Reviewer 1

In this study, the authors investigate the sources of water absorbed by Willow trees using both bulk and in situ monitoring of water isotopic signature in the soil and stems of two of these trees. A Bayesian mixing modelling approach to water sourcing suggests that the vast majority of water is absorbed from the top soil (above 40 cm depth), except at the end of the summer period when a fraction of the xylem water could originate from deeper soil layers (40 cm and 100 cm).

General comments:

From a general point of view, I believe the manuscript is sound, well-written and structured. It is also very well adapted to the audience of Hydrology and Earth System Sciences. On the one hand, I found the comparison of soil water content and in situ monitoring of water signatures particularly interesting and think the discussion could emphasize the related potential implications a bit more. On the other hand, the conclusion on the production of reliable 2-hourly resolution data for water signatures I think would need further support in the main text or should be tuned down. My main concern is about the quality of the figures, which would require substantial work. Yet, I think the manuscript is overall in a good shape and could be considered for publication in HESS after minor revisions.

Response to General Comments

The authors thank the reviewer for their constructive comments. In the revision, we will emphasize more the relationship of in-situ and other ecohydrological parameters like soil water content in the discussion. In turn, the discussion about the reliability of 2-hourly resolution data will be shortened. All figures will be uploaded with a higher resolution and larger axis labels to ensure high quality.

We add the all the revised figures already into this document here below (which addressed all the constructive suggestions by the reviewer).

Specific comments:

Lines 81-*84 (L81-84): Besides radial growth, doesn't the swelling / shrinking response of tree stem diameter reflect stem water tension, which results from the imbalance between canopy transpiration and root water uptake? This kind of formulation would seem more intuitive to me. Was any such swelling / shrinking fluctuation observed daily? If not, this part of the introduction could as well be removed to make the manuscript more concise.*

Response to L81-84: The authors thank the reviewer for their suggestion. Yes, there was daily swelling / shrinking observed in the tree stems of the willows. The authors will emphasize the daily cycle of shrinking and swelling in the results section. We will also make clearer that stem water tension due to swelling / shrinking occurred.

L85: *I think this part needs a bit of rephrasing too as transpiration is considered as the product of VPD and canopy conductance (often termed gs). Canopy conductance is considered to react to multiple factors like limited light, limited export of products of photosynthesis and limited water availability. Therefore, I think writing that sap flow usually reflects VPD is a bit strong. Maybe clarify "in absence of (…) limitations (…)".*

Response to L85: The authors will clarify the relationship of sap flow and VPD during the revision.

L88-89: *Here I would suggest to re-order the terms. It could be more intuitive to mention decreased sap flow after low leaf water potential, turgor, and stomata closing, as it would come last in a temporal sequence of observations.*

Response to L88-89: Great suggestion. The authors will change the order accordingly.

Figure 1: *There seems to be something wrong with the coordinates along the frame of panel A. Could you check if that is the case? Also, several elements in the figure are very small and hard to read. For*

instance, the labels in panel B, the dots marking soil pits and probes, whose colours are hard to distinguish for such small points. Could you modify the figure to facilitate its readability?
Response to Figure 1: The coordinates along the frame are set automatically by using QGIS. The coordinate systems is World Geodetic System 1984 used in GPS (EPSG:4326 - WGS 84) and will be updated to Universal Transverse Mercator (UTM). The figure will be modified to improve its readability.

L143-144: *Can the eddy flux covariance measurement be considered as representative of the Willow trees transpiration? If not, could you clarify already at this point how the data is intended to be used?*
Response to L143-144: No. For this the footprint of the Eddyflux is too big. However, the measured ET data were used as normalized data in terms the variability in the ET dynamics rather than the absolute values of Willow transpiration. The authors will clarify the usage of Eddy Flux data in terms of using the Eddy flux data as an indicator of dynamics / variability in ET (rather than absolute values).

Figure 3 (and others): *The use of the symbol "/" preceding units in the y-axis label is a bit confusing. Could you use another symbol, like the pair of brackets for instance?*
Response to Figure 3 (and others): We used the axis description after ISO/IEC 80000. However, for improved readability we will change this for all units in the revision.

