Final author comments

Comments from Referees

Reviewer Comment 1

This paper attempts to demonstrate trends in different hydroclimatic variables and how they may relate to recent droughts in eastern Africa. While it is laudable that many (mostly model-based) time series have been used to address uncertainties, the documentation of these data is somewhat confusing, the presentation of results lack clarity and the interpretation/discussion of findings is rather vague. In general, the material is presented in a way that makes it hard to follow the implications of the chosen synthesis method, the differences among models and regions, and the overall conclusions; also,rather long-term trends than droughts are being analysed.

Response: Thank you for your insightful review. We have made substantial changes to the structure and have hopefully improved the readability and clarity of the manuscript.

Main comments:

Already the title is somewhat misleading, as you do not really analyse droughts but(modelled) annual soil moisture and climatic trends. I understand the argument that soil moisture and PET may be proxies for agricultural drought, but the connection of this analysis to drought and even food security is too vague. Moreover, the analysis of long-term annual trends probably tells little about the (shorter-term) droughts. The attempt to interpret the recent drought years as part of the overall trends is too limited. I'd suggest to rephrase the setup in terms of that you analyse hydroclimatic trends over the study region rather than suggesting the analysis is on droughts. Moreover, the study may better fit a specialised hydrological or climatological journal.

Response: Thank you for your comment. We now see that the title does not reflect the paper content well enough. In general 'droughts' are defined as 'below normal water', and when examining soil moisture this can be referred to as 'agricultural droughts'. From this perspective, we think the use of the word drought in the title and study is

appropriate, but we propose to also add the word ‘trends’ to the title as indeed the study concerns trends in drought rather than drought itself or specific episodes. We will also emphasize that while shorter-term drought can be severe, we choose to analyse trends in longer-term (annual) drought as the impacts are far reaching. Concerning the connection to food security, we will edit the paragraph at the top of page 3 - see the response to the following reviewer point. Section 3.2 entitled ‘Relating the results to recent droughts’ (now renamed as “Illustrative examples”) is actually intended as a presentation of illustrative examples of the method and not as an interpretation of the results. We want to show that known recent droughts that local readers will be familiar with do stand out (have multiple year return periods) following the annual averaging and that the starting month of the annual averaging has no influence on the resulting trend detected. We propose to change the subsection title and will also explain better the purpose of the subsection.

Changes:
Manuscript Title: Impact of precipitation and increasing temperatures on drought **trends** in eastern Africa
In the abstract, p.1 L3: In the current study we focus on **trends in long-term** agricultural drought
Section 3.2:

title changed to: Illustrative examples
See also responses below to questions on Methods section.

Introduction: lenghty, with some passages that do not straightforwardly lead to the study‘s objectives or promise too much. Specifically, I think, the statements on food production (p. 3 lines 3-15) are not needed; it could be much more straight forwardly said that you analyse four variables without attempting to construct such an argument(which you instead could shortly point to in the Discussion/Conclusions).

Response: Thank you, we see your point and we will reduce the length of the introduction and make it more focussed on the study’s objectives.
Changes: We will remove lines 5-11 from p.3. The paragraph will then be re-written. The purpose of the paragraph will be to briefly motivate the choice of study variables and express the wish to align the study variables as closely as possible with one of the major impacts of drought - reduced food security.

The paragraphs on p. 4 also belong rather to the Discussion. The study‘s leading questions should be much more concrete, and focused on the East African region.

Response: We agree. The section from p.3 L29 - p4. L15 is indeed lacking focus and contains elements that would be better in the discussion. Reviewer 2 is of the opinion that p.4 L6-14 (probably s/he means L6-15) are not closely related to the topic of the manuscript and we agree to remove that paragraph.

Changes: The paragraph will be used to highlight the use of PET and soil moisture in previous drought attribution/trend studies, with less or no detail on the outcomes of individual studies. p.4 L6-15 will be removed. The rest of this chunk will be moved to the discussion and also condensed.
The leading questions have been edited, as specified below.

