Summary:

This manuscript presents a comparison of the vegetation transition in the Sahara between the Holocene and the Last Interglacial (LIG), using an earth system model of intermediate complexity (iLOVECLIM) in combination with two dynamical vegetation models, VECODE and LPJ-GUESS. The Holocene simulations (8.5 ka to 1 ka BP) were run for this study, and the LIG simulations were the same as Li et al. (2020) that focused on simulated climate-vegetation interactions in North Africa during the LIG. These two periods have been adopted as the CMIP6/PMIP4 target period, and this study help to understand the role of changes in insolation, one of major forcing agents in the Earth system, as well as climate-vegetation interactions. I am not sure whether the three questions were appropriate, a research design was developed or not, and a reasonable amount of analysis was conducted to answer them. However, the presentation of the experiments and results is a bit vague/unclear and lacks necessary information in its current manuscript because the authors have written too briefly.

I am outlining my main concerns below, and the paper can be accepted after moderate revisions if the research questions are pertinent.

Major comments:

(1) The authors could do a better job of motivating this study. It is not clear why it is necessary to simply compare these two time periods regarding Green Sahara (vegetation-precipitation transition in the Sahara). We understand some studies already modified to simulate Green Sahara reproducing intensification and geographical expansion of the West African monsoon (e.g., Pausata et al. 2016; Hopcroft and Valdes 2021), but many Paleo-
modelling still fails to simulate it (Tierney et al. 2017). Therefore, it would be better to have a clear motivation, for example, to obtain clues (regarding vegetation-climate interactions) to modify the Holocene simulation by comparing the two periods. Alternatively, it would be interesting to have a new fact (not known from the data alone) that can be obtained through the comparison.

https://www.science.org/doi/10.1126/sciadv.1601503


(2) It might be better to describe what is already known in the Introduction section and show more new results in this paper. For example, the spatial and temporal complexity of the termination of the African Humid Period (AHP) have already known (Shanahan et al. 2015; Tierney et al. 2017; Dallmeyer et al. 2020), and data regarding the abruptness of precipitation/vegetation decline is a local- or regional-scale feature (Brovkin and Claussen 2008), not the whole Sahara. In this study, one of the main analyses is to investigate changes in climate (particularly surface temperature and precipitation) and vegetation cover in the whole Sahara or North Africa, including the Sahel, during the LIG and Holocene. However, it is better to analyse the western and eastern parts of the Sahara separately rather than the whole Sahara.

(3) It seems that the data already show that the Green Sahara happened in the two periods, but what else do we know from the data, especially about differences? Finally, is there a reason for no quantitative model-data comparison, in particular, the Holocene has been made at all? We cannot decide whether these simulations are good or bad at all.

Not directly related to this study, but compared to the data (e.g., Hoffman et al. 2017; Capron et al. 2017; Scussolini et al. 2019) how good do the LIG simulation (Li et al. 2020) reproduce global-scale surface temperature and precipitation?

https://www.science.org/doi/10.1126/science.aai8464


https://www.science.org/doi/10.1126/sciadv.aax7047

**Line comments:**
11: “Hand” to “Hans”

43: “the rate of this transition remains controversial” – The authors would be better to clarify whether these are differences among data or discrepancies between data and model.

60: Since Liu et al. (2006), many studies have already been implemented on this topic, but why do the authors not mention any recent studies? -- Liu et al. (2006): SC, Claussen et al. (1999): UC, and how about the other recent studies?

60: “during the termination of the AHP” to “during the AHP”

67 and 71: I do not understand the mechanism well. Compared to the desert, the vegetated area has lower albedo and absorbs more SW radiation leading to a warmer surface. However, a vegetated surface produces more latent heat and cools the surface (thus warming the atmosphere above). Thus, changes in surface temperature should be determined by the balance between warming due to increased absorption of SW radiation and cooling due to increased surface latent heat.

81: Which specific period is the LIG here? For example, from about 129 ka BP to 120 ka BP.

82-93: After all, are those simulations quantitatively consistent with the data? Does the LIG data also show a fact of “nonlinear response of the African monsoon to orbital forcing” and “the spatial heterogeneity of the response” as well as the MH data?

I found a qualitative LIG precipitation data (Scussolini et al. 2019), but do we have any quantitative LIG precipitation data?

93-95: Since these two sentences are new topics, they could be moved to a new paragraph. Also, is the issue the authors point out here limited to iLOVECLIM, or does it involve other GCMs as well?

103: Could the authors also use these two vegetation models under the same conditions? In other words, can VECODE also be simulated asynchronously with iLOCECLIM? Is it technically impossible?
The scientific significance of the first and second questions is a little unclear to me. Could the authors please elaborate a bit more on why these questions are important?

The authors can describe a little more about cloudiness, humidity, and precipitation of ECBilt because ECBilt is somewhat different from AGCMs. I understand that ECBilt uses the prescribed/fixed cloud cover based on the modern condition throughout the paleo-simulation.

LPJ-GUESS adopts a simple two-layer bucket model (with prescribed percolation rate and water holding capacity), but is VECODE the same/similar structure? If they differ, it would be better to describe the difference.

LPJ "standard" version (Sitch et al. 2003) has 10 PFTs, but what is the other PFT? Do the authors count bare ground as a PFT?