Figure 4: *The right-side y-label would be clearer specifying "stem diameter variation" as it seems the stem size was not close to zero in May 2020. Northern and Southern Willow labels directly in the Figure would also be convenient. The legend indicates "Sap flow N" in both panels but it is unclear what "N" stands for. Could you clarify? Sap flow rates and ET would also gain at being overlain all in panel 4a for easier comparison, while stem diameter variation could be displayed in panel 4b.*
Response to Figure 4: Thanks for the comments. We will update figure 4 accordingly. There was a typo causing the beginning of the measurement to be above 0 for the stem diameter variation. This has been corrected for in the updated version of the Figure. Further, "Sap flow N" will be clarified as the sap flow measurement from the northern site of the stem.

L290: *The subtitle could also mention "and transpiration".*
Response to L290: The authors will update the subtitle.

Figure 5: *The figure is particularly hard to read as symbols are small, several of them have similar colours and/or are hidden under geometrical elements. If the ordering of the legend is conserved in the bar plots, it is confusing that precipitation is associated to light-grey bars but black dots, surface water medium-grey bars but white dots, etc. Could you work on a new version of this figure that is easier to read?*
Response to Figure 5: The authors will increase the transparency of the polygons and precipitation, surface water, and groundwater will be plotted with dots matching the color of their bar plot counterparts.

L307: *Here and at other places in the manuscript, the use of the "-" sign to specify ranges of values while these values are negative is not ideal. Could you replace the "-" sign by the word "to"?*
Response to L307: This is a good suggestion. The authors will replace "en dashes" used for ranges by "to" in the revised manuscript.

L309: *I am relatively new to water isotopic studies, so the concept of "lc-excess" I had to look up. I think for the audience of HESS it would be worth carefully defining the lc-excess with the associated equation and possibly a one-sentence example to make it easier to grasp.*
Response to L309: The authors thank the reviewer for this comment and will clarify lc-excess in the methods section incl the equation.

L312: *The expression "generally similar to" is a bit vague given the wealth of data available in this study. Were in situ and bulk measurements significantly different? Is there any indicator of their similarity (e.g. R-square, …)?*
Response to L312: The authors will give a similarity analysis either with R^2 or Euclidean distance or Mahalanobis distance.

L314: *Please clarify the type of variability (space, time, …).*
Response to L314: The authors will update this section by stating that the higher variability of Pit A compared to Pit B over time is caused by the special differences of the two soil pits.

L339: *Could you comment on what could make responses more marked in covered pit A than pit B?*
Response to L339: Pit A was directly under one of the Willow trees and higher interception seems to have resulted in much drier soil conditions at Pit A compared to Pit B. Thus, Pit A relatively to B had smaller water amount present in the soil. Due to these generally smaller water amounts incoming precipitation varied the isotopic composition of the top soil water of Pit A more intense than it did at Pit B. We will clarify this section in the revision.

Figure 7: *In panel (a), the location of the "zero" differs between vertical axes, which complicates the visualization of the results. Could you fix this? Some of the panels (b) to (e) could as well be merged to facilitate the comparison of the results and the visualization overall as vertical axes are currently very small.*
Response to Figure 7: Thank you for pointing this out. The authors will bring the zero-lines from panel (a) x-axis on one line. Further, panel (b) and (c) as well as (d) and (e) will be merged to one graph.

Figure 9: *Several of the box plots are hard to distinguish from the border of the panels. Could you make them easier to distinguish for instance with a lighter panel border colour? The legend for soil moisture lacks the dotted / dashed aspects. Could you make the legend more consistent with the content of the figure?*
Response to Figure 9: These are good suggestions. We will improve the clarity of the box plots. Further, the legend will be extended to also include the dotted and dashed line.

