

Response to Anonymous Referee #1

We would like to thank both Anonymous Referees for their review and suggestions for improvements.

This study used a model to investigate the impact of high and low precipitation years on ecosystem carbon and water cycling in the Sahel region, Africa. Annual precipitation anomalies were simulated by changing rainfall intensity, event duration, or rainy season length and holding the remaining two variables constant. The resultant fluxes were then analyzed with respect to contributions from the three major regional plant functional types. In this framework, grasses demonstrated a flashier response to all precipitation scenarios than trees, and changing the length of the rainy season had the greatest impact on vegetation due to changes in runoff and surface evaporation associated with precipitation intensity and duration. No single-year precipitation treatments had a significant effect on the long-term (decadal) simulated carbon or water fluxes.

Response: We would like to clarify that we did not study the anomalies caused by changing event duration, i.e. the average number of consecutive rainy days, but rather the event frequency, i.e. the average number of consecutive dry days between rain events. Of course, these are related when the length of the rainy season is kept constant.

General comments

I thought this research was creative and well designed, and work from understudied regions like Africa is especially valuable. However, the stated research questions (L91-93) are left mostly unanswered by the study that presents data from only one site in its current form. I don't understand or agree with the decision to exclude three of the four sites from the main manuscript text. I realize that these results are available in the supplement (without interpretation or discussion), but they're needed in the main text insofar as they're key to meaningful interpretation of your work and its scope. To me, the impact of this manuscript, and thus the suitability for the journal, depends on its applicability to other dryland systems. To establish this, (1) model performance must be evaluated at all sites, and (2) systematic biases (or lack thereof) subsequently analyzed with respect to relevant climate/soil/vegetation, etc. characteristics, in order to generate transferable biogeoscientific insight.

Response: We thank the reviewer for her/his appreciation for our study design and focus on the Sahel region. We agree that it would be better to analyse the results of all four study sites in the main text, as this can indeed further contribute to the knowledge on how global drylands respond to extremes in seasonal characteristics. We initially chose to focus on one site in the main text, mainly to make the manuscript not too lengthy.

We hope that it would be a good compromise to include a few figures which summarize the results from all sites, only showing the maximum amplitude and the legacy for each scenario. Such a figure is given below, where we summarize the scenario impacts on leaf area index (LAI) for each PFT (Figure R1). As it can be seen, there are some contrasts in both amplitude and legacy between the different sites and we will elaborate further on these differences in the Results and Discussion sections. Of course, such a summarised presentation has its limitations, as for example overshoots are neglected, but the full time-series can still be found in the supplementary materials.

