Referee 3 (Andrew Guswa)

'In this work, the authors compare results from the HTESSEL land-surface model with vegetation root depth determined in two different ways. The control case (CTR) comprises root depths (and, correspondingly, root-zone storage capacity, Sr) determined via soil depth; in another formulation (MD), Sr is determined via the memory method, which is based on the concept that plant roots adjust to mitigate droughts with certain return periods. These models are implemented for 15 catchments in Australia with tropical, temperate, and Mediterranean climates.

Results reveal that the root-zone storage capacities determined via the memory method are shallower and more variable across the watersheds than those from the control cases. This is in contrast to what others have suggested – that root depth in LSMs may be too small. The changes in Sr (from the CRT to MD approach) manifest as improvements in the variability and seasonality of streamflow. Long-term water balances and ET are relatively insensitive to changes in Sr. The paper is well-written, and the methods and results are well-explained and valuable. I offer a few suggestions and comments below, and I recommend publication with minor revision.'

We would like thank dr. Guswa for the comments. We appreciate the time and effort taken to read our manuscript in detail and to provide us the very useful and interesting thoughts on our research. We will take the comments into account in revising the manuscript.

We have separated the different comments (shown in *italic*) and have written our replies below. Text in the original manuscript is shown in *'italic'* and revised text in '***bold***'. Wherever line numbers are mentioned in our reply they refer to the original manuscript version.

Comment 3.1

Low sensitivity of ET

'The authors note that model estimates of ET do not appear very sensitive to the differences in Sr. I agree with Hageman that the low sensitivity may be due in part to the fact that the values of Sr,MD are less than Sr,CTR. The catchments in question are arid/water-limited, and ET dominates the water balance, with annual ET being 4-16 times the annual streamflow. One of the proclaimed advantages of the memory method is that it facilitates deep and/or expansive roots so as to maintain ET during periods of drought. In this work, it seems that the soil-based root depths (Sr,CTR) were already sufficiently large, such that the limits of the root zone storage were not being approached. An analogy would be a water-supply reservoir that never fills – if you make it a little smaller,you do not affect the available supply, since you were starting with excess capacity. The authors may wish to expand their discussion in the paper along these lines.'

We investigated this issue (also based on the comments 2.1 and 2.2 by dr. Hagemann). Indeed, the limited sensitivity of evaporation to changes in soil depth can be partly explained by $S_{\mathrm{r,CTR}}$ being larger than $S_{\mathrm{r,MD}}$. It was found that modelled evaporation is sensitive during dry periods only, when the MD soil moisture in layer 4 is smaller than CTR. However, most of the time soil moisture is in a wet, insensitive regime, and the evaporation does not approach the limits of soil moisture storage. This is illustrated by Figure C3, that shows modelled transpiration and soil moisture in the four separate soil layers in a tropical (a), temperate (b) and Mediterranean (c) catchment. In order to further explain the evaporation (in)sensitivity we will consider a wet period (mid 1990) and a dry period (start 1991) in the temperate catchment (Fig. C3b).

During the wet periods soil moisture in the upper three layers is above or close to θ_{cap} and little differences are observed between CTR and MD. However, in layer 4, MD soil moisture is larger than CTR soil moisture. In this case evaporation is not moisture limited and is controlled by the top three layers because of the larger root distribution in these layers (eq. 14 and 15). Therefore, the modelled transpiration is not sensitive to the increase in layer 4 soil moisture in MD.

During the transition from a wet to a dry period, the upper three layers dry out first, as there is a reduction in precipitation input. As these layers are dry, evaporation is controlled by the fourth layer. Layer 4 soil moisture in MD also reduces to values close to θ_{pwp}, while in CTR layer 4 remains wet. This difference causes the sensitivity of transpiration in MD during this wet to dry transition.

Most of the time the modelled soil moisture is in the wet and insensitive regime, and, therefore, the overall effects of MD on modelled evaporation are small.

We will add this explanation and Figure C3 to the revised manuscript in the discussion 4.1.