L358-359: *The prediction of water uptake almost solely from the top soil in the Bayesian modelling output is quite interesting. If you had to explain the remaining differences between water isotopic signatures in the stem and at 10 cm depth, what would be the other sources that you would consider as necessary complements? Could you discuss the results also from this perspective? From a quantitative point of view, I think the authors could as well have argued (but don't have to) on the possibility to have root water uptake solely from the top 40 cm of soil by comparing the cumulative tree transpiration and rainfall to the soil water storage change. My estimations suggest that, even when neglecting water capillary rise from deeper layers, the water balance in the top soil seems reasonable if the tree roots extend about 10 meters away from the stem.*
Response to L358-359: With the setup of the in-situ site we considered groundwater, lake, stream as well as soil water to be the main sources of root water uptake. The groundwater, due to the adjacent lake, is relatively shallow with ~2 m below soil surface and well mixed with the lake water. Still, we don't know about the soil water isotopic composition below 1 m and the groundwater surface. Further, the IGB garden is generally irrigated which was done ~8 m away from the site. During the setup period we did not consider this, because we assumed a much smaller horizonal root distribution. The irrigation was done with lake water, but it's possible that the irrigation, being a lot of water poured on a relatively warm soil during hot summer days caused enriched water input into the soil that was later used by the more distant roots of the willows. It's quite interesting that quantitative the trees could also have used the water at 40 cm and that's a good point to discuss since it shows the importance of stable water isotopes as tracers for root water uptake.

L376: *I found the phrasing "soil moisture availability was (…) not less than (…)" a bit odd. Do you mean that it is higher (as suggested by Fig. 9) but not significantly?*
Response to L376: The soil moisture in September was not significantly different in September compared to hotter summer month like July and August. We will clarify this in the revision.

L394-397: *This part of the discussion is quite interesting. In the results I kept wondering why deeper water uptake occurred just as the upper soil was rewetting, which is counterintuitive for a specialist of hydrodynamics, as passive water flow from shallower soil layers to roots is supposed to increase as the water potential in these layers increases (a process called "root water uptake compensation"). So, your observations seem to go against the second law of thermodynamics (how exciting!). You do mention the possible explanation that the stem water signature could be the integral of past and present water uptake (thereby possibly reconciliating your observations with thermodynamics laws). If that is the case and if the volume of the stem water pool mixing with newly absorbed water is large enough, one would expect to see the signature of deep water in stems weeks after deep water uptake occurred. Obviously, this calls for more investigations of the mixing of soil and stem water pools ideally under controlled conditions and with labelled water, as well as for new versions of Bayesian mixing models that account for such a temporal integration of water signatures (or do they exist already?). The temporal integration of water signatures would also alter the ability to infer on water uptake profiles at high-temporal resolution, unless one was able to "deconvolute" the temporal dynamics of stem water signatures. Could you discuss this a bit more in depth in the manuscript?*
Response to L394-397: Thank you for these constructive comments. It's true that our assumption goes against the second law of thermodynamics, but since trees are living organisms, it's possible they found a way around this problem (like giraffes found a way to breath with an air tube volume larger than the one of their lungs… physically impossible). We think about this possibility, because other studies found willows to prefer water from drier soil parts even though it could be accessed more easily in stream or groundwater (cf. Martilla et al. 2017). If biological activity is behind this phenomenon, it will be fascinating to identify the physical/biological processes behind it. However, other processes like mixture of "new" and "older" water in the xylem could explain our results as well and are already discussed in detail the study from Mennekes et al. (2021). We agree that further condition-controlled experiments are required to determine the processes behind our results in detail. Further, a mixing model addressing the temporal integration is also favorable for stable water isotope investigations in xylem water. Moving-window analysis might be an easy, but useful approach to evaluate the temporal variability. We will revise this in further detail.

L418-440: *Here diurnal fluctuations of the signals are discussed in depth, which I found a bit odd as I did not find supporting data and results in the main text and figures (I guess they are in appendices). I think if they are to be discussed in depth, they should be presented in the results and figures for the audience to have a grasp on what was observed at such a temporal resolution. It is striking in particular to conclude (L453-454) on the production of "reliable high-frequency (…) 2-hour resolution" observations while I think only daily-averaged data is presented in the body of the manuscript. I think it would be more consistent to present the high-resolution data and demonstrate confidence in its quality, or not to conclude on the production of such high-resolution reliable data.*
Response to L418-440: You are correct. We will reduce the discussion about sub-daily data (see also our response to your comment further above).

L463-464: *This sentence is unclear to me. Could you rephrase it?*
Response to L463-464: The authors will clarify that ~90% of root water uptake was estimated by the mixing model to be linked to the first 10 cm of the soil.

Typos:
L25: *I think the expression "uptaking" is incorrect and the correct version is "taking up".*
Response to L25: The authors will use the correct term of "taking up".

L326: *The signs "?:" seem to be written in place of " for".*
Response to L326: The authors will remove the signs "?:" from the text.

