**Anonymous Referee #1**

**Review of: "Accurate and high-resolution measurements of UTLS water vapor by balloon-borne laser absorption spectroscopy" by Brunamonti et al.**

**Content**

The manuscript presents the first in-flight validation of ALBATROSS, a balloon-borne mid-infrared laser spectrometer for UTLS water vapour, based on seven soundings from Payerne and Lindenberg (2022-2026). Each flight carried the Vaisala RS41 and a CFH as reference instruments. The authors report tropospheric excellent agreement in comparison to the two reference instruments, and use two balloon-wake "spikes" to argue that ALBATROSS has a low response time of about 1s. A substantial part of the paper documents the identification and iterative mitigation of internal contamination pathways and the in-flight CO2 purge system developed to suppress them.

**Overall impression and rating**

This is a valuable and timely contribution and, in my view should be clearly published in AMT. Independent, fast-response, spectroscopically traceable balloon hygrometer is exactly what the community needs given the phase-out of the cooling agent R23 for the CFH/FPH operation, the transition to new chilled-mirror concepts, and the limited future satellite coverage. The engineering work is careful and honestly reported and I particularly appreciate that the authors document the failures rather than only the successes. The F6 profiles are a genuinely impressive result. The manuscript is well written and well organised and I therefore recommend publication in AMT after some minor revisions stated below.

**Authors:** We thank the Referee for the positive and constructive feedback that helped us to improve our manuscript. Below are the individual comments from the Referee (in black) and the replies from the Authors (in blue). Page and line numbers refer to the revised manuscript.

**Main comments:**

Measurement technique: Which pressure and temperature are used to convert retrieved H2O number density (or line area) to mixing ratio? RS41? An in-cell sensor? The SC-MPC body temperature? This is the single most consequential retrieval input and it is not stated. And on a related point: How homogeneous are the temperature and pressure fields in the Albatross inlet, since that also affects retrieval?

**Authors:** The $H_2O$ mixing ratios retrieved from ALBATROSS use the air pressure and temperature measured by the RS41 radiosonde. This is now stated in Section 4:

"*The $H_2O$ mixing ratios retrieved from ALBATROSS use the air pressure and temperature measured by the RS41 radiosonde, interpolated to the time grid of the recorded spectra.*" (page 10, lines 6-7)

The homogeneity of temperature and pressure is ensured by the large flow rate of air in the intake tube of the instrument. In particular, the ALBATROSS retrieval is only sensitive to the temperature and pressure fields within the thin slice of air interrogated by the laser beam along

the optical plane (thickness < 3 mm), with a volume of < 28 $cm^3$ (cell diameter = 10.88 cm). With an inlet tube diameter of 7.9 cm and an ascent rate of 5 m $s^{-1}$, the total flow rate of air passing through the instrument during flight is approximately 0.025 $m^3$ $s^{-1}$ (i.e., 25 L $s^{-1}$). Any residual temperature gradients are expected to be negligible compared to the uncertainty associated with the RS41 measurements, owing to the extremely short residence time of air within the sensing volume (<2 ms). The following clarification sentence was added to the manuscript:

"*The large flow rate of air passing through the instrument (~25 L $s^{-1}$ at 5 m $s^{-1}$ ascent rate) ensures homogeneous temperature and pressure fields in the sensing volume.*" (page 4, lines 16-18)

Do corrected data enter the statistics? Table S1 lists the five F3 intervals as flagged ("partly corrected by diffusion model"), while Fig. 6 presents the corrected points among "valid measurements". Section 4.4 defines flagged points as those excluded. Please state unambiguously whether the corrected points are included in Table 2 and Fig. 8, and report the statistics with and without them. As it stands, a reader cannot tell whether the stratospheric 3 +/- 10 % includes model-corrected data.

**Authors:** Yes, the corrected ALBATROSS data of flight F3 enter the statistics, up to a pressure of 40 hPa (above which, they are flagged due to external moisture sources: see Table S1). This is now explicitly mentioned in Section 4.3:

"*For the statistical comparison (Section 4.4), the corrected ALBATROSS measurements shown in Fig. 6a are considered as valid data up to 40 hPa pressure, above which they are flagged due to external moisture sources (see Table S1).*" (page 16, lines 11-13)

The statistics have been recalculated after excluding the data points affected by the correction (total 7 data points between altitude levels 14.75–22.25 km at 500 m averaging) and the result (4 ± 10 %) is now stated in Section 4.4:

"*Note that the mean deviation becomes 4 ± 10 % excluding the ALBATROSS data of flight F3 corrected by the diffusion model (7 data points at 500 m averaging). This demonstrates that the overall conclusions are insensitive to the inclusion of model-corrected measurements.*" (page 20, lines 2-4)

Purging system: Does CO2 reach the optical cell, and if so what does it do to the retrieval? The connectors are held at slight overpressure with CO2. If CO2 leaks into the cell, two possible effects could arise: first any absorption or continuum contribution near 1662 cm-1 line, and second more importantly, a change in the collisional broadening of the H2O line, since CO2-broadening coefficients differ from air-broadening. Please comment on that, if this could be potentially affect the retrieval.

