Interactive comment on "Asymmetric Responses of Primary Productivity to Altered Precipitation Simulated by Ecosystem Models across Three Long-term Grassland Sites" *by* Donghai Wu et al.

Responses to the Comments of the Anonymous Referee #1

Please note that our responses are shown in black while the comments of the reviewers are in blue.

This manuscript presents a study of modeled and observed grassland NPP variability across years and three sites (Konza Prairie, Stubai Valley and the Central Plains Experimental Range). Fourteen terrestrial ecosystem models are used to simulate these three sites and the analysis focuses on modeled and observed responses of NPP to precipitation variability, and the asymmetry of these responses, i.e. different magnitude NPP responses for equivalent increases and decreases in precipitation. To assess model responses to precipitation variability in more detail, simulations that alter precipitation across a range of values at all three sites are also conducted.

In general the manuscript is well written and the simulations are extensive and well executed. The introduction is strong and well written with many references and fairly clear definition of goals. Given the organization of the abstract and the introduction, I find the presentation of results is not well structured. Also, the application of statistics could be improved. The interpretation of the results in the discussion is not well executed and the conclusions regarding mechanism are not well tied to a strong understanding of the mechanisms encoded in the models.

Responses: We greatly appreciate the reviewer's pertinent feedback and valuable comments on our study. We thank you for your time and effort in helping us to improve this paper. In our revised version, we conducted a lot of work in responses to all these profound suggestions. In particular, we have (1) reorganized the main results in the abstract part; (2) reorganized the four specific objectives in the introduction part; (3) clarified the metrics of the response of productivity to precipitation changes following the four specific objectives in method part; (4) reorganized and strengthened the result part following the four specific objectives; (5) reorganized and strengthened the discussion part following the four specific objectives; and (6) rewrote the conclusion part to be more focused on our main results. We now consistently present first the comparison between spatial slopes and temporal slopes; which is followed by the asymmetric responses of productivities to precipitation under normal and extreme conditions using two indices (asymmetry index from inter-annual productivity and precipitation, and sensitivity of productivity to altered rainfall conditions). The curvilinear responses of productivities to altered precipitation across the three sites by each model are presented last. The first three specific objectives follow the structure used in Knapp et al. (2017), who have established a conceptual model for the precipitation-productivity relationships.

In addition, we have redrawn the figures to make them clearer, we have added uncertainty estimates for observation-based

asymmetry index, and we have also added uncertainty estimates by models. Detailed responses are as follows under each of your comments.

Knapp, A. K., Ciais, P., and Smith, M. D.: Reconciling inconsistencies in precipitation-productivity relationships: implications for climate change, New Phytologist, 214, 41-47, doi:10.1111/nph.14381, 2017.

The abstract and introduction are structured around the asymmetry of responses and spatial versus temporal differences in responses. Yet the results are organized first with the simulations that alter precipitation (Figs 1 and 2); then the analysis of modeled and observed responses to spatial and temporal variation in precipitation (Figs 3 and 4); and then back to the model results of altered precipitation (Fig 5). I suggest leading with Figs 3 and 4, the comparison of model results and observations. Then follow with the altered precipitation results. The results from altered precipitation simulations should be used to interpret the model and observation comparison.

Responses: We agree. In our revised version, we have reorganized and strengthened the results following the same logical structure as the introductory section. We now consistently present first the comparison between spatial slopes and temporal slopes; which is followed by the asymmetric responses of productivities to precipitation under normal and extreme conditions using two indices (asymmetry index from inter-annual productivity and precipitation, and sensitivity of productivity to altered rainfall conditions). The curvilinear responses of productivities to altered precipitation across the three sites by each model are presented last. The results have been improved in our revised version (page 8|line 8 to page 10|line 16).

Uncertainty and statistics must be presented for the observations. Figs 3 and 4 need uncertainty bars on the observations. Why were stats done on the asymmetry of the model ensemble and not the observations. Statistical analysis of asymmetry in the observations must be done. In my view the stats on the model ensemble are unnecessary.

Responses: We agree. In our revised version, we have added the uncertainty estimates of observation-based asymmetry index. The method has been clarified in our method part (page 6|line 24 to page 7|line 18).

In addition, we have added the uncertainty estimates of observation-based temporal slopes using a bootstrap sampling method with 1000 replicates of the ANPP and precipitation time series. We can obtain 1000 temporal slopes for each replicate and present the observed uncertainty ranges using interquartile spread of the temporal slopes between individual replicates (10^{th} and 90^{th} percentiles).

The discussion is poorly organized and in many cases only tangentially linked to the results presented in the study. The authors should rewrite the discussion, trying to avoid generality and link into their specific results.

