This study by Smith et al. presents a novel combination of in-situ, high-frequency measurements of micrometeorological variables, water fluxes, stores and stable isotopes in soil and xylem together with a process-based modelling approach, in order to identify the dynamics of water partitioning under 2 willow trees and a neighbouring grass patch over a growing season.

The increased perspective on soil-plant water dynamics brought by this intensive monitoring, further presented in another manuscript (Landgraf et al., 2021) is used as a multi-data calibration and evaluation of the ecohydrological outputs provided by the Ech2O-iso model. The authors use this baseline to then evaluate a new conceptualization of water uptake and transport along a vertically-and-laterally-distributed root profile, in order to understand the relation between soil and xylem water dynamics and signatures.

The topic addressed by this study is highly relevant to the ecohydrological research community, with an impressive experimental setup combined with a state-of-the-art modelling approach. The performance of the model with respect to diverse ecohydrological observations at contrasting plots (willow and grass) provides a multi-faceted evaluation, strengthening a baseline for further hypothesis testing. However, I am quite concerned with the way the modelling development at the core of the second part of this study, i.e. the “distance-based mixing” component along the root-xylem system, is presented, applied, and its overall performance. By contrast with the stated rationale, in my view the methodology does not consistently address the question of better identifying the spatial heterogeneities among water pools sustaining plant development. For these reasons, among other ones further detailed below (including modelling setup), I think a thorough revision of the study design and presentation is needed before this manuscript can be considered for publication in Biogeosciences.

General comments

- The rationale for the distance-based mixing development is to mimic the capacity of the root system to tap water pools at various depths and that may be laterally distant, in a spatially- and time-explicit manner. The authors take good care in considering the time domain, and describe in Sect. 3.3.2 how the non-zero length of the root system
translates into root-scale transit times distributions. In the spatial domain, it seems that the modelling approach links xylem water to same-pixel (6x6 m$^2$) soil water, both in terms of root uptake and signature (isotopic content or ages). My understanding is that transpiration in ECH2O-iso uses same-pixel water content, and the distance-based mixing application makes no clear mention of which simulation pixel is considered. Section 3.2.1 mentions that the proportion of “potential root-uptake from outside model cells containing vegetation”, I find it confusing that no explicit mention of how this is actually taken account is further made, and Fig. 6 suggests that same-pixel signatures (soil and xylem) are compared. However, it is clearly stated in the Discussion that “small-scale [vertical] variations, as well as the large spatial differences from the soils below the willows and below the grass, and between different soil layers (Fig 3) reveal the significant heterogeneity of the site despite relatively immature soils and the local spatial scales” (L505-507) and then, crucially, that “around half of the [water] uptake (by root length and water availability) estimated to occur outside of the willow [pixels]”(L518-520). It is then likely that a significant part of the isotopic signal found in the xylem of Willow 2 originates from water pools in neighboring, dynamically-distinct vegetation patches, in particular the grass patch. It makes it difficult to then assess the added value of this “distance-based mixing” model, which seems to essentially add a lag-based component to water mixing in along the root-stem continuum, while fine-scale spatial patterns may play a crucial role. This inference is only based on the main text though, as the source code for root-stem mixing does not seem to be part of the main EcH2O-iso repository referenced in this manuscript (if that is correct, it would seem appropriate that the authors publish the full source code used in this study). As such, this approach resembles a conceptualization adopted in an earlier study published by some of the authors, cited in this manuscript, where a tree storage component was shown to improve modelled xylem isotopic signature at a coarser spatial scale where lateral contributions may cancel out (Knighton et al., 2020).

Given the above, I encourage the authors to clarify throughout the manuscript what water pools (in particular, “laterally” speaking) are considered when quantifying root water uptake and associated isotopic signatures and transit times. If these are indeed limited to the local (same-pixel) scale, then the scope of this paper becomes more limited, and I suggest to discuss much more thoroughly the limitations of this study, beyond merely stating “the potential influence of spatio-temporal variability of source waters on xylem isotopes” (L.520-521), including a potential rejection of the adopted root-xylem model conceptualization.

