

## Authors' responses to the comments of reviewer#2

Thank you for reviewing and commenting on our manuscript, "Technical note: A fast and objective autosampler for direct vapor equilibration isotope measurements". We find your advice constructive and will incorporate your suggestions into our revision. We responded to each comment individually below.

Thank you for letting me review this manuscript. The authors present an autosampler for measuring the water isotope composition of water (and potentially, soil and plant) samples using the direct vapor equilibration method. I know and appreciate the work of the group and I am user, hence, very familiar with the DVE-LS method as well. Providing an open-source autosampler for increasing sample throughput and minimizing labor and cost is simply amazing and I applaud the others for developing and providing all codes publicly. However - and I am sorry to say this - the presented study has severe inconsistencies and shortcomings, which simply need to be addressed before publishing. In brief, I could summarize these as follows: i.) for a convincing proof-of-concept, the results for measuring 21 samples is simply insufficient; ii.), the ms. refers to soil samples, CVD and artifacts throughout, but does not actually report any results for using the autosampler on soil sample equilibrations, which would be much more interesting and relevant for the actual application of the method; and iii.) potentially relevant factors were not tested for nor commented on (e.g., the effect of leaving the samples to be measured pinched by the canules). Additionally, the structure of the ms is currently confusing and is missing a red line (mainly due to ii.) I believe). I hope my comments and suggestions help in improving the ms. Having that said, I totally support publication of the autosampler, after the data foundation for a true objective evaluation is improved.

General/Main comments:

Were only 21 samples measured in total? This is very low for assessing accuracy/uncertainty and makes me seriously wonder – the authors have the autosampler, why would they rely on such a small number of samples for proofing its accuracy and uncertainties?

We have already measured 1000+ soil samples with the VapAuSa. However we did not report on these samples as we did not co-measure them by hand. These samples originated from mountainous catchment soil profiles where the seasonality- and altitude effect of precipitation is difficult to assess. Therefore, assessing the measurement uncertainty would be nearly impossible.

However, to address this and the following point, we will take more soil samples and co-measure them by hand and autosampler and report on the findings in the next manuscript revision.

The manuscript mentions soil samples, CVD, and artifacts introduced by extraction multiple times and quite prominent; however, no soil data is presented in the manuscript at all. This is confusing me extremely. Firstly, CVD is not of relevance for the autosampler at all for the purpose

of proof-of-concept. Secondly, I would've loved to see comparisons between manual injections and autosampler results for soil samples! I don't understand why this is not part of the study.

As stated before, we will take soil samples and measure them manually and by autosampler.

155-165: This chapter is completely unrelated to the actual study, which was done using liquid water, not soil samples. Also, the 'bashing' of CVD and explanation of the differences or not of any importance for the presented study. I strongly suggest to focus on comparing and testing the accuracy and uncertainty of the data obtained with the autosampler vs. manual injection and liquid water measurements for comparable matrices, i.e., liquid, non-extracted water.

We will discard the whole section regarding CVD and focus on the DVE-LS method improvements by the autosampler.

Same as above is true for chapter 4.2: The study did not compare data obtain from CVD with data obtained by DVE-LS – why is this discussed in the manuscript? I suggest again to focus on the actual purpose of the study: Provide an autosampler for water vapor isotope measurements and proof its accuracy, benefits and caveats. The methodological artifacts introduced by such a study must be eliminated, i.e., by only using water samples.

We think this point is already addressed by our water-to-water extraction of DVE-LS, done manually, by autosampler as well as liquid measurement. However we will try to improve the clarity of the section so the suggested points are addressed more clearly.

Avoid colloquial language and vague statements ("some", "lots of", e.g., l.195)

We will rephrase the content to scientific language.

Minor comments:

Abstract:

Vapor or vapour? à use consistent BE or AE

Thank you for pointing this out, we will revise it to the AE version

12: the reported uncertainty of both manual and automated measurement for d2H seems pretty high, compared to liquid injections. It would be great to see some elaborations on the reasons and/or recommendations to improve this somewhere in the ms.

Yes, the uncertainties are pretty high, we will address this in the revised manuscript

Please report here on the methods also (how was the reported accuracy etc. elaborated – how many samples of which matrix were tested)

Regarding this and the following point: We will rework the accuracy section to incorporate measurement repeatability (as standard deviation) as well as giving measurement bias (average deviation from measurement to "true" liquid measurement). This will then also be part of the abstract.

