

Response to reviewer 2 (Prof. Ansgar Kahmen)

The reviewer's comments are in normal font and the authors' responses are in bold font.

This manuscript by Scott Allen and James Kirchner evaluates the effects caused by biased xylem water d2H measurements in the assessment of plant water sources. The manuscript is motivated by several recent reports suggesting biases in the isotopic composition of water that are associated with the cryogenic extraction of water from soils and plant materials. In particular, the authors refer to a recent paper by Chen et al. 2020 who suggest using d2H values obtained for cryogenically extracted xylem water should be corrected for an average bias of -8.1 per mil.

The paper starts off with an explanation of error propagation and lays down in mathematical terms how a systematic bias in d2H values of xylem water affects the application and uncertainty in two-pool mixing models. In essence, the authors conceptually demonstrate with this exercise that the described bias of -8.1 per mil may be of concern if endmembers in mixing models are not sufficiently resolved but that this bias is of less concern, if the end members are clearly distinct in their hydrogen isotope composition. The authors then demonstrate their case by re-evaluating several "classic" water sourcing papers and show how a -8.1 per mil bias challenges the conclusions on plant water sources in studies where mixing model endmembers are isotopically similar but has little effect in studies that use mixing model end members with larger isotopic differences.

I enjoyed reading this manuscript. It addresses a timely topic and puts the relevance of recent reports on measurement or extraction biases carefully into critical perspective. I think that such a contribution is clearly needed in a debate that starts to become overheated and influences scientific progress where it should not. At the same time the authors caution that there are indeed experimental settings, where uncertainties in the data don't allow robust conclusions regarding distinct tree water sources.

From my perspective this manuscript is an important contribution and it can pretty much be published as it is. There are a few thoughts that came to my mind but these are rather ideas or suggestions that the authors may or may not consider in their revisions:

We thank Prof. Kahmen for these comments. We are glad to hear that he sees the timeliness of the messages conveyed in our paper.

1) I found the choice of sample studies a bit redundant. Maybe, the number of studies discussed could be reduced or substituted with studies that are slightly different in scope. In particular, I would suggest to also discuss work that is less focused on quantifying the absolute water sources of plants but rather seeks to identify species specific differences in water sources. In plant ecology this has been and still is an important topic. In these cases it is the relative difference of water sources among plants that is of interest. Irrespective of the isotopic resolution of the end members in a potential model, an extraction bias would affect all species and thus introduce a systematic error but would not affect the identified differences among species (given a consistent bias for all species ...).

This is an important point that we will elaborate upon in a revised version. If biases shift all points equally, the relative values will be reasonably robust to any such artifacts (or artifact corrections). We will emphasize this point when discussing both Case D and E, because these studies involve inter-species comparisons. However, we will also (more thoroughly) make the point that the extraction artifacts may vary among individuals and species, expanding on where we already make this point (line 247) to more explicitly show the range of potential bias correction values (i.e., both the range discussed in Chen et al., as well as the range discussed in other studies) and their impacts on comparisons among species or individuals. Figure 1 will be updated to show ranges of offset values, informed by several other studies. By emphasizing the range of potential corrections in inter-species analyses, we will also address Prof. Kahmen's following comment, as well as comments made by the other two reviewers (to the effect that we should more thoroughly discuss variations in extraction artifacts).

2) The authors focus on the -8.1 per mil bias in d2H values of cryogenic extracted soil water that was recently suggested by Chen et al. The authors correctly mention that other biases in d2H values occurring during the extraction of soil water and/or water uptake by the roots exist and that biases have also been reported for d18O values. To put the evaluation of a -8.1 per mil bias in xylem water d2H into perspective, it might be good to report a few values from other studies indicating the order of magnitude in biases e.g. in soil water extractions of for d18O that other studies have reported and that would cause similar issues as discussed in the current manuscript.

This is a useful suggestion. Our revised manuscript will include more detailed discussion of the ranges of values reported elsewhere, which we have compiled at <https://doi.org/10.5281/zenodo.4832899>. Those values will be introduced in section 1 and then further iterated upon in section 4.

3) I like that the authors stress in the beginning of the manuscript that all sampling and analyses are associated with error. This brings up a point on analytical precision and accuracy that is increasingly forgotten in the discussion of cryogenic artefacts. In particular accuracy of measurements is rarely reported in plant water papers but accuracy can easily be off by several permil (for H) between labs and thus produce similar biases as cryogenic artefacts. This is in particular relevant when data from different labs (e.g using xylem water data from one lab that are referenced to precip data from another lab (e.g. GNIP) or instrument) are related. Interestingly, few authors and referees seem to care about this. I acknowledge that this is not really the scope of this manuscript but the authors may want to mention this with a sentence or two to put cryogenic artefacts into perspective.

This advice is appreciated and we agree that this is a valuable point to make. It could naturally follow the sentence that is currently on lines 76-77. Although parallel to the scope of this current manuscript, it will fall within the scope of the revised version, which will be broadened to address comments by the other two reviewers. We agree that this potential source of error – combining data from different labs – is often neglected.