Original referee comments in blue, with author response in black.

1) L. 4: not so clear to me what is meant by "which can have a vertical structure", do you mean the fluxes or the atmosphere.

Rephrased to: “However, gathering observations in harsh environments still poses challenging, particularly in regard to observing the small-scale exchanges taking place between surface and atmosphere. It is especially important to resolve the vertical structure of these processes.”

2) L. 5: "lowermost level of the surface layer": do you mean a layer or a specific level? Maybe make this clear by mentioning the first 2 m above the surface, which is what you designed your system for.

Rephrased to: “We have designed the ISE-CUBE system as a modular CRDS deployment system for profiling stable water isotopes in the surface layer, specifically the lowermost 2 m above the surface.”

3) L. 25: "structures" -> infrastructure?

Yes, infrastructure is a more accurate term. Changed accordingly

4) L. 28: robustness of the sampled gradient?

Changed accordingly

5) L. 44: Arctic with capital A

Changed accordingly

6) L. 54: induce minimal flow disturbance

Changed to: “The entire system should also induce minimal disturbance to the flow around the measurement site.”

7) L. 71: add a manufacturer and a serial number for the check valve for completeness, since you gave this information for all other components (I know that you have it in the table in the Appendix, but for consistency with the rest of the text, I would add it).

Added accordingly

8) L. 75: remove one "the"

Changed accordingly

9) L. 86: remove "now" this is not the first instrument with higher flow rates

Changed to: “... is a custom modification of the standard L2130-i, which enables higher flow rates, similar to the analyzer described in …”

10) L. 99: does Inlet really need a capital I (here and elsewhere)

Changed accordingly, except in cases where a reference is made to text in a figure

11) L. 106: "assign a spatial dimension to these measurements" sounds a bit confusing. How about something along the lines of "We will next describe a profiling system, that enables high-resolution vertical profiling of the lowermost atmosphere", or similar. But make clear it is about sampling vertical gradients.

Changed to: “We will next describe the Profiling module, which enables high-resolution vertical profiling of the lowermost levels of the surface layer.”

12) L. 118: "The height of the inlet..."

Changed accordingly

13) L. 187: The last sentence is a bit confusing, your statistical analysis does not allow you to quantify the extremeness of the full period. I would simply state: "Over the time period of the deployment, we encountered several episodes of extreme cold conditions", or something along this line of thought.

Rephrased to: “Additionally, Dahlke et al. (2022) have identified Feb and Mar 2020 as having some of the strongest marine cold-air outbreaks of the last 42 years. Overall, with several episodes of extreme cold, our deployment was a formidable testing ground for the system.”

14) L. 213: "to be compared" sounds a bit too qualitative for my taste. How about writing what you actually do, namely you normalise your data to the VSMOW-SLAP scale? Also you should mention what the primary standard is somewhere because you mention that you calibrate with secondary standards, which can be confusing for an unexperienced reader.

Changed to: “Isotopic measurements are calibrated on the VSMOW2-SLAP2 (Vienna Mean Ocean Water 2 - Standard Light Antarctic Precipitation 2) scale, composed of international, primary standards. The use of the scale allows for the relative ratios of heavy to light isotopes (R_{sample}) measured in CRDS analyzers to be normalized to a standard ($R_{standard}$), as described in Eq. 1 (Craig, 1961).

…

The value of the resulting δ^i (with i representing one of the heavy SWIs) is expressed in permil (per thousand, ‰). For our calibrations, we employed two secondary standards, whose isotopic values on the VSMOW2-SLAP2 scale have been established in the laboratory.”

15) L. 237: mention the 20° already in the second sentence: well-controlled laboratory setting at room temperature of approximately 20°C. Remove "was operating with ambient room temperatures of approx 20°C" below.

Rephrased to: “The first represents the optimal operating environment for the analyzer, a well-controlled laboratory setting with ambient room temperatures of approximately 20 °C. This first period runs from June–July 2020 when the same analyzer was used at FARLAB, University of Bergen, Norway. During this time, the analyzer was routinely sampling standard vapor with mixing ratios down to 0.155 g/kg, comparable to humidity minimums encountered in the field.”

