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 δ^2H and δ^{18}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.

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

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 ‰.

Specific Comments:

1. Line 25: "waters' " written as plural possessive pronoun is weird and unprecedented. Consider rephrasing. Delete "isotopic".

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

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 Orlowski et al.*, 2016a]. Yet, the authors are choosing to ignore that. 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. 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.

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.

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. 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.

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].

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 δ^{18}O values than δ^{2}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. Orlowski 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.

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 δ^{2}H and δ^{18}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.

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 δ^2H or δ^{18}O that are more negative than those ever observed in precipitation. For these reasons I struggle to understand the relevance of this 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 δ^2H data, but have ignored the fact that an incomplete distillation could also bias the δ^{18}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?

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.

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.