Response to reviewer #2

General comments

The authors used the STEMMUS-SCOPE Model to analyze how the revegetation of shrubs in a desert steppe modifies the water, energy and GPP dynamics during wet and dry years. Since the paper is not a model development study, I will focus on its presentation aspects. Nonetheless, I think the paper is quite straightforward, but can be improved by taking care of my suggestions.

Thanks for the kind words! Your comments are immensely valuable in enhancing the manuscript's quality.

The title: Since the only carbon flux analyzed is GPP, I recommend the title be changed to "Understanding the Effect of Revegetated Shrubs on fluxes of Energy, Water and Gross Primary productivity in a Desert Steppe Ecosystem Using STEMMUS-SCOPE Model". Additionally, if the EC tower measured CO_2, how about also showing the net ecosystem exchange as one indicating effect of different vegetation? Or it is not possible to differentiate the vegetation effect from NEE?

We agree with the title modification because we only presented the simulation of Gross Primary Productivity (GPP). It's true that the EC tower measured Net Ecosystem Exchange (NEE) and the GPP was partitioned from the measured NEE. In contrast, the model usually calculates photosynthesis (i.e., GPP) first and then determines the Net Ecosystem Exchange (NEE) by subtracting ecosystem respiration (Reco) from GPP. In STEMMUS-SCOPE, the Reco is calculated using a simple linear equation, Reco = 2.5 + 0.054375*$T_{surface}$, where $T_{surface}$ represents the surface temperature of leaves or soil. To avoid error propagation from the calculation of Reco in the model, we decided to directly compare GPP, making the comparison more direct and accurate.

Specific comments

Comment 1: Equation (1), what does C_shrub mean?

Response 1: Thank you for your attentive review. We have added an explanation of this variable in the revised manuscript.

$$GPP = C_{shrub}\,GPP_{shrub} + (1 - C_{shrub})\,GPP_{grass} \quad (1)$$

C_{shrub} is the contribution of shrubland to the overall flux, and its value is 58.33 % while the contribution of grassland is $(1 - C_{shrub})$ = 41.67 % (Table 1, Line 125).

Comment 2: From the method description, it appears LAI is used as model input. Does this mean the model actually does not simulate the carbon cycle, aka it is a prescribed phenology simulation?

Response 2: The STEMMUS-SCOPE model uses leaf area index (LAI) as an input and focuses on carbon assimilation, excluding carbon allocation. We are aware of this issue and will address it in our upcoming study by coupling the STEMMUS-SCOPE model with the WOFOST crop growth model. This coupling will allow us to simulate LAI and plant growth processes effectively. The evaluation of the coupled STEMMUS-SCOPE-WOFOST model, performed with various plant functional types, showed successful simulation of vegetative dynamics, including LAI and plant height. We intend to submit the study of the STEMMUS-SCOPE-WOFOST model to a journal soon.

However, in this manuscript, our primary focus lies on simulating ecosystem flux, and we use MODIS LAI as a reference due to its availability and relatively high temporal resolution. The MODIS LAI proved more sensitive to dense vegetation and might not be representative in the arid region with sparse vegetation (Fensholt et al., 2004; Fensholt and Sandholt, 2005). Therefore, we have corrected the MODIS LAI using field observations and relevant literature values, in order to ensure this critical input as representative as possible.

Comment 3: Figure 2, what is the uncertainty of the observed LAI? Can you also show it in the Figure?

