

Response to Reviewer 3

The reviewer's comments are in normal font, our responses are in bold.

Recommendation to the Editor:

I recommend this manuscript be rejected for publication at this time. I recommend that it could be resubmitted after major revision.

The authors discuss the potential errors associated with 2H and $\delta^{18}\text{O}$ values obtained from water samples liberated from plant tissue via distillation and capture in a freeze trap. This has been documented in many studies by now, where the underlying reasons for those errors may include poor sample handling and storage, incomplete distillation, complete distillation within a leaky system, and/or incomplete condensation in the freeze trap. There are other error sources, including the confounding effect of some volatile organic compounds on the measured photon adsorption within laser spectrometers. Despite these varied mechanisms, the authors adopt a single value for the potential error (which they associate entirely, and somewhat ambiguously, with "cryogenic extraction"). The value is -8.1‰ , based on an average error reported in a previous paper.

We of course understand that there are many possible sources of errors in these isotope determinations, and it was not our purpose to catalogue them all. As the reviewer points out, these have been documented in many previous studies. Our purpose was instead to show how one specific bias, as reported by Chen et al., would potentially affect inferences that are made about plant water sources. We were careful to make clear that the value of -8.1‰ was chosen for purposes of illustration, and to make clear that the actual biases could be larger or smaller than this value in individual cases. We have made a compilation of studies that have measured cryogenic extraction biases in xylem water samples, using a variety of methods (<https://doi.org/10.5281/zenodo.4832899>). The average measured bias in the studies that we judged to be most reliable was $-6.1 \pm 0.6\text{‰}$ $\delta^2\text{H}$ (mean \pm 1 SE), similar in magnitude to Chen et al.'s results for deuterium.

Most studies based on xylem water deuterium values have not attempted to measure or account for this extraction bias. The point of our paper is to show the possible impact of this bias, and thus to illustrate the importance of better accounting for it (including quantifying how it may depend on plant characteristics and laboratory procedures).

The authors present a model of how this assumed error propagates within a two-component-mixing calculation in equations 1-6. With regard to studies that consider water in soil pores as a possible end member, however, they have completely ignored the fact that some error could exist with the delta values representing the soil-water end member (which is also commonly obtained by distillation and capture by freeze trap). As such, their equations 1-6 seem irrelevant to the real problem at hand.

The purpose of our analysis was not to catalogue all of the potential biases in isotopic determinations of plant source waters, or even all of the potential biases due to cryogenic extraction. Our purpose was to look at how one specific bias – the directional 2H effect introduced by cryogenic extraction of xylem water as recently quantified by Chen et al. and others – could potentially affect conclusions drawn from plant water isotope data. This is distinct from the cryogenic extraction biases in soil water samples – which is a substantial issue, but not the issue that we are analyzing here.

We can revise the introduction to explicitly mention these many other sources of error and bias, and to explicitly mention that we are looking at the potential consequences of one specific bias, not the full chain of potential errors and biases that could be involved in soil/plant/water studies.

If the editor desires, we could of course also extend equations 3-6 to handle biases in the measurements of the potential end-members as well as the xylem (mixture). What we cannot do, however, is to apply such an equation for typical soil water extraction biases, because the literature does not exhibit a consensus about the direction, let alone the magnitude, of extraction biases in the case of soil water (see further comments below). Thus, extending equations 3-6 to handle biases in the soil water end-members seems to us to be premature and unnecessarily complicated at this stage.

They retrospectively add 8.1 ‰ to some data values obtained from previous publications and discuss how that addition could alter the interpretation of results from those studies. At present, this is too cursory for publication in *HESS*. Throughout the manuscript they have ignored that the error associated with the method would be variable, and perhaps absent, as practiced by different lab workers. They have ignored the reality that the error could be variable among plant species.

The reviewer's assertions here are not correct. We specifically mention the possibility that different labs may have different biases (lines 90-91 and 245-250, 257-259), and we discuss at some length the implications of possible variations in extraction biases among plant species (lines 213-215, 232-250, and 257-259). Taken together, these lines comprise about 10% of the entire text.

