Dear editor,

We sincerely appreciate your valuable comments and constructive suggestions. We have carefully addressed all the points raised and have revised the manuscript accordingly.

• Specify the units for all variables in the manuscript (e.g., TD, N, s, L, …).
Response: Added.

• L60: "As hydrostatic and wet mapping functions differ": What do you mean? Please reword.
Response: The slant tropospheric delay, as is shown in equation (5), can be expressed as:

$$TD=ZHD\cdot mf_h+ZWD\cdot mf_w+\Delta T_{grad} \quad (5)$$

where mf_h and mf_w are the mapping functions for the hydrostatic (ZHD) and non-hydrostatic (ZWD) part of the tropospheric delay. "As hydrostatic and wet mapping functions differ" means that mf_h and mf_w have different values. When fixing ZHD to estimate ZWD, since there is a difference between mf_h and mf_w, the ZHD error will not be completely absorbed by the estimated ZWD, which in turn affects the accuracy of the ZTD estimation (the sum of ZHD and ZWD).

• L73: "in-site meteorological … water vapor": This sentence seems contradictory. Why do in-situ meteorological measurements fail to measure dynamic nature of water vapor?

Response:

Sorry for using such a seemingly contradictory expression.

The original draft of this sentence is "Although the ZWD can also be calculated in a similar way, i.e., using an empirical model, e.g., the Askne-Nordius model (Askne and Nordius, 1987), together with in-situ meteorological measurements, its accuracy is not as good as that of the Saastamoinen model due to the dynamic nature of water vapor (Chen and Liu, 2016)."

Explanation: Askne-Nordius model is an empirical model for calculating ZWD, which needs in-situ water vapor pressure as input parameter. The water vapor pressure can be measured by in-situ meteorological sensors, however, the Askne-Nordius model itself is not a very accurate model because the water vapor changes rapidly in both spatial and temporal domain.

The revised version is "Although the ZWD can also be calculated using empirical models such as the Askne-Nordius model (Askne and Nordius, 1987), which relies on in-situ meteorological measurements (e.g., water vapor pressure), its accuracy is generally lower than that of the Saastamoinen model. This is because water vapor exhibits high spatiotemporal variability, and the Askne-Nordius model's empirical formulation cannot fully capture these rapid fluctuations, even when precise in-situ measurements are available (Chen and Liu, 2016)."

• L76-77: Provide the official names of NWM, UNB3m, and GPT.

Response:

NWM: numerical weather model;

UNB3m is one version of UNB Neutral Atmosphere Delay Model, where UNB denotes University of New Brunswick, "3m" is the number of the model version.

GPT: Global Pressure and Temperature model.

The revised text is: "*Since most GNSS stations are not mounted with meteorological sensors and it is complex for real-time GNSS users to process forecasted NWM (numerical weather model) data to obtain the atmospheric parameters, empirical tropospheric delay models like UNB3m (Leandro et al., 2006) developed by University of New Brunswick (UNB) and GPT (Global Pressure and Temperature) models developed by Vienna University of Technology (TU Wien)*"

• L78-81: I am unsure about the references to empirical models here, maybe short review for the studies might be necessary for explaining limitations of previous studies.

Response: The manuscript has been revised:

"In such empirical models, spatiotemporal variations of the atmospheric parameters are modelled, and then the atmospheric parameters can be predicted directly. Incorporating advanced height correction model is an effective method to improve the modeling accuracy of atmospheric parameters or the tropospheric delays (Huang et al., 2023; Jiang et al., 2024; Li et al., 2018; Sun et al., 2023). However, while easy to use, these models have limited accuracy due to rapid variation of the troposphere (Wang et al., 2017; Xia et al., 2023), as these models can only capture the mean status of the annual, semi-annual and diurnal variations of the parameters."

• L90: Provide the official names of GFZ-VMF1 and UNB-VMF1.

Response: Both GFZ and UNB are institution names. We have revised the manuscript: "In addition, some other VMF1-like products are also publicly available for users (Santos, 2011; Zus et al., 2015)."

• L95: "a difference of 1.83 cm mean root-mean-square"?

Response: corrected.

• L108-113: Provide the references for each sentence.

Response: While ZTD derived from VMF1/VMF3 products generally exhibit superior accuracy compared to those derived from empirical tropospheric models, discrepancies in ZHD and ZWD have been documented in certain studies (Sun et al., 2021b; Yang et al., 2021; Yao et al., 2018b). Specifically, the RMSE of ZHD estimated by grid-wise VMF1/VMF3 using the

recommended interpolation method can reach 5 cm when compared with reference ZHD values obtained from radiosonde measurements in some regions (Sun et al., 2021b). Similarly, the RMSE of ZWD can also attain substantial magnitudes (Sun et al., 2021a; Yang et al., 2021).

• L111: "Similarly, … magnitudes" Provide the reason for enlargement of "the RMSE of ZWD".

Response: This is a good question.

Firstly, the accuracy of the source data for the VMF1/VMF3 grid products—namely, the atmospheric profile outputs from ECMWF's numerical weather models—is not uniformly distributed globally. Secondly, since the VMF1/VMF3 grid products only provide surface values at grid points, when there is a significant elevation difference between the target location and the surrounding grid points, an imprecise ZWD vertical correction model can lead to substantial interpolation errors in ZWD.

