Major comments

We would like to thank the referee for the many constructive comments on our manuscript, which will help us to significantly improve the paper.

This is an interesting and very well written paper, that introduces an novel approach, that of transferring intact lysimeters from a wet cool climate to a much drier and somewhat warmer climate in Germany, to study the effects of climate (change) on the water balance and growth of grass vegetation. The paper has long introduction and materials and methods sections that already reveal a substantial amount of results and bring in a lot of literature references. At times it feels more like a review, and some of these references could perhaps have been saved for the discussion. The methods used are overall sound, although it was not fully clear to me why they had not selected a more recent and complete grass hydrology/growth model.

We had a rather specific purpose with the modelling. We built a relatively simple model to try to understand the reasons for differences in the water balance and grassland growth found at the two sites. Existing crop models include many other processes that are not relevant to our question, and using them would also have been much less transparent.

We will make this objective clearer in the revised version.

Why was interception not included? This would not have caused much computational burden.

It didn't seem to be necessary to account for interception. As noted in the paper, the net evaporative loss caused by interception is usually quite small for short vegetation, since transpiration is reduced by a nearly equivalent amount.

What frustrated me a little was the fact the Results and Discussion section is rather short. After all this pre-amble on how the model works, the reader is still left with quite a few questions.

For example, what are the reasons for the upward flow in the drier lysimeters, even though LAI and Dry Matter are lower? Is this real? Is it the deeper root depth for these lysimeters that could have caused it. I would like to see the authors elaborate a bit more hear, for example by going back to the findings of their sensitivity analyses.

Yes, the net upward flow at Selhausen is definitely real (it has been measured). Yes, the deeper roots may contribute to this, together with the hydrogeological setting and the topographical position of the site on a low-lying flood plain. It seems likely that lateral groundwater flow from the surrounding higher land maintains the supply of water to the drying root zone.

We will add some text to elaborate on this.

Minor comments

Line 195-196: “It is also striking that the actual evapotranspiration slightly exceeds precipitation at Selhausen, so that the net percolation at the base of the lysimeters is negative (i.e. an upwards directed flow; Table 2).”.... Does this mean that the percolation was not measured separately? You say that the lysimeters enable the measurement of a complete (closed) water (line 113)? What is causing this negative percolation (capillary rise?).

Percolation was measured separately. We will make this clearer in the revised version (at line 173). The reasons for the negative percolation are mentioned in our response above under “Major comments”. We will elaborate on this in the revised version.

Chapter 2, Materials and Methods already contains a lot of results on the lysimeter water balance, AG biomass etc. Should that be moved to the results?

We prefer to keep this in the M&M section, as it sets the scene for the model development (why we model)

Line 214: water use efficiency (WUE) of the grassland in the drier climate was lower than that of the wet climate. Would we have expected the opposite (plants becoming more efficient under dry conditions)? So, is it really the leaf-level WUE hat has changed, or is this the result of other factors?

Yes, a lot depends on how WUE is defined. Note that we defined it from the data we have as the harvested yield divided by total evapotranspiration. The modelling that is described later explains the decrease in WUE as a consequence of an increased allocation of assimilates to the root system and a reduction in above-ground growth. So, this is not leaf-level WUE.

It also seems to depend on the crop in question. We identified for arable lysimeters with a slightly different definition of WUE, that WUE increases after a move to a drier climate (see Groh et al. 2021: https://hess.copernicus.org/articles/24/1211/2020/

How exactly is infiltration calculated in the model?

This is explained at lines 292-297

And the flow at the bottom boundary of the soil profile?

This is explained at lines 290-292

What about Runoff?

There was no surface runoff, as the soil infiltration capacity was never exceeded. We will add a sentence at line 297 to explain this.

In Line 424 and further you talk about feedbacks from the plant growth model to the hydrological model. You mention the effect of LAI and height on the aerodynamic resistances and hence on the ET fluxes. Surely LAI also affects radiation extinction, and therefore the energy available for ET. This could be mentioned too?

Yes, this is correct. We will mention this in the revised version

Also, the aerodynamic effects will have been relatively low. From that point of view would it not have been better to also consider interception (as it is also affected by LAI)? Although the values for grass are low, they are comparable to winter values of ET, and the equations required are straightforward?

It didn't seem to be necessary to account for interception. As the referee notes here, the net evaporative loss in interception is usually quite small for short vegetation.

Lines 489-492 You say: "The measurements from the matric potential sensors installed in the uppermost soil horizon (0-24 cm depth) appeared to be unreliable. We therefore also used the HYPRES pedotransfer functions to estimate the shape parameter n in the topsoil, while α was set equal to the same value as the deeper horizons. First of all: why where these measurements unreliable? Was the soil too dry?

The pressure potential at 10 cm depth was measured by an MPS 1 sensor and not as in the other depths, by tensiometer. The MPS 1 sensor is relatively good at measuring pressure heads in dry soil, but in comparison with classical tensiometers, it is known to give unreliable values for the range between -200 cm until saturation. The problem here is that this range of pressure heads is important for the definition of the Mualem-van Genuchten parameters (especially theta s and alpha). Thus, in this study we didn´t used data from the MPS 1 sensor at 10 cm depth, and relied on the available tensiometer data.