L413: *I think N2 requires a subscript for the number 2.*
Response to L413: This will be corrected.

Revised figures:

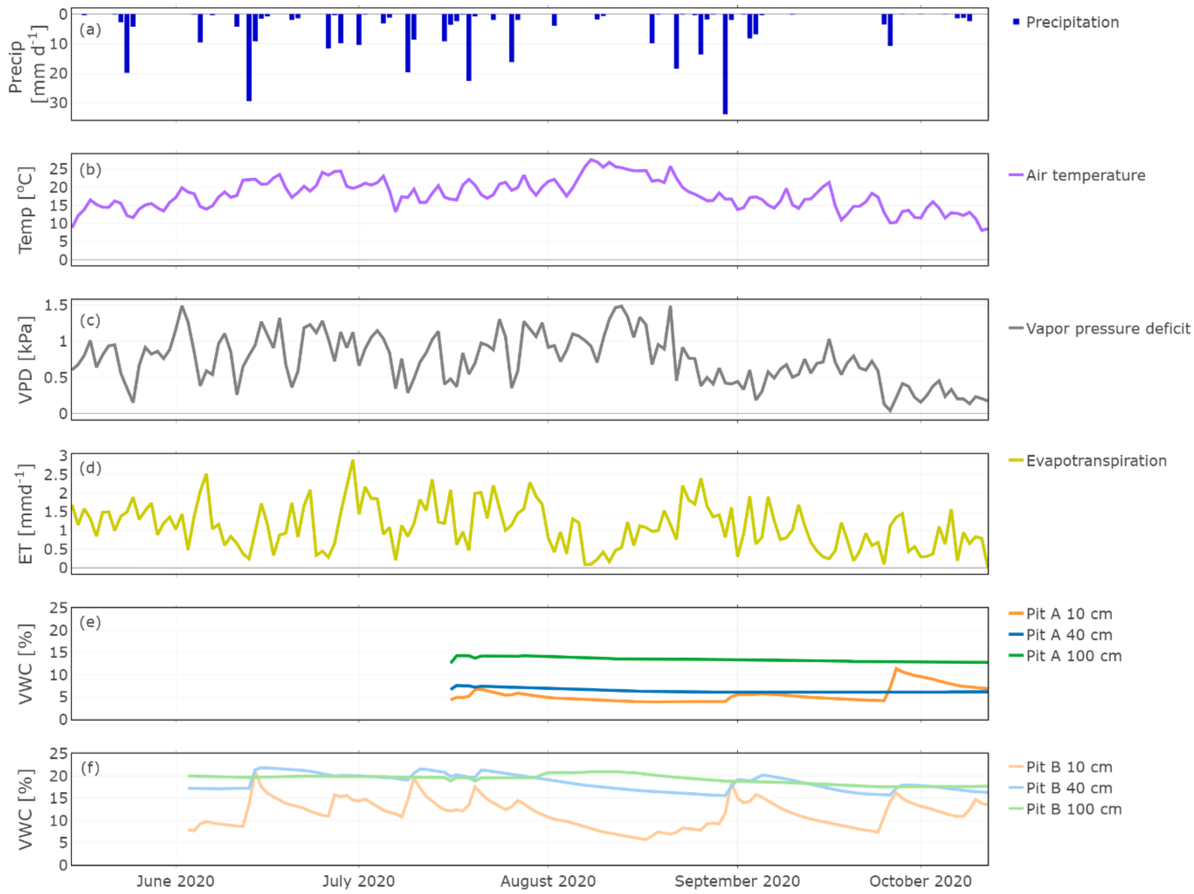

Figure 3. Hydroclimatic conditions showing daily precipitation (a), air temperature (b), vapor pressure deficit (c), evapotranspiration (d), and soil moisture for Pit A (e) and Pit B (f).