There is an initial paragraph missing which summarizes the key result(s). The first paragraph of the discussion is mostly irrelevant, only the penultimate sentence relates to this particular study. The second paragraph of the discussion belongs in the

results. The third paragraph is reasonable but should come later when trying to interpret the discrepancies between modeled and observed results. Section 4.2 is reasonable and should lead the discussion, after the first paragraph suggested above.

Responses: We find the criticism of the reviewer fair and we have rewrote the discussion section in the revised manuscript according to the comments raised. We firstly discussed the modeled and observed responses of productivity to altered precipitation, including spatial slopes and temporal slopes, and asymmetric responses of productivities to precipitation; we then discussed the curvilinear responses of productivities to altered precipitation by models; at last, we discussed the uncertainties of the two indices, knowledge gaps and suggestions of further work. We have removed the irrelevant discussion. The discussion have been improved in our revised version (page 10|line 17 to page 13|line 24).

p12 |28 – were these simulations done without using soil texture form the sites? That is a major oversight. It is clear that soil texture is a key driver of water availability in these models and could be a major reason for the discrepancy. If this were a single model study I would insist, but the logistics of a model intercomparison are such that I won't insist. Nevertheless, the authors should consider redoing the simulations with the actual soil textures. If they opt not to do this then it needs to be made very clear in the methods that site soil textures were not used in the study and the discussion should interpret the results in this context in more detail.

Responses: Our previous wording in the manuscript created a misunderstanding by the reviewer. Model simulations were carried out using soil texture properties measured at each site. All models used the same properties as reported in Table 1. We have clarified that in the revised manuscript. We totally agree that without soil textures from the sites the study would be weak, which is not the case.

Section 4.3 is not really related to the study. None of the key conclusion in this section have been teased out from the analysis. Most of the recommendations are based on the literature cited in the section. What, in this study specifically, have you found the models are lacking and how can that be addressed? The authors need to do a much better job of identifying the short-falls in the models compared with the observations and providing a logical understanding of the causes of these shortfalls. Do some models perform better than others? If so, why?

Responses: We agree with the referee and have thoroughly revised Section 4.3. We first discussed the uncertainties of the two indices (asymmetry index from inter-annual productivity and precipitation, and sensitivity of productivity to altered rainfall conditions) to study the asymmetric responses of productivities to precipitation under normal conditions. Then we indicate important knowledge gaps that should be considered in our following model-experiment interaction studies. This is the first study where a large group of modelers simulated the response of grassland primary productivity to precipitation using long-term observations for evaluating the asymmetry responses to altered precipitation. Hence, a lot of additional work is needed to deal with the very specific problems that became clear from our exercise. The Section 4.3 have been improved in our revised

version (page 13|line 1 to page 13|line 24).

The conclusions are poor. The first sentence stating novelty is unnecessary. The first half of the second sentence is not what I take from the results. Fig 4 shows the models do a bad job of capturing asymmetry at SGS but not at KNZ or STU. The second half of the second sentence is primarily speculation. The third sentence is a throwaway and is unnecessary. The fourth sentence seems to suggest that the collaboration between site investigators and modelers in this study was not very strong. The fifth sentence is about extremes, which this study only tangentially addresses, responses to "normal" variability are the focus of the study.

Responses: In our revised version, we have rewritten the conclusion part to tight it much closer to our main results. The conclusion has been improved in our revised version (page 13|line 27 to page 14|line 11).

Minor comments:

p2 |3 – Unclear. Do you mean changes in variability or just variability?

Responses: This sentence has been removed in our revised version.

P2 |3-5 – Suggest moving this sentence to after the following one.

Responses: This sentence stated the asymmetric responses of ANPP to precipitation under normal and extreme conditions observed from previous field measurements is a topic sentence in this paper. It also introduced the subject of this study here. Thus, we have left it unchanged.

P2 |5, |15, |20 etc – What do you mean normally variable?

Responses: In this study, normally variable means conditions of normal variation (not extreme) in precipitation.

P2 |8 – you switch between using asymmetrical and non-symmetrical, pick one and stick with it

Responses: This expression has been revised to asymmetric.

p2 |13-15 – awkward sentence, rephrase

Responses: This sentence has been revised to: "The spatial slopes derived from modeled primary productivity and precipitation across sites were steeper than the temporal slopes obtained from inter-annual variations, which was consistent with empirical data."

p2 |17 – what do you mean consistently here, across what? Sites?

Responses: "consistently" has been removed in our revised version.

P2 |19-21 – Be more precise in what you mean here, what do you mean "extent of negative drought effects" and "impacts of increased precipitation" ? By "extent" and "impacts" do you mean different things?