Why is the later mentioned accuracy not reported in the abstract? (and also, repeatability)

Text:

20: obtain estimate

We will change this in the revision

24: distribution composition

We will make this alteration in the manuscript

26: partitioning effects à unclear what is meant with this, please rephrase or explain.

Suggestion: water partitioning

Thank you for the suggestion, we will change it

29/30: this sentence is unclear, it is very much possible to extract water from soils and plants; isotope analyzers are also not built to measure specifically liquid water; in fact, the lasers all measure water vapor. The autosamplers though are build for liquid water, I guess this is what is meant here. Please rephrase.

Yes, we wanted to state that the isotope analyzers (CRDS and OA-ICOS) will need water-vapor as an input to measure. Therefore an extraction and vaporization or an equilibration is necessary.

We will rephrase the content for more clarity.

31: DVE is not really an extraction, it's an equilibration

We will clarify this.

32-34: the explanation of CVD is explained very sloppy; from the current explanation it is absolutely not clear how CVD works. Also, the citation used is very old – the method has been much improved meanwhile (classic citation Koeniger et al., 2011, but also many others)

As stated above, we will discard the whole section regarding CVD and focus on the DVE-LS method improvements by the autosampler.

38: For other users, a reference to the company and size of the laminated bags would be useful here

Thank you for pointing this out, we will name the manufacturer and sizes used in the application

49: volume similar (space missing)

We will change this

49/50: CVD requires much less soil compared to DVE-LS, only 10g are sufficient for one replicate in many instances. Having that said, it should be stated somewhere how much soil is needed. Also, the constraint for high spatial and temporal sampling is often the digging, not the sample storage or amount.

We will change this paragraph to more clearly state the requirements and discard the CVD comparison. In our experience, the drilling was not as limiting as the sample processing and

analysis. However we acknowledge that both can be limiting depending on your available field- and laboratory-staff and equipment. Therefore we will rephrase this statement.

74: please provide details of the cannulas (diameter) – it might be a constraint for providing the 35ml/l to the Picarro and induce errors if diameter is too small

Thank you for pointing this important fact out! We will state the diameter (2.1 mm) in the next revision

127: It seems like the cannulas all have to be connected to the bags in the beginning of the measurement cycle. This means that the later they are in the sampling line, the more potential exists for contaminating the sample because the sample was pinched. It would be great to have some comments on this. Also, silicone might affect the absorption spectra of the measurements – could this be explanations for the relatively high uncertainties of the method? It would be nice to see some elaborations on this.

We will add this to the manuscript. So far what the data of the 21 bags tell us is that there is no effect on later samples by the pinching. However we only sampled for 5 hours. With larger systems changes might get more pronounced. Therefore we will also do a longer test and report on it in the revision

Regarding the silicone, we think this effect is minimal as other studies which also applied the silicone had lower uncertainties (e.g. Gralher et al., 2021)

129: Just referring to the table is not sufficient, please explain what the stability criteria actually checks.

We will improve this in the revised manuscript

135: next to uncertainty, the accuracy of the method should also be reported in the abstract

We will add this.

152/153: Perhaps the authors are right, but this needs to be further discussed. As stated before, there are potential influencing factors which could cause a difference of the manual vs. automated injections; e.g., the time until a sampling is analyzed (intrusion of atmospheric air?) and the potential contamination with silicone (could affect only one isotope and hence, explain different behavior of d2H and d18O).

As stated before, we will assess the effect of time until sampling and don't think silicone will be such a contaminating source (after all, the Picarro liquid vial caps are also lined with silicone). We will further discuss this uncertainty in the revision.

173/174: I agree that ONLY reporting repeatability is masking the actual outcomes, but not checking for repeatability is equally questionable when providing a proof-of-concept. Why not report both?

As stated above, we will rework the accuracy section to incorporate repeatability as well as measurement bias. The measurement bias is  $<0.1\text{e-}13$  for both isotopes, therefore we did not

think about reporting it. But we agree that giving this might be beneficial, therefore we will add it to the revised manuscript

General suggestion for the discussion: Try to stay objective in the statements. Reading the discussion, I felt that the authors are trying to convince me how good the proposed method is, rather than objectively reporting and discussing the pros and cons.

We will rephrase the content to be more objective.

4.3: The elaborations here are good, but please comment on the potential effect of pinched bags during the measurement cycle

We will add this to the manuscript.

216: may be is ;)

Thank you, we will rephrase it!

219/220: "So far we tested the system with liquid and soil samples only" Why are these results not presented here?

As stated, we will add this in the manuscript