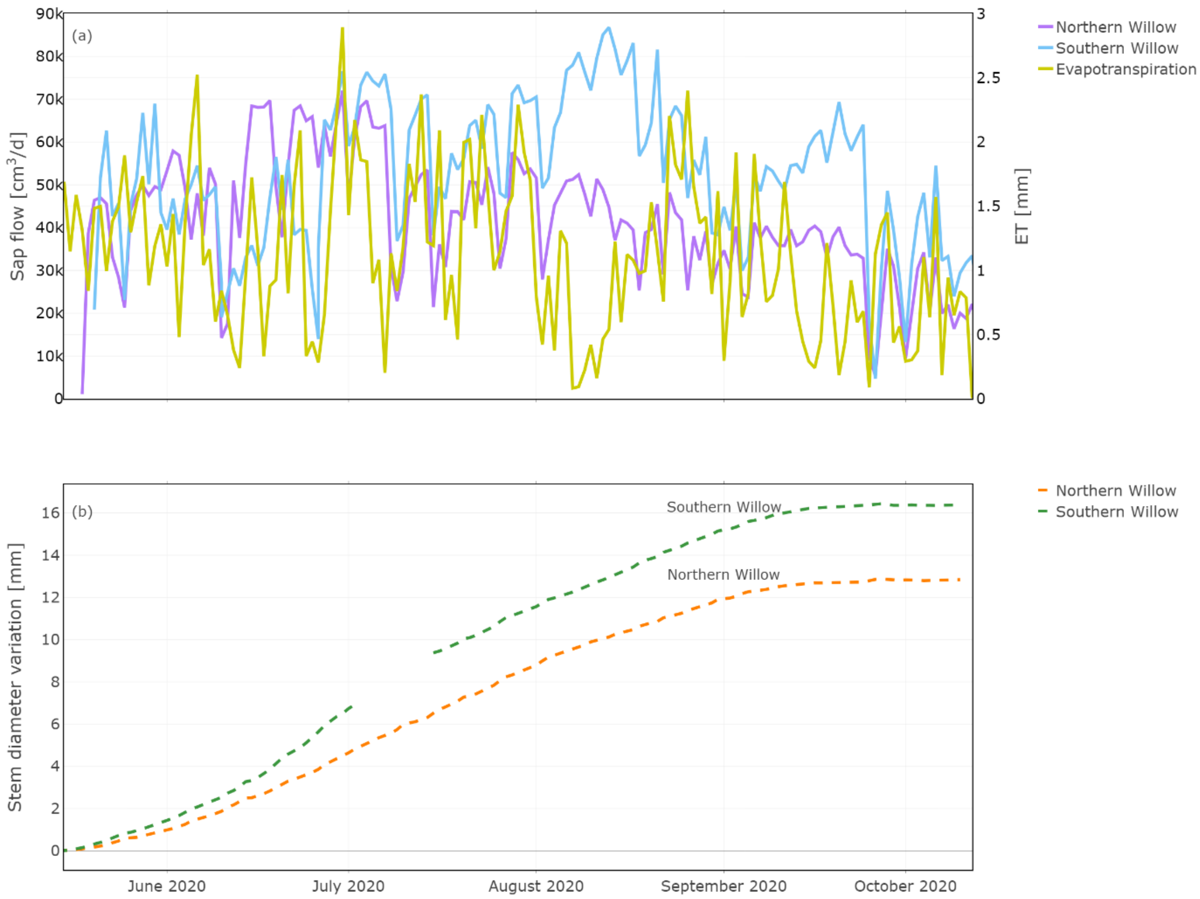

Figure 4. Daily total sap flow measured at the northern sap flow sensor and stem diameter variation measured for the stem-radius. Plot a) shows the results of sap flow together with Evapotranspiration for comparison and b) the results of the stem diameter variation during the measuring period.

