Response to Reviewers

We sincerely appreciate the insightful comments and constructive suggestions from both reviewers, which have significantly improved the clarity and robustness of our methodology. We also thank the reviewers for their positive and encouraging feedback.

In this response letter, we have addressed each comment in detail below. Our responses are highlighted in blue, with corresponding revisions in the manuscript also indicated in *blue italics*. All line numbers in this document refer to the updated version of the manuscript.

Thank you for your time and consideration.

With best regards,
Di WANG
On behalf of all authors

AC2: 'Comment on amt-2024-151', 16 Dec 2024

Wang and others have developed a theoretical model to describe water vapour diffusion through the surface of a sampling bag. They calibrated the model's parameters using laboratory experiments. This model allows for the reconstruction of the initial isotopic composition of the sampled vapour by using measurements taken from both inside the bag and the surrounding environment. I believe it is an important work but not well explained and supported in several sections, and needs a through revision.

We appreciate your positive comment. We also appreciate your detailed suggestions.

L67 Distinguish between aerial mobile measurements for water vapour isotopes and mobile water isotope measurements, i.e. picarro mounted on a van. The latter is quite common and does not require sample storage anymore.

We have incorporated your suggestion by distinguishing aerial from mobile water isotope measurements (lines 64-67):

"However, their heavy instrumentation, substantial power requirements, and limited mobility restrict their usability in certain situations, particularly for aerial water vapor isotope measurements, which require lightweight and flexible sampling approaches."

L88-90 check usage; this should be isotopologues

We have replaced 'isotopes' with 'isotopologues'.

L104 What diverse conditions?

We have replaced 'diverse conditions' with 'varying humidity and isotopic composition differences between the inside and outside of the air bag' (line 106)

L110-111 Picarro direct observations - as in measurement on the Picarro? And satellite data of what?

To provide a more precise description of the data used, we have revised lines 112–117 as follows:

"The corrected near-surface drone-based measurements using our diffusion model show consistency with direct, in-situ surface-level measurements using the Picarro analyzer. Similarly, at two mid-tropospheric levels, the corrected drone measurements align with IASI satellite observations of water vapor isotopic composition, further confirming the model's theoretical and practical reliability in applications."

L125 State the boundary conditions and assumptions clearly under which equation 1 is valid

We appreciate your advice and have added the boundary conditions and assumptions for Equation 1 and 2 (lines 139-142):

"The validity of Equation (1) and (2) relies on the assumptions that internal and external pressures remain equal to atmospheric pressure, ensuring no pressure gradient across the bag membrane, that the internal vapor is well-mixed, and that the exchange rate follows a first-order process. Additionally, if the temperature remains constant, k *and* k_i *are assumed to be constant."*

L130 Equation 2 should be based on isotope notations, and the isotopologue for which they apply can be in subscript. This is not clear yet.

We appreciate the reviewer's advice. While δ-values provide a clearer visualization of isotope variations, their direct use in equations obscures the model's physical basis and complicates interpretation of fractionation processes. To maintain mathematical consistency and align with standard fractionation factor definitions, we use isotope ratios (R) instead of δ-values.

To ensure the correct use of R and δ-values in different contexts, we have now explicitly stated how R and δ-values are used in the manuscript (lines 212-216):

" For mathematical clarity and consistency, isotopic ratios (R) are used in the equations presented in previous sections. Replacing R with δ-values would only shift the physical basis without affecting the mathematical validity of the equations or the estimation of α, as the standard ratio cancels out. For clearer visualization, δ-values are used for numerical applications and in the subsequent figures and tables."

L202 Convert the flow rate to volume per unit time

We have added the following clarification to explicitly state the flow rate (230-233):

"*In the measurement procedure, we first activated the dry air cylinder and adjusted the pressure reducing valve to 2 psi (pounds of force per square inch), within the Picarro water isotope analyzer's recommended range of 2–4 psi for carrier gas. The instrument's built-in flow regulation maintains a gas flow rate of 30–50 mL/min,* *ensuring stable sample delivery."*

L228 Provide more details on the experimental setup, including the type of airbags used and the specific model of the Picarro analyzer

We have provided the type of airbags used and the specific model of the Picarro analyzer (lines 219-221):

"In this study, we used 0.5 L and 4L Teflon air bags produced by Dalian Hede Technologies Co., Ltd to collect and store vapor, and measured the vapor isotopes using a Picarro 2130i water isotope analyzer."