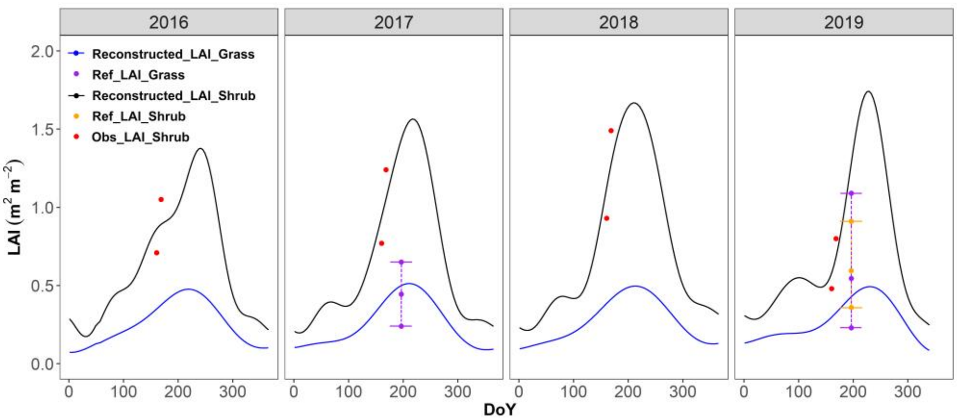


Figure 1. Reconstructed LAI of shrubland and grassland from 2016 to 2019.

Response 3: Apologies for any confusion in the manuscript. The 'observed' LAI values are not subject to uncertainty because they are 'dummy' LAI. The red dots (Obs_LAI_Shrub) are 'observed' LAI of shrubs on DOY 160.5 and DOY 168.5, where we did not have real measurement, but we calculated them using a linear relationship derived from the measurements and MODIS LAI in 2022 (Table S4 in the Supplement). We will modify the caption of Figure 2 to make this clarification.

Comment 4: 2.4.1 Sensitivity analysis, it was mentioned that SA was only done for parameters of the shrubland simulation, why is it not done for grassland as well?

Response 4: In *Section 2.4.1*, we conducted a sensitivity analysis by running 160 sets of parameters for the shrubland simulation, while a fixed parametrization was used for the grassland simulation. We chose to perform the sensitivity analysis only for shrubland due to the following reasons: First, the primary objective of the sensitivity analysis was to evaluate the impact of parameters on the dominant land cover, which in our case is shrubs. Second, running the model for one set of data would take approximately 40 minutes, making it computationally intensive for 160 sets. Therefore, to optimize computational resources and focus on the primary objective, the sensitivity analysis was limited to the shrubland simulation.

Comment 5: 3.2 Model performance, in no place I found model spin up was mentioned. Do the model simulations include model spinup? How is it done?

Response 5: No spin-up was performed in the STEMMUS-SCOPE model simulation. Spin-up is typically used to initialize the model with stable initial conditions. Generally, STEMMUS-SCOPE utilized observed soil moisture and soil temperature as initial conditions, which proved to be effective and stable for initializing the simulation (Wang et al., 2021). Additionally, considering computational efficiency and time constraints, we opted not to conduct a model spin-up for this investigation.

Comment 6: Figure 4c, there are a few points aligning as a vertical line. Does it indicate there are some problems with the data or the model?

Response 6: Figure 4 (c) on the right-hand side compares the observed sensible heat flux (H_obs) and simulated sensible heat flux (H_sim). The points aligning as a vertical line indicate that the observed H were zero but the simulated H were not. A sensible heat flux of zero means no **net** heat transfer through conduction and convection between the Earth's surface and the atmosphere. Moreover, this only happened in the data of 2017 among year 2016-2019.

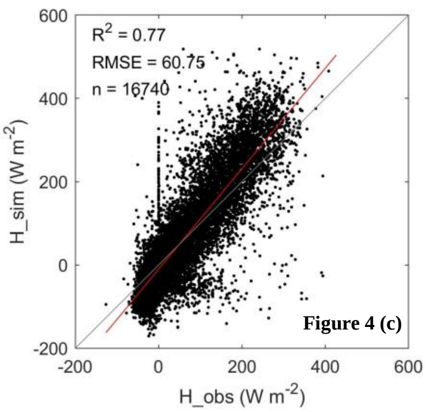


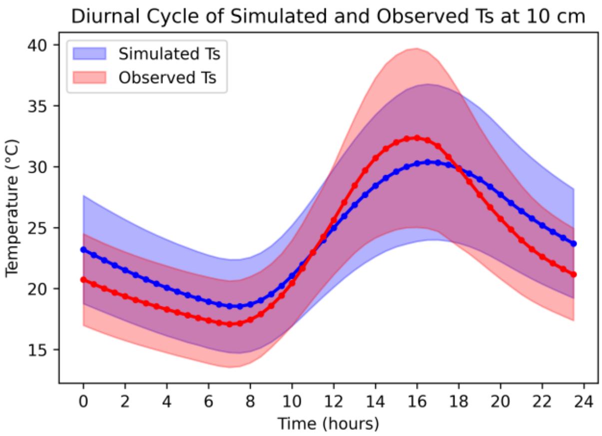
Figure 4 (c)

We acknowledge that there might be issues with the observed data, particularly in 2017, as we did not filter the data in terms of energy balance closure (Line 65 in the Supplement). The lack of surface ground heat flux data in 2017 made it impossible to conduct an energy balance closure assessment and apply the filtering process for that specific year.