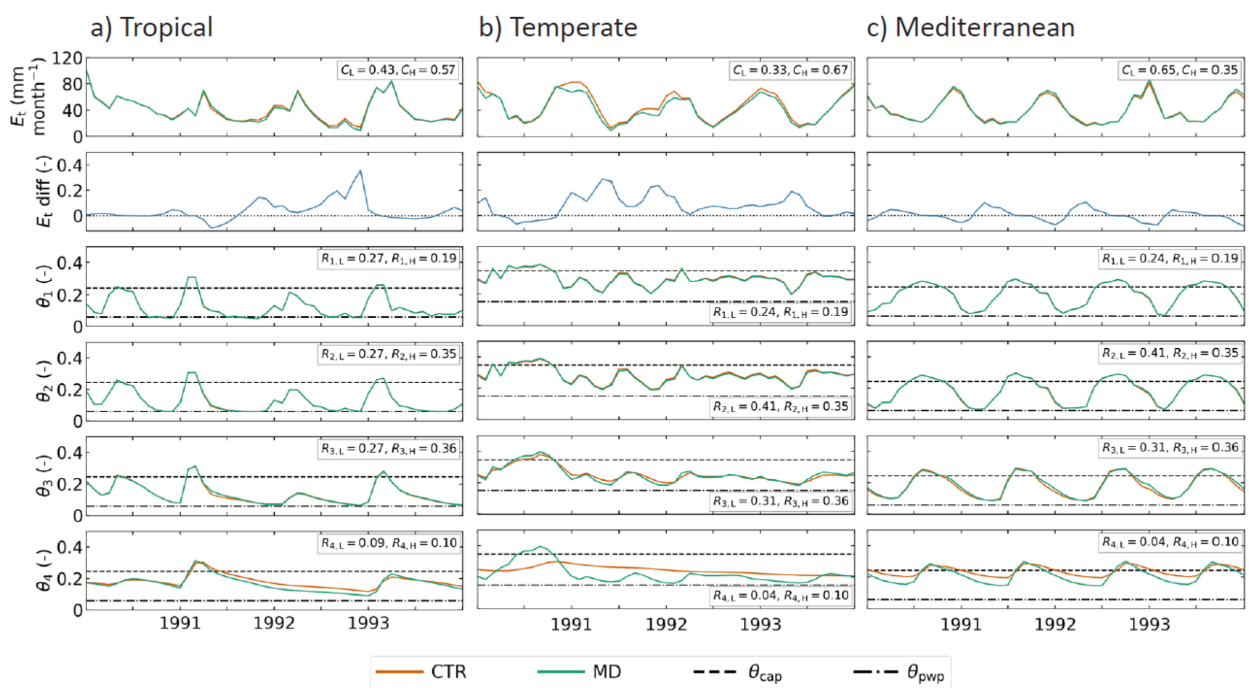

Figure C3 (new Figure 8). Modelled transpiration and soil moisture with CTR and MD models in a (a) tropical (Mi), (b) temperate (Na) and (c) Mediterranean (K) catchment. From top to bottom: transpiration, relative difference between CTR and MD transpiration $\left(\frac{E_{t,CTR} - E_{t,MD}}{E_{t,CTR}}\right)$*, soil moisture layer 1, soil moisture layer 2, soil moisture layer 3, soil moisture layer 4. Additionally, the vegetation coverage (*C_L *and* C_H*) and the relative rooting distribution (*R_k*) for the dominant high and low vegetation types are presented.*

Comment 3.2

'Additionally, the authors may also wish to acknowledge (again) the uncertainty in the ETobservations. As they point out in lines 110-115, the estimates from the water balances are lower than the estimates from the FLUXCOM data by 20%, and it may be worth reminding the reader of this in the discussion.'

