

Replies to comments of Reviewer 2

Authors would like to express sincere gratitude to an anonymous reviewer for his/her valuable comments and suggestions. Our manuscript has greatly benefited from these insights. We have carefully revised the manuscript, taking all comments into account. Our responses to the reviewer's comments are provided below.

General comment

This study provides practical insights for estimating PWV from the spectral measurements of solar global, direct, and diffuse irradiances. The retrieval of PWV under all-sky conditions is useful for monitoring the atmospheric conditions. However, concerns exist regarding the following points: The input data for the DNN model includes “Day number of year” and “solar zenith angle”. The model may have learned seasonal characteristics of the observation site. If the instrument is relocated to another site, would retraining be necessary?

→ Thank you for this valuable comment. In our approach, solar zenith angle (SZA) is not used as an independent input feature. Instead, all spectral irradiances are first normalized by the cosine of SZA, so that the geometric effect of solar incidence is already incorporated into the transformed radiometric features before they are provided to the DNN. Thus, the model does not receive SZA as a separate predictor; its influence is implicitly captured through the SZA-normalized spectra.

To assess whether the model nevertheless relies on seasonal or site-specific characteristics—particularly through the inclusion of DOY—we evaluated the contribution of all input variables using the SHAP framework (Lundberg and Lee, 2017). SHAP provides a unified, theoretically grounded framework based on cooperative game theory to quantify the contribution of each input feature to the model output by computing Shapley values, defined as the average marginal contribution of a feature across all possible feature combinations. The method has been widely used in recent ML-based atmospheric and remote-sensing studies to assess the relative influence of input variables on model predictions (e.g., Lundberg et al., 2020; Zhao et al., 2024). Tables R1-R3 below summarize the SHAP values obtained for all input features under different modeling strategies (i.e., using global, direct, and diffuse irradiances jointly, or using only global or only direct irradiances), as described in the manuscript. The SHAP values for the radiative components shown in the table represent averages across all wavelengths, given that wavelength-by-wavelength SHAP reporting would be excessively lengthy.

Table R1. Relative SHAP values (in %) of the input features for the DNN models trained using spectral direct, diffuse, and global irradiances of different absorption bands, together with the direct-to-diffuse irradiance ratio (Ratio) and the day of year (DOY)

Centered band/Features	Direct(%)	Diffuse(%)	Global(%)	Ratio(%)	DOY(%)
940 nm	23.8	17.89	35.67	22.42	0.22
820 nm	18.3	19.2	48.39	13.97	0.14
720 nm	20.03	15.24	44.69	19.95	0.08
Combined (all above)	27.18	19.43	32.92	20.35	0.11

Table R2. Relative SHAP values (in %) of the input features for the DNN models trained using spectral global irradiances of different absorption bands together with day of year (DOY)

Centered band/Features	Global(%)	DOY(%)
940 nm	98.65	1.35
820 nm	99.67	0.33
720 nm	99.51	0.49
Combined (all above)	99.54	0.46

Table R3. Relative SHAP values (in %) of the input features for the DNN models trained using spectral direct irradiances of different absorption bands together with day of year (DOY)

Centered band/Features	Global(%)	DOY(%)
940 nm	97.84	2.16
820 nm	99.19	0.81
720 nm	99.15	0.85
Combined (all above)	98.49	1.51

Across all model configurations (Tables R1–R3), spectral irradiances account for more than 95% of the total feature importance. DOY contributes less than 2% (often below 0.3%), indicating that the model’s predictive power is drawn almost entirely from the radiometric information. Because the irradiances are normalized by SZA, and because spectral absorption and scattering signatures inherently vary with atmospheric path length, the effects of SZA and seasonal variability are already encoded in the spectral measurements themselves. This explains the negligible contribution of DOY and shows that the model is not learning seasonal characteristics of the training site.

We agree that the model does not explicitly distinguish site-dependent and instrument-dependent factors; it learns statistical relationships between irradiances and PWV. As a result, relocating the same instrument to a site with markedly different atmospheric or surface conditions may reduce retrieval accuracy if the model is applied without adaptation. Nevertheless, this does not limit the broader applicability of the approach, as the model can be effectively extended to new locations either by retraining with locally available PWV references (e.g., radiosonde or GNSS PWV) or by applying transfer learning, in which a pre-trained model is fine-tuned using a relatively small amount of local irradiance–PWV data (Pan et al., 2010; Weiss et al., 2016). Recent studies (e.g., Chen et al., 2025; Dong et al., 2024; Gupta et al., 2024) demonstrate that such adaptation strategies successfully maintain retrieval performance when transferring ML-based remote-sensing models across sites.

We have summarized above discussion in the revised manuscript by including Table R1 in the revised manuscript as Table 2, and then adding following sentences (Lines 459-474).

Since DOY is included as one of the input features in our DNN models, it is important to examine whether the agreement between predicted and true values shown in Figure 5 could have been

influenced by climatological patterns encoded in DOY. To assess the relative importance of DOY, together with other input features, we applied the SHAP (SHapley Additive exPlanations) method (Lundberg and Lee, 2017) by computing Shapley values, which represent the average marginal contribution of a feature across all possible feature combinations. This method has been widely used in ML-based atmospheric and remote-sensing studies to evaluate the relative influence of input variables on model predictions (e.g., Lundberg et al., 2020; Zhao et al., 2024). Table 2 summarizes the SHAP values (in %) for the input features—global, direct, and diffuse irradiances, the direct-to-diffuse irradiance ratio, and DOY—for individual absorption bands and for their combined dataset. The SHAP values for the radiative components in Table 2 represent averages over all wavelengths, as reporting wavelength-specific SHAP values would be excessively lengthy. The results clearly show that when global, direct, and diffuse irradiances are used jointly, the global irradiance component consistently exhibits the highest relative SHAP importance (32–48%), followed by the direct and diffuse components. The direct-to-diffuse irradiance ratio contributes a moderate but meaningful amount (approximately 14–22%). In contrast, DOY contributes less than 0.3%, indicating that the seasonal patterns observed in Figure 5 are dominated by spectral-irradiance-based features rather than by DOY input feature. DOY can have very low importance because it only provides indirect seasonal information, whereas the spectral irradiances directly capture the actual atmospheric state (water-vapor absorption, scattering, SZA effects, etc.).