Responses: Here, we meant the same thing. This sentence has been revised to: "Our results indicated that most models overestimate the negative drought effects and/or underestimate the positive effects of increased precipitation on primary productivity under normal climate conditions, highlighting the need for improving eco-hydrological processes in those models in the future."

P3 |10 and many other places – "P-ANPP" sensitivities, you are analyzing ANPP in response to precipitation, it is conventional therefore to put ANPP first. Change to "ANPP-P".

Responses: This expression has been revised to "ANPP-P".

p3 |17-20 – this is a reasonable argument but depends on time-scale

Responses: We agree, and we also expressed the idea here. To make it more clear, this sentence has been revised to: "For projecting the effect of climate change on grassland productivity in near to mid-term (coming decades), inter-annual relationships are arguably more informative than spatial relationships because spatial relationships reflect long-term adaptation of ecosystems, and because ANPP-P relationships from spatial gradients are confounded by the co-variation of gradients in other environmental variables (e.g. temperature and radiation) and soil properties (Estiarte et al., 2016; Knapp et al., 2017b)."

p4 |7 – quantify these rainfall regimes by adding MAP

Responses: This sentence has been revised to: "In this study, we aim to evaluate the responses of simulated productivity to altered precipitation from fourteen ecosystem models at three sites representing dry (304 ±118 mm yr^{-1}), mesic (827 ±175 mm yr^{-1}), and moist (1429 ±198 mm yr^{-1}) rainfall regimes."

p4 |13-15 – this sentence is unnecessary, delete it or move to the discussion

Responses: This sentence has been removed.

p4 |15-16 – this sentence is unnecessary, delete it

Responses: This sentence has been removed.

p8 |15 – must add uncertainty to the observations

Responses: Uncertainty for observation-based results has been added in our revised version.

p12 |8-10 – this sentence comes from nowhere, delete it

Responses: This sentence has been removed.

p12 |28 – this is unclear, do you mean use measured SWC as an input?

Responses: In this sentence, we mean that models should simulate SWC in the same soil layer as experiments in following studies, so that we could evaluate the modeled SWC compared to observations. This sentence has been revised to: "Models should report SWC at the same depth of experiments and experimental data should be made available for better comparisons in following studies. This can provide insights into the bias of modeled sensitivities to precipitation and check explicitly the sensitivity of vegetation productivity to change in SWC"

P12 |28 – rephrase "This will help in figuring out"

Responses: This expression has been removed.

Figure 1. Differences in x-axis scales should be noted in the caption.

Responses: This has been added in our revised version (page 26).

Figure 2. As above. Observations should be added to the ANPP plots.

Responses: This has been added in our revised version (page 27). Observations are shown in our first figure (temporal slope vs. spatial slope) in our revised version, and the curvilinear responses of productivities to altered precipitation by models have been moved to the last part.

Figure 3. Ho do you calculate a "mean slope"? Need uncertainty for observations. Technically standard deviation is not a measure of uncertainty, it is a measure of variability. I think your shading represents model variability.

Responses: Here, for each site, we firstly calculated the temporal slopes for each model under ambient simulation relating inter-annual variability in precipitation to inter-annual variability in the productivities using linear regression analysis. Then, we could calculate a mean temporal slope. In addition, we have presented the model uncertainty ranges using interquartile spread of the temporal slopes between individual simulations (10^{th} and 90^{th} percentiles). The uncertainty estimates for the observation have been added in our revised version (page 23).

Figure 4. The grey boxes are unnecessary, just use a black line if you want to show the mean/median. Change "pulses" to "gains." Need obs uncertainty.

Responses: The grey boxes are removed in our revised figure. Here, we chose similar expression as that in Knapp and Smith (2001) (relative ANPP pulse and relative ANPP decline). Thus, we have left it unchanged. The uncertainty estimates for the observation have been added in our revised version (page 24).

Knapp, A. K., and Smith, M. D.: Variation among Biomes in Temporal Dynamics of Aboveground Primary Production, Science, 291, 481-484, doi:10.1126/science.291.5503.481, 2001.

Figure 5. "1 sigma ~ 17 %" was this the same across sites? Again, do you mean uncertainty or variability?

Responses: In our revised figure, we have presented the variability of the three sites individually (1 sigma, 2 sigma and 3 sigma). In addition, we have presented the model uncertainty ranges using interquartile spread of the sensitivities between individual simulations (10^{th} and 90^{th} percentiles) (page 25).

Table 1. Add variability (standard deviation) to MAT, MAP, and ANPP. If soil texture was not used in the simulations make this clear in the caption.

Responses: The variability has been added in our revised version. In ecosystem model simulations, models have used the soil textures from the three sites as described in Table 1 (page 28).