical model for slant delays in the troposphere (GPT2w). GPS solutions 19, 433-41.

Collins, J.P., Langley, R.B. A tropospheric delay model for the user of the wide area augmentation system. Department of Geodesy and Geomatics Engineering, University of New Brunswick …, 1997.

Davis, J., Herring, T., Shapiro, I., et al., 1985. Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length. Radio science 20, 1593-607.

Dee, D.P., Uppala, S.M., Simmons, A.J., et al., 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the royal meteorological society 137, 553-97.

Dehvari, M., Karimi, S., Farzaneh, S., et al., 2023. Improving IRI-2016 Global Total Electron Content Maps Using ELM Neural Network. Advances in Space Research.

Desportes, C., Obligis, E., Eymard, L., 2007. On the wet tropospheric correction for altimetry in coastal regions. IEEE transactions on geoscience and remote sensing 45, 2139-49.

Dogan, A.H., Erdogan, B., 2022. A new empirical troposphere model using ERA5's monthly averaged hourly dataset. Journal of Atmospheric and Solar-Terrestrial Physics 232, 105865.

Du, Z., Zhao, Q., Yao, W., et al., 2020. Improved GPT2w (IGPT2w) model for site specific zenith tropospheric delay estimation in China. Journal of Atmospheric and Solar-Terrestrial Physics 198, 105202.

Dumont, J., Rosmorduc, V., Carrere, L., et al., 2016. Jason-3 products handbook. Technical representative, CNES and EUMETSAT and JPL and NOAA/NESDIS.

Forootan, E., Dehvari, M., Farzaneh, S., et al., 2023. Improving the Wet Refractivity Estimation Using the Extremely Learning Machine (ELM) Technique. Atmosphere 14, 112.

Forootan, E., Dehvari, M., Farzaneh, S., et al., 2021. A functional modelling approach for reconstructing 3 and 4 dimensional wet refractivity fields in the lower atmosphere using GNSS measurements. Advances in Space Research 68, 4024-38.

Gong, Y., Liu, Z., 2020. Evaluating the accuracy of Jason-3 water vapor product using PWV data from global radiosonde and GNSS stations. IEEE Transactions on Geoscience and Remote Sensing 59, 4008-17.

Hersbach, H., Bell, B., Berrisford, P., et al., 2020. The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146, 1999-2049.

Heydarizadeh Shali, H., Sampietro, D., Safari, A., et al., 2020. Fast collocation for Moho estimation from GOCE gravity data: the Iran case study. Geophysical Journal International 221, 651-64.

Hopfield, H., 1969. Two-quartic tropospheric refractivity profile for correcting satellite data. Journal of Geophysical research 74, 4487-99.

Kalita, J., Rzepecka, Z., 2017. Impact of the initial tropospheric zenith path delay on precise point positioning convergence during active conditions. Measurement Science and Technology 28, 045102.

Landskron, D., Böhm, J., 2018. VMF3/GPT3: refined discrete and empirical troposphere mapping functions. Journal of Geodesy 92, 349-60.

Li, W., Yuan, Y., Ou, J., et al., 2012. A new global zenith tropospheric delay model IGGtrop for GNSS applications. Chinese science bulletin 57, 2132-9.

Limberger, M., 2015. Ionosphere modeling from GPS radio occultations and complementary data based on B-splines. Technische Universität München.

Ma, Y., Liu, H., Xu, G., et al., 2021. Empirical orthogonal function analysis and modeling of global tropospheric delay spherical harmonic coefficients. Remote Sensing 13, 4385.

Ma, Y., Liu, T., Chen, P., et al., 2022. Global tropospheric delay grid modeling based on Anti-Leakage Least-Squares Spectral Analysis and its validation. Journal of Atmospheric and Solar-Terrestrial Physics 229, 105829.

Mateus, P., Catalão, J., Mendes, V.B., et al., 2020. An ERA5-based hourly global pressure and temperature (HGPT) model. Remote Sensing 12, 1098.

Mateus, P., Mendes, V.B., Plecha, S.M., 2021. HGPT2: An ERA5-based global model to estimate relative humidity. Remote Sensing 13, 2179.

Nohutcu, M., Karslioglu, M., Schmidt, M., 2010. B-spline modeling of VTEC over Turkey using GPS observations. Journal of Atmospheric and Solar-Terrestrial Physics 72, 617-24.

Penna, N., Dodson, A., Chen, W., 2001. Assessment of EGNOS tropospheric correction model. The Journal of Navigation 54, 37-55.

Plag, H.-P., Rothacher, M., Pearlman, M., et al., 2009. The global geodetic observing system. Advances in Geosciences: Volume 13: Solid Earth (SE). World Scientific, 105-27.

Razin, M.R.G., Voosoghi, B., 2017. Regional ionosphere modeling using spherical cap harmonics and empirical orthogonal functions over Iran. Acta Geodaetica et Geophysica 52, 19-33.

Saastamoinen, J., 1972. Atmospheric correction for the troposphere and stratosphere in radio ranging satellites. The use of artificial satellites for geodesy 15, 247-51.

Schmidt, M., Bilitza, D., Shum, C., et al., 2008. Regional 4-D modeling of the ionospheric electron density. Advances in Space Research 42, 782-90.

Subirana, J.S., Hernandez-Pajares, M., Zornoza, J.e.M.J. GNSS Data Processing: Fundamentals and Algorithms. European space agency, 2013.

Sun, J., Wu, Z., Yin, Z., et al., 2017a. A simplified GNSS tropospheric delay model based on the nonlinear hypothesis. GPS Solutions 21, 1735-45.

Sun, L., Chen, P., Wei, E., et al., 2017b. Global model of zenith tropospheric delay proposed based on EOF analysis. Advances in Space Research 60, 187-98.

Sun, Z., Zhang, B., Yao, Y., 2019. A global model for estimating tropospheric delay and weighted mean temperature developed with atmospheric reanalysis data from 1979 to 2017. Remote Sensing 11, 1893.

Tregoning, P., Herring, T.A., 2006. Impact of a priori zenith hydrostatic delay errors on GPS estimates of station heights and zenith total delays. Geophysical Research Letters 33.

Tunalı, E., Özlüdemir, M.T., 2019. GNSS PPP with different troposphere models during severe weather conditions. GPS Solutions 23, 82.

Vedel, H., Huang, X.-Y., 2004. Impact of ground based GPS data on numerical weather prediction. Journal of the Meteorological Society of Japan. Ser. II 82, 459-72.

Vuille, M., Werner, M., Bradley, R., et al., 2005. Stable isotopes in precipitation in the Asian monsoon region. Journal of Geophysical Research: Atmospheres 110.

Waskom, M., Botvinnik, O., Hobson, P., et al., 2014. Seaborn: V0. 5.0 (November 2014). Zenodo.

Yan, X., Ducrocq, V., Poli, P., et al., 2009. Impact of GPS zenith delay assimilation on convective-scale prediction of Mediterranean heavy rainfall. Journal of Geophysical Research: Atmospheres 114.

YAO, Y.-B., HE, C.-Y., Zhang, B., et al., 2013. A new global zenith tropospheric delay model GZTD. Chinese Journal of Geophysics 56, 2218-27.

Yao, Y., Hu, Y., 2018. An empirical zenith wet delay correction model using piecewise height functions. Annales Geophysicae. Copernicus GmbH, pp. 1507-19.

Yao, Y., Hu, Y., Yu, C., et al., 2016. An improved global zenith tropospheric delay model GZTD2 considering diurnal variations. Nonlinear processes in geophysics 23, 127-36.