Accordingly, we will clarify and modify *4.1 Discussion Line 367* in the revision. **Improved version of Line 367:** (iii) The quality of gap-filled forcing data and in-situ measurements are the basis for a valid comparison between simulations and observations. For example, some observed H were zero while the simulated values were not in 2017 (Figure 4c). Although we diligently conducted quality control for EC data and filtered the fluxes based on energy balance closure (Eq. S10 in Supplement), we were unable to filter the data for the years 2016 and 2017 due to the unavailability of surface ground heat flux for energy balance closure assessment.

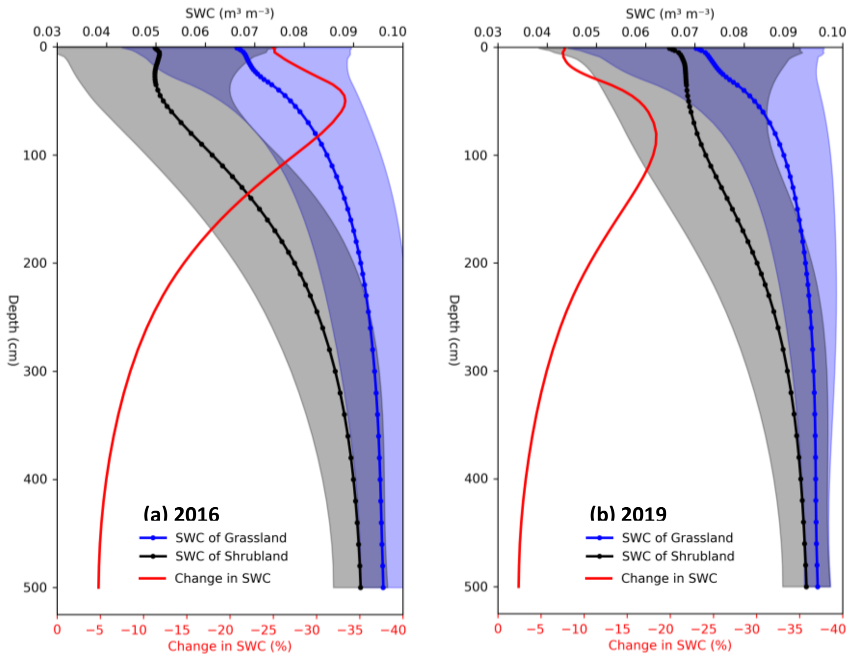
Comment 7: Figure 5c, I don't quite understand the diurnal pattern. If you are average over a certain time period, please show the mean and as well as the variability (i.e., standard deviation).

Response 7: The figure on the right-hand side now includes the variability. The data points represent the average values of every half-hour in a day over a certain time period. For instance, the mean (± standard deviation) of simulated soil temperature under grassland (blue line) was 23.19 ℃ (± 4.43 ℃) at 0 o'clock, spanning from DOY 121 to DOY 273 in the year 2019.



Comment 8: Figure 6, could you also show the variability?

Response 8: The variability is added in the figure below. The soil water content (SWC) is the average value of each soil layer over a certain time period. For example, as shown in Figure (b) below, the mean (± standard deviation) of SWC of grassland was 0.07 $m^3\ m^{-3}$ (±0.03 $m^3\ m^{-3}$) at 1 cm depth, over DOY 121 to DOY 273 in the year 2019.



Comment 9: Section 3.3.3 on GPP, I don't think the interpretation of midday depression is accurate. It may involve factors more than radiation. Perhaps you can show the diurnal pattern of leaf temperature, and temperature dependence of carboxylation as well.