The reviewer's claim that xylem water cryogenic extraction errors would be "perhaps absent, as practiced by different lab workers" is not, to our knowledge, supported by any studies that combine both steady-state transpired vapor and rewetting experiments. Among five studies within the past five years that quantified xylem water extraction offsets (Poca et al., 2019, Newberry et al., 2017, Zhao et al., 2016, Chen et al., 2020, Barbeta et al., 2020), only one did not find that extracted xylem water was significantly lower than its source in $\delta^2\text{H}$ (Newberry et al., 2017).

In some instances they make inferences about the relationship between $\delta^2\text{H}$ and $\delta^{18}\text{O}$ for specific water samples, but they only consider the effect of the bias on $\delta^2\text{H}$, not $\delta^{18}\text{O}$.

We do this because the available evidence indicates that the bias in $\delta^{18}\text{O}$ due to xylem water extraction is small – indeed, small enough that it cannot be reliably quantified. Of those four previously mentioned studies that found significant differences between xylem water and source water $\delta^2\text{H}$, only Poca et al. found significant differences in $\delta^{18}\text{O}$ and that was only in arbuscular mycorrhizal treatment samples; $\delta^{18}\text{O}$ differences were not significant in the other samples. In addition to those four, Ellsworth and Williams (2007) evaluated $\delta^{18}\text{O}$ offsets in one species and also concluded that $\delta^{18}\text{O}$ offsets were not statistically significant, unlike their $\delta^2\text{H}$ values. These values can be seen here: <https://doi.org/10.5281/zenodo.4832899>. Given that $\delta^{18}\text{O}$ values are not consistently lower in xylem water than they are in the source water, it may be more appropriate to treat $\delta^{18}\text{O}$ extraction artifacts as uncertainties rather than as biases. This will now be discussed the revised manuscript (where we will also make clear that to date, there is no consistent evidence of extraction biases in xylem $\delta^{18}\text{O}$).

As stated before, they have also ignored the fact that the bias would also affect data from water samples distilled from soils, and that effect would be variable among soil types and studies. This article will inevitably be interpreted, and cited, as an indictment of those pre-existing works that are used as case studies here. I don't think that is a fair impression to make based on this simplified analysis.

Again, the reviewer's statement is simply incorrect. We have not "ignored" biases from cryogenic extraction of soil waters, but instead we have specifically mentioned them and excluded them, as

explained on lines 87-89: "We do not address the potential consequences of cryogenic extraction biases in soilwater samples, because the current literature lacks a sufficiently clear consensus on the magnitude, or even the direction, of such potential biases in soilwater extractions."

Moreover, Thielemann et al., 2019 have shown that the extraction biases measured in soil rewetting experiments can be largely explained by the strong fractionation of the residual water that is left behind by the oven-drying procedure. Once that artifact is taken into account, it is less clear how much soil water extraction bias remains to be explained. By contrast, Chen et al. found consistent xylem water deuterium extraction biases using two independent methods, and similar biases have been found in xylem water deuterium by several other independent studies (<https://doi.org/10.5281/zenodo.4832899>).

In a revised manuscript, 1) we will more explicitly point out that xylem ^{18}O biases have been found to be small, 2) we will briefly summarize the literature on soil extraction biases, to more clearly justify our exclusion of those biases from our analysis, and 3) we will better emphasize the range in xylem deuterium biases measured in other plants and labs.

However, the reviewer's statements about what we have 'ignored' are factually incorrect, as shown by even a cursory reading of lines 85, 88, 91, 92, 183-185, 187, 213-215, 243, 246-250, 258, and Figure 2A, which are just a subset of examples in our paper where we address each of the points that the reviewer incorrectly claims that we ignored.

I think a productive path forward would be for the authors to (1) reformulate equations 1-6 to properly represent all potential biases in the mixing model calculations, (2) consider the bias for both isotope ratios in question, and (3) present a quantitative analysis of how this revised error model influences their previous results. I think their commentary about the other previous studies is interesting and could be utilized within the introduction or discussion of that revised paper.

But it must be emphasized that the retrospective analysis is speculative. The real error from those past studies is unfortunately unknown. I don't think it is useful to just assume it was 8.1 ‰.