Non-exclusively, a stronger case for the development of the distance-based mixing approach could be made using a case where the contribution of soil pools within root radial extent are considered in calculating xylem water ages and isotopic signatures (e.g. extrapolating from grass-patches values, since Landgraf et al. (2021) suggest that Willow 2 is surrounded by Willow1 and grass patches otherwise?). Ideally the water fluxes should also be factored in when calculating transpiration ; if it requires a heavier development of the ECH2O-iso code, the associated limitation should again be thoroughly discussed, as a bare minimum.

The general concern described above also arose because it does not seem that the “distance-based” model significantly outperforms the default “instant mixing” approach (Figure 6 and Table 4), contrary to what is stated in the Discussion (L515-516). In evaluating the two mixing approaches, the authors took a very welcome step in comparing, in both approaches, the cases where “transit time and xylem isotopes were calibrated 1) using modelled soil isotopic compositions and sap flow, and 2) using measured soil isotopes and sap flow” as “The use of measured soil isotopes and sap flow tests the maximum potential for how each model performs and is not limited to the performance of EcH2O-iso for sap flow or soil isotope” (L273-276). In the end, I can only agree with the authors that “seasonal magnitudes of xylem isotope dynamics
were predominantly due to differences in simulated v. measured soil isotopes in the shallow soils [rather than differences between mixing approaches]” (L527-528), and it also seems that AIC and KGE values, in the case of using measured soil isotopes and sap flow, are rather close between “instantaneously mixing” and “distance-mixing” cases, with even KGE values slightly higher in the former case (Table 4). On a side note, it seems somewhat surprising that these higher KGE values translate into slightly worse AIC values given that the “distance-mixing” requires 4 additional parameters as compared to the “instantaneously mixing”.

Specific comments

- **L31**: The 80-90% T/ET estimate by Jasechko et al. (2013) is often thought to be overestimated; maybe the “updated” estimate Schlesinger & Jasechko (2014) would be more appropriate for citation.

- **L36**: Please considering citing the original, peer-reviewed publication by Zink et al. (2017)

- **L36-37**: I am not sure what is meant by “beyond vegetation uptake during the growing season”, please rephrase.

- **L43**: Rather than “small or larger scales”, please consider providing indicative scale (e.g. plot to stand)

- **L65**: Appropriate citations of ecohydrological modelling advances may also include Maneta et al. (2013) and Fatichi et al. (2012).

- **L81-82**: The stated achievements are rather general; additionally it would preferable to have this section turned this into research questions and/or testable hypotheses (it is not clear to me what these are), to further detail the general goal described L79-80. In this process, rather than “exploring” achievements/question #2 should better state the adopted strategy regarding root-mixing development and its evaluation/rejection (see General Comments)

- **Fig. 1**: In connexion with the General Comments regarding the rooting system, it would be welcome to have a visual description of the land patches neighbouring the study plots (e.g. in Fig 1b or c, as in Fig. 1c in Landgraf et al., 2021), since the main text (L90-91) only describes what is at least 20m away from the plots.

- **L117**: Did the author mean “Köppen Index Cfb”?

- **L147-155**: I could not find a description of how in-situ LAI measurements are carried out, although such data is presented in Fig. 5, could the authors clarify?
L177-183: It seems from the text that the version of code used in this study uses the SPAC module developed by Simeone et al. (2019), if so the authors should acknowledge and cite this work.

L206-207: Is it a full mixing in the whole soil domain? Or some compartments are differentiated?

L217: The 100 “best” simulations have not been defined yet, please refer to Sect. 3.4.2.

L240: I do not understand the synchrony between the proposed descriptino of rooting length and SPAC, as the latter module is mostly focused on tree mortality (roots included).

Eq. (1): I am not sure how this equation was derived from Sperry et al. (2016). I am guessing it combines the cumulative root proportion provided in Eq. (6) in the above reference, the use of center-of-biomass depth, and layer depths in EcH2O-iso, but the intermediate steps to Eq. (1) escape me. In addition, I am confused so as if the beta factor here is the same beta found in Sperry et al. (2016) and its relation to the exponential factor kroot, also because the value of 0.995 is also found (for beta) in Sperry et al. (2016) Also, in calculating the vertical length, shouldn’t one add the height-above-ground at which xylem measurement are made (here, 1 meter)?