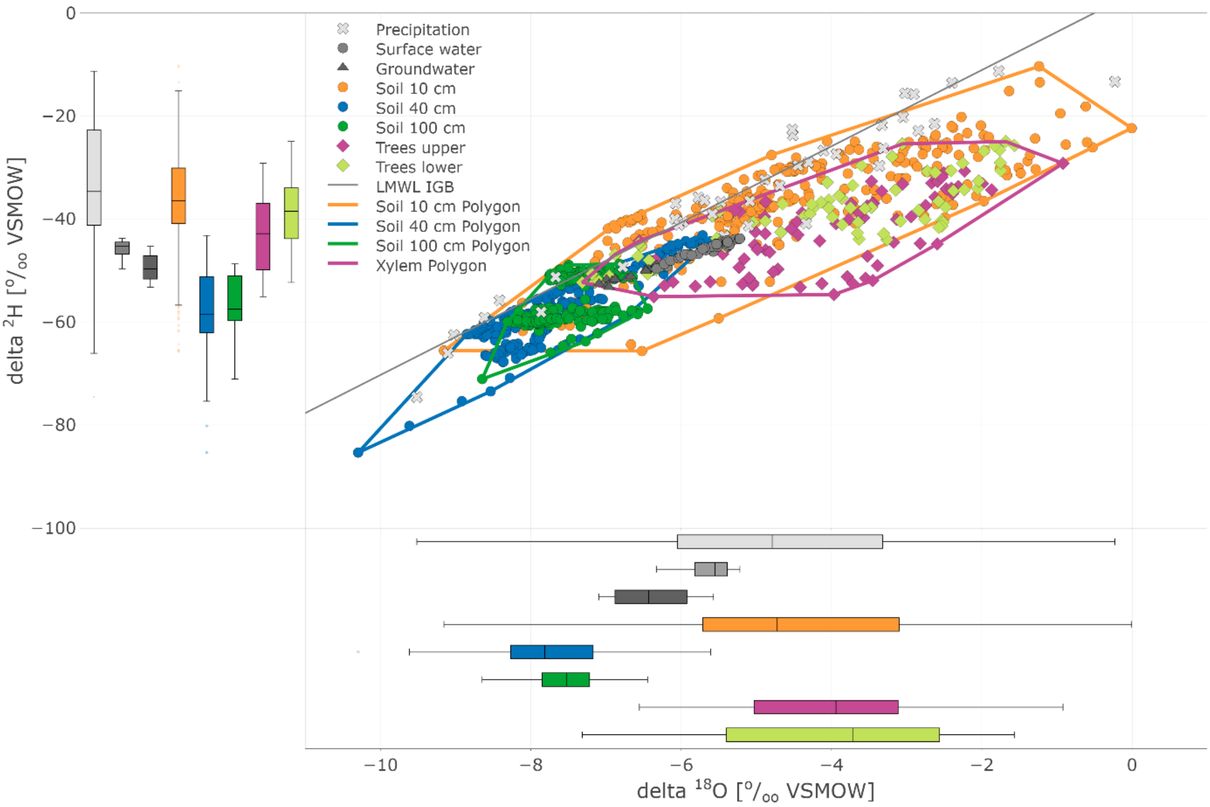

Figure 5. Dual isotope plot of in situ (daily) soil and xylem as well as precipitation (daily), surface and groundwater (weekly) sampling. Soil and tree data are highlighted with boundary polygons for 10 cm, 40 cm, 100 cm and tree (upper and lower results joined) clusters. Additional boxplots show the sample distribution of the data sets.