L252 Provide more details on the experimental setup, such as the specific amounts of water injected and the isotopic values used, to give a clearer picture of the conditions tested.

We have revised the text to include the specific amounts of water injected and the isotopic values used (lines 294-301):

*"To validate the diffusion model under diverse conditions and evaluate its uncertainties, we repeated Experiment No. 2, but injected different amounts of water with known isotopic values to achieve a range of humidities from approximately 1/8 * q_e to q_e. Using the method described in Experiment No. 2, we injected 6 to 50 μL of reference water into a 4L air bag filled with dry air to achieve the desired humidity range. Additionally, we repeated the experiment using two reference waters with distinct isotopic compositions, specifically $\delta^{18}O$ = -58.07‰, $\delta^{2}H$ = -447.41‰ and $\delta^{18}O$ = -29.84‰, $\delta^{2}H$ = -222.84‰. To assess extended-duration variations, we also lengthened the time interval to 24 hours."*

L277 Drone flight path and sampling strategy need to be better explained. Also, samples aren't measured in situ.

We appreciate the reviewer's suggestion and have provided more details about the drone flight path and sampling strategy (lines 328-337):

"We collected water vapor samples every 500 meters, starting from near the surface along the vertical profile. To optimize sampling across different altitude ranges, we deployed UAVs designed for varying flight altitudes. Generally, the UAV operating at lower altitudes collected samples at seven heights from 4,000 to 7,000 meters in a single flight. The mid-altitude UAV collected samples at four heights from 7,500 to 9,000 meters in one flight, while the high-altitude UAV collected samples at four heights from 9,500 to 11,000 meters in two flights. Each flight took approximately 20~30 minutes. In case of any disruptions during sampling, we repeated the process until a complete vertical profile was obtained. At the beginning of the experiment, we also collected replicate samples at each height to ensure data consistency."

To clarify our analysis procedure, we revised Line 338-340 as follows:

"By integrating high-altitude drone sampling with subsequent water vapor isotope analysis using the Picarro analyzer at the surface, we obtained vapor isotopic profiles up to an altitude of 11 km."

L305 I would suggest a net uncertainty or error propagation of some kind to be calculated and reported for
λ_surface, α_δ, λ_alt. Currently, this section does not explain the uncertainties (only how they are calculated) or how they affect the results.

We have revised the manuscript to explicitly describe how uncertainties from all sources were combined to calculate the net uncertainty. First, we clarified that as the pressure and sampling time are well known, the uncertainty in λ_{alt} is fully propagated from the uncertainties in $\lambda_{\text{\_surface}}$ and pump efficiency (ε), rather than being treated as an independent source, as shown in Equation 15 (Section 3.4.1):

$$\lambda_{\_alt} = \lambda_{\_surface} * \frac{P_{\_surface}}{P_{\_alt}} * \frac{Sampling\ time_{\_alt}}{Sampling\ time_{\_surface}} * \varepsilon$$

We have ensured that Section 3.4.2 ("The Method of Uncertainty Estimation") details the calculation methods for $\lambda_{\text{\_surface}}$, α, and pump efficiency (ε), while Section 4 ("Discussion") explicitly discusses how these uncertainties influence the results.

We also expanded the discussion in Sections 4.3 and 4.5 to explicitly report the net uncertainties for δ^{18}O and d-excess across all altitudes (lines 569-570):

"The combined uncertainty from all sources, including $\lambda_{\_surface}$, α, pump efficiency (ε), and model-experiment mismatches, results in a total uncertainty of approximately 1‰ for $\delta^{18}O$ and 8‰ for d-excess across 98% of the data."

Fig 4 How are the equilibriation lines intersecting? And mention which experiment generates this.

We now provide an interpretation of this result. To achieve this, we slightly modified the final form of Eq.8 to isolate two terms: the first term drives $R_{(t)}$ towards R_e at a constant rate, while the second term drives $R_{(t)}$ towards R_e/α at a rate dependent on $(q_e - q)$ (line 206).

$$\frac{dR(t)}{dt} = \frac{\lambda}{\alpha} * \left(R_e - R(t)\right) + \frac{\lambda}{q(t)} * \left(q_e - q(t)\right) * \left(\frac{R_e}{\alpha} - R(t)\right) \qquad (8)$$

We have added the following discussion in the manuscript (lines 459-466):

