Comment on hess-2020-680
Anonymous Referee #2

In the manuscript entitled, “Insights into isotopic mismatch between soil water and Salix matsudana Koidz xylem water from root water stable isotope measurements”, results from a field study show isotopic differences between different soil water and plant water pools. The dataset is interesting and results from intensive measurements. I think some of the findings are potentially interesting and warrant publication. However, the current presentation has some weaknesses and ambiguities that need to be addressed. Ultimately, the imprecise use of jargon obscures the interpretation and implications of the study.

Throughout, references are made to ecohydrologic separation and the two-water-worlds (TWW) hypothesis. However, I am not sure how the authors define these phenomena; their definitions seem different from my own, different from the literature that they are citing, and they may change throughout. At times, it seems that these terms just mean “there are isotopic differences”, which is not particularly novel to identify. The paper would greatly benefit from less use of jargon. It should be more explicitly stated what is being tested. From my reading of this, the key questions of this paper are “Does root water isotopic composition match that of soils’ bulk, mobile, and bound water fractions of soils at the same depths”. Then there is a second question “How does choice of stem sample (from four heights) or choice of soil versus roots influence inferences of water uptake depths”. Neither of these questions is especially related to TWW or ecohydrologic separation. Frankly, the duration of the study is too short to assess either TWW or ecohydrologic separation because any observed differences between plant water and groundwater isotope ratios might be a product of lags, rather than the use of fundamentally different sources. The process Brooks et al referred to as ecohydrologic separation was only observable through using measurements across multiple seasons.

The isotopic differences between root water and soil water is key to the conclusions made in this paper. The authors seem to suggest that the water in roots should match the water in soils around them. They did not, and this was interpreted as potential fractionation. However, roots can transport water from different locations. I would only expect similarity between roots and surrounding soils if fine roots were sampled. For coarser roots, as used in this study, I would expect those roots to transport water from much deeper depths and integrate large volumes of soil water. Thus, it is not clear that “a combination of plant fractionation and TWW-type separation” (which, again, needs to be clarified) is needed to explain the observations here. This needs to be further discussed. Potentially, additional excavations may be warranted to identify whether the size of root samples could include fine roots that extend substantially deeper.
14, 18, 77: I think the author is referring to ecohydrologic separation, not the "two water worlds hypothesis", which is about streamflow. However, I am generally unclear on this.

20: From the abstract, it is not clear why this is "in conclusion".

21-23: It is better to say what that contribution is, and what those insights are, rather than just mentioning that they exist.

41-44 If movement alone creates the change, I’d argue that this statement violates laws of mass conservation. Any changes must be matched by equal and opposite changes elsewhere. Ecohydrological separation is not a change from soil to root to xylem.

46-48 See comments above. Please re-read Brooks et al.

48-49 How would the hypothesis be supported by groundwater and streams? This does not make sense to me. The hypothesis is related to plant-available soil water.

51: What does “related to infiltration” mean? Is “tightly bound” water not related to infiltration?

120-124: This is a short period for assessing ecohydrologic separation, especially in a dry-climate region. Identifying ecohydrologic separation phenomena requires detecting bypassing of stored waters.

180-182 Given the interest in fractionation upon uptake, why not set it to a value more consistent with others in the literature.

193-205 I do not understand why slopes in dual-isotope space are being used here. I don’t think they are the most effective way to make the comparisons that are being made. First, were these slopes fit orthogonally? They should be. Second, when one line is compared to another, is that the result of an ANCOVA test? Third, are the p values the fits of the lines, or p values for the comparisons being described? I do not understand why fitted lines, rather than the actual values, are being compared; it seems that the interpretations relate to the actual values.
What are these uncertainties: standard deviations? All depths are lumped together for soil water?

Ecohydrologic separation was a process that could only be revealed from a long duration of sampling. How can this short period of measurement show ecohydrologic separation?

What is the statistical test used?

Does “always” mean every single value was different? I do not understand how one would test that.

Section 3.3 Are these fine roots only? If not, I see no reason to think that roots should match soils at the same depths. A larger root is almost certainly going to integrate waters from zones larger than the soil surrounding it.

I do not understand this “horizontal homogeneity” discussion. Figure 3 shows that mobile water and bulk water at a given depth differ. That implies heterogeneity.

So if “separation” here is not the same as the “separation” that has been described throughout, it would be useful for the authors to use a more literal descriptor of the process that they are investigating.

At this point, I’ve become overly confused with regard to what the authors consider “TWW” to be.

What does “mask” mean? Even if there is evaporation, fractionation upon uptake would result in different values between the roots and soils, regardless of the background (soil water) signal.

The authors should probably specify that this is the duration of the detectable label and not the mean or median of the residence time distribution. Is that correct?

what is “this interpretation” referring to?

It is best to not use “enriched” or “unenriched” unless specifying what they are
enriched in (e.g., deuterium) and what they are enriched relative to (e.g., precipitation).

How is it known that root water is an accurate approach? This is important. It seems like root water tells you a different thing than soil water: the depth of roots contributing to transpiration, rather than the depth of soils contributing to transpiration.

Please be more specific.