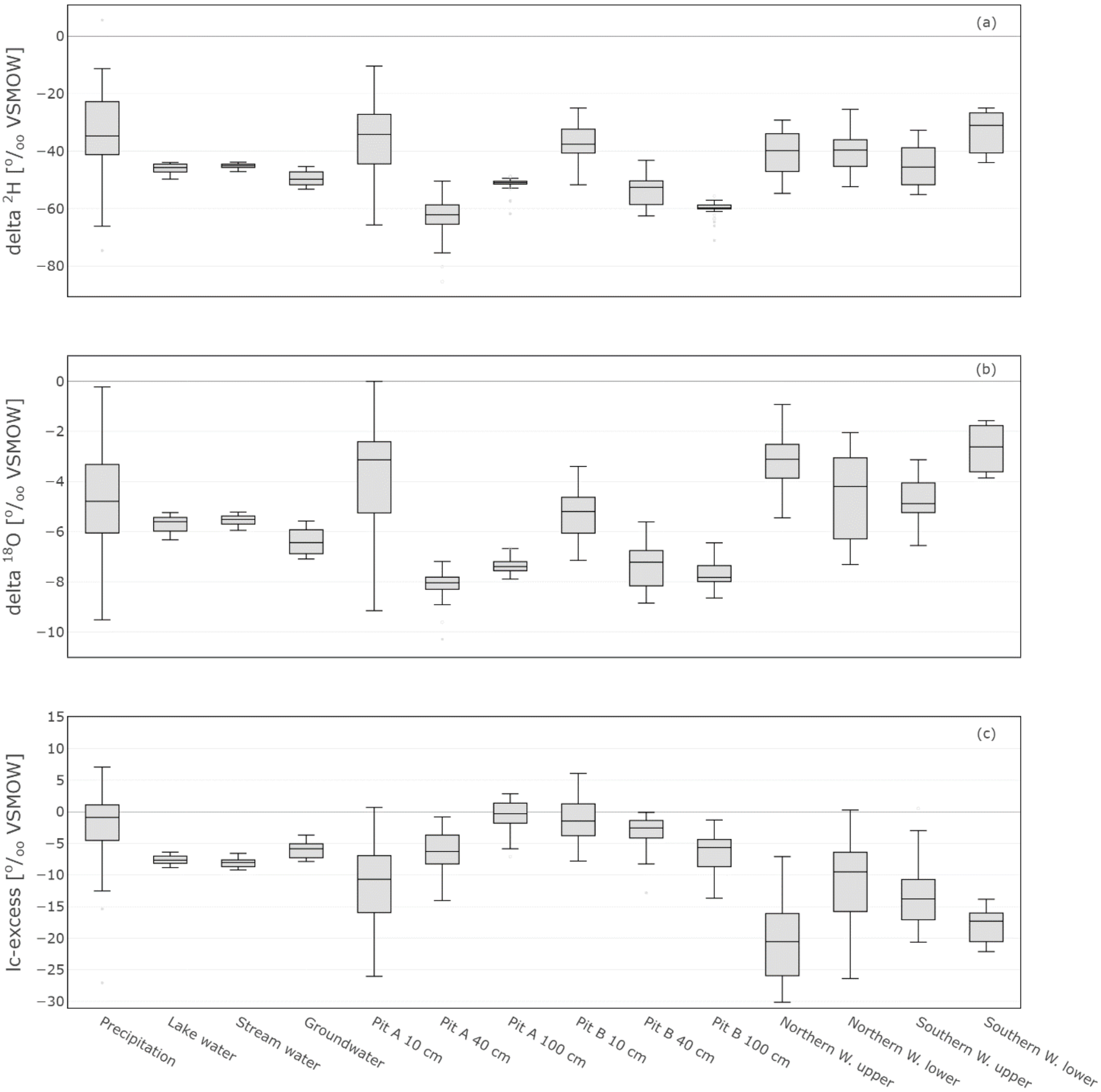

Figure 6. Box plots showing the isotopic composition of daily precipitation sampling, weekly sampled groundwater, lake and stream water, and in situ sampled soil, and xylem water.

Figure 7. In situ time series of daily δ2H in: Precipitation (a), Pit A and Pit B (b), and Northern Willow and Southern Willow (c).