Study region: It is not clear how the three criteria were applied: homogeneous precipitation (is it really homogeneous by the way, at what time scale?), livelihood zones, expert judgment? And is it really so that the final results are only aggregated for these6 zones, and only based on annual data? This should be said clearly early on, as it limits the scope of the analysis (while arguably increasing robustness).

Response: We have to strike a balance between the size of the regions and their homogeneity and did so under guidance of local experts. The annual mean precipitation as well as the seasonal cycle in precipitation is used to assess homogeneity in precipitation. For example, the WE box coincides with the wettest part of Ethiopia and is clearly distinct from the EE box in annual precipitation. Its southern boundary is fixed however at 7N rather than further south because south of this latitude the form of the seasonal cycle changes from a single to double peak in precipitation. The broad livelihood zones displayed in Fig. 1b were then consulted to check that the land use in the chosen regions was predominantly the same kind(s) and that the boundaries make sense. Box EE is the least homogenic, but we decided to keep this the same as in a separate already published study on Ethiopia (Philip et al., 2015), also because it was discussed with local experts. Experts from the National Meteorological Agency of Ethiopia (in Philip et al., 2015) and FEWS NET also reviewed and discussed our chosen regions in terms of homogeneity.
The final results are indeed based on annual time scales, with conclusions drawn for the six study regions individually. Averaging over large regions indeed makes results more robust. In our study region, averaging over smaller regions would result in too much

uncertainty in the results and climate models would be less able to capture the small areas. Longer droughts, spanning more than one growing season, have an impact on food security, and therefore we average over the year. This again is a compromise - the more growing seasons affected by drought, the larger the impact, but averaging over multiple years reduces the length of the data series and increases confidence intervals. Therefore we indeed aggregate over a time interval of one year and over each chosen region. We will make this clearer in the text.

Changes: We selected six regions based on precipitation zones *in which the seasonal cycle is homogeneous (Fig. 1a),* livelihood zones *(see Fig. 1b)*, and *discussions with local experts from Kenya Meteorological Department, and the National Meteorological Agency (NMA) of Ethiopia and* the Famine Early Warning Systems Network (*FEWS NET)***.**
Added: Data is annually averaged and spatially averaged over the study regions.

Datasets: It would be very helpful if there was a summary of the methodological approach in the very beginning of the Methods or as part of the Introduction. Figure 2 and the following text is not easy to follow; a well-structured and annotated table showing all data, acronyms, time periods and references would be way better.

Response: we agree it would increase readability if we mention the methodological approach further towards the beginning of the paper, in the introduction. Furthermore we moved some text following Figure 2 to the supplement as it is not relevant for understanding the results, and we turned the list below Figure 2 into a table. We however received positive comments on the figure itself and prefer to keep both the figure and table of data including references. The figure has the advantage that the connections between the datasets, i.e. which driving model/data sets feed which hydrological/impact models, as well as the number of runs, are easily visualised. We agree however that the description of the data below the table decreases readability and therefore does not belong to the main paper. We moved this to the supplement.

Changes:

- We added a paragraph before the current last paragraph of the introduction that reads:
  We have a stepwise approach to answer the attribution question:

1. Definition of the event and explanation of the study regions
2. Description of observational data and detection of trends in observations
3. Model evaluation including description of the models
4. Attribution of trends in models
5. Synthesis of the results

- In the methods section we added tables for the observational and model data sets instead of the itemized list, in which we will specify the short and full name of each data set, the time period and reference:

Observational dataset	**Full Name**	**Time period**	**Reference**
Gridded data set / reanalysis			
CenTrends (prcp)	...	...	...
...			
Observation-driven hydro/impact model			
LPJmL-WFDEI (soil moisture)	...	...	...
...			

Model data set	**Full name**	**Time period**	**Reference**
GCM/RCM			
GFDL (temp, prcp)	...	...	...
...			
Hydro/impact model			
H08 (soil moisture, PET)	...	...	...
...			

- To the first paragraph of the methods section we added “Furthermore, the two subsections describe (i) the assumptions and decisions that are made concerning the data/model setup and (ii) an example of how the method is applied to real data.”