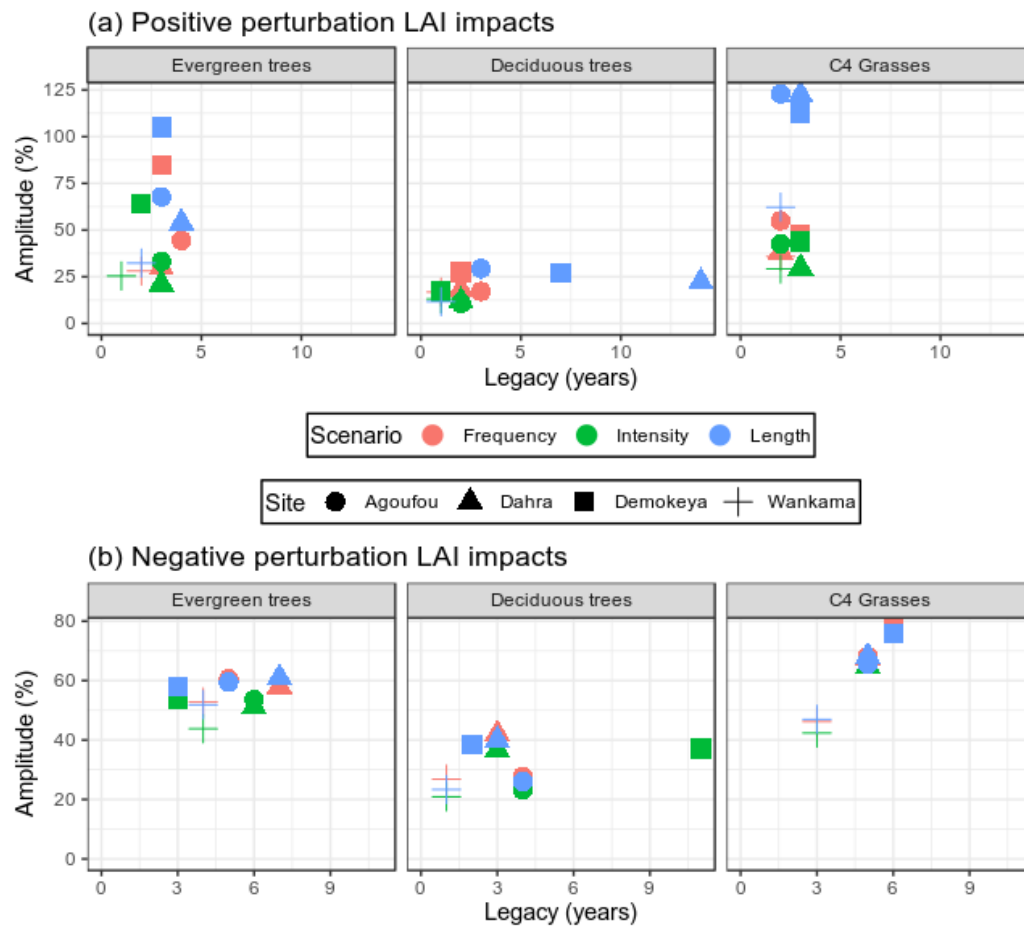


Figure R1. Absolute amplitude and legacy of the vegetation response to (a) positive and (b) negative rainfall perturbations due to (a) increased and (b) decreased event frequency, intensity or season length. The legacy was defined as the last year for which the ensemble average (absolute) response is larger than the ensemble standard deviation, while the amplitude is the maximum response, relative to the reference run values. Results shown for all Sahel fluxtower sites.

Specific comments

L16-19: Are extremely high precipitation years also forecast for the Sahel? Only drought extremes are invoked on L17.

Response: We will look into this again and add relevant references to the Introduction.

L20: “The rainy season” is confusing as you have yet to establish precipitation seasonality. Also “signature” could be more specific e.g., “the impact of the rainy season on the ecosystem carbon balance” or similar.

Response: We will rephrase this line so that it will be more clear.

L24: Please clarify what’s meant by “meteorological consistency” in this context. Reading ahead, I see it now, but if you wish to invoke this term/concept in the abstract, it should stand alone.

Response: We will remove this statement, as it is not of central importance to mention this in the abstract.

L29: Is it your intention to use “semi-arid” and “dryland” interchangeably? Maybe better to just choose one term and stick with it, at least in the abstract.

Response: Agreed and the manuscript will be modified accordingly.

L30: After reading the entire manuscript, I disagree with the use of “long-term” here.

Response: We agree that this term can be interpreted in many ways, depending on the background of the reader. Our (unwritten) definition of long-term was “more than an order of magnitude longer than the duration of the disturbance”, in which case its use is justified. Nevertheless, we will clarify this sentence so that it is less open for interpretation.

L54: And presumably in other dryland areas right? It’s to your advantage to keep the results as broadly applicable as possible.

Response: Agreed.

L71-72: Shouldn’t mean annual precipitation (MAP) be less in arid areas than in semi-arid areas? Also 650 mm described as the cutoff for mesic (L45) but then 700 mm cited as a cutoff for “less arid” (L73). Please establish a clear precipitation classification scheme.

Response: We used the definitions from the respective papers here, but we agree that it would be good to translate them into a consistent precipitation classification scheme for our manuscript, and we will do so accordingly.

Table 1: Does “fallow bush” imply 0% tree cover at the Niger site?

*Response: No, this fallow site does contain a few trees, but the dominant vegetation is a mixture of shrubs (mainly *Guiera senegalensis*) and annual grasses (Boulain et al., 2009). We will clarify this in the table.*

L103: Most readers will be unfamiliar with this area and would benefit from a site map.

Response: Agreed and will be added.

L108: I think including the measurement years for each of the flux tower sites would be warranted here or in Table 1. They’re used to validate the model, so the length of the instrumental record is particularly relevant. Also it bothers me when eddy covariance measurements are simply presented as “the truth”; please indicate at minimum the mean uncertainty associated with the flux observations.

Response: Agreed. The measurement years will be added to the table, and the uncertainty on the eddy covariance measurements will be quantified and discussed.

L140-141: This is unexpected given “soil control on surface water balance” (L28) and the different soil classifications in Table 1.

Response: All of the sites consist of Arenosol, which are all sandy soil types. The model does not take into account differences in lower-level soil classification (i.e. Luvisol, Ferralic or Cambic). We will add this information to the discussion.

L158: How was the length (1 to 37 years) of any particular meteorological cycle determined?

Response: The input meteorological dataset consists of 37 years (1979-2016). We marked all extreme (dry or wet) years and perturbed the remaining years one by one, represented by the different branches marked as “experiments” in Figure 2. Each branch is simulated until the year that is perturbed, and is then followed by a spin-down run.