L246-253: This approach differs from Sperry et al. (2016), where the volume of roots is calculated in the first layer, using radial length in the first layer, and then radial in others layers is estimated by assuming that each layer has the same volume of root. It is likely not the case here because layer depth is fixed but kroot seems to be calibrated and differs between simulations. So I am guessing the authors used total root volume, implying that Eq. (2) uses total rooting depth (rather than $d_1$ as currently written) and then use Eq. (3) as a custom-made formula to reach the radial lengths in each layer?

L249: According Sperry et al. (2016), D should be the maximum rooting depth, not the total soil depth.

L252-263: While the principles of root-length-based transit times is nicely described, it is quite frustrating not to see the calculated values for the rooting length (radial, vertical, total) in the results section or elsewhere in the manuscript. This could be a supplementary figure or table, at a minimum.

L264-265: At first glance, this no-cavitation hypothesis seems inconsistent with the integration of the SPAC module, whose purpose is precisely to describe occurrence of cavitation using plant hydraulics. Did the authors found evidence that no cavitation occurred during the simulated growing season?

L288: Do the authors mean that the bottom depth of each layer in the model is fixed to correspond to 10, 40, 100cm, with effective layer “thickness” of 10, 30, and 70 cm,
respectively? This information is provided in Table 3’s caption, but it would be handy to have it earlier in the manuscript.

- **L293-295:** How is the grouping done for vegetation parameters? This is quite unclear, all the more that the type of information on calibrated parameters in Table S1 is not provided for vegetation parameters, could the authors provide a similar table? In addition, the SPAC module requires further parameterization that was carefully constrained in Simeone et al. (2019), but no mention is made on this topic, nor associated parameters, in the manuscript. Overall, it seriously limits the reader’s understanding of the modelling setup used in this study.

- **L305:** Have the authors looked at the additional information brought by lc-excess? This could further help analyzing contribution from shallow/deep soil horizons, and further fractionation effects (or lack thereof) during root-stem transport.

- **L307:** By “split calibration”, do the authors mean using a calibration period and a validation period? Or a calibration period for one step, and another period for the other step? A combination of both? Please clarify.

- **L309-315:** I am confused so as to how this step-wise calibration was performed. First, I am interpreting L309-310 as having a first step using isotopes, energy fluxes and water balance data as a constraint, and then a second step using biomass data; or rather, 4 steps for each data group? Please be more explicit, and possibly add this information to Table 2 as well. Secondly, since each steps use 100,000 samples, I am guessing that step $i+1$ does not use a subsampling from calibration step $i$; how were the calibration steps connected? Overall, this section needs a substantial rewriting to understand how calibration was actually performed; under which hypotheses regarding parameter space, total number of parameters, etc. Consider adding additional supplementary tables with information on calibration ranges at each step, resampling procedures, etc.

- **L325-328:** How was the sub-discretization done? Also, why not trying to change the thickness of the first layer so that the measurement depths fall within the model layers, not at interfaces between model layers (e.g. layer 1 could be 20cm-deep)? Adding the same red line to L1 moisture under grassland could be informative in checking for percolation; from these figures it seems that infiltration-percolation under the grass patch is underestimated.

- **L330:** Another obvious isotopic feature is the much higher ~week-scale variability in 10cm isotopic at site A (Fig 3a) as compared to site B (Fig 3b). This is reasonably differentiated in the simulations cells although 1. simulations at site A are too dampened and 2. there an unrealistic depletion in October at site B. While the former is briefly mentioned in Sect 4.3, I suggest to add these descriptions here and discuss them further on in the Discussion.

- **Figure 3:** Are isotopic datasets daily-averaged in this figure? If so, it should be stated somewhere in the main text.
L344: The model description states that there are two thermal layers in EcH2O-iso (without providing the depth of each), can the authors briefly describe how they extrapolated the modelled soil temperatures at three depths?

L345: Although the scales in Fig. 4 (Site B, latent & sensible heat fluxes) are quite squeezed (please consider expanding them), it is apparent that latent heat is overestimated throughout the growing season.

Figure 4: How was modelled grass transpiration converted into sapflow? It would be informative to see the transpiration rate (mm/d) in the second row, perhaps using a secondary y-axis on the right?

L358-364: Could the authors precise which MODIS LAI product was used? These products usually have a much larger spatial resolution (500m-1km) then the modelled domain of this study. Can the authors develop on their methodology and assumptions made to distinguish willow and grass patches?