Response 9: Thanks for your hints and please see the diurnal pattern of leaf temperature in the next paragraph. Indeed, the high radiation is more like one of the inducements to the close of stomata, represented as middy depression. Figure S10 in Line 260 of *Supplement* shows midday depression of GPP and stomatal conductance.

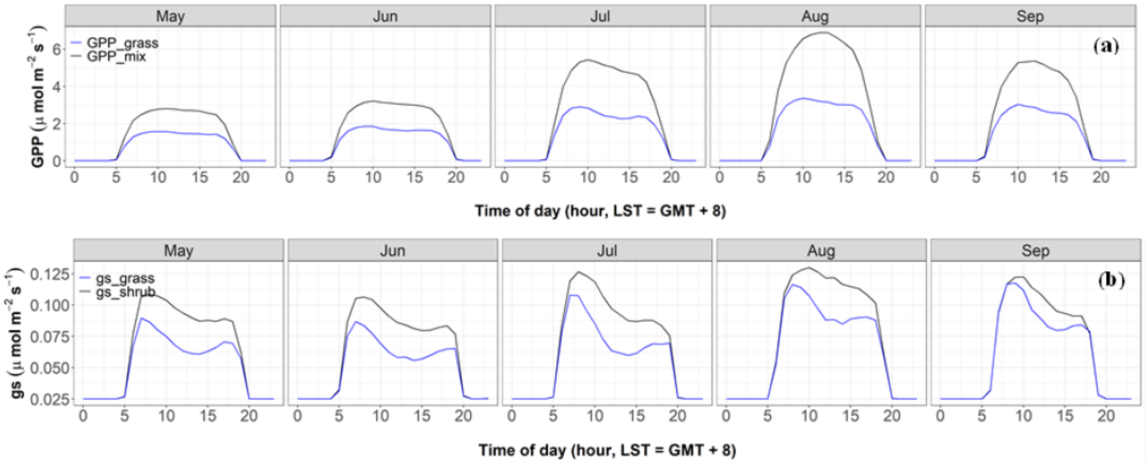
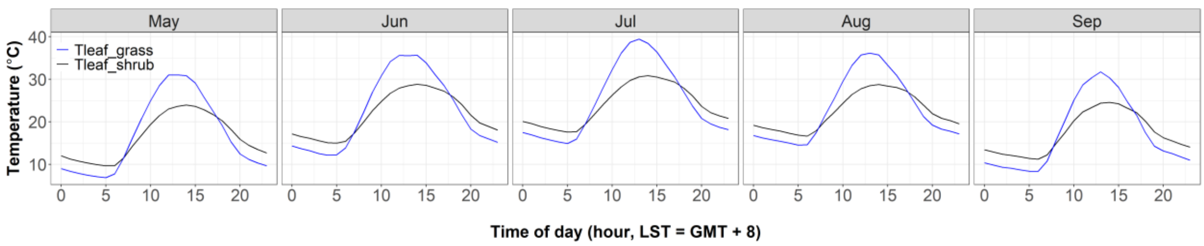


Figure S10. Diurnal courses of simulated (a) Gross Primary Productivity (GPP) and (b) stomatal conductance (gs) of two ecosystems during May–September in 2016 and 2019.

As shown in the figure below, we observed higher midday leaf temperatures in grasses (Tleaf_grass) compared to shrubs (Tleaf_shrub). The increase in leaf temperature could potentially lead to stomatal closure in the plants to regulate water loss, as supported by previous studies (McDowell et al., 2019; Deans et al., 2020; Chen et al., 2014). To enhance the discussion, we will include this figure in Figure S10 and make corresponding clarifications in *Section 3.3.3 On GPP* regarding the impact of leaf temperature in the revision.



We agree that the temperature dependence of carboxylation plays a role because the leaf temperature at the shaded side is used in calculating Vcmax in STEMMUS-SCOPE (Eq. (S23) in the Supplement).

Comment 10: Discussion 4.2.1, paragraph 2, why don't you use the Bowen ratio as an indicator, which is likely more informative than the ratio of LE/Rn here.