Here again the reviewer misrepresents the paper that we actually wrote. We did not "just assume it was 8.1 ‰." Lines 89-93: "Although further research is also needed to determine how cryogenic extraction biases in xylem water vary among species and among laboratories, we use -8.1 ‰ *for purposes of illustration* (emphasis added). We emphasize that the actual bias for any given species and any given laboratory's extraction procedures could differ substantially, in either direction, from this illustrative number." Lines 256-259: "Further work to quantify potential extraction artifacts will hopefully yield correction factors for a range of species and extraction protocols. *These are urgently needed because evidence does not support using 8.1 ‰ $\delta^2\text{H}$ as a universal correction factor, although it was adopted here for illustrative purposes* (emphasis added)." And almost everywhere that we mention the 8.1 per mil value, we reiterate that it is a "possible" or "potential" or "presumed" extraction bias.

Nonetheless, we will follow this guidance when revising the manuscript because it is clear that our message was not conveyed clearly enough to this reviewer. We will add a range of correction factors to the studies shown in Figure 1. This range will be informed by the range in values in Chen et al., as well as the range in values described in past studies (<https://doi.org/10.5281/zenodo.4832899>). While we cannot know the true error – a point we repeatedly made throughout the paper – we can use later studies to estimate a range of possible errors. The reviewer may disagree with this approach, but we continue to believe that this is the best course of action at present.

Specific Comments:

1. Line 25: “waters’ ” written as plural possessive pronoun is weird and unprecedented. Consider rephrasing. Delete “isotopic”.

We can rephrase this.

2. Lines 75-76: An important point. I suggest making it a stand-alone sentence rather than the weaker parenthetical statement.

Thank you for the suggestion, we will rewrite this.

3. Lines 88-89: There seems to be no mechanistic explanation for the error source. It is casually attributed to cryogenic extraction. If that is the case, then water extracted from soil pores using the same method will also be significantly biased, as has been shown in several of the studies the authors have already cited [e.g., *Natalie Orłowski et al.*, 2016a]. Yet, the authors are choosing to ignore that.

The work of Orłowski et al. (2018) and related studies (e.g., Oerter et al., 2014), considered together, show uncertainties but they do not show consistent directional biases. As discussed in our paper, those need to be treated differently. We can further expand on this in a revision.

Our work concerns the effects of xylem water cryogenic extraction biases, not soil water extraction biases (because, as indicated above – and as we explicitly said in the paper – there is not clear evidence of consistent directional biases in soil water extractions). We will revise the introduction to further clarify that we are focusing just on the effects of xylem water extraction biases (and uncertainties in those biases).

Despite the conditional statements provided here by the authors, the results of this retrospective analysis will inevitably be interpreted and cited as indictments of some of the studies chosen. Leaving that impression is quite unfair to the authors of those previous studies considering that the whole error model just presented in the preceding equations is known, at the outset, to be fundamentally inadequate.

We do not agree that one must know everything about all possible sources of error or bias before one can say anything about the implications of individual errors or biases. We don't think it is "unfair" to assess the implications of probable methodological biases, either in prior studies or current ones. All scientific results are necessarily tentative and subject to reconsideration when better data become available (concerning either the natural systems themselves, or the methods used to study them). We do not see how our analysis is “unfair”, particularly given that no assumptions or data processing steps are hidden, and everything is straightforward enough to be quickly reproduced by readers.

There is of course the other problem that analysis of water volumes extracted from plant tissue via laser spectroscopy may be further confounded due to presence of volatile ethanol and methanol gases [*Leen et al.*, 2012]. This would be less of an issue, or not at all an issue, for many water samples extracted from soils if their dissolved organic carbon content was markedly lower than found in plant tissues.

Only one of the five studies we considered was based on laser spectroscopy methods. The others were based on conventional mass spectrometry.

4. Lines 91-92: These sentences only emphasize that the analysis is quite indefinite. I think this would perhaps be better as a technical note, with the scope limited to the retrospective analysis of the authors’

own work. Further, the use of the single value 8.1 % as the correction factor is proclaimed by the authors to be even a simplification of the real range of errors reported by Chen.