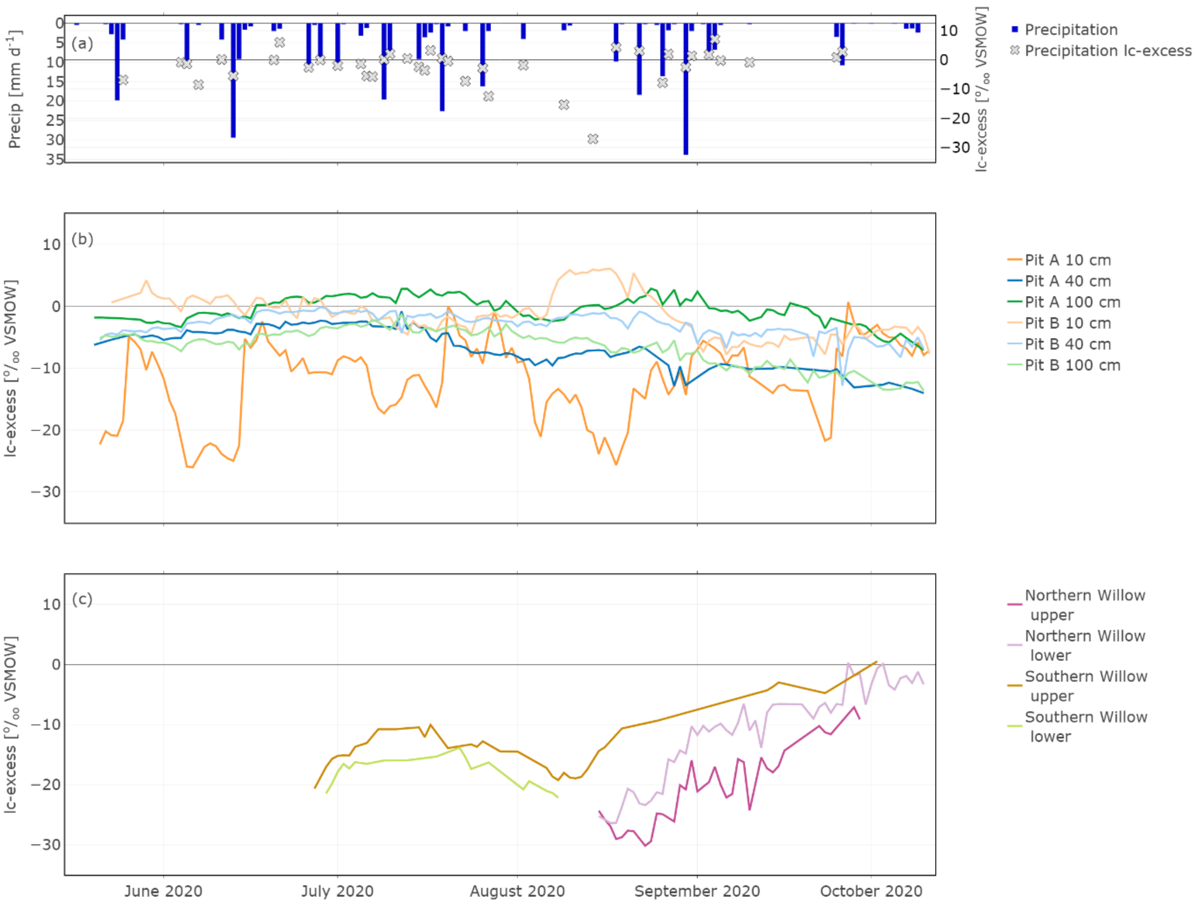

Figure 8. In situ time series of daily lc-excess in: Precipitation (a), Pit A and Pit B (b), and Northern Willow and Southern Willow (c).

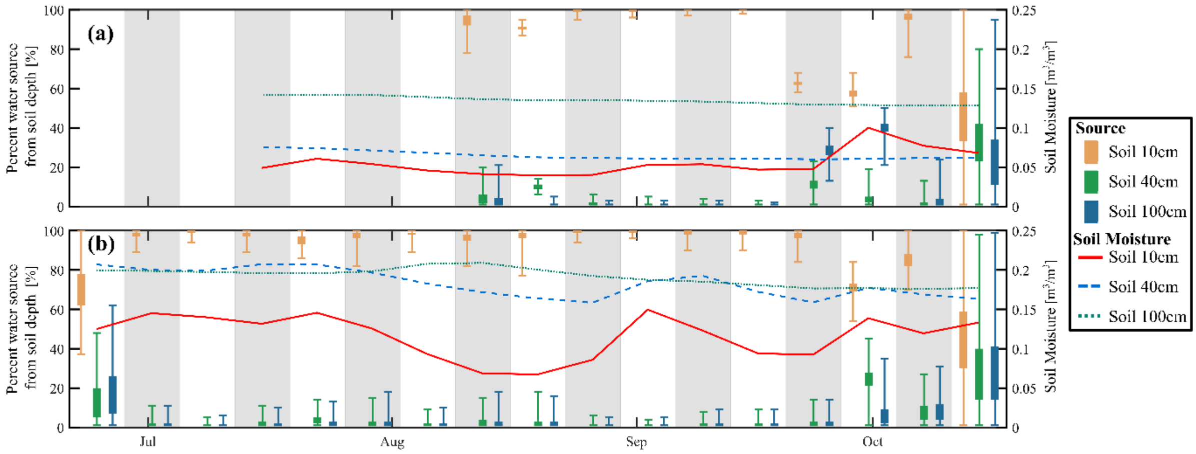

Figure 9. Estimated percentage of willow water from each soil depth in (a) Northern Willow and (b) Southern Willow. The box plots for each source show the 25th and 75th percentiles, with whiskers extending to the minimum and maximum estimated proportion. Gray and white bands show the division of each week. The right-side y-axis shows the soil moisture for each soil depth (weekly average).