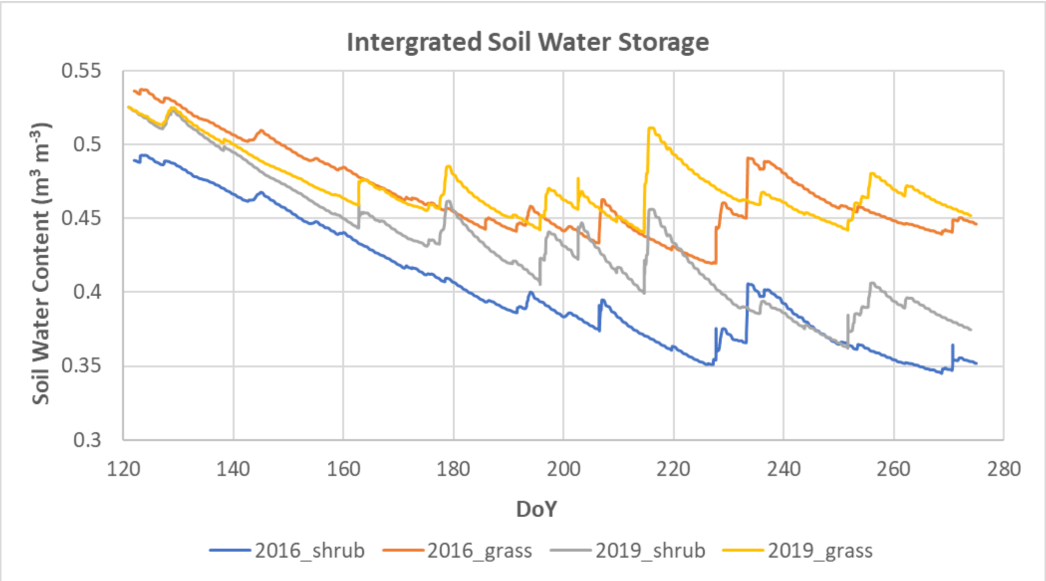
Response 10: We think this is a great idea, and will implement it in the revision.

Comment 11: Section 4.2.3, on water fluxes. I am wondering if the model can do a good job in overall water mass balance. I am thinking changing from grass to shrub will lead to difference in column integrated water mass, could you show time series of total water storage in the model? Also, the drainage flux? How far the rainfall infiltration could go? I am then also wondering how sensitive is the model simulation to the represented soil depth, given the gravitational drainage condition is used at the lower boundary. Could you elaborate?

Response 11: This is a very nice conclusive question, and we'll break it down into the following 5 sub-points to respond, in terms of the integrated water storage, infiltration flux, drainage flux, overall water balance closure and lower boundary condition.

1. **Integrated Soil Water Storage**

   We did calculate the integrated soil water storage in the root zone (i.e., 0-200 cm depth) (Eq. S15 in the Supplement). Indeed, the replacement of shrubs decreased root zone water storage in 2016 (-27 %) and 2019 (-11 %), respectively. As you suggested, we made a time-series of total soil water storage in the whole soil column (i.e., 0-5 m depth) (see figure below). Similarly, in both years, the total water storage under shrubland was less than that of grassland.



