Responses to specific comments of Reviewer #1

I do not recommend using the word "loss". In hydrology there is no loss. Use interception evaporation instead of interception loss, as is done in line 16 of p8. But do it throughout.

Agreed. We will replace the three occurrences of "interception loss" with "interception evaporation".

Please don't use the concocted word evapotranspiration. You managed to avoid it almost everywhere and correctly used the term "total evaporation" or just "evaporation" instead. But it still remains in a few places: line 12 of p8, line 18 of p.9, line 13 of p22.

We also agree with this comment and will replace "evapotranspiration" with "evaporation".

Responses to specific comments of Reviewer #2

Abstract: "the concept of a single rooting zone storage capacity was appropriate at most temperate and cold sites" This conclusion seems too strong/general. Can e.g., parametrisation, data uncertainty, or model structures not be the reason given the research design and the scope of the performed analyses?

Indeed, this statement is outside the scope of the research question, and not necessarily supported by the analysis.

Abstract: "mismatched were attributed to...[]...oxygen stress and low soil temperature". It is not clear to me how the attribution was made. Please consider providing searchable key words that make it easier to locate the related analyses. (I searched for "oxygen" and "attrib" without finding any related analyses).

The term "attribution" might be misleading here – this statement refers to a possible (but untested) explanation for the mismatch at high-elevation spruce sites. This is perhaps given too much weight in the abstract. In a revised version, we will simply state that in some situations, factors other than carbon uptake may control rooting depth, such as cavitation risk, low soil temperature or oxygen stress.

Abstract: "Nevertheless, the overall good agreement suggests that this model may be useful for generating estimates of rooting zone storage capacity for both hydrological and ecological applications. Another potential use is the dynamic parameterization of the rooting zone in process-based models, which greatly increases the reliability of transient climate-impact assessment studies." These are not key conclusions from the study, and rather speculative. I would suggest removing these statements.

As we write in our comment AC1 (Response to reviewer #1), assessing the potential of the G10 method for implementation in a dynamic model was our primary motivation for this study. As we have reformulated our research goal and questions, it makes more sense to mention this at the beginning of the abstract. In our revised research questions

(see AC1), we state the criteria by which we assess the suitability of the G10 model for this purpose. Therefore, the sentences quoted above can be removed, and replaced by more specific statements on the research questions.

P3L4-5: "Yang et al. (2016) identified the approach proposed by Guswa (2008) as the most meaningful from a hydrological and ecological point of view." This sentence suggests that Yang et al (2016) made a comparison between all aforementioned approaches, which was not the case. The word meaningful is also vague – do you for example mean that this approach yields best performance in both hydrological and ecological modelling or that their approach captures the most major hydrological and ecological drivers of Ze?

We agree- this formulation is potentially misleading. We will adapt it accordingly. The word "meaningful" (in the second sense) provides a link to the points discussed in RC1 and AC1 (comparison with data-driven approaches)

P12-Table3: "LAI". Do you mean "maximum LAI"? Where is it described how LAI is varied?

This is indeed the maximum LAI. We assume no inter-annual variation of maximum LAI. The description of intra-annual variations of LAI in deciduous and mixed stands is indeed incomplete, and we will add it in a revised version.

P18-Fig5 caption: "There is a relatively narrow range of Sr leading to Paretooptimal scores". The black Sr dots appear to range between approx. 50 and 280 mm. I would be hesitant to refer to this as a narrow range.

Agreed.

P18-Fig5 caption: "conducted using the optimal parameter sets" Please be specific and add cross reference. It is not entirely clear which optimal parameter set is considered. The suggested overview table (see General comments) of simulation settings/parameter combinations would be helpful to cross refer to.

Agreed – we will include a table and reformulate this sentence accordingly.

P18 Fig 5 (and SI figures): Please consider changing the line color and style. At first sight, one might think that the black colors share some common point, which is not the case.

We agree that this will improve the legibility of these figures.

P21-Fig 7: Possibly consider collapsing the two subplot columns G08 and G10 into one column, and use colour coding or other visual cues for identifying the model approach used.

This could also be helpful. Furthermore, as we will replace the rough parameter sensitivity analysis with a more formal uncertainty analysis (see AC2), the two bottom panels can be removed.

P23: The cross reference to Fig 4 seems to be wrong.

Yes, the correct reference is Fig. 7.

P25: "The results of this study suggest that G10 better captures the behavior of forests under energy-limited conditions". Please consider to add a cross reference. I have difficulties understanding how the analyses and results support this statement.

This should refer to the discussion in the first paragraph of Section 4.2. However, instead of a cross-reference, it is perhaps preferable to merge these two paragraphs (first paragraph of Sect 4.2 and last paragraph of Sect 4.3).

P25L17: "suggesting that the use of a bulk Sr is inappropriate at these locations". I struggle to understand how this claim is supported by the performed analyses. In my view, to be able to make such as claim would require a comparison between a model structure with bulk Sr and a model structure without bulk Sr (e.g., some other structure hypothesised for Mediterranean conditions), and this comparison would need to show that the model structure without bulk Sr performs better than the other one. It seems to me that current analyses only suggest that FORHYTM as a whole does not appear appropriate for modelling evaporation in Mediterranean conditions.

We suggest replacing this statement with the following :

"By contrast, FORHYTM failed to reproduce local water balance properly under Mediterranean climates and on dune soils. This raises the question whether the use of a bulk Sr is appropriate at these locations.

FORHYTM combines three distinct sub-models: the energy partitioning scheme of Guan and Wilson (2009), the Jarvis-type model of canopy resistance, and the soil water balance routine of the hydrological model HBV (Bergström, 1992). Energy partitioning is physically quite well constrained, and the scheme of Guan and Wilson (2009) has been tested under various climates (Lu et al., 2014). On the other hand, previous studies suggest that the two other sub-models may face severe limitations under Mediterranean conditions. For example, Poyatos et al. (2007) calibrated a stand-level evaporation model, including a Jarvis-type parameterization of canopy conductance, in a sub-Mediterranean *Pinus sylvestris* forest. Despite satisfactory calibration efficiency, the model performed poorly during the calibration period. The authors explained this with variations in hydraulic conductance, possibly due to xylem embolism. Recently, Bai et al. (2017) compared different Penman-Monteith based water balance models at Mediterranean eddy covariance sites. Models with a multilayer soil representation performed better than single-layer models. Therefore, under Mediterranean conditions, transpiration may be more sensitive to the vertical distribution of soil moisture and roots. While it is not possible to determine to what extent the canopy resistance or soil water balance submodels contributed to the poor performance of FORHYTM at Mediterranean sites, it is likely that a multi-layer soil model is more appropriate at these sites."

To avoid redundancy, the first paragraph of Section 4.1, where the Bai et al. (2017) article is discussed, should be simplified accordingly.

Additional references

- Lu, H., Liu, T., Yang, Y., Yao, D., 2014. A Hybrid Dual-Source Model of Estimating Evapotranspiration over Different Ecosystems and Implications for Satellite-Based Approaches. Remote Sens. 6, 8359–8386. https://doi.org/10.3390/rs6098359
- Poyatos, R., Villagarcía, L., Domingo, F., Piñol, J., Llorens, P., 2007. Modelling evapotranspiration in a Scots pine stand under Mediterranean mountain climate using the GLUE methodology. Agricultural and Forest Meteorology 146, 13–28. https://doi.org/10.1016/j.agrformet.2007.05.003