Our objective was to put forward a framework for dealing with these biases, and to illustrate its potential application. It would misrepresent the state of the science to treat the extraction bias as being more definite than it is. Our manuscript focuses on signal to noise ratios and discusses how these biases and uncertainties will have little effect if the end members are sufficiently distinct, but can make results inconclusive when the end members are not sufficiently distinct. Discussions with colleagues have convinced us that this is a useful framework for thinking about these issues, even if the exact magnitude of the extraction bias remains poorly quantified at this time.

5. Lines 107-146: The error in results from two-component-mixing calculations attributed to putative bias from cryogenic extraction might even be viewed as marginal here in comparison to the uncertainty that results from the spatial variability (vertical and horizontal) of delta values in soils. If the uncertainty around the average delta value at any depth is integrated into the two-end-member mixing model [e.g., as demonstrated by *Genereux*, 1998], then the results are made even more ambiguous than they are due to this putative effect of cryogenic distillation.

We agree that there may be other uncertainties and that these may be large. But uncertainties and biases are two different things; uncertainties can potentially be reduced by more or better sampling, but biases cannot. Spatial variability is not a bias. Furthermore, we do not think it can be said that “the results are made even more ambiguous than they are due to the putative effect of cryogenic distillation” as a universal rule (although surely the additive effects of the two on the overall uncertainty is bigger than the effect of either in isolation).

Honestly, we just need to stop talking about plants preferentially using water from 20-cm depth, or 70-cm depth, or any other. Woody stemmed plants grow root systems that extend meters below ground [*Schenk and Jackson*, 2002]. They invest perhaps 30% or more of the carbon they assimilate through photosynthesis into the revitalization and maintenance of the fine root systems. Yet, are we supposed to think that it became evolutionarily-advantageous to these plants to make use of only a tiny fraction of that root system that exists at one depth or another? It is a silly notion that is only further propagated when hydrologists continue to write sentences like these implying that it is even plausible. Unless the vertical variation in delta values of soil water is characterized in detail, and unless the delta values from xylem samples fall at one or the other extreme ends of that range, then there are myriad possible mixtures of water than can yield any particular delta value measured in water from xylem cells.

The reviewer's concern here is tangential to our paper (and we would not advocate for such a literal interpretation of mixing model analyses). If the reviewer wants to complain about "silly notions" underlying mixing models in general, we respectfully submit that this is not the place to raise this issue.

We do agree that no one should using mixing model analyses to conclude that plants, for example, exclusively extract water from any two specific depth slices of a soil profile. While there are many ambiguous cases, mixing models can be useful in cases where shallow and deep soils are isotopically distinct and xylem water is more similar to one than the other. It provides a tool to quantify where on that scale a xylem water sample falls, and to quantify our confidence in that value (when accounting for uncertainties and errors; eq 1-6); indeed, when values are intermediate, they are not easily interpreted. Regardless, the point of our paper is (among other things) to improve the accounting of uncertainties and errors in mixing models, because they do still find widespread use in the field of isotope ecohydrology.

6. Line 168: Or you could/should cite [Dansgaard, 1964] since they demonstrate the same facts almost 60 years ago, or [Rozanski *et al.*, 1993].

That is correct. We will do this.

7. Lines 182 – 183: Perhaps more reliable, but I think this is pure speculation. Exactly the opposite is true with regard to soil samples. Cryogenic distillation tends to more strongly influence $\delta^{18}\text{O}$ values than $\delta^2\text{H}$, possibly due to oxygen exchange with carbonate minerals or with hydroxyl groups in clay minerals [Meibner *et al.*, 2014]. And the effect is variable depending on clay content [N. Orłowski *et al.*, 2016b]. But these effects are entirely ignored here, which I think leaves equations 1-6 being essentially irrelevant to the real problem at hand, and the retrospective assertions made about this set of case studies as just very, very crude approximations.

The reviewer again misrepresents our comment, which is clearly not a comment about soils. It is not speculative because this sentence, “cryogenic extraction biases appear to be much smaller for $\delta^{18}\text{O}$ than for $\delta^2\text{H}$ in xylem water”, is about xylem water. We cite evidence (Chen *et al.*, 2020), and we can cite further evidence of this strong effect in ^2H but not ^{18}O from Zhao *et al.* (2016), Ellsworth and Williams (2007), and Barbeta *et al.*, (2020). This will be done in the revised version.