2. Infiltration

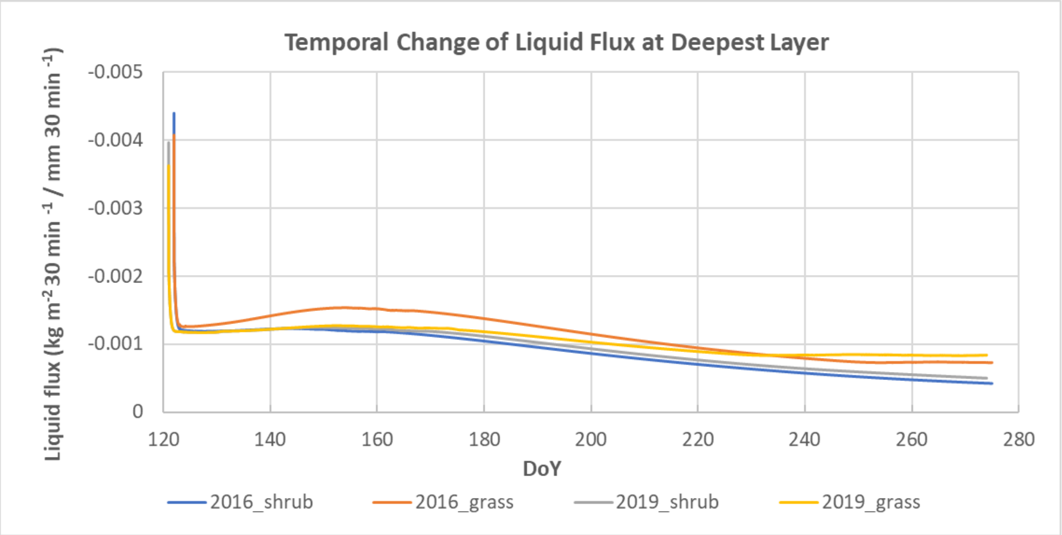
In the study area, the precipitation-induced surface runoff is rare after the canopy retention or direct infiltration, and the latter is the major process. In the model, the infiltration flux is calculated as the surface moisture flux remaining after surface runoff has been removed from precipitation (refer to CLM model, Niu et al., 2005; Oleson et al., 2004):

$$Infiltration = Precipitation \times (1 - fmax \times e^{-0.5 \times fover \times \frac{Tot_{Depth}}{100}})$$

where $Precipitation$ is the model input [mm]. $fmax$ is the maximum fractional saturated area, which is the percent of area whose topographic index is larger than or equal to the mean topographic index in a grid cell (Oleson et al., 2004). We extracted $fmax$ (= 0.3694) from a global dataset (Niu et al., 2005; Oleson et al., 2004). $fover$ is a decay factor (= 0.5 m^{-1}) (Oleson et al., 2004). Tot_{Depth} is the water table depth (= 5 m). It appears that the infiltration amount is closely related to precipitation, and the infiltration rate is related to the saturated water conductivity (Ks).

3. Drainage

Given the gravity drainage as lower boundary condition, the temporal changes of liquid flux at the deepest layer (i.e., layer 54^{th} at 5 m depth) indicates the water exchange at the boundary layer. As shown in the below figure, (1) the negative drainage flux indicates the water flow downwards to the boundary layer, with a very small value ranges from 0 to -0.005 [mm 30 min^{-1}]; (2) The drainage from the grasses was greater than that from shrubs. These observations are in line with the field situation according to the local expert. In this semi-arid area, characterized by low precipitation and deep groundwater level, there is minimal interaction between precipitation and groundwater.

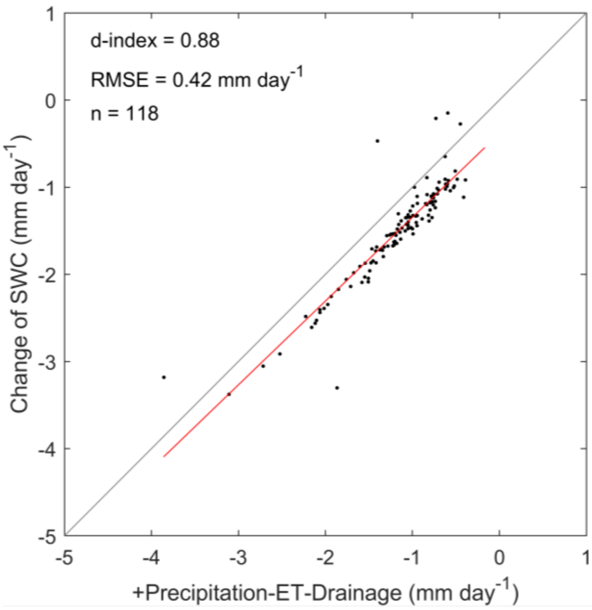


4. Water Balance Closure

The water balance closure was evaluated by comparing soil water storage and the difference between water input (i.e., precipitation) and outflow (i.e., simulated evapotranspiration and drainage):

Change of Integrated soil water storage = +Precipitation - Evapotranspiration - Drainage

On the right-hand side, we present an evaluation of water balance closure in the grassland simulation of the days without rain in 2019 (i.e., 118 of 153 days). Good agreement was found with values for the RMSE and the index of agreement (*d*-index) equaling 0.42 mm day^{-1} and 0.88, respectively.