It seems that the content here is not a model description, but an experimental design.

The last two sentences in the paragraph are a little unclear for me. Could the authors explain the experimental design for the asynchronously coupled version, ECBilt-CLIO_LPJ-GUESS with a chart?

1850 AD, not 850 AD (for prescribed pre-industrial condition) typo(?)

The description of LBM should be moved before the paragraph on each of the three-model configuration.

LPJ-GUESS has 11 PFTs, but did those PFTs simply convert into 3 types (trees, grasses, and desert) for the LBM?

This paragraph is a bit confusing. Does it mean that soil hydrology calculated in LPJ-GUESS does not directly affect ECBilt, but has some indirectly influence through vegetation type?

At each time-slice simulation (HOL_LPJ, LIG_LPJ), how many model years did the authors run the model and how many years of the output were used in the analyses?
L.213: What is the range of the target area (latitudes and longitudes) for North Africa or
the Sahara here?

L.223-225: How about the recent (CMIP6/PMIP4) simulations about? Comparison with past simulations is important, but comparison with recent simulations as well as data is also important.

L.225: Why is the LIG_FIX temperature trend positive?

L.228: Because LIG_LPJ does not show large changes in surface temperature in North Africa, are changes in surface temperature and desertification (vegetation cover) less relevant in this simulation?

L.235: Fischer and Jungclaus (2010) analysed time-slice simulations, not transient ones. So that may not be an appropriate reference here. Moreover, according to Brovkin and Claussen (2008), which is also cited in this paper, Francus et al. (2013) may not be an appropriate reference either, because the individual data represent local responses and are not representative of the whole North Africa.

L.235~237: Figure 1f shows that magnitude of precipitation decline in HOL_LPJ is similar to one in HOL_FIX, and this sentence may not be appropriate.

L.245: Change Fig. 2f to Fig. 1f or Fig. 2b(?) Anyway, we cannot consider "the simulated vegetation distribution and spatial divergence in North Africa" from this figure, I think.

L.253: It seems that surface temperature trend in HOL_LPJ is similar to one in HOL_FIX.

L.257: Can we check the ratio of trees to grasses in North Africa? Vegetation-induced changes in albedo and surface evaporation may also depend on the surface conditions between trees and grasses.

L.262: What is the reason for the sharp decline in vegetation cover, especially from 123ka o 121 ka in the LIG_VEC simulation? Moreover, why is that trend not seen in HOL_VEC?
L.263: Can the authors check how much the ratio of trees to grass in North Africa varies from model to model? Looking at the vegetation area fraction anomalies (Fig. 3), there may be considerable differences between the two models in terms of the proportion of trees and grasses.

About the different vegetation diversity between the two models, unlike Claussen et al. (2013) VECODE and LPJ-GUESS are completely different process-based DGVMs, and there must be many differences besides diversity.


L.272 and L.347: Yu et al. (2017) proposed the observed positive vegetation feedback on precipitation in the Sahel (not the Sahara) by a moisture recycling mechanism rather than the classic albedo-based mechanism. Messori et al. (2018) have a similar idea for the Holocene Green Sahara. Does this concept apply to the authors’ experiments (the Sahara in the LIG and Holocene)?

https://www.nature.com/articles/s41467-017-02021-1


Fig.3: PFT fraction anomalies for LIG simulations (in particular 120k ~ 118k) show shrinking vegetated areas extend much southerly, but why does that feature not happen in Holocene simulations?

L.278: What do the authors think caused the decline with large error bars at 5 ka and 4 ka in the HOL_LPJ simulation? Moreover, why does that feature not catch in the LIG_LPJ simulation?

L.279~284: The authors should describe the spatial heterogeneity in the Introduction section, not here because this is a known fact. Furthermore, based on this, from the beginning, the region should be divided into East and West for analysis, I think.

L.284: As mentioned before, the authors can confirm this by making VECODE asynchronous with iLOVECLIM.
L.292: What are the grid points for both western and eastern North Africa, respectively?

L.298: It’s hard to see the differences between West and East Sahara from Figure 4. Moreover, what exactly is “A spatial and temporal complexity of the termination of the AHP”?

L.302: About the sentence “the magnitude of our vegetation decline is much weaker than in their study”, which study/value matches the available data?

L 302: Is “the differences in model complexity” simply about the vegetation models between Lie et al. (2007) and this study?

L.321-323: Does any data also support the changes in climate and vegetation in the LIG are stronger than ones in the Holocene?

L.328: Is around 125 ka BP and around 8.5 ka BP each peak of insolation at 20N for the LIG and Holocene respectively?

L.338: Did Shanahan et al. (2015) discuss the vegetation cover and vegetation-albedo feedbacks using TraCE-21 simulation?

L.341-344: Does any data also support the changes in climate and vegetation in the LIG are stronger than ones in the Holocene? Or, will the results of the LIG simulation help to improve Holocene simulation?

L.361: In this study, “the fractional surface albedo of trees, grassland, and desert are seasonally fixed”. Could this setting also be relevant?

L.380: It seems that the section 3.4 is not what the authors found out through comparison between the LIG and Holocene simulations. How about discussing at least one issue that arose through comparison?