L370: A reference to Table 3 would be useful.

Table 3: This table shows a lot of information. It might be much more reader-friendly if transformed to a multi-panel plot, either using bar or points with errorbar, e.g. keeping the row and column organization with facets and a color code for time periods. In addition, the third grouped-row (RU-L*) might be more intuitive if instead of layer number, depth ranges were used (e.g. RU[0-10cm]).

Figure 6: My understanding is that soil isotopes are measured in-situ at three depths, as reported in Fig. 3; why then are there not 3 solid lines in the diurnal plots, instead of 1 (panels a) and c)) or none (panels b) and d)), and why is the solid (measurement) line flat, as if there no high-frequency information? Additionnally, given the high-frequency dynamics, readability would be improved by making this figure wider, e.g. having Willow 1 and Willow 2 panels on top of each other.

L412: A reference to Fig 3a (in addition to Fig 6a & b) would be helpful.

Table 4: I am assuming the values between brackets give KGE variability among best runs? If so, why isn’t the same number given for AIC? Consider using a plot rather than a table (althoug less critical than for Table 3).

L440-449: In my view this labelling by “contributing month of the year to current store/flux” rather provides a very nice perspective, equally important and intituive as the “time elapsed since arrival” reported above ; it directly replies to the question “what precipitation period is most important for plant water use?” ; I would suggest to move key Fig. S3 to the main text.

Fig. 7: “Time in xylem” (panel g) is somewhat misleading, as the transit time considered integrate transit along root and xylem? Besides, my impression was that
transit length (and thus time) in the xylem was neglected when computing $v(i)$ in Eq. (1) (see related comments above)?

**L450:** “an increase of zero days” seems somewhat odd, maybe rephrase: “Since intantaneous(ly?) mixing equates xylem water age to that of where water is taken up (reaches 1m instantly), transit times along the root-xylem system are only shown...”.

**L479-496:** The underestimation in modelled willow transpiration (or rather, th sapflow, see a comment above) at the end of the growing season is quite interesting, as perhaps not as “minor” as stated here ; the model-data discrepancy exceeds the dispersion among best runs. That would deserve further discussion, as the current ones somewhat circumvent the issue with more general considerations. Besides, the concommitent overestimation of modelled L1 moisture (and possibly L2’s, and thus percolation, Fig. 3a) suggests that it’s not necessarily due to missed contributions from adjacent cells or a short-term reliance on deeper stores (which would have been interesting as a drought-protection process!), but merely that there is something wrong with evaporative demand when the energy balance is computed ; is it something due modelled energy fluxes and/or to forcings? In other words, is a process being missed?

**L480-481:** Is this sentence suggesting that EcH2O-iso account for off-cell contribution to calculate root water uptake? And associated transit times?

**L492-494:** From the ‘slight decrease’ I am wondering if the authors meant “was under stress”? Besides, it could be informative to further have the absolute biomass in each compartment (in addition to biomass allocation) reported somewhere, perhaps as time series over the growing season, to check if the potential decay rates exceed (or not) allocation, and where.

**L514-521:** I assume this part of the discussion will be substantially revised (see General Comments)

**L536-539:** If the measurement uncertainty is known, it would be highly informative to add it as error bars on any related plots presented in this manuscript. Actually, it should be common practice, helping to tamper interpretations where inferred dynamics are commensurate with uncertainties.

**L546-553:** This issue could be explored with the different tree storage mixing types presented in Knighton et al. (2020), it could help the current discussion and open avenues for further development?

**L552:** Are the authors referring to measured basal diameter?

**L561-563:** Maybe further precise “across Switzerland” after “Allen et al. (2019)” and “in the study region” before “(Miguez-Macho and Fan, 2021)”?

**L575-578:** This is also potentially due to the fact that other studies considered root-to-
shoot transit times (Meinzer et al., 2006) while this study “stops” at 1m height.

L579: Essential or indispensable?

L580-591: Again, it is quite surprising not to see any references to Knighton et al. (2020), a study the authors contributed to, and which precisely studies this issue of tree water storage and mixing.

L589: Mennekes et al. (2021) and Benettin et al. (2021) are recent studies on this topic, albeit in semi-controlled conditions.

Conclusion: Having the Conclusion framed as Summary (L596-608) seems a bit redundant with the abstract and the main text. Rather, discussing high-level limitations, insights and potential avenues would more efficient.