The closer to 1 of the *d*-index, the better the performance of water balance closure. The simulated temporal change in SWC was underestimated. The uncertainties might come from (1) the method to calculate the integrated soil water storage (Summation of SWC over all the layers v.s. Inversion of the water balance equation) (Yu et al., 2016); (2) the bias in the simulated LE (i.e., ET) (See the paragraph at Line 390 in manuscript).

5. Lower boundary condition

We used gravity drainage as the lower bottom condition in a 0-5 m soil column because the groundwater level in the study area is quite deep (> 6 m). Besides, STEMMMUS-SCOPE model mainly focuses on simulating the soil-vegetation interactions and hasn't included any groundwater modules. Based on the available information on groundwater level and rooting depth of shrub, we think the settings of 0-5 m soil column and gravity drainage are reasonable. Sorry that by now we don't have the answers to *"how sensitive is the model simulation to soil depth and the bottom conditions"*. We can explore adjusting the soil column depth and modifying boundary conditions to address this question. However, we believe these factors may not be our main focus in comparing simulations for two scenarios. Nevertheless, it remains an interesting question to explore in future applications of the model in diverse study areas with varying groundwater conditions.

References

Chen, L., Zhang, Z., Zeppel, M., Liu, C., Guo, J., Zhu, J., Zhang, X., Zhang, J., and Zha, T.: Response of transpiration to rain pulses for two tree species in a semiarid plantation, Int. J. Biometeorol., 58, 1569–1581, https://doi.org/10.1007/S00484-013-0761-9, 2014.

Deans, R. M., Brodribb, T. J., Busch, F. A., and Farquhar, G. D.: Optimization can provide the fundamental link between leaf photosynthesis, gas exchange and water relations, Nat. plants, 6, 1116–1125, https://doi.org/10.1038/S41477-020-00760-6, 2020.

Fensholt, R. and Sandholt, I.: Evaluation of MODIS and NOAA AVHRR vegetation indices with in situ measurements in a semi-arid environment, Int. J. Remote Sens., 26, 2561–2594, https://doi.org/10.1080/01431160500033724, 2005.

Fensholt, R., Sandholt, I., and Rasmussen, M. S.: Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements, Remote Sens. Environ., 91, 490–507, https://doi.org/10.1016/j.rse.2004.04.009, 2004.

McDowell, N. G., Brodribb, T. J., and Nardini, A.: Hydraulics in the 21st century, New Phytol., 224, 537–542, https://doi.org/10.1111/nph.16151, 2019.

Niu, G. Y., Yang, Z. L., Dickinson, R. E., and Gulden, L. E.: A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, J. Geophys. Res. Atmos., 110, 1–15, https://doi.org/10.1029/2005JD006111, 2005.

Oleson, K., Dai, Y., Bonan, B., Bosilovichm, M., Dickinson, R., Dirmeyer, P., Levis, S., Hoffman, F., Houser, P., Niu, G., Thornton, P., Vertenstein, M., Yang, Z., and Zeng, X.: Technical Description of the Community Land Model (CLM), https://doi.org/10.5065/D6N877R0, 2004.

Wang, Y., Zeng, Y., Yu, L., Yang, P., Van der Tol, C., Yu, Q., Lü, X., Cai, H., and Su, Z.: Integrated modeling of canopy photosynthesis, fluorescence, and the transfer of energy, mass, and momentum in the soil-plant-Atmosphere continuum (STEMMUS-SCOPE v1.0.0), Geosci. Model Dev., 14, 1379–1407, https://doi.org/10.5194/gmd-14-1379-2021, 2021.

Yu, L., Zeng, Y., Su, Z., Cai, H., and Zheng, Z.: The effect of different evapotranspiration methods on portraying soil water dynamics and et partitioning in a semi-arid environment in Northwest China, Hydrol. Earth Syst. Sci., 20, 975–990, https://doi.org/10.5194/HESS-20-975-2016, 2016.