"We observe that some curves unexpectedly intersect, which can be understood by analyzing Eq. (8). The first term $(\frac{\lambda}{\alpha}(R_e - R_{(t)}))$ continuously drives $R_{(t)}$ towards R_e, while the second term ($(\frac{\lambda}{q(t)}(q_e - q_{(t)})(\frac{Re}{\alpha} - R_{(t)})$ modulates the rate of change. Initially, for simulations with lower q_0 values (e.g., the red and orange curves), $q_e - q_{(t)}$ is large, making the second term significant and positive, thereby increasing $R_{(t)}$ more rapidly. However, as $R_{(t)}$ exceeds R_e/α, the sign of this term reverses, slowing down the increase in $R_{(t)}$ compared to other curves. In contrast, curves with higher initial q_0 values (e.g., blue curve) experience a steadier growth and eventually surpass the initially faster-growing curves, leading to the observed crossing."

We have specified which experiment generated the data (line 448):

"To validate the model, we used Experiment No.3 described in Subsection 3.2."

L441 explain why <1 permil is unrealistic

We have removed the following sentences from the manuscript and retained data points with d-excess values less than 1‰:

"In this dataset, acquired from the drone observations and subsequently corrected using the diffusion modeling, data points with d-excess values less than 1‰ were omitted, as these values are unrealistic and likely result from overcorrection of the δ-values. This resulted in the exclusion of 6 out of 1039 samples."

This revision only affects a very small number of data points (6 out of 1039), and including these data does not alter the overall results and conclusions.

L460 I would explain this in the methods and bring it back in the discussion as a model sensitivity to its parameters

We appreciate the reviewer's advice. The method for error calculation is detailed in Section 3.4.2 ("The Method of Uncertainty Estimation"), while Sections 4.3 to 4.4 focus on interpreting the results. The full range of uncertainties listed in Table 1 is incorporated into the analysis, inherently capturing the model's sensitivity to its parameters. The final uncertainties in Subsections 4.3 to 4.5 reflect this combined maximum error.

To clarify this, we have revised the manuscript as follows (lines 388-391):

"The maximum discrepancy across all calibration results—using the full uncertainty range for $\lambda_{surface}$*, α, and pump efficiency (ε)—was determined. The model-experiment mismatch was then added as an independent error component. The final uncertainty estimates, reported in Subsections 4.3 to 4.5, account for all potential error."*

Additionally, we have expanded the discussion to explicitly highlight the model's sensitivity to its parameters (lines 570-575):

"Among these sources, ε contributes the largest uncertainty, particularly at higher altitudes (Figure 8 and 9), likely due to the conservative uncertainty range we applied to account for potential reductions in collected air mass at high altitudes. Additionally, the fractionation coefficient (α) also contributes considerably to the total uncertainty. To mitigate this, we recommend conducting multiple measurements to obtain an averaged value and performing repeated parameter validation to ensure robustness."

L485 How different are these storage times to really affect the measurements? Can this be incorporated as part of the correction in the model?

The storage duration of air bags typically ranges from 10 minutes to 2 hours. The actual storage time was recorded for each sample and incorporated as a variable parameter in the model. This ensures that the effect of varying storage times on the final measurements is explicitly accounted for in the correction process.

To clarify, we added the following explanation and equation (lines 198-211):

"The constants (λ, $\alpha\_{}^{18}O$*,* $\alpha\_{}^{2}H$*) can be determined through laboratory experiments and Equations 10 and 13 (see Subsection 3.2 and 4.1). If we know the initial values within the air bag (*q_0*,* $R\_{}^{18}O_0$*,* $R\_{}^{2}H_0$*), the ambient values (*q_e*,* $R\_{}^{18}O_e$*,* $R\_{}^{2}H_e$*), and the storage time (*$T_{storage}$*) of the sampling bag, we are able to simulate the*

variations in humidity and isotopic ratios inside the air bag according to Eqs. 5 and 8. Similarly, if we know $T\_{}_{storage}$*, the humidity and isotopic values at time* $t = T\_{}_{storage}$ ($q_{(T\_storage)}$, $R\_{}^{18}O_{(T\_storage)}$, $R\_{}^{2}H_{(T\_storage)}$ *in the air bag, and the ambient values, we can deduce the initial values in the air bag at* $t = 0$ *by back-calculating. The equation used for reconstructing the initial isotope ratio* (R_0) *is:*

$$R_0 = R_{measured} - \int_0^{T_{storage}} \frac{dR(t)}{dt} dt$$

$$= R_{measured} - \int_0^{T_{storage}} \left(\frac{\lambda}{\alpha} * \left(R_e - R(t)\right) + \frac{\lambda}{q(t)} * \left(q_e - q(t)\right) * \left(\frac{R_e}{\alpha} - R(t)\right)\right) dt \quad (14)$$

where R_0 *represents the initial isotopic ratio we want to reconstruct,* $R\_{}_{measured}$ *is the observed isotopic ratio after* $T\_{}_{storage}$*, and* $\frac{dR(t)}{dt}$ *is defined in Eq.8.*