• L113: Specify the geographical and atmospheric conditions that discrepancies in ZHD and ZWD occur.

Response: This is a very good question too. As is mentioned above, since the VMF1/VMF3 grid products only provide surface values at grid points, when there is a significant elevation difference between the target location and the surrounding grid points, an imprecise ZHD/ZWD vertical correction model can lead to substantial interpolation errors in ZHD/ZWD.

• L133: "394 IGS stations were selected" -> out of how many sites?

Response: The IGS routinely produce ZTD products of the IGS stations. But the number of the stations provided is not fixed. For example, the number for 31/12/2022 is 418, for 01/01/2020 is 382. ZTD products in the 3-year period of 2020–2022 at 394 IGS stations were selected using the criteria mentioned in the manuscript.

• L137: Provide the method how 5-min time series were resampled to 2-hr ones.

Response: The IGS ZTD product provide ZTD values at fixed epochs with 5 min interval: 00:00, 00:05, 00:10…01:00…, thus we can resample the data directly by picking up specific epochs.

• L156: ZHD0 in (8) and ZWD0 in (10) -> ZHD0 in (7) and ZWD0 in (9)

Response: Corrected.

• L158, 165, 192, 228: Saastamoinen model -> Eq. (6)

Response: Corrected.

• L163: What do you intend by "S (A'~D')"? Consider rewording this sentence.

Response: As is shown in Figure 2, Points A, B, C, and D are the surface grid points of VMF1/VMF3, and S is the target location. The horizontal plane passing through S intersects the vertical lines of A, B, C, and D at points A'–D', meaning that the elevation of S is the same as that of A'–D'. First, we need to correct the tropospheric delay values at A, B, C, and D along the elevation direction to obtain the tropospheric delay values at A'–D'. Then, we perform horizontal interpolation to derive the tropospheric delay parameters at point S.

• L164: What are "target" and "reference"? Add detailed information.

Response: The manuscript has been revised: Δh is the difference between the target and reference heights, i.e. $\Delta h_{AA'}$, $\Delta h_{BB'}$, $\Delta h_{CC'}$, $\Delta h_{DD'}$

• L166: What is "product-provided ZWD"? Add explanation.

Response: Product-provided ZWD means the grid-wise ground surface ZWD values provided by VMF1/VMF3 product. The manuscript has been revised:"[6]. Using ZWD0 (i.e., ZWDs at A to D) and the following model to obtain ZWDs for A'~D':"

• L166: Is "reduction model" correct phrase in English?

Response: The manuscript has been revised:"[6]. Using ZWD_0 (i.e., ZWDs at A to D) and the following model to obtain ZWDs for A'~D':"

• L238: "These results…" I suspect this results as self-evident because both the reference and Scheme 3 (τ and r) are estimated using ERA5 data. I recommend applying different reanalysis data for estimates of the reference and Scheme 3.

Response:

This is a very good suggestion. Both the reference and Scheme 3 are based on ERA5 data, however, the reference data are ERA5 hourly data from 2020 to 2022, while the Scheme 3 (an empirical model for lifting surface ZTD to the target height) was developed using ERA5 monthly mean data from 2010 to 2019. The VMF1/VMF3 surface ZTD product are developed using ECMWF OPERATIONAL NWM data and FORECAST NWM data. The most important reason for using ERA 5 in the previous manuscript is that the error embedded in VMF1/VMF3 surface ZTD product is also a major error source when interpolating ZTD at the target position, using reanalysis data provided by ECMWF may eliminate the discrepancy resulting from other reanalysis data, such as MERRA-2 and JRA55.

To further evaluate the performance of these three methods, radiosonde data from 2020 to 2022 at 608 stations are tested and the results demonstrates similar conclusions, i.e., the new model developed in this research is a good method for improving the ZTD interpolation accuracy of VMF1/VMF3 surface ZTD product. The manuscript has been revised.

• L 248: What is "IQQE station"? Why is this station selected for the analysis? Does this

station guarantee representativeness of 394 IGS station? Provide detailed explanation.

Response:

This is a good question.

What is "IQQE station"? IQQE is the name of an IGS station.

Why is this station selected for the analysis? The ZTD time series of this station was selected to prove that for the stations with substantial height disparities from adjacent VMF grid points, the accuracy of the interpolated ZTD can be significantly improved by utilizing the new model proposed in this paper.

Does this station guarantee representativeness of 394 IGS station? When interpolating *a priori* ZTD at the GNSS station using VMF1/VMF3 product, the height differences between the station and the adjacent VMF grid points should be considered. Such height differences are small for part of these stations, which means that the scheme 1 (official one) may not leads to large ZTD prediction errors. However, for the stations with substantial height disparities from adjacent VMF grid points, the accuracy of the interpolated ZTD can be significantly improved by the new method, such as IQQE station presented in the manuscript. In fact, there are thousands of other geodetic GNSS stations running on the earth surface, continuously observing signals broadcast by GNSS satellites (GPS, Galileo, QZSS, etc.), and large amount of civilian GNSS receivers are mounted on rover objects like Unmanned Aerial Vehicle (UAV), cars, and mobile phones, this implies that large elevation differences between the GNSS receiver and the VMF1/VMF3 grid points are highly likely.