16) L. 239: Merge with the next sentence: "During this time the analyser was routinely sampling standard vapour with mixing ratios down to ... "

Rephrased (see comment above)

17) L. 240: The second period,

Changed accordingly

18) L. 257: I think referring to the different results sections in this first paragraphs would be more intuitive than referring back to the methods section.

Agreed. The paragraph now reads: “We now detail how the field conditions influenced analyzer performance and thus data quality, using the laboratory and observatory periods as performance benchmarks. Thereby, we focus first on temperature and pressure conditions of the analyzer (Sect. 4.1), before evaluating the impact of field conditions on the water isotope measurements (Sect. 4.2). Then we detail the performance of the profiling module, especially the capability of the module to deliver sample to the analyzer for the purpose of resolving vertical profiles (Sect. 4.3). Finally, the performance of the cold trap expansion module is briefly presented (Sect. 4.4).”

19) L. 268: TDAS at L. 82 you have Tdas, be consistent with the variable names, and consider putting variables in italics, using $T_{\rm{DAS}}$, $p_{\rm{DAS}}$,...

Removed T_{das} on L.82, and changed variables to italics

20) Very nice shorting of Section 4.1, the information is now nicely condensed and this part is interesting to read.

Thank you. The comments of the referees in the previous round really helped us pin-point the utility of this section.

21) L. 284 and elsewhere: not sure the correlation value of 0.000 is needed, no correlation is clear enough.

Removed accordingly

22) The y-axes of Fig. 7 are labelled in a slightly disturbing way: it should be sigma for all of them, right? I find d18O,sigma confusing.

Rearranged labels accordingly. Now of the form: σ, $\delta^{18}O$ (‰)

23) Fig. 7 caption: Differences between bin means from field and laboratory…

Changed accordingly

24) L. 368: this corresponds to a variability increase of 30% in the field compared to the laboratory benchmark.

Rephrased to “Averaged across humidity bins, this corresponds to a variability increase of 30 % in the field compared to the laboratory benchmark, though the largest relative increase occurs at the higher humidities.”

25) L. 380: remove one "the"

Changed accordingly

26) L. 426: I find it surprising (and unlikely) that the dD should be so strongly affected by non-equilibrium fractionation effects, while d18O is not... Could it be that the difference comes from the normalisation of the liquid measurements (bias in the depleted standard used), or due to some interaction between the sampled vapour and the tubing (which is likely stronger for dD, but I wouldn't call that a "kinetic effect").

We’ve removed specific reference to kinetic, but keep mention of fractionation effects in general. The very low temperatures of the collection vial have an impact on the difference between fractionation factors of dD and d18O, which can manifest during incomplete freeze-out. Additionally, we also emphasize that sample loss is a major factor in the comparison disagreement. We have also now included a caption for Table 4.

Rephrased to: “This is likely due to a combination of deficiencies involving sample collection, inconsistent flow regulation, and possibly fractionation effects at the low temperatures in the glass sample vial. These fractionation effects might arise from incomplete freeze out of the vapor, induced by insufficient thermal stability of the glass collection vial. We additionally observed substantial ice crystal formation in the neck of the collection vials, which inhibited and decreased flow during multiple collection periods. This ice formation also compromised sample recovery during vial exchange, causing frozen sample to fall out of the vial during collection. Therefore, the disagreement between sample and analyzer measurements is not unexpected, especially as the collected sample would no longer correspond with the integrated time period being compared with.”

27) Very nice that the authors added a comparison between the on-line vapour measurements from the CRDS system and the cold trap sample.

Thank you. We believe it was worth the effort to support our qualitative assessment with quantitative numbers.

28) Fig. 10 and L. 447: I wondered why some of the higher elevation measurement points have such a small error bar in the dexcess. Shouldn't these error bars also represent the precision of the measurements that were quantified in Fig. 7 to be of about +/-2 permil in dexcess at 0.8 g/kg (3 m height). I think the error bars should represent total uncertainty, even more so since this is how the authors argue at L. 375-377 they would proceed to use their results of Section 4.2.

