Review of the Manuscript

"Enhancing GNSS Water Vapor Retrieval via Synergistic Microwave Radiometry: Thermodynamic Error Diagnosis and Bias Correction"

by Avinash N. Parde, Christina Oikonomou, and Haris Haralambous.

General Comments

GNSS is a well-established technique for observing precipitable water vapor (PWV). However, GNSS provides zenith total delay (ZTD) data, which must be combined with auxiliary data, specifically, zenith hydrostatic delay (ZHD) and weighted mean temperature (Tm), to derive PWV.

Various methods and data sources have been used to estimate Tm. Empirical regression models (e.g., Bevis et al., 1992) relate Tm to surface temperature (Ts) and thus require Ts as auxiliary data. Static climatological models (e.g., GPT2w and HGPT2) are standalone and cannot capture short-term variations in Tm. Only numerical weather prediction (NWP) models can provide real-time Tm estimates, albeit with limited spatiotemporal resolution.

This manuscript proposes to use ground-based MWR temperature- and moisture-profile measurements as an alternative data source for deriving Tm. The idea is original and the approach is worth exploring. However, the study suffers from methodological and interpretative deficiencies that must be addressed before publication.

Specific comments

1. Overemphasis on Tm as the Main Source of Uncertainty

The manuscript claims that Tm is the primary source of uncertainty in GNSS PWV data. While reducing Tm errors would improve PWV retrieval, it is well established that Tm is **not** the main error source (e.g., Ning et al., 2016). The exaggerated emphasis on Tm should be revised throughout the manuscript.

2. Error Propagation Analysis

The discussion of error propagation from ZTD to PWV is fragmented and inconsistent (Introduction, Sections 2.4, 3.3, 3.4, and 3.5). The quoted error values often do not align with standard error propagation rules and overestimate the impact of Tm errors on PWV. For example:

- A 1 K error in Tm would propagate to ~0.37% in PWV, or ~0.18 mm if PWV = 50 mm (an exceptionally high value for the Mediterranean region), **not** 0.3–0.5 mm (lines 54–60).
- A 1–2% error in Tm cannot translate to a 1–2 mm error in PWV.

A comprehensive formulation using partial derivatives should be introduced, with derivatives quantified. This is essential for assessing the proportional contributions of each error source.

3. Systematic vs. Random Errors

The manuscript must clearly distinguish between systematic and random errors. For example:

- Errors in refractivity constants introduce a **bias** in derived PWV. Their associated uncertainty cannot be directly compared to Tm-related uncertainty, which likely combines both systematic and random components (Figure 9a).

- Figure 9b may contain a factor-of-10 error.

The discussion also omits the uncertainty due to the refractivity constant k1 in ZHD calculations.

Recommendation: Revisit the seminal works by Bevis (1992, 1994), Wang (2005), Healy (2011), and Bock (2021) on Tm modeling and refractivity constants.

4. MWR Retrieval Bias

The initial MWR retrievals used in this study exhibit a large bias, suggesting inadequate neural network training. This issue is common with MWR profilers and is acknowledged later in the manuscript (lines 247 and 268), where bias correction is introduced. While bias correction improves accuracy, the conventional approach would be to refine the MWR retrieval algorithm through additional training with radiosonde data.

5. Water Scale Height (Hv) Analysis

The analysis of the water scale height (Hv) is imprecise and irrelevant to this study for several reasons:

- Hv is a crude representation of the moisture profile.
- The discussion on upper cutoff altitude highlights the ambiguity in estimating Hv from MWR data, as reflected in the large scatter in MWR results (Figure 5).
- The uncertainty in estimated Hv values (quantifiable via least-squares regression) is not considered when comparing MWR and radiosonde (RS) estimates.

6. Comparison with HGPT2

The manuscript argues that MWR-derived Tm is superior to HGPT2, which is unsurprising. HGPT2 is a static climatological model and cannot compete with real-time data sources like MWR (Figures 6 and 9). Due to this intrinsic limitation, static models are less relevant for meteorological applications compared to empirical and NWP models. Additionally, the Tm variable in HGPT2 is derived via regression on Ts, which is known to introduce spurious diurnal cycles into Tm (Wang, 2005; Bock, 2021).

Recommendation: Remove HGPT2 from the study and instead compare MWR-based Tm to operational NWP-derived Tm estimates (e.g., ERA5 with 1-hourly resolution).

7. Terminology and Clarity

The manuscript contains awkward phrasing and non-standard terminology. Examples include:

- “Critical error culmination”
- “Thermodynamic severity”
- “Noon ballooning”
- “Thermodynamic path delays”
- “Thermodynamic errors”

Concepts such as **uncertainty**, **error**, and **bias** are used inconsistently and confusingly throughout the manuscript.

8. Scientific Rigor in Sections 3.5, 4, and 5

- **Section 3.5:** Focuses excessively on the deficiencies of HGPT2 and static models, which are well known. These issues could be avoided by excluding HGPT2 from the analysis.
- **Section 4:** The discussion on planetary boundary layer (PBL) and free-troposphere "decoupling" lacks evidence. Radiosonde (RS) data could help document this climatic feature, but it is unclear how MWR profiles would capture it effectively.
- **Section 5 (Conclusions):** Requires revision based on the above comments.

9. Manuscript Organization

The manuscript's structure is confusing. For example:

- The capacity of MWR to retrieve PWV directly is not mentioned until Section 3.2.
- The derivation of GNSS PWV in Section 3.2 is unclear.
- Tm estimates are evaluated in Section 3.3, **after** the first GNSS PWV estimates are assessed (Section 3.2).
- MWR calibration is introduced in Section 3.4, **after** MWR Tm data are used.

Recommendations for Revision

Before publication, I recommend the following:

1. **Improve MWR retrievals** through complementary neural network training.
2. **Replace HGPT2** with a high-resolution NWP model (e.g., ERA5) as an alternative Tm source. Compare MWR-derived Tm to both NWP and RS Tm.
3. **Assess GNSS PWV** against MWR and RS PWV.
4. **Remove the Hv discussion**, as it is not relevant to the study.
5. **Conduct a proper error propagation analysis** (Ning, 2016) to contextualize Tm errors relative to other sources in the ZTD-to-PWV processing chain.
6. **Adhere to standard scientific terminology** and ensure consistent use of terms like *uncertainty*, *error*, and *bias*.

References

Wang 2005: doi:10.1029/2005JD006215

Healy 2011: doi:10.1029/2010JD014013

Ning 2016: doi:10.5194/amt-9-79-2016

Bock 2021: https://doi.org/10.5194/essd-13-2407-2021

Final Assessment

I do not recommend publishing this manuscript in its current form. The authors should address the methodological, organizational, and terminological issues outlined above. A revised version, incorporating these suggestions, would significantly improve the manuscript's scientific rigor and clarity.