How do you know that the deeper sensors could be deemed reliable?

This type of sensor was not installed in the deeper layers (only in the topsoil)

Also, can I ask why you did not use the VG parameters for the medium-layers where you did have measurements, instead of having to revert to generic HYPRES PTFs? Were these horizons too different?

Yes, that's right. There are large differences in texture between horizons (see lines 141-143)

Your alpha parameter in Table 3 appears to be the same throughout the entire profile, yet n varies considerably. Is this realistic?

Yes, it appears to be reasonably realistic (see figure S3 and table S1)

How come that theta_s in the first soil layer is so much higher in Rollesbroich?

Yes, this is interesting. This could be the result of chance spatial variation. But it is also possible that the physical properties of the uppermost layers of the Rollesbroich soil have changed following the move of the lysimeters to Selhausen. One plausible explanation is that the drier soil conditions at Selhausen have led to increased mineralization rates of soil organic matter, leading to a decline in soil organic matter content and consequently, an increase in bulk density (i.e. a loss of porosity). It may also be the case that the drier surface soil conditions at Selhausen have reduced soil wettability.

We will add some text in the revised version on this.

In Table 5 you talk about post-priori parameter ranges, whereas in the text (line 579-580) you mention that the posterior uncertainty ranges are much smaller than the prior uncertainty ranges. In figure 3 you talk about posterior distributions of the four parameters. Where exactly are you showing the prior uncertainty ranges? I find this all a little confusing.

We apologize for the confusion. We will change post-priori in table 5 to posterior.

The prior uncertainty ranges are shown in the first column of table 5.

Lines 595-596: You say: “The simulations suggest that the maximum root depth at Selhausen has increased to ca. 80 cm,while the maximum stomatal conductance has roughly doubled”. Were you not able to measure the root depth? I guess this would have caused destruction of the lysimeter core.

That’s right. We would need to destroy the lysimeters to measure the root depth. This would not be in agreement with the goals of TERENO-SOILCan to provide long-term observations. However, we did observe the initial root depth at the time of sampling (see lines 147-149)

Also, you say that stomatal conductance has doubled, but could it be that the parameter had assimilated aerodynamic effects to changes in vegetation structure? You hint at this perhaps in the following sentences, but it is not clear.

This seems unlikely. Aerodynamic resistance is calculated from plant height, which in turn is estimated from LAI. No differences in the relationships between plant height (and above-ground biomass) and LAI are apparent at the two sites (see figure S6), which suggests that the vegetation structure is similar.

In Figure 4 you need to make it clear which set of 3 graphs is representing which site (the same goes for Fig. 5). It is hidden in the legend but should be more explicit.

Yes, we will do this

Also, based on these plots it is surprising that you opted for a theta_s of 0.55 for Selhausen (estimated “by eye”). I understand that these are daily averages (?), but during the winter months there much have been saturated conditions?

This is a misunderstanding. Theta_s was 0.45 at Selhausen not 0.55 (see table 3)

Water contents do get close to saturation during winter, yes (see figure 4)

In Figure 5 it is quite hard to see what is going on. Would it be possible to separate the years somehow? Or perhaps make cumulative plots?

Yes, we can include cumulative plots (in supplementary)

Line 611-612: You say that the “model performs very well, matching the temporal dynamics in the high-time resolution data on state variables and fluxes as well as reproducing the differences in the overall water balances at the two sites”. I am not sure that Table 6 reflects that statement? While model efficiencies are high the values for ET are much lower and values for LAI are negative (with LAI being a crucial variable in many equations), which makes me wonder whether Figure 5 hides a multitude of sins..?

This statement only refers to the hydrological simulations (not the simulation of LAI). Figures 4 to 6 and table 6 undeniably show that the statement is justified.

I find the discussion around Fig. 7 somewhat incomplete. While soil evaporation clearly depends on soil moisture content, windspeed etc. it also depends on incoming radiation (which would have been lower in Ro and higher in Se), and radiation reaching the soil through the canopy. Seeing the DM was lower in Se (and therefore LAI, see also figure 9) I would not necessarily have expected soil evaporation to have been lower for the Se site.

Yes, it is interesting that evaporation was apparently smaller at Selhausen despite higher radiation inputs. We will add some text on this point

Line 628: You say: At Rollesbroich” grassland is harvested 3-4 times during the growing season” Was this not the case at the Selhausen site? This makes comparison between the two sites difficult?

The grassland is also cut at Selhausen 3-4 times per season (see lines 150-153)

The discussion around water stress, with a focus on Figure 8, seems to ignore the fact that one of your earlier figures indicated that the rooting depth at Selhausen was much deeper (80 cm) than at Rollesbroich. Is that mot the main reason for the relatively modest water stress experienced at this site?

Yes, this is good point. We will add this to the discussion

Line 692-693: Are some of these ME values less than excellent if the best value is 1? Can you provide a scale for what consititues poor, good, excellent etc. in the methods where you introduce ME?

Yes, we agree that this description is not really warranted. We will change “excellent” to “satisfactory”. With ME, there is only one objective cut-off: an ME value of less than zero defines poor simulations, see line 559.