Comment on hess-2020-640
Anonymous Referee #2

Referee comment with author replies on "Quantifying the effects of urban green space on water partitioning and ages using an isotope-based ecohydrological model" by Mikael Gillefalk et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-640-RC1, 2021

Author replies in red

The manuscript "Quantifying the effects of urban green space on water partitioning and ages using an isotope-based ecohydrological model" written by Gillefalk *et al.* provides a set of insights for water partitioning in a complex urban landscape. They incorporated the use of water stable isotopes in precipitation and soil to verify the model capacity for partitioning water fluxes. Also, they use eddy flux and sap flow data to evaluate the model results. Despite their meticulous work, there are some concerns about data collection and applicability.

We thank Reviewer 2 for their suggestions and complementing our "meticulous" work. We address the data concerns that the reviewer raises below.

Major Comments

Flux tower:

Authors mentioned in lines 147-150 the use of another urban flux tower for a portion of the sampling period. It is important to highlight the fact that despite their similarities as "Urban Environments", the proportion of green spaces/buildings can affect considerably the model outputs. Also, the authors did not show a consistency analysis or a comparison for the period June – November in which both towers could be operating.

This is fundamental to considerer the fluxes as similar, equal, or different. Fluxes such as outgoing longwave and shortwave radiation, as well as water vapor can be affected and give different proportions. As an example, the 2-fold overestimation showed in Figure 7 with respect to sap flow data for April and May can be linked to the differences between flux towers without counting on the constant LAI effect.

The authors should ensure that this data set can be used by this other location.

* How different are the land covers within the tower's footprints?

* Are the fluxes for the period June-November equal/proportional/different?

* Do both towers have the same setup in terms of instrumentation?

We do clearly state the periods of each eddy flux tower at the beginning of section 2.2. From the summer 2018 until the end of the of the modelling period in Nov 2019 we use the data from the tower on site. Therefore, no differences shown in figure 7 can be explained by the use of data from the second tower. The tower located 6 km north of the study site is only used during the first part of the spin-up and any difference in land covers should therefore be negligible for the calibration period.

Calibration:

The manuscript is based on the application of a model which requires a calibration period. The authors mention the application of this procedure (Section 2.4). Despite the detailed description of the calibration procedure, two main questions remain unanswered:

* Which data (period and source) was used for the calibration?

* Did the authors apply a spin-up procedure (how long) or not?

This issue is important to assess possible trends or initial effects in the flux initial values.

We are slightly surprised by this comment as both these questions are clearly answered in the manuscript, specifically lines 158-162 and lines 203-204.

LAI and sap flow:

During the modeling procedure, the authors used a constant Leaf Area Index per cover. This can be true for the grassland depending on the species but the effects in trees and shrubs fluxes can be important. The application of this assumption triggered important consequences for the model results which end up with the overestimation of transpiration fluxes during the first part of the year (Figure 7 – April and May).

However, the lack of sap flow data in shrubs affects the reliability of the fluxes from this cover.

As noted in the response to Reviewer #1, including the sap flow measurements was not meant to be a quantification of transpiration, what we wanted was a qualitative comparison of the variability. The overestimation only is visible in April and is exaggerated by our use of an average sap flux measurement from all sensors. As suggested below by the Reviewer, we will show the range of measured variability in the revision (suggested figure shown below, next to the recommendation).

Urban karst

Along with the discussion, the authors mention the term "urban karst" given by Bonneau et al. (2017) which affects the water fluxes and redistribution by the preferential flow. Taking into account the heterogeneity of the subsurface on the sampling area (Lines 134-135) where the "subsurface is heavily impacted by human activities, and in places has an added layer of up to 50–180 cm of debris", how does the preferential path flow form by these debris affects or potentially affects the soil water age estimations?

Our comments were perhaps misleading here. The study site includes significant areas of made ground with some subsurface building rubble, though the upper 1-2 m of soil is generally now quite well developed. We will include a comment in the revision as to how this is likely to affect subsurface drainage and water ages.

Tree with run-on

How does the "tree with run on water" compare against the transpiration of the tree(s) sap flow measured in those pixels? This will support the affirmation given by the authors about the model performance and side effects of nearby impermeable land covers (e.g, buildings, pathways). Unfortunately, sap flow was not measured in the pixels where we explored effects of tree with run-on, so we are unable to do this.

Minor Comments

Does the soil water content measurements were calibrated with soil samples along the sampling period? Yes, the TDR probes were calibrated to local data at installation.

The authors mention the use of a German Weather Station that "records essentially the same rainfall" (Line 156). Can the authors provide the values? Yes, we will clarify on revision.