**Authors:** The $CO_2$ purge gas does not reach the optical cell, since the SC-MPC is sealed at both apertures by nitrile O-rings and tilted $CaF_2$ windows (see Graf et al., 2018). The purge gas stays within the volume of the two connectors installed between the cell and the laser/detector, which are kept at slight overpressure, and slowly leaks out towards the inner volume of the instrument's enclosure. Furthermore, no absorption features of $CO_2$ occur in the selected

spectral window (~1 $cm^{-1}$ around 1662.809 $cm^{-1}$). The following sentence has been added to Section 2.2 to clarify this issue:

"*The SC-MPC is sealed at both apertures by nitrile O-rings and tilted $CaF_2$ windows (Graf et al., 2018), preventing the purge gas from reaching the optical cell*" (page 6, lines 13-14)

Missing concentration dependence analysis: You have done a great job of illustrating the differences between the instruments on a height scale and thus also on a pressure scale, but in validation studies, a comparison as a function of H2O is typically used as well. This can be done either with a regression of ALBATROSS against CFH (log-log or orthogonal regression, with slope, intercept and confidence intervals) over the full 5-5000 ppm range or just with the difference in H2O (detlaH2O) along the reference. A slope test is the standard way to demonstrate that agreement holds across three orders of magnitude in H2O and a addition figure showing this would be a very good addition to the paper.

**Authors:** We thank the reviewer for this valuable comment. Following this suggestion, we have performed and integrated in the manuscript a concentration dependency analysis of all ALBATROSS measurements against both RS41 and CFH. This includes a new figure (Fig. 9), showing scatter plots of the $H_2O$ mixing ratios from ALBATROSS against CFH and RS41, and a log-log regression model applied to quantify the linearity of the two relationships.

The text below was added to Section 4.4 and the figure can be found in the revised version of the manuscript. We agree that this analysis strengthens the validation and have therefore incorporated it into the revised manuscript.

"*To further investigate potential concentration-dependent biases in the ALBATROSS retrievals, Fig. 9 compares altitude-averaged $H_2O$ mixing ratios measured by ALBATROSS against RS41 (panel a) and CFH (panel b), for all valid data included in the intercomparison. The corresponding relative differences ($\Delta H_2O$) as a function of H2O mixing ratio measured by RS41 and CFH are shown in Fig.9c-d.*

*The measurements demonstrate a high degree of linearity of the ALBATROSS retrievals over more than three orders of magnitude in H2O mixing ratio (~3–5000 ppm), with no evidence of significant concentration-dependent deviations. This behavior was quantified using a log-log regression analysis of type $y = Bx^a$ (solid lines in Fig. 9). The fitted values (95 % confidence intervals) of the regression coefficients are a = 0.973 (0.967–0.980), B = 1.134 (1.093–1.177) ppm for the ALBATROSS vs. RS41, and a = 0.989 (0.985–0.993), B = 1.049 (1.026–1.073) ppm for ALBATROSS vs. CFH. The corresponding coefficients of determination ($R^2$) are 0.9964 and 0.9980, respectively. These results confirm the absence of systematic trends in $\Delta H_2O$ across the investigated concentration range and, indicate that the ALBATROSS retrieval remains linear and accurate from the humid lower troposphere to the very dry lower stratosphere.*" (page 21, lines 5-16)

Please note that, while performing this additional analysis, the following two issues were identified and corrected in the revised manuscript:

- Three individual data points, previously considered as valid, were measured at $H_2O$ mixing ratios between 6000–9000 ppm (according to CFH), which is above the ~5000

ppm limit beyond which the ALBATROSS absorption line becomes saturated. Therefore, these data points were excluded from the statistical comparison. This affects, namely, altitude level 4.375 km of flight F2 (ascent), 3.875 km of flight F3 (ascent), and 3.75 km of flight F6 (ascent). All these data points are found in the lowermost altitude levels of the profile and their removal has no effect on the overall conclusion of the paper.

- A plotting error was found in Fig. 4 (panels h-n) and Fig. 5 (panels b, d), where the standard deviations (error bars) of the ALBATROSS – RS41 differences were mistakenly applied to the ALBATROSS – CFH data points. No numerical values reported in the manuscript were affected. The figures have now been revised by applying the correct standard deviation bars to the ALBATROSS – CFH data points.

