Reply to Review #1

We would like to thank the reviewer for the thorough and constructive review of our manuscript. The well-considered comments and suggestions will help us clarify some points that might otherwise have been misunderstood or that simply benefit from further elaboration. Our detailed response to the reviewer's comments is provided below. For clarity, the reviewer's comments are shown in black and our response in blue.

This manuscript provides a highly rigorous quantitative framework to evaluate the termination of the African Holocene Humid Period (AHHP). The authors successfully move the long-standing debate from a qualitative discussion of "collapse vs. gradual decline" to a standardized, metric-driven assessment. The integration of transient simulations from the MPI-ESM with carefully aggregated pollen groups and marine dust flux records is solid. By applying a changepoint detection method to linear regressions, the authors effectively demonstrate that the end of the Green Sahara was characterized by a complex spatial and temporal mosaic: a gradual expansion of desert areas on a macro scale, accompanied by abrupt, step-like transitions within specific PFTs and regional ecosystems. The discussion of the limitations of the single model study (such as the challenge of representing savanna biomes in JSBACH), and the uneven distribution of proxy data availability provides a clear path forward. The manuscript is a pleasure to read, and my comments below are offered as minor (and mostly technical) suggestions to further polish an already outstanding paper.

We are very pleased with this statement, and we appreciate very much the concise questions and comments which we respond to below.

1. As the authors noted, the complete omission of the Sahelian group from the model comparison—due to JSBACH's known inability to accurately simulate savanna ecosystems—is a significant limitation. Given that the Sahelian transition is central to the aridification of the Sahara, the authors should explicitly outline how the lack of a savanna biome representation might skew the modeled rapidity of the bare-surface expansion. If grass and trees cannot properly co-exist as savanna in the model, does the transition to bare soil appear artificially abrupt?

Our statement (line 184) that "the land-surface model JSBACH is known to not properly describe any type of savanna (Groner et al., 2018) " is misleading in this context. First we apologize for the erroneous citation: not Groner et al. (2018), but Dallmeyer et al. (2021) showed that the simulation we analyzed in our study captures the present-day global patterns of the savanna biome with the poorest skill compared to all other biomes in terms of best neighbor scores (see Dallmeyer et al., 2021; Table 2). Dallmeyer et al. (2021) mention that mapping savannas in terms of biomes is limited in JSBACH, because savannas are affected more by perturbations like fire and windbreak and to a lesser extent by climate. In addition, the various types of savanna mainly differ by their proportions of tree.

However, JSBACH does describe a savanna; the grid boxes in the SW of the Lake Yoa (Fig. 4, see also our answer in response to comment 3) can be taken as an example. At the beginning of the simulation around 8 ka, the MPI-ESM produces a lush green savanna dominated by grass in this region. Around 5 ka, the grass fraction considerably declines while the tree fraction is almost constant. After around 3 ka, the tree PFT starts to vanish, and the bare surface fraction in the MPI-ESM increases well above 0.8 after 2 ka. Thus, the rapid increase in simulated bare surface fraction is not caused by the absence of a savanna, but by a change in the type of savanna, from a lush grass savanna to a meager tree savanna, and finally, to a desert.

