Biogeosciences Discuss., referee comment RC3
https://doi.org/10.5194/bg-2021-278-RC3, 2021

Comment on bg-2021-278

Anonymous Referee #2

Referee comment on "Modelling temporal variability of in situ soil water and vegetation isotopes reveals ecohydrological couplings in a riparian willow plot" by Aaron Smith et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2021-278-RC3, 2021

The study by Smith et al. (bg-2021-278) presents a novel combination of in situ temporally high-resolution 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 neighboring 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 a for 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 of this study is highly relevant and timely. The ecohydrological community is ‘on alert’ at present, with novel opportunities arising from in situ, higher-frequency isotope measurements in soil and plants. At the same time many new discoveries related to methodological issues measuring water isotopes in these compartments arise steadily. Both aspects provide opportunities, but also a number of challenges related to modeling these datasets.

The presented study is a complex and well-conducted investigation on how to combine multi-facetted datasets into a joint modeling framework. This is certainly something I applaud the authors for. With multiple years conducting in situ isotope and ecohydrological measurements in several environments, it is simply great to see how such datasets can be put into one modeling framework. Having that said, I find it crucial to also implement measurement uncertainty into modeling frameworks. All the recently discovered isotope effects certainly increased the measurement uncertainty, and this – in my opinion – also increases model uncertainties? Can we even make reliable quantitative statements considering both? I know this goes farther than this publication, but I think it is necessary to have this in mind. Hence, the way this modeling exercise was carried out is excellent and has great potential for using such models for other, recently recorded, in situ datasets. However, the quantitative estimates of this study only incorporate modeling uncertainties. The lack of replication, uncertainty of soil and plant water isotope measurements, and spatial variability of ecohydrological measurements makes the quantitative value of the modeled data at least questionable. While it is probably impossible to address this in the presented study, this should certainly be on the future agenda. However, an honest evaluation and interpretation of the modeled data in that regard would benefit the manuscript in my opinion. At the same time, the manuscript could be shortened by putting less emphasis on the quantitative results and more on the modeling framework, strengths and also weaknesses.

In summary, the study definitely deserves to be published in BG, but requires thorough revision.

Specific comments:

l.75: I would leave out importantly. It is important, but doesn't need this explicitly here

Fig. 1: Figure caption is incomplete, in particular d) what are the blue and red bars? What is the grey box?

l.144: Sensors were installed until 1 m soil depth. Is that the maximum rooting depth for both willow trees and grass? This is crucial for root water uptake depth determination

l. 145-160: Even though I understand the method is described in Landgraf 2021, the information on how isotope standards were prepared and measured would be good here. Also, referencing the borehole method because of the short description herein should be considered.

l. 178-180 and chapt. 3.2.2: how were these parameters determined/calibrated?

L.214: the last part of the sentence is unclear, please rephrase and clarify

L.216: calibration? How was it calibrated?

L.216-223: this approach is interesting, was this used somewhere before? (citation?). It appears like such an approach would completely neglect preferential flow, am I correct? If yes, this should be stated somewhere (‘does not account for pref. flow’)

L.229: assumed root distributions…this is a BIG assumption. How were they assumed?

3.3.1.: How were the root parameters determined/approximated?

3.3.2.: For someone who does not model every day, the explanation on root length determination should be clearer. Coming from the field side of things, I wonder ‘how is maximum rooting depth implemented?’; which measured parameters does one actually need (precipitation and sap flow?). I also wonder, if the general root distribution in the model always has the same shape? This is a large simplification that is definitely not true for any given vegetation species. How does it look like if we have a deep-rooter, for instance?

How was the fact handled that there very likely were willow roots present underneath the grass, affecting soil water contents and hence, the modeling efforts?

L.277: this is an interesting point, but it should be noted that there is not only an error in simulating, but also measuring soil water isotopes. I am not saying that it should be, but is there a way to include this in such simulations?

L.288-291: Maximum rooting depth is constrained to 100 cm. This needs to be proven/backed up. Stating another paper under review/discussion (here and in many other instances) is sort of cheating, to me. Root water uptake depths shift over a year and it cannot be assumed for the time of experiment (~3 months) that 100 cm max. rooting depth are a given. Please clarify this; I do believe the authors and a quick search tells me that willow trees are generally shallow rooted. However, another citation would help.

L.305 please explain thoroughly, why 18O was not used in calibration

L.306: What is meant with ‘the values for 18O were not greatly different from 2H’? First off, these values are usually very different. Second, the dual-isotope space provides an excellent way of validating the effect of kinetic fractionation. Third, I feel like a comparison of measured and modelled values in dual-isotope space would greatly benefit the trust in the model, apart from the statistical parameters.

Table 2: Calibration data: Why is only sap flow of 1 tree used? Likewise, Surface Temp and latent heat only from site B? This seems subjectively chosen and is not explained in

the text.

L.324: …starting from likely, it belongs to discussion

Results: the subjective phrase like ‘adequate’ or ‘slightly different’ should be backed up by some objective measures in the results section.

L.335-338: Just to clarify: The heating cables were not put inside the soil profile, or were they? I am asking this because we did this mistake once in my group and it turned out the cables heated the surrounding soil, hence, producing a heating of the area around the soil gas probes and tdr probes. As a result, one would calibrate data on a totally non-representative dataset that is highly influenced by the heating cables and not representative for the stand.

Figure 3: This looks nice indeed, in particular for Site B! However, I repeat my statement from before that the dual-isotope space allows for a more precise evaluation of model performance and further interpretations such as root water uptake depth or kinetic fractionation. Another thing: There is definitely an uncertainty in the in situ isotope measurements, which is almost never incorporated into modeling. However, modeling always incorporates uncertainty in calibration results. I find this odd and not necessarily correct.

The complete section 4.1 does not make use of any goodness-of-fit criteria and uses subjective and biased statements throughout. For instance, the calibrated sap flow data is judged as ‘adequately captured by the model’. If I look at Fig.4 I (subjectively) see that the dynamics are OK (Site A) while the magnitudes are sometimes. For site B, there are no measured values for sap flow. This is not convincing to me. I strongly recommend adding goodness-of-fit criteria here.

L.343: ‘quite’ well…objective measure?

L.396: simulated day-to-day variability could not reproduce the measured values

4.3: I find this section well-written and less subjective/biased. The general dynamics are met, but it needs to be said that an offset of 10 in d2H is already a large deviation (in isotope space). Now is that because of a non-perfect model fit or, and I am sure that it also plays a role, uncertainty in the in situ measurements. I feel like including some statements/metrics in regard to the measurement part of the second paper submitted by the authors could benefit the interpretation here. I find the aspect of the time-steps quite interesting: Why temporal resolution do we actually need? In isotope-space, daily is already a great resolution.

L.479/480: ‘with only minor under-estimation of the transpiration in the willows toward the end of the growing season’…I do not agree that the deviation is minor (>50%) nor that the fit is great for the rest of the period. The dynamics fit, but the magnitudes often do not. And at site B, no comparison is provided.

(Recent) literature that might be of interest:

Beyer, M. and Penna, D.: On the Spatio-Temporal Under-Representation of Isotopic Data in Ecohydrological Studies, Front. Water, 3, 643013, doi:10.3389/frwa.2021.643013, 2021.

Kühnhammer, K., Dahlmann, A., Iraheta, A., Gerchow, M., Birkel, C., Marshall, J. D. and Beyer, M.: Continuous in situ measurements of water stable isotopes in soils, tree trunk and root xylem: field approval, Rapid Commun. Mass Spectrom., e9232, doi:10.1002/RCM.9232, 2021.