Figures 4, 5, 8 and Table 2 of the manuscript were updated accordingly. The effect of these corrections is very minor and limited to the lower tropospheric part of the intercomparison, leaving the overall results of the paper unchanged.

**Further Specific comments/questions:**

Abstract, line. 30 and Sect. 5: Please state that valid stratospheric data were obtained in four of seven flights, and that the 12-28 km statistics rest on two independent profiles.

**Authors:** The abstract was modified as follows:

"*In the stratosphere, valid data were obtained in four out of seven flights and show that ALBATROSS achieves a mean agreement of 3 ± 10 % relative to CFH between 12 and 28 km altitude*" (page 1, lines 29-30)

Please note that the 12–28 km statistics are derived from four independent stratospheric profiles (F2, F3, F5, and F6), since flights F2 and F5 also provide valid data above 12 km (besides F3 and F6).

Page 2, line 15: I think there are also satellite measurements with higher resolution. For example, ACE-MAESTRO provides an altitude resolution of about 1 km (e.g. Weaver et al. 2019). I would suggest expanding the range to 1-3 km and adding the phrase "depending on the satellite."

**Authors:** Sentence modified accordingly:

"*Satellite instruments provide global observations of UTLS $H_2O$ with extensive spatial and temporal coverage and a typical vertical resolution of 1–3 km, depending on instrument (e.g., Hegglin et al., 2013; Santee et al., 2017; Weaver et al., 2019)*" (page 2, lines 14-15).

Page 6, line 15-17: The purge budget doesn't close. 25g CO2 = 0.568 mol approx. equals 12.7 sL which result in approx. 85-127 min at the stated 0.10-0.15 slpm, not the "about 3 hours" claimed. Can you please comment on it ?

**Authors:** We thank the reviewer for this comment. This statement was indeed overly simplified and potentially leading to confusion. In fact, the flow rate is nearly constant at 0.1 slpm for about 2 hours (as long as liquid $CO_2$ is present in the cartridge) and then decreases slowly with

time (as the remaining pressure inside the cartridge decreases), resulting in a total release time of about 3 hours above a threshold of 0.02 slpm (empirically determined minimum flow rate required to suppress outside-cell absorption effects).

The description has been now improved as follows:

"*This solution provides a nearly constant flow rate of ~0.1 slpm for about 2 hours, as long as liquid $CO_2$ is present in the cartridge and the internal pressure is maintained at the $CO_2$ saturation vapor pressure (~57 bar at 20 °C), followed by a gradual decrease of the flow rate once the liquid CO2 has fully evaporated, resulting in an effective release time of about 3 hours.*" (page 6, lines 19-22)

Page 7: Please state how ALBATROSS, CFH and RS41 time bases were synchronised, and the residual uncertainty as this is very essential for the later comparison.

**Authors:** The ALBATROSS and RS41 time grids are synchronized using the time of launch of the balloon (i.e., onset of pressure decrease), detected by both RS41 and the spectrometer's internal pressure sensor (MS8607), as a reference. This is now mentioned in Section 4, along with specifications on the residual uncertainty of this procedure:

"*The two signals are synchronized using the time of launch of the balloon (i.e., onset of pressure decrease), detected by both RS41 and the spectrometer's internal pressure sensor (MS8607), as a reference. The uncertainty associated with the synchronization is approximately ±1 s, which is negligible compared to the 50–100 s averaging intervals used for the statistical comparison (see Section 4.4).*" (page 10, lines 7-10)

For the comparison with CFH in tandem flights (F0-F3), the data are simply averaged in altitude bins, and no time synchronization is performed (since the instruments are flown on separate balloons with different ascent rates). In co-located flights (F4-F6), CFH data share the same time grid as RS41 data.

Page 8: Flight F2 used a dry-ice/alcohol cryogen mixture rather than R23 for the CFH. Since cryogen choice is one of the motivations for the paper, does F2's performance differ systematically from the R23 flights? Even a sentence would be of interest to GRUAN readers.