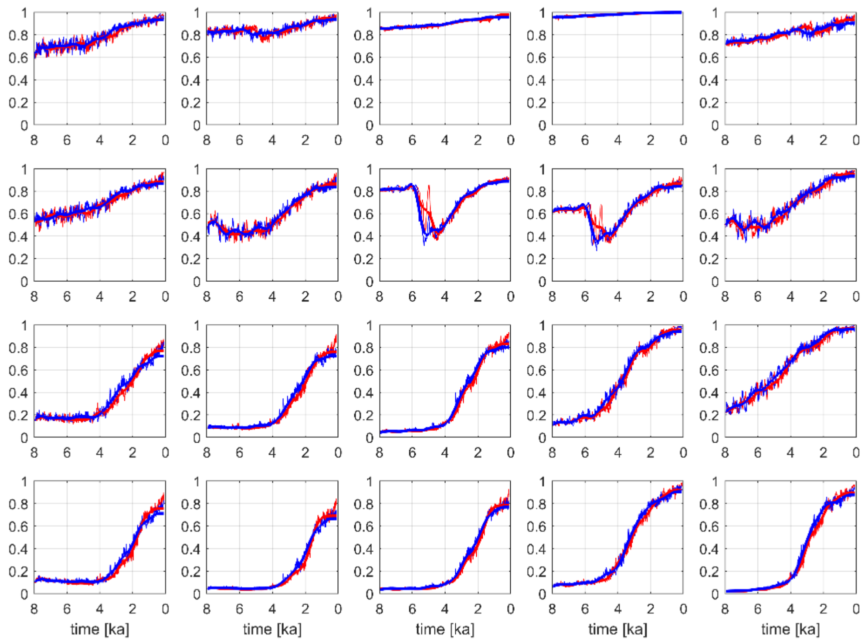
What we actually wanted to say – and will say in a revised manuscript – is the following. Sudanian and Guinean taxa are clearly dominated by the abundance of tree species and, hence, can be more or less directly compared with the fraction of simulated tropical tree PFT. The Sahelian group is defined by a mixture of grass, shrub and tree species, depending on the topography, the occurrence of bush fire and climate. This group can thus not realistically be compared with the limited numbers of PFT simulated in JSBACH (Dallmeyer et al., 2021).

As a side remark in response to this comment only: We have tried to compare the Sahelian group with some flexible proportions of simulated PFTs or by using some non-linear function of the proportions of simulated woody types. But we were not successful so far.

2. Orbital forcing vs. internal variability. The definition of abruptness relies on changes occurring faster than the linear trend in orbital forcing. While this is a solid baseline, it would be beneficial to elaborate on the potential role of unforced, centennial-scale internal climate variability. Does the MPI-ESM transient simulation show periods where internal variability can coincidentally push the vegetation over a bioclimatic threshold or beyond the defined abruptness threshold, accelerating the transition independently of the orbital decline?

In the MPI-ESM, the coherent patterns of largest abruptness in northern Africa are found in the region of the largest vegetation changes in a strip around 20°N in the west and 15°N in the east. This pattern of large vegetation changes seems to be mainly dominated by the orbital forcing rather than by internal climate variability. We can show this by comparing two different simulations from the same ensemble of our transient Holocene simulations, but with different initial conditions and with a different volcanic forcing.

The figure below depicts the bare land fraction in the western Sahara region (see Fig. 3 in our manuscript) for the simulation used in our study (blue line) and for a simulation with a presumably unrealistically strong Holocene volcanic forcing (red line). Please note that we switched to the more conventional notation ‘ka’ = 1000 years before present as suggested by Reviewer#2. Both simulations differ with respect to their internal climate variability. But the vegetation changes seem to differ only on very short time scales, but not on much longer time scales and not on the long-term trend. Only the “cumbersome grid boxes” (second row, 3rd and 4th column) discussed in our Appendix B as model artefacts show a somewhat larger deviation between simulations. Such grid boxes are rare exceptions from the rule when considering all grid boxes of our study area.