This approach allows us to correct for diffusion-induced isotopic shifts and reconstruct the original vapor composition."

L490 This section has been introduced several times in the paper but is not discussed enough here. I would expect some prior information about why they may be different based on the remote sensing method but necessary to fit wider regions or global models. I would also expect the authors to mention other such repositories like TES and SCIAMACHY.

We appreciate the reviewer's suggestion and have revised the manuscript to provide additional context on the differences between satellite-derived and in-situ measurements. We now clarify that (lines 408-420):

"Satellite measurements, particularly for vertical profiles of water vapor isotopes, are inherently different from direct sampling, they represent a vertical average over layers determined by the averaging kernels (Rodgers and Connor, 2003; Worden et al., 2006). Therefore, their comparability with ground-based or drone-based observations, which provide high-resolution local data, is limited. In this study, we use the MUSICA retrievals from the IASI satellite instrument (Diekmann et al 2021), which provides water vapor isotope data at three altitude levels: 1-3 km in the lower troposphere, 4-7 km in the mid-troposphere, and 8-12 km in the upper troposphere. Given that our study started at an altitude of 3856 m, we used the retrieved δ^2H *data for the 4–7 km and 8–12 km levels. However, these measurements represent a vertical average over layers determined by the averaging kernels (Rodgers and Connor, 2003; Worden et al., 2006). While using averaging kernels to smooth the observed profile could facilitate a more quantitative analysis, we simply averaged the observations for the corresponding altitudes. Consequently, the comparison remains mainly qualitative."*

Additionally, we have expanded the discussion to acknowledge other satellite retrievals, including TES and SCIAMACHY, as follows (lines 394-402):

"Several satellite missions have contributed to water vapor isotope observations,

including the Tropospheric Emission Spectrometer (TES) onboard Aura (2004–2018) (Worden et al., 2006), the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) onboard Envisat (2002–2012), the Atmospheric Infrared Sounder (AIRS) onboard Aqua (since 2002) (Worden et al., 2019), and the Tropospheric Monitoring Instrument (TROPOMI) onboard Sentinel 5 Precursor (since 2017) (Schneider et al., 2022). In this study, we use the MUSICA retrievals from the Infrared Atmospheric Sounding Interferometer (IASI) onboard METOP due to its broad spatiotemporal coverage, vertical profiling capability, and the availability and accessibility of its dataset (Diekmann et al., 2021).”

Fig 10 The left panels are of d2H, but the figure caption and subsequent discussion on d18O. I expect the satellite data to be that of d2H. What am I missing here?

We exclusively compared and discussed δ^2H in this section, as satellite data is only available for δ^2H. I have revised the discussion section accordingly. Thank you for the correction.

Fig 10e Explain why, for higher elevation samples, the satellite dD differs more with measured/corrected data than other altitudes.

We have added an explanation in the manuscript to clarify this discrepancy (lines 543-551):

“Errors due to uncertainties in pump efficiency (ε) are the main source, exhibiting the largest spread (Fig. 8) and increasing with altitude (Fig. 9). Errors derived from $\lambda_{\_surface}$ *and αalso increase with altitude (Figs. 8 and 9). As a result, at higher elevations, the satellite* δ^2*H differs more from the measured and corrected data than at lower altitudes (Fig. 10). This pattern arises because* $\lambda_{\_alt}$ *deviates more from* $\lambda_{\_surface}$ *at higher elevations (Eq.16), primarily due to increased errors in estimating* $M_{\_alt}$*, amplifying correction errors. Moreover, the humidity and isotopic disparity between the air captured in the air bag and lower-altitude ambient air widens with altitude, requiring more intensive corrections. Consequently, both the uncertainty (Figs. 8 and 9) and the magnitude of the diffusion correction (Fig. 7) increase with altitude.”*