L160: Just the rainy season or the entire year (as implied by the Figure 1 y-axis)? This may seem common knowledge to you, but most readers will be unfamiliar with the climatology of the Sahel. Maybe a “percent of MAP during the rainy season” metric could be added to Table 1?

Response: Indeed, we have not mentioned that the climate in the Sahel has a high degree of seasonality. We will elaborate on this and add a seasonality metric in the table or in the main text.

Figure 1: I’m unsure how to interpret the frequency variable; would a value of 0.5 day⁻¹ correspond to one rain event every two days? Perhaps an explanatory sentence could be added to the caption (or just a pointer to Table 3)? How was the length of the rainy season determined quantitatively? I’m more familiar with growing season length calculations that can greatly affect interpretation of results. Reading ahead I see a definition in terms of “climatological anomalous accumulation” in Table 3, does this imply a typical lack of any precipitation outside of the rainy season? If so, please state clearly in the site description.

Response: The frequency characteristic is the inverse of the number of days between two rain events, regardless of how long these events last. There might be a small amount of rainfall outside the season, but the climate has a high degree of seasonality.

Table 4: It’s taken me a while to interpret the “rainfall disturbance” column. I think it means the net change in rainy season precipitation due to perturbation of one of the three precipitation metrics. Perhaps change to “change in rainy season precipitation” or similar to clarify this? And/or list the relevant perturbation first and then the resulting effect second?

Response: We will rearrange the table columns and rename “Rainfall disturbance”.

Figure 3: Can you speculate as to why there is a much larger difference between modeled and observed data at Dahra in 2010 compared to any other year? This could inform your other results (see general comments).

Response: One factor that is not included in our model simulations is livestock grazing, which was found to have a positive impact on both gross primary productivity (GPP) and ecosystem respiration (ER), although further studies are needed to fully understand this mechanism (Tagesson et al., 2016). Cattle density was the highest during 2010, potentially explaining the higher discrepancy with the model that year (Tagesson et al., 2016). We will elaborate further on the differences between modelled and measured fluxes, and their potential causes, in the revised manuscript.

Figure 4: Many people (including myself) find Taylor diagrams difficult to interpret. Why not just show an analog to Figure 3 for all three sites? This would provide ecohydrological information about each site, as well as highlight periods of relative agreement and disagreement between the modeled and observed data i.e., my previous comment, which could yield additional insights. Why RMSE on the previous figure and RMSD here?

Response: Taylor diagrams provide a compact way to directly visualize an evaluation of model performance, based on different metrics. In this diagram we evaluate the results for four sites, each based on three different meteorological drivers, which would lead to at least eight time-series plots if we include evapotranspiration as well. We refrained from showing all these in the main text due to space limitation. However, as noted by Anonymous Referee #2, the current diagram mainly evaluates how well the model captures the seasonality of the growing season, rather than the fluxes during the season. Therefore we will add a second Taylor diagram based on the growing season separately, and we will make sure to use consistent terminology throughout the manuscript (ie. RMSE).

L231: I'm unclear as to how exactly +122% and +54% are meant to be interpreted. For example, is it that increased rainy season length stimulated the C4 grass LAI 122% more than both similarly increased frequency and intensity? Perhaps including some of the actual values would clarify this.

Response: All relative impacts are given with respect to the reference runs, see next point.

Figure 5: How to interpret (a-c)? Is it simply the aggregate of (d-f)? If so, what information is meant to be taken away from (a-c)? What about the other three sites? Given that this is a recurring framework, please treat this as a major comment.

Response: Plots (a-c) represent the reference runs, or “control runs” if this term would be more clear. Each of these reference runs, i.e. each ensemble member, is driven by exactly the same meteorological sequence as its partner experiment run, but without the perturbation applied. In other words, these plots show what the LAI of the PFTs would be if the anomalous year did not take place. Plots (d-f) then show the relative difference with these runs, i.e. the absolute difference divided by the reference value. Therefore, showing these reference values is not only justified, but necessary in order to correctly interpret the impacts shown in (d-f). This answers the previous question on L231 as well.

L253-254: I see no mention of autotrophic respiration, do you mean GPP? How was it calculated? See Chapin et al. 2006 Reconciling Carbon-Cycle Concepts, Terminology, and Methods for relevant definitions.

Response: We did mean the NPP, although we indeed didn't mention autotrophic respiration in the text. We will clarify this.

L255: Excluding changes in respiration, how can it be less than 1x the reference NPP if there's additional photosynthesis?

Response: All results given here are relative to the reference value. The cumulative impact (d-f) is obtained as follows. First the impact on the yearly NPP is calculated by subtracting the reference yearly NPP from the total NPP. Then the cumulative sum is taken over the resulting values, to incorporate respiration changes as well. We agree that the manuscript may benefit from further clarifications here, and we will add them accordingly.

Figure 6: Same comment as Figure 5.