Also, importantly, we remind the reviewer that not all studies of plant-water sources use soil-water values as end members (see Case E and a subset of the analysis in Case A). So, the comment on “equations 1-6 being essentially irrelevant to the real problem at hand” comes from a perspective that does not capture the full scope of the problem addressed in our paper. The reviewer again focuses on one problem (soil isotopic uncertainties) as the “real problem at hand” but it is not the primary focus of this manuscript (nor does the literature present any clear consensus that it is indeed the “real problem”).

8. Lines 193-194: Why do you think that the data points would be limited to plotting along lines with these slopes? For water in soil undergoing evaporation in the absence of precipitation, sequentially sampling the soil water day after day and measuring $\delta^2\text{H}$ and $\delta^{18}\text{O}$ would yield a set of data points that trend somewhat to the right, and somewhat above, the data point representing the original liquid water. As a Rayleigh distillation process with a humidity-dependent exponent (the humidity-dependent fractionation factor) the trajectory of those data points would depend on the humidity of the soil pore space just above the evaporating surface (i.e., in the theoretical semi-turbulent layer above the boundary layer), and the temporal duration of the evaporation process (as quantified by the f coefficient in the Rayleigh equation, representing the remaining fraction of the original water mass). Considering a line fit through those hypothetical data points, a broader range of slopes is plausible than what is stated here.

We are simply repeating the methodology that is already used in our prior study. Any other approach would be confusing and lead readers away from understanding the point about biases in dual-isotope space that we are making. The reviewer should note that variations in assumed slopes were used in the original paper; however, it makes more sense to isolate the effects of the extraction bias rather than let those effects be obscured by introducing numerous other uncertainties that are not the focus of this article.

9. Lines 195-196: I have seen this done, for example, in the paper by Evaristo *et al.* cited by the authors. I have never seen any theoretical justification for this line fitting exercise. It ignores the reality that a water sample collected from soil on any given day contains collections of molecules that have drastically different residence times in the soil, have undergone different rates of evaporation, and have experienced

different degrees of mixing (due to diffusion and dispersion) with other collections of water molecules that were sourced from different precipitation events. I have to believe that in some cases, if the line fitting is done via least-squares regression, that the intercept point on the Global Meteoric Water Line could be completely unrealistic, like at corresponding values of $\delta^2\text{H}$ or $\delta^{18}\text{O}$ that are more negative than those ever observed in precipitation. For these reasons I struggle to understand the relevance of this commentary.

The point-fitting procedure described here (e.g., as used by Evaristo et al.) is not something we use in our analysis. It is unclear why the issues described by the reviewer (which are also described in Benettin et al. 2018 and Bowen et al. 2018) detract from the relevance of our commentary.

10. Lines 208-213: What are we really supposed to infer from this, and do you really want to cast doubt on your own previous work based on this cursory analysis? You have made this correction for a presumed bias in the $\delta^2\text{H}$ data, but have ignored the fact that an incomplete distillation could also bias the $\delta^{18}\text{O}$, which influences the plotting position of these data points. And again, you have also ignored the potential bias that an incomplete distillation introduces into the quantification of the soil-end-member tracer concentration. The inference here has to be confounded by these omissions, no?

These are not soil-water end-member values, they are precipitation, as we clearly stated (lines 165-166), so "quantification of the soil-end-member tracer concentration" has fundamentally nothing to do with either our original xylem water analysis, or our reanalysis presented here. It should also be noted that end-member standard errors are quantified and shown (Figure 3; the green and cyan lines represent the sources +/- 2 standard errors); this will be more clearly stated in a revised version. So, no, our inference is not likely to have been confounded by what the reviewer incorrectly thought was an omission.

We do not see this as casting doubt on our previous work. We are happy to have the opportunity to revise our previous findings based on new findings by others. This is scientific progress, and so we would not see it as any indictment of our own work.

11. Line 230: I don't think you should ignore this work [Genereux, 1998], which is an important contribution and precedes all the other works you have cited.

Thank you for this suggestion. We will be sure to give appropriate credit to Genereux in a revised manuscript. We will also cite Gauss 1823, which predates Genereux by nearly two centuries.