What are the urban tree species sampled for this manuscript (Lines 159-161)? Can the authors provide more information about the individual trees sampled (e.g, diameter, species, height, etc)? The tree species sampled were maple, elm, plane and oak, we will clarify on revision. Full info can be found in the now published open access article Kuhlemann et al., 2021, *HESS* (in the submitted manuscript this was referred as Kuhlemann et al. (2020b), since then the preprint has been accepted and published).

The paragraph between lines 167 to 175 describes the results obtained from the data collection described in the previous methodological sections. Consequently, this should be in Results and not in Methods and Material. Again, the collected data are part of a related data-driven study (Kuhlemann et al. 2021, *HESS*) and hence need to be presented as background context for this modelling study, rather than new results.

The authors mention the use of Nash-Sutcliffe efficiency (NSE) as objective functions (Line 216). However, across the manuscript, there is only one reference to NSE in a broader context (Line 234) with no reference to the results of this analysis and neither in the supplemental material. The authors only mention Kling-Gupta Efficiency in detail (e.g, Line 272, 282). What happened with the NSE analysis? We used NSE for screening early runs but for later stages KGE was deemed better suited for evaluating model performance of soil water content, as there is less primacy on the simulation of peak values alone. We will add this information to method section.

- The authors should follow the recommendations given by Knoben *et al.* (2019) when using Kling-Gupta Efficiency analysis in models. During the modelling the aim was to maximise KGE, no benchmark was explicitly set.

- The authors should add more information about the results using the NSE. Also, adding the respective equations as for KGE. See above for comment on NSE.

- It is necessary to add more information about the sampling processing (precipitation and soil samples). The following questions must be answered:

* How were soil samples collected?

* When collected, how long after a rain event the soil samples were taken?

* Which soil water extraction procedure was applied?

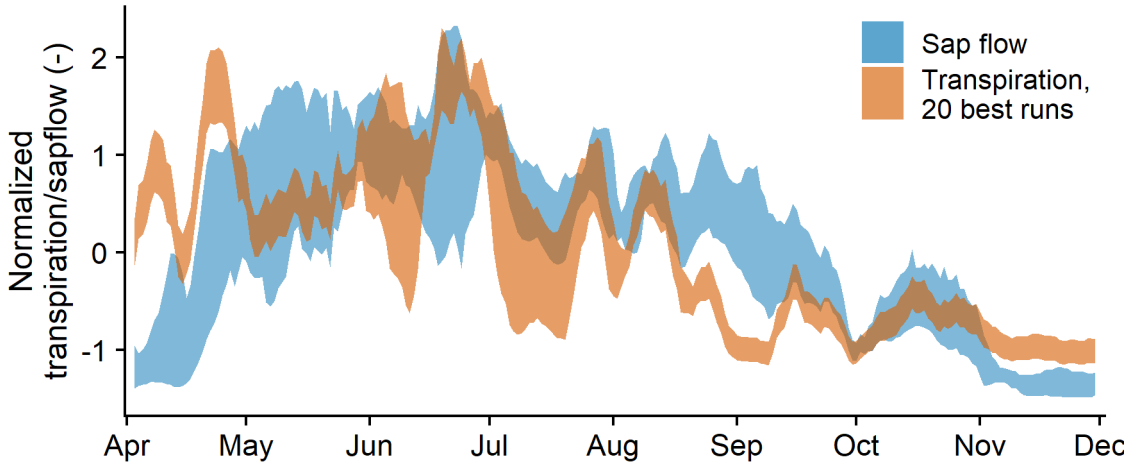
* What method/equipment/laboratory performed the stable water isotope analysis?

We will add more details in the supplementary material to make our paper more standalone. However, full details are available with in Kuhlemann et al. (2021), *HESS*.

Recommendations

The authors can use boxplots in "Figure 10: Soil layers" instead of bars. In this way, the reader can have a better idea of the data distribution for each layer/cover. We will consider this suggestion for the revision, but we note Reviewer #1 liked the summary plot as it stands.

The authors can add the transpiration envelop in Figure 7. This will allow the readers to have a better notion of the temporal flux variability. We will add an envelope for the measured sap flow, ranging from min to max values for each day (see below).



References

Bonneau, J., Fletcher, T. D., Costelloe, J. F. and Burns, M. J.: Stormwater infiltration and the 'urban karst' – A review, J.Hydrol., 552, 141–150, 10.1016/j.jhydrol.2017.06.043, 2017.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrol. Earth Syst. Sci., 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Kuhlemann, L.-M., Tetzlaff, D., Smith, A., Kleinschmit, B., and Soulsby, C.: Using soil water isotopes to infer the influence of contrasting urban green space on ecohydrological partitioning, Hydrol. Earth Syst. Sci., 25, 927–943, https://doi.org/10.5194/hess-25-927-2021, 2021.