**Authors:** The performance of CFH operated with a dry ice/alcohol (DIA) cooling mixture in flight F2 shows no systematic differences relative to the R23 soundings. A preliminary analysis of the *Swiss $H_2O$ Hub* project dataset shows that, besides some small differences (e.g., longer "heating cycles"), DIA-CFH measurements are generally characterized by the same noise level in the reflectivity signal and same mirror temperature calibration accuracy as the R23-CFH measurements (Poltera et al., 2024). A systematic investigation of the CFH performance with the two cooling agents is in preparation and will be presented in a separate manuscript. The following text, including a reference to Poltera et al. (2024), was added to provide more insights on this topic:

"*The performance of CFH in this flight showed no systematic differences compared to R23 soundings, and a preliminary analysis of the Swiss H2O Hub dataset indicates that DIA-CFH measurements were generally characterized by the same reflectivity signal noise level and mirror*

*temperature calibration accuracy as the R23-CFH measurements (Poltera et al., 2024).*" (page 8, lines 20-23)

Reference: Poltera, Y., Wienhold, F. G., Peter, T., Brossi, S., Brossi, T., and Romanens, G.: Reference Humidity Retrievals with 'Dry Ice' CFH during the Swiss H2O Hub Summer 2023 campaign, 15th GRUAN Implementation and Coordination Meeting (ICM-15), Bern, Switzerland, 11-15 March 2024, https://www.gruan.org/gruan/editor/documents/meetings/icm-15/pres/pres_0606_Poltera_Dry-Ice-CFH.pdf (last access: 23 October 2026), 2024.

Page 10, line 26: I guess the low ascent rate cannot be explained by the parachute. The drag by the balloon itself is much larger than the one of the parachute. For me it looks like the balloon had a small leakage and thus lost slowly helium this would explain, why the ascent rate even further decrease with time (see your Fig. 2).

**Authors:** The exact cause of the anomalously low ascent rate of flight F2 cannot be established retrospectively, due to the lack of quantitative evidence. A small leak of the balloon lifting gas ($H_2$) represents a plausible explanation and therefore has been added to the manuscript alongside enhanced drag effects (see below). The analysis of ascent rate data by RS41 does not allow to clearly confirm nor exclude any of the two hypotheses.

"*[…] possibly due to a balloon filling ($H_2$) leak or enhanced drag effects by the parachute in a strong wind field*" (page 11, line 8)

Figure 4, Section 4.2: Is there a reason why the Albatross measurements in the troposphere don't start at higher pressures than about 750 hPa or lower? Is the selected water vapor line already visually dense due to the increased concentration? I think it would be good to at least mention this in this section.

**Authors:** Yes, the ALBATROSS retrieval typically starts between 2 and 4 km altitude because the absorption line becomes saturated at the high $H_2O$ mixing ratios encountered at these altitudes. This was already stated in Section 4.2 (page 13, lines 4-5) and has been further clarified as follows:

"*Note that the ALBATROSS retrieval typically starts between 2 and 4 km altitude, as the absorption line becomes saturated at the high $H_2O$ mixing ratios (above ~5000 ppm) encountered near the surface (see Graf et al., 2021).*" (page 13, lines 9-10)

Figure 6a: Above approx. 30 hPa the corrected (turquoise) points scatter between roughly 8 and 20 ppm against CFH values of 6-7 ppm, which is more than the "slight moist bias of 1-2 ppm" stated in the text. Please reconcile the text with the figure, or restrict the statement to the altitude range where it holds.

**Authors:** Thanks for this comment. The text was revised accordingly:

"*Above ~24 km altitude, ALBATROSS shows a moist bias of up to ~10 ppm (increasing with altitude) relative to CFH*" (page 16, lines 9-10)

**Technical comments/questions:**

Page 8, lines 21-22: "European Centre for Medium-Range Weather Forecasts (ECMFW)" --> ECMWF.

**Authors:** Corrected accordingly (page 8, line 31)

Page 12, line 31: "F3 reached an altitude of 35 km (approx. 6 mbar)"; Please use hPa consistently.

**Authors:** Corrected accordingly (page 11, line 13)

Page 13, line 1: "flights F1 (Fig. 2i) and F4 (Fig. 2l)" should be Fig. 4i and Fig. 4l.

**Authors:** Corrected accordingly (page 13, line 14)

Page 17, line 13: "wake of the ballon" --> balloon.

**Authors:** Corrected accordingly (page 18, line 10)

Page 18, lines 1-2: "5 m for ALBATROSS and and 15-20 m for CFH", Please remove the duplicate.

**Authors:** Corrected accordingly (page 18, line 20)

**References**


- Weaver, D., Strong, K., Walker, K. A., Sioris, C., Schneider, M., McElroy, C. T., Vömel, H., Sommer, M., Weigel, K., Rozanov, A., Burrows, J. P., Read, W. G., Fishbein, E., and Stiller, G.: Comparison of ground-based and satellite measurements of water vapour vertical profiles over Ellesmere Island, Nunavut, Atmos. Meas. Tech., 12, 4039–4063, https://doi.org/10.5194/amt-12-4039-2019, 2019.