Cited References:

Dansgaard, W. (1964), Stable isotopes in precipitation, *Tellus*, 16(4), 436-468, doi:10.3402/tellusa.v16i4.8993.

Genereux, D. (1998), Quantifying uncertainty in tracer-based hydrograph separations, *Water Resources Research*, 34(4), 915-919.

Leen, J. B., E. S. F. Berman, L. Liebson, and M. Gupta (2012), Spectral contaminant identifier for off-axis integrated cavity output spectroscopy measurements of liquid water isotopes, *Rev. Sci. Instrum.*, 83(4), doi:044305
10.1063/1.4704843.

Meibner, M., M. Kohler, L. Schwendenmann, D. Holscher, and J. Dyckmans (2014), Soil water uptake by trees using water stable isotopes--a method test regarding soil moisture, texture and carbonate, *Plant Soil*, 376, 327-335.

Orlowski, N., D. L. Pratt, and J. J. McDonnell (2016a), Intercomparison of soil pore water extraction methods for stable isotope analysis, *Hydrol. Process.*, n/a-n/a, doi:10.1002/hyp.10870.

Orlowski, N., D. L. Pratt, and J. J. McDonnell (2016b), Intercomparison of soil pore water extraction methods for stable isotope analysis, *Hydrol. Process.*, 30(19), 3434-3449, doi:10.1002/hyp.10870.
Rozanski, K., L. Araguas-Araguas, and R. Gonfiantini (1993), Isotopic patterns in modern global precipitation, American Geophysical Union Monographs, 78, 1-36.

Schenk, H. J., and R. B. Jackson (2002), The global biogeography of roots, *Ecol. Monogr.*, 72(3), 311-328, doi:10.2307/3100092.

Barbeta A, Gimeno TE, Clavé L, Fréjaville B, Jones SP, Delvigne C, Wingate L, Ogée J. 2020. An explanation for the isotopic offset between soil and stem water in a temperate tree species. *New Phytologist* 227: 766–779.

Ellsworth PZ, Williams DG. 2007. Hydrogen isotope fractionation during water uptake by woody xerophytes. *Plant and Soil* 291: 93–107.

Gauss, C. F. (1823), Theory of the combination of observations least subject to error: part one, part two, supplement (Theoria combinationis observationum erroribus minimus obnoxiae: pars prior, pars posterior supplementum), English translation by G.W. Stewart (ed. Golub, G. H.), Society for Industrial and Applied Mathematics, Philadelphia, 1995.

Newberry SL, Nelson DB, Kahmen A. 2017. Cryogenic vacuum artifacts do not affect plant water-uptake studies using stable isotope analysis. *Ecohydrology* 10: e1892.

Orlowski, N., Breuer, L., Angeli, N., Boeckx, P., Brumbt, C., Cook, C. S., Dubbert, M., Dyckmans, J., Gallagher, B., Gralher, B., Herbstritt, B., Hervé-Fernández, P., Hissler, C., Koeniger, P., Legout, A., Macdonald, C. J., Oyarzún, C., Redelstein, R., Seidler, C., Siegwolf, R., Stumpp, C., Thomsen, S., Weiler, M., Werner, C., and McDonnell, J. J.: Inter-laboratory comparison of cryogenic water extraction systems for stable isotope analysis of soil water, *Hydrol. Earth Syst. Sci.*, 22, 3619–3637, <https://doi.org/10.5194/hess-22-3619-2018>, 2018.

Poca M, Coomans O, Urcelay C, Zeballos SR, Bodé S, Boeckx P. 2019. Isotope fractionation during root water uptake by *Acacia caven* is enhanced by arbuscular mycorrhizas. *Plant and Soil* 441: 485–497.

Thielemann, L, Gerjets, R, Dyckmans, J. Effects of soil-bound water exchange on the recovery of spike water by cryogenic water extraction. *Rapid Commun Mass Spectrom.* 2019; 33: 405–410. <https://doi.org/10.1002/rcm.8348>

Zhao L, Wang L, Cernusak LA, Liu X, Xiao H, Zhou M, Zhang S. 2016. Significant Difference in Hydrogen Isotope Composition Between Xylem and Tissue Water in *Populus Euphratica*. *Plant, Cell & Environment* 39: 1848–1857.