We will add the following sentence on the uncertainty of the E_t observations in lines 321 to remind the reader of this:

'… E was very limited in all climate regions (Table 3; Table 4; Table S6; Fig. S2). ***As stated before, the reliability of the FLUXCOM E is questionable in our study catchments (Fig. 3). Although the model performance with respect to E fluxes is uncertain,*** *the lack of evaporation sensitivity to S_r was unexpected and requires more in depth evaluation of HTESSEL.'*

Comment 3.3

'Lines 110-121 and Figure 1: The authors may wish to address whether the lower estimates of ET via the water balance method might be affected by or attributable to deep groundwater drainage that does not get recorded by the stream gauges.'

We assumed that the catchments are large enough that deep groundwater drainage does not play a major role in the catchment water balance. Therefore, this would not explain the lower estimates of E-WB compared to E-FLUXCOM.

We will change line 115 as follows:

'… an integrated catchment scale estimate as it is derived from observations of Q ***assuming that the catchments are large enough to neglect deep groundwater drainage to or from other catchments****.'*

Comment 3.4

'Lines 155-eq 7: Like the other reviewers, I am a bit confused by the approximation of "Catchment Sr,MM" by a weighted average of the Sr values for trees and grasses. As drainage below the root zone is a non-linear process, this will affect the results (e.g., drainage below the grass portion may be non-zero, whereas drainage below an "average" root depth may be zero). I understand that the model is limited in its resolution, and one has to make some concessions. It may be worth an expanded comment, however.'

The memory method allows us to make a separation in S_r for high and low vegetation. We decided to combine S_r values for high and low vegetation to get one catchment representative S_r and consequently one model soil depth in MD. This approach was chosen because the CTR model parameterization does not allow to change soil depths differently for different vegetation types, as HTESSEL has one soil discretization for the entire grid cell (Fig. 4b). However, we acknowledge that we could get a separation between high and low vegetation S_r by modification of the root distribution separately for high and low vegetation, as mentioned in the reply to referee 1 (comment 1.3). However, in this study we did not want to change multiple model parameters at the same time, to avoid difficulties in identifying the differences between CTR and MD model output. Modification of both rooting distribution and soil layer depth was therefore not desired.

L 155: '...maximum storage deficits. ***Theoretically we could treat S_r separately for high and low vegetation in HTESSEL, however, this would require changing the root distributions (see section 2.4), which we decided not to do as we did not want to change multiple parameters at the same time.'***

Comment 3.5

'Figure 6 is a nice figure with a high information density. Like reviewer 1, however, I found the differences in y-axis scales to make interpretation challenging, e.g., the differences in the Temperate and Mediterranean Q appear (visually) to be much greater than those in E. Perhaps it would be worth showing the six subplots, first with all the same y-scale, and to then provide second versions of 6b and 6c that are more zoomed in. I think seeing how very small the runoff is from the Temperate site, before diving into the differences among the models and observations, would help the reader.'

For visibility of Fig. 6 we will follow your suggestion to use the same y-axis scale in all subplots, and additionally provide 'zoomed' versions of Fig. 6b and 6c. Figure C2 is the new version of Fig. 6. (see also comment 1.9 by referee 1)