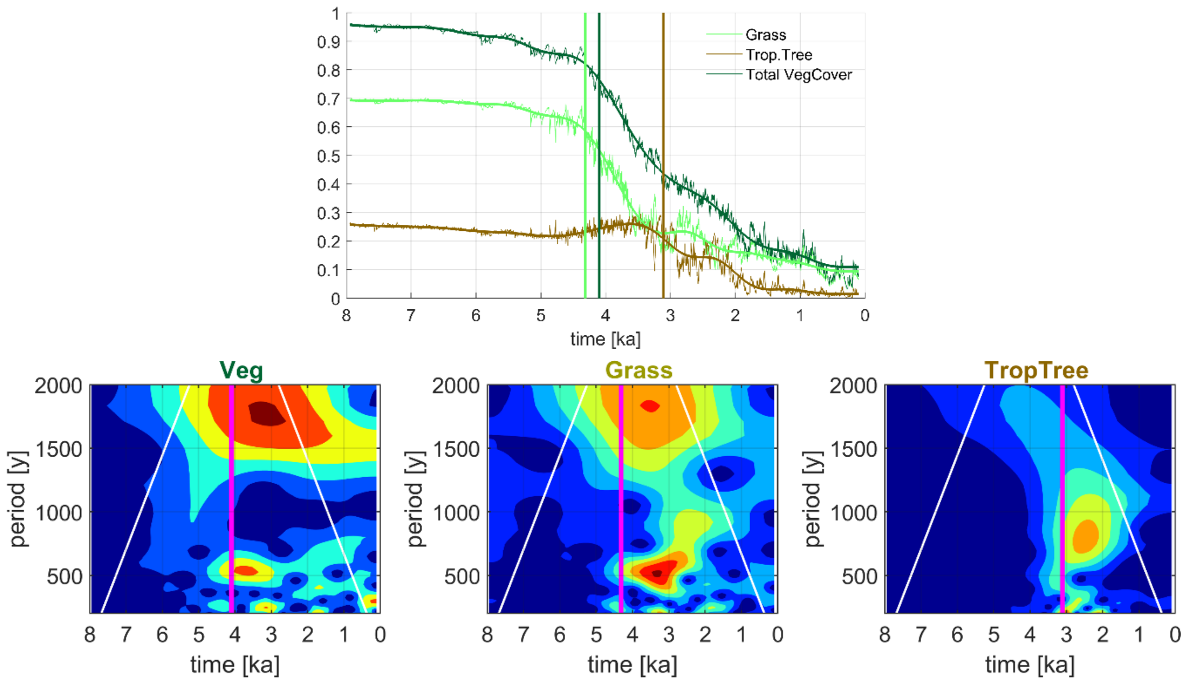


Thus, we suppose that the coherent pattern of largest vegetation changes, and hence largest abruptness, is a robust feature, mainly triggered by changes in orbital forcing. We agree with the reviewer that a brief statement on this issue is sensible to add.

3. To restrict the changepoint detection a priori to a single transition in linear regression is a logical choice to isolate the primary step-like termination of the AHHP. However, some of the highly variable proxy records (and perhaps some regional grid boxes in the transient simulation) might naturally exhibit multi-stage transitions or "flickering" before settling into a desert state. The authors might consider adding brief discussion on whether releasing this single-changepoint constraint in future studies might reveal secondary and smaller transitional phases.

We have excluded a discussion of multi-stage transition and flickering from our study, because we wanted to focus on the abruptness of step-like transitions which is a “logical choice” as stated by the reviewer. Many simulated time series and proxy records considered in our analysis of the last 8000 years are characterized by a big transition either as a step change or as a change in the trend. But there are grid boxes and proxy records that reveal small long-term trends with large fluctuations. If we would consider multiple changepoints to analyze the grid boxes characterized by large variability, then the quantitative definition of abruptness would become a challenge. Would the number of changepoints be a measure of abruptness or the strongest abruptness of the various changepoints in a grid box? Or something else? We will mention this point in the discussion. For other records and simulations, an analysis of multiple changepoints makes perfectly sense, of course, as for example for the long Chew Bahir record (Trauth et al., 2018).

As for the “flickering”, we have begun to look into this topic. Flickering does occur in our simulation. The figure below shows the time series and wavelet analysis of time series of grass PFT, tropical tree PFT and total vegetation cover in a grid box in the southwestern part of the Lake Yoa region.

