Glotfelty et al. test the applicability of different LSMs incorporated in WRF RCM for Sub-Saharan Africa and their suitability to simulate the effects of land use changes on the regional climate conditions adequately. The results show that surface albedo, leaf area index and surface roughness are not accurately represented in the default models. Therefore, a new version of WRF coupled to the CLM LSM is developed, adjusted to the specific surface conditions in Sub-Saharan Africa.
The topic of the study is within the scope of GMD and relevant for the large WRF modeling community in Africa and even beyond in the context of land use change effects on the regional climate in Africa. The manuscript is well structured and comprehensively written. The motivation of the paper is clear and the methods are well documented. Nevertheless, I have some concerns regarding the experimental setup, which need to be addressed by the authors before the manuscript is suitable for publication.

Major comments:

1) In a first step, a model validation experiment is performed, in which simulations with five different models are conducted for one year. In general, I would say one year simulations are rather short to validate model performances, especially when only annual averages are presented. The authors argue that they chose the year 2013 because its a neutral year for the El Nino Southern Oscilation, but they do not really explain why this is the ideal boundary condition to validate the LSM performances. In any case, it is a single year that cannot represent the whole climate variability. Deviating atmospheric circulation conditions can considerably affect the impact of the land surface conditions on the regional climate conditions. Moreover, due to the exclusive consideration of annual averages, seasonal conditions are excluded in which the land surface conditions have larger impacts on the regional climate (e.g. dry conditions). Therefore, I recommend to extend the simulation period, or at least, to consider seasonal effects.
In a second step, it is intended to quantify the impact of land use changes on the simulation results with different LSMs. For this, the results of climate simulations for the period 2001-2010 with static land use conditions are compared to results of climate simulations for the period 2010-2015, including observed land use changes. But differences between two simulations with different land use conditions do not have to be caused inevitably/exclusively by the different land use conditions, in the case of deviating simulation periods. The different atmospheric circulation conditions in both periods can have certain impacts on the simulation results. Thus, from my point of view, identical simulation periods would have been preferable (2001-2015). If it is not possible to perform these simulations with respect to computing time, one could eventually reduce the number of LSMs based on the results of the validation experiment. By the way, I do not really understand why Noah Sat is included in the study, if one cannot use it for land use change scenarios at all. The authors should at least discuss the potential effects of the different simulation periods.

2) It is very difficult to assess the differences between the different LSMs in the validation experiment based on the shown figures. It is therefore very difficult to compare these differences to the changes caused by land use changes. Plots of the differences to observations as shown in section 7 would help a lot.

3) To be able to understand the results of the validation experiment comprehensively, an assessment of the sensible and latent heat fluxes is necessary.

Minor comments:

1) The biogeophysical effects of the surface roughness on the climate impacts of land use changes is not considered. For instance, this impact can be seen for the deforestation regions. The model results consistently show a warming with deforestation. In the manuscript, this is explained by

reduced evapotranspiration rates and an associated reduced latent cooling. But if the reduced latent heat fluxes are the reason for the increased near-surface temperatures, accordingly the sensible heat fluxes should be increased (due to the increased temperature gradient between the land surface and the atmosphere). But this is not the case, the sensible heat fluxes are also reduced. Therefore, I suppose that the efficiency of the deforested land surface to transform the incoming solar energy in turbulent heat is reduced due to the reduced surface roughness, resulting in a warming of the surface (e.g. Winckler et al., 2019; Breil et al., 2020).

2) simulation results show that the cloud cover is consistently overestimated in the validation experiment. At the same time, downward short-wave radiation (swdown?) is also overestimated. How does that fit together?

3) Several abbreviations are used which are not explained in the text (e.g. SWDOWN, GLW, OLR, SWUPT).

References:

Breil, M., and Coauthors, (2020). The Opposing Effects of Reforestation and Afforestation on the Diurnal Temperature Cycle at the Surface and in the Lowest Atmospheric Model Level in the European Summer. *Journal of Climate*, *33*(21), 9159-9179.

Winckler, J., and Coauthors, 2019: Different response of surface temperature and air temperature to deforestation in climate models. Earth Syst. Dyn., 10, 473–484, https://doi.org/10.5194/esd-10-473-2019.