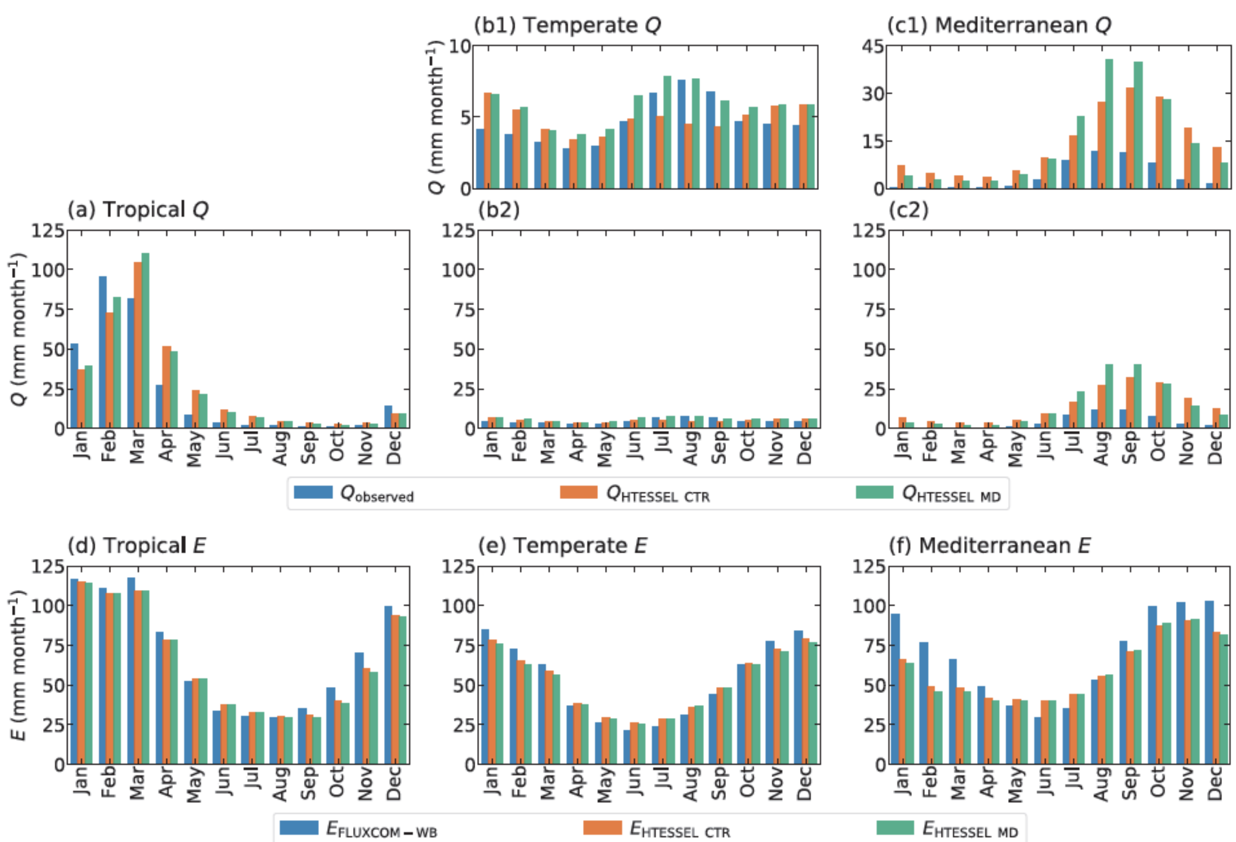

Figure C1 (new Figure 6). Monthly seasonal climatology of observed discharge (Q) (top) and FLUXCOM-WB evaporation ($E_{FLUXCOM-WB}$) (bottom) and modeled values in the HTESSEL CTR and MD versions, averaged for the tropical (a, d), temperate (b, e) and Mediterranean (c, f) catchments for the time series 1975-2010. b1 and c1 represent the same data as b2 and c2, but with a different y-axis. Similar figures for the individual catchments are shown in Fig. S1 (Q) and Fig. S2 (E).

Comment 3.6

‘Lines 298-302: The authors make a point to acknowledge that Sr need not be synonymous with root depth, which is fair enough. In this work, however, it IS synonymous, and I found this introduction to the discussion a bit odd. I recommend that the second sentence of the first paragraph be dropped.’

It is correct that in HTESSEL the root depth is synonymous with the soil depth, as roots are present over the four model soil layers. However, it should be noted that in S_{r} represents a conceptual water volume in the model world, without the assumption where this volume is in reality. The model soil depth and root depth are model parameters that are required to schematize the S_{r} in the model, but are also not found in nature in the way they are schematized in the model.

We will change line 298-302 as follows:

‘… in simulated evaporation (Kleidon and Heimann, 1998; Pan et al., 2020; Sakschewski et al., 2020). ***However, S_r represents a conceptual water volume that is accessible to roots, without the assumption where this volume is in reality.*** *Therefore, S_r is not necessarily proportional to root depth as a small S_r does not preclude the presence of deep roots, as illustrated in Fig. 4 in Singh et al. (2020).’*