Again, we have added the following sentences to clarify the query related to the application over a new environment (Lines 264-270).

Furthermore, the model learns statistical relationships between spectral irradiances and PWV, which can be influenced by local atmospheric and surface conditions. Relocating the instrument to a site with substantially different conditions may reduce accuracy. However, the model can be adapted to new locations through retraining with local reference PWV data or, more efficiently, via transfer learning (Pan et al., 2010; Weiss et al., 2016), in which a pre-trained model is fine-tuned using a relatively small amount of site-specific data. Recent studies (e.g., Chen et al., 2025; Dong et al., 2024; Gupta et al., 2024) show that such adaptation strategies effectively preserve retrieval performance when transferring ML-based remote-sensing models across sites.

The methodology details and results are clearly described. This paper is recommended after minor revision.

Specific comments

- L50-53: GNSS precise point positioning is also used to retrieve PWV in worldwide (e.g., Zhang et al., 2019, DOI:10.1109/JSTARS.2019.2906950).

→Thank you for providing this valuable information.

We have updated our manuscript by adding the following sentences in the revised manuscript (Lines 72-75).

In parallel, the International GNSS Service (IGS) Real-Time Pilot Project has enhanced global access to high-precision GNSS data streams, enabling near-real-time estimation of PWV through precise point positioning (PPP) techniques at dense ground-based networks worldwide (Zhang et

al., 2019).

- Figure 2: How did you obtain the calibration constants at each wavelength? Is the radiometric calibration necessary in the retrieval method of this study?

→ The calibration constants at each wavelength were obtained through the factory calibration conducted by EKO Instruments using a NIST-traceable standard lamp. These calibration coefficients are embedded in the instrument's calibration file and are automatically applied by the operating software to convert raw detector outputs into physical spectral irradiance values ($\text{Wm}^{-2}\mu\text{m}^{-1}$). Therefore, since the measurements are already provided in calibrated physical units, no additional radiometric calibration is required in the retrieval method used in this study.

We have added the following sentence in the revised manuscript to add clarity (Lines 145-147)

These spectral irradiances, in units of $\text{Wm}^{-2}\mu\text{m}^{-1}$, are automatically generated by the observational software from the raw signals after applying the factory calibration constants.

- L220-229: There are many technical terms. If possible, include the references.

→ Thank you for this important suggestion.

We have added relevant references in the revised manuscript (Lines 229-245).

- L493-494: Could the discrepancies in Fig. 8b from July to September be attributed to the limited amount of training data available during the wet season?

→ Yes, the discrepancies observed in Figure 8b (Figure 9b in the revised manuscript) from July to September can indeed be attributed to the limited amount of training data available during the wet season. During this period, the frequency of cloud cover is substantially higher (Khatri and Takamura, 2009), which significantly reduces the availability of direct solar irradiance measurements required for training. Consequently, the model may have encountered fewer representative samples under clear-sky conditions, leading to reduced predictive accuracy and greater variability in the retrieval results for those months.

We have added the following sentence in the revised manuscript to add clarity (Lines 604-606)

In particular, the discrepancies in Figure 9b from July to September likely stem from the limited training data during the wet season, when persistent cloud cover greatly reduces the number of usable direct irradiance measurements (Khatri and Takamura, 2009).

References

- Chen, D., Guo, H., Gu, X., Wang, J., Liu, Y., Li, Y. & Wu, Y.: Physical-guided transfer deep neural network for high-resolution AOD retrieval, *Remote Sens.*, 17, 3606, 2025.
- Dong, S., Li, Y., Zhang, Z., Gou, T. & Xie, M.: A transfer-learning-based windspeed estimation on the ocean surface: Implication for the spatial-spectral resolution of remote sensors, *Appl. Intell.*, 54, 7603–7620, 2024.
- Gupta, S., et al.: Spatial transfer learning for estimating PM_{2.5} in data-sparse regions, *Proc. ACM*, 2024.
- Lundberg, S. M. and Lee, S.-I.: A unified approach to interpreting model predictions, *Adv. Neural Inf. Process. Syst.*, 30, 2017.
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I.: From local explanations to global understanding with explainable AI for trees, *Nat. Mach. Intell.*, 2, 56–67, 2020.
- Khatri, P. & Takamura, T.: An algorithm to screen cloud-affected data for sky radiometer data analysis, *J. Meteorol. Soc. Japan*, 87, 189–204, 2009.
- Pan, S. J. and Yang, Q.: A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.*, 22, 1345–1359, 2010.
- Weiss, K., Khoshgoftaar, T. M., and Wang, D.: A survey of transfer learning, *J. Big Data*, 3, 9, 2016.
- Zhang, H., Yuan, Y., Li, W., and Zhang, B.: A real-time precipitable water vapor monitoring system using the national GNSS network of China: method and preliminary results, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 12, 1587–1598, 2019.