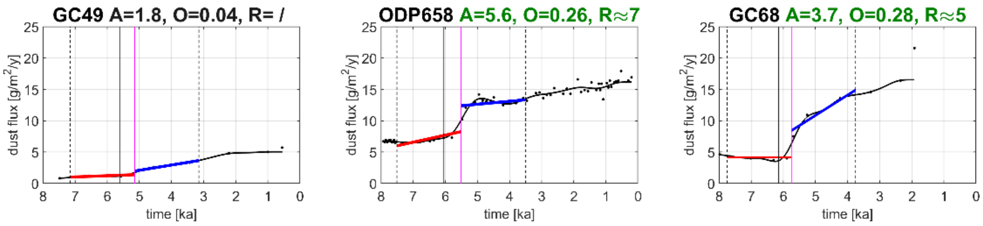


It is interesting to see that flickering starts prior to the respective changepoints (vertical straight lines). Please note that the times series were detrended by using high-degree polynomial before applying the spectral analysis. With some imagination, even a shift towards larger frequencies /

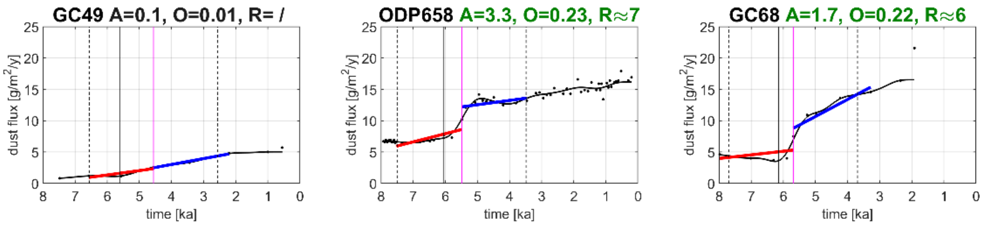
longer periods can be found in the grass and the tropical tree PFTs, but after the changepoint, which could point at an early warning signal as the PFTs approach a bifurcation point. This, however, is work in progress, not mature enough to be discussed in our present manuscript.

4. To ensure consistency between the high-resolution simulation output and the proxy data, the authors interpolate pollen and dust records to annual values before applying a 500-year Butterworth filter. Because the temporal resolution of the proxy records varies considerably (e.g., from roughly 50 years to nearly 400 years for pollen, and up to 990 years for the GC49 marine core), the authors might briefly clarify if the interpolation process introduces any minor spectral artifacts or aliasing, particularly for the coarsest records. A minor note assuring the reader that the changepoint detection is insensitive to this specific interpolation step would be a helpful technical addition.

This is an interesting point which we will briefly mention in the revised version of our manuscript. We have considered using the raw proxy data without any interpolation. It appears that the use of raw data does not affect the detection of changepoints, except for “marginal changepoints” as shown in the case of the marine record GC37 depicted in the figures below. However, the values of abruptness differ between the analysis of interpolated and non-interpolated time series. With apparently similar offsets and similar regression lines, we can attribute the reduction in the A values mainly to an increase in standard deviation of data from the linear regression lines in the non-interpolated time series. Hence, we find it sensible to interpolate the proxy data to the same resolution as the simulated data, thereby putting the abruptness statistics of data and model on the same footing.



(analysis with interpolation of data)



(analysis without interpolation of data)

In a revised version of our manuscript, we will briefly mention this effect of interpolation on the detection of changepoints and evaluation of A without showing the figures.

Typos:

Thanks a lot for spotting the typos which we will eliminate in a revised version of our manuscript

References

Dallmeyer, A., Claussen, M., Lorenz, S. J., Sigl, M., Toohey, M., and Herzschuh, U.: Holocene vegetation transitions and their climatic drivers in MPI-ESM1.2, Clim. Past, 17, 2481–2513, https://doi.org/10.5194/cp-17-2481-2021, 2021.

Groner, V. P., Raddatz, T., Reick, C. H., and Claussen, M.: Plant functional diversity affects climate–vegetation interaction, Biogeosciences, 15, 1947–1968, https://doi.org/10.5194/bg-15-1947-2018, 2018.

Trauth, M. H., Foerster, V., Junginger, A., Asrat, A., Lamb, H. F., and Schaebitz, F.: Abrupt or gradual? Change point analysis of the late Pleistocene–Holocene climate record from Chew Bahir, southern Ethiopia, Quat. Res., 90, 321–330, https://doi.org/10.1017/qua.2018.30, 2018.