Figure S1: The channel is not represented in Fig 1b, and the similar color code for snowpack/channel may confuse the reader).

Code availability: The statement is somewhat incomplete, as the post-processing model to compute root geometry and and associated transit times does not seem to be on the referenced repository.

Data availability: Again, this statement is misleading, first because “open-access” is incompatible with password-protection. Secondly, not all the data used in this manuscript is archived in the provided link; only sapflow, stem variation and in-situ isotopes data are listed, while neither eddy-covariance energy fluxes, micrometeorological measurements, in situ LAI, and soil moisture can be found. I would strongly encourage to have all datasets published, or at a minimum have them listed along with their open-access metadata on FRED so that potential users can make informed queries to the curators.

Technical comments

L34: “seasons” instead of “seasonally”?

L80: “using” instead of “by”?

Table 1: Precipitation is in mm (not mm/year), given that column 4 reports the value over the growing season only.
L155: I am guessing that “any” refers to wounding effects, but the formulation seems odd. Consider rephrasing.

L334: “to measured δ²H”?

L440-441: Rather than “discretized”, maybe “aggregated” is more accurate?

L455: “[...] from the tip of the roots to 1m [...]”

L545 & L552: It seems this Fig. S5 is missing, and that the current Fig. S5 is the one referred to as “Fig. S6”

**Figure S1:** “Maneta (2021)” does not seem to exist, did you mean “Maneta et al. (2013)”? In addition, there are two “Smith et al. (2020)” cited in this manuscript, but I could find a similar figure in neither of these references...please clarify.

References

Benettin, P., Nehemy, M. F., Asadollahi, M., Pratt, D., Bensimon, M., McDonnell, J. J., & Rinaldo, A. (2021). Tracing and closing the water balance in a vegetated lysimeter. *Water Resources Research, 57*(4), e2020WR029049.

Fatichi, S., Ivanov, V. Y., & Caporali, E. (2012). A mechanistic ecohydrological model to investigate complex interactions in cold and warm water-controlled environments: 1. Theoretical framework and plot-scale analysis. *Journal of Advances in Modeling Earth Systems, 4*(2).

Jasechko, S., Sharp, Z. D., Gibson, J. J., Birks, S. J., Yi, Y., & Fawcett, P. J. (2013). Terrestrial water fluxes dominated by transpiration. *Nature, 496*(7445), 347-350.

Knighton, J., Kuppel, S., Smith, A., Soulsby, C., Sprenger, M., and Tetzlaff, D.: Using isotopes to incorporate tree water storage and mixing dynamics into a distributed ecohydrologic modelling framework, Ecohydrology, 13, 2020.

Landgraf, J., Tetzlaff, D., Dubbert, M., Dubbert, D., Smith, A., and Soulsby, C.: Xylem water in riparian Willow trees (Salix alba) reveals shallow sources of root water uptake by in situ monitoring of stable water isotopes, *Hydrol. Earth Syst. Sci. Discuss.*, https://doi.org/10.5194/hess-2021-456, in review, 2021.

Maneta, M. P. and Silverman, N. L.: A Spatially Distributed Model to Simulate Water, Energy, and Vegetation Dynamics Using Information from Regional Climate Models, *Earth Interactions*, 17, 1-44, 2013.

Mennekes, D., Rinderer, M., Seeger, S., & Orlowski, N. (2021). Ecohydrological travel times derived from in situ stable water isotope measurements in trees during a semi-controlled pot experiment. *Hydrology and Earth System Sciences Discussions*, 1-34.

Schlesinger, W. H., & Jasechko, S. (2014). Transpiration in the global water cycle.
Agricultural and Forest Meteorology, 189, 115-117.

Sperry, J. S., Wang, Y., Wolfe, B. T., Mackay, D. S., Anderegg, W. R., McDowell, N. G., & Pockman, W. T. (2016). Pragmatic hydraulic theory predicts stomatal responses to climatic water deficits. New Phytologist, 212(3), 577-589.

Zink, M., Kumar, R., Cuntz, M., and Samaniego, L.: A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty, Hydrol. Earth Syst. Sci., 21, 1769–1790, https://doi.org/10.5194/hess-21-1769-2017, 2017.