Response: See previous answers.

L272-276: Suggest rearranging the text (or figure) to introduce figure panels in order.

Response: Agreed, we will rearrange the text.

L287: Four sites were introduced in the methods, but so far, we've only seen results from Dahra. This is confusing and requires explanation and justification for your rationale.

Response: See response to the general comments above.

L290-292: This harkens back to my comment on L140-141, perhaps the effect of soil type/texture on the results is discussed in the next section...

Response: This falls outside the scope of this study, as the different sites used the same soil type in the model. In the manuscript we already discuss that a sensitivity test on different soil types may be an interesting follow-up research paper.

L298-302: Right on cue, here's a partial response to my comment on L287. In short, I'm not satisfied with your treatment of the other sites. The vast majority of readers will not reference the SI, and even those that do must have model interpretation expertise to gain useful information. This functionally excludes three of the four sites from your analysis. Without these other sites, you're left with a site-specific tuned model that will be of little interest to other dryland or even Sahel researchers. To say it a different way, in my eyes, the strength of your work is the development of a transferrable model that could be used by dryland savanna researchers globally. This doesn't have to be a fatal flaw as I think you have the data to support a broadly applicable presentation/contribution, they just have to be worked into the manuscript. Especially in light of my comments about the "reference" panels in Figures 5-7, I think you have the space for this. And it's okay if the model performance wasn't as good at the other sites (not saying this is the case), simply presenting the data and discussing potential reasons for periods of better and worse model/data agreement would speak to process-based information and benefit the research community.

Response: The results from the other sites will be brought into the main manuscript, see response to the general comments above.

L317: You've discussed both NPP and NEP, so the "total carbon flux" is unclear. Also, fires emit carbon dioxide, so do you mean to convey that they reduce (not add to) the carbon flux by this amount?

Response: We agree that the terminology is a bit mixed up and leaves too much room for interpretation, for example regarding which convention of positive flux is used. Indeed, fires of course reduce the net productivity by adding to the emission of carbon. We will clarify this accordingly in the text.

L326-328: Suggest adding an initial discussion paragraph before section 4.1 that summarizes the principal results and why they matter. This is particularly important given the lack of a conclusion section.

Response: Agreed, we will add this.

L355-356: Is there evidence for hydraulic redistribution in these Acacias? If so, the potential for impacts on grass-tree facilitation vs competition must be addressed.

Response: We are not aware of hydraulic redistribution in these trees, but we can shortly discuss this phenomenon and its potential impacts. In either case, our vegetation model is not able to simulate such detailed plant hydraulics.

367-368: This is an important result that could be elevated to the abstract and/or first discussion paragraph.

Response: Agreed, we will add it.

L372: Not currently corroborated at all four sites.

Response: We will discuss the results of the other sites into more detail.

L374-376: Recent work by Dannenberg et al. also showed negative asymmetry (with respect to precipitation variability) for trees in semiarid areas: MP Dannenberg et al. 2019. Reduced tree growth in the semiarid United States due to asymmetric responses to intensifying precipitation extremes. Science Advances.

Response: Thank you, we will look into this paper and add it to the discussion.

L389-392: Some kind of sensitivity analysis would go a long way toward reducing the uncertainty associated with this caveat e.g., a comparison of results from model runs where the texture was varied by increments of ~5% within some permissible range.

Response: Indeed, this would be a highly interesting topic for a follow-up study. In the current manuscript we already discuss several dimensions of variation, so including changes in soil texture would add even more results that need to be presented.

L393: “Variations between the different sites” have not been shown.

Response: We will discuss the results of the other sites, and the variations between them, into more detail (see previous answers).

L402-403: A citation would strengthen this claim

Response: This claim follows from a discussion with field-experts. As they are included as co-authors, this claim is for now given as such, but we will try to find relevant papers and add them.

References

Boulain, N., Cappelaere, B., Ramier, D., Issoufou, H. B. A., Halilou, O., Seghier, J., Guillemin, F., Oï, M., Gignoux, J. and Timouk, F.: Towards an understanding of coupled physical and biological processes in the cultivated Sahel - 2. Vegetation and carbon dynamics, *J. Hydrol.*, 375(1–2), 190–203, doi:10.1016/j.jhydrol.2008.11.045, 2009.

Tagesson, T., Ardö, J., Guiraud, I., Cropley, F., Mbow, C., Horion, S., Ehammer, A., Mougin, E., Delon, C., Galy-Lacaux, C. and Fensholt, R.: Very high CO₂ exchange fluxes at the peak of the rainy season in a West African grazed semi-arid savanna ecosystem, *Geogr. Tidsskr. - Danish J. Geogr.*, 116(2), 93–109, doi:10.1080/00167223.2016.1178072, 2016.