As we utilize the measurement means to obtain the gradients, we believe that it is more relevant that the errorbars show the uncertainty in that mean. We have now included a variability scale for each profile, denoting a representative standard deviation of any particular height. Sect. 4.5 and Fig. 10 have been revised that reflect this change.

We have also corrected L.375-377, as the increased uncertainty is moreso inherent to our measurement system, rather than something that needs to be added to it, which is how L.375 reads now. This sentence (L.380) has been changed to: "Nonetheless, we will show that though this decreased precision is inherent to the observations, it does not hinder useful measurements, in particular since the measurement precision is quantified."

29) Excellent analysis in Section 4.5! Very nice to see!

Thank you. Just as with shortening Section 4.1, the referees in the previous round highlighted the necessity to expand on the performance of the profiling arm.

30) L. 456: measuring,

I don't think a comma is necessary here as "being able to select heights of interest while measuring" is the subject of the sentence

31) L. 465: context in which, they are taking place.

Changed accordingly

32) L. 466: I think it would be very important to mention that fixed height profiling systems with only few unadjustable inlet heights still have the advantage of allowing to sample automatically over much longer time periods, thereby obtaining a different statistics than with the deployment of a profiling system like the ISE-CUBE. To me the ISE-CUBE is by design operating punctually in a campaign-based setting and it needs an operator.

Changed to: "Finally, it would be quite possible to deploy alongside a tower with fixed height inlets, as these towers have the advantage for automated sampling over much longer time periods."

33) L. 470: could be used without…

Changed accordingly

34) L. 471: I don't understand "a point of diminishing returns"

Rephrased to: "While the Analyzer and Pump modules could be used without the Profiling module for an extended deployment, at some duration the benefits of a more conventional enclosure would begin to warrant additional effort during installation. A larger enclosure (such as a pre-existing building) offers a level of security that the ISE-CUBEs cannot provide for long-term measurement efforts… "

35) L. 480: I don't really understand the use of this last paragraph. I think it is very unlikely the case that the quality of the measurements would be substantially improved in low flow mode and it would require longer sampling times at a given height.

This point was brought up during the interactive discussion, and we thought it relevant to include it as a possibility.

36) L. 484: most rapid -> fast

Replaced accordingly

37) L. 490: replace in the system by ISE-CUBE

Replaced accordingly

38) L. 498: Due to the high vertical resolution in the profiles, the observed gradients are robust.

Changed accordingly

39) L. 509: yes but unlikely optimised for high flow-rates... isn't this the major challenge here?

The calibration device described in Leroy-Dos Santos (2021) provides flow up to 0.6 LPM, which would be enough to calibrate the analyzer, even when it is in high flow mode. The ~10 LPM flow rate discussed in Sect 4.3 pertains only to inlet characterization (Profiling module performance), not analyzer calibration.

40) Consider moving part of the appendix figures and text into supplements (which would be accessible in separate PDFs and would avoid charging the paper PDF unnecessarily. Of course Appendix A should stay an appendix.

We have moved Appendices B and C into the Supplement.

References

Craig, H.: Isotopic Variations in Meteoric Waters, Science, 133, 1702–1703, 1961.

Dahlke, S., Solbès, A., and Maturilli, M.: Cold Air Outbreaks in Fram Strait: Climatology, Trends, and Observations During an Extreme Season in 2020, Journal of Geophysical Research: Atmospheres, 127, 1–18, https://doi.org/10.1029/2021JD035741, https://onlinelibrary.wiley.com/doi/10.1029/2021JD035741, 2022

Leroy-Dos Santos, C., Casado, M., Prié, F., Jossoud, O., Kerstel, E., Farradèche, M., Kassi, S., Fourré, E., and Landais, A.: A dedicated robust instrument for water vapor generation at low humidity for use with a laser water isotope analyzer in cold and dry polar regions, Atmospheric Measurement Techniques, 14, 2907–2918, https://doi.org/10.5194/amt-14-2907-2021, https://amt.copernicus.org/articles/14/2907/2021/, 2021.