Methods: Not clear to me why "global“ temperature is used and what the purpose of this analysis is. Is it done for all time series, and why not just use the original data? What "validation tests“ were done, and if they were more or less qualitative you may still have applied a quasi-objective criterion of whether the seasonal cycle "resembles“ the observational data (which actually). What is a return period for a specific year, 2018(e.g. Fig. 3)? Why does w@h require no fitting?

Response:

- we use GMST as a measure of anthropogenic climate change rather than just a trend over time. This is a common approach in attribution science. We added this to the text.
- The text reads "we check that the seasonal cycle resembles that of at least one of the observational datasets, in both the number and the timing of peaks." We thus require that the models broadly reflect the observed seasonal cycle. As there are sometimes also differences between observational data sets and, unless we have very good reasons, we do not rank one better than another, we do not use more objective sophisticated tests on the seasonal cycle than this. The fit parameters for the fit to GMST are assessed more objectively, as explained in the text.
- Concerning the comment on return period, in the method section, we explain that we evaluate the fitted distribution for the years 1900 and 2018, which means that for any threshold we can calculate a return period for the climate of 1900 and the climate of 2018. We added a sentence in the paper to explain this.
- The w@h does not require fitting as the large amount of data available for that model permits a direct estimation of the trend. This is already written on p 10 lines 12-14.

Changed: "We use global mean surface temperature (GMST) as a measure for anthropogenic climate change for calculating trends. We calculate trends for..."
Furthermore, after the sentence "In each case, the fitted distribution is evaluated twice: once for the year 1900 and once for the year 2018." we added "This allows us to calculate the return period of an event if it would have happened in the year 1900 or in the year 2018."

Page 13 point 4: So different time periods are mixed in your synthesis product? Doesn't that produce biases or at least merit discussion?

Response: Indeed the data sets do not all have the same time periods, but the data is first extrapolated onto the same time period (1900-2018), as mentioned here, before it is synthesized. However, there is no best way to tackle the problem of mixed products: models are framed differently and observational data has different lengths. Alternatives would be to restrict to the longest data set or use data with a common (shorter) time period, or not synthesize results at all. But this goes against our ethos. Our goal is to produce an overarching statement representing what we can conclude from a representative range of different available methods and data - i.e. methods which could

each have been used by others individually to report potentially conflicting messages in response to the same attribution (research) question. We want to use as much information as possible; only using the longest dataset or chosing just one framing in order to avoid this would not lead to more robust results, it would rather lead to an incomplete attribution result. We added this to item 4.

Added: We consider the use of all available observational and reanalysis data despite different model framings to lead to a more complete and robust attribution statement.

Section 3.2 is weak; I do not see a convincing approach to drought analysis here, and why is but one illustrative example explained which also only says that there is a marginally significant trend over the whole time period? This seems to be also something that should go to the Results.

Response: this section is added to illustrate the method and show an example. All other data sets are analysed following similar steps. In case this was not become clear in this section, we added a sentence on this. The trend results from these examples are already present in the results section, contained in the synthesis figures. The synthesis of observed and modelled trends is the main result and basis for the conclusions. Therefore we do not think the illustrative examples here are better moved to the results section.

Changes: This section will be retitled "Illustrative examples"
Added: In this section we show an example to illustrate the method of detection of trends in precipitation data, as droughts are often initially experienced as reduced or failed rainy seasons.

Results: I am not sure if this is the best selection of figures to portray results, and whether the set can be extended (note, the methods part has more figures than the main text's results part). Are also maps possible? Why focus text on the SS region only? In any case, the results section is way too short, and the reader gets lost on what figures, tables, data you refer to in the Results' text. A clearer presentation of key findings is needed, plus a more academic style (terms like "looking at" etc. should be more precise analytically). The order of presentation also need improvement, maybe variable after variable.

Response:

- We thought carefully about the selection of figures before submission and now have done so again, but we still think our selection of figures provides a good balance between simply showing the information needed to understand our

method along with the final results, and showing an overwhelming number of figures including intermediate results. We chose to explain the method (including figures) in the main text rather than referring to the supplement, which would in our opinion reduce readability. A description of the chosen figures is as follows:
    - Figure 1: regions and arguments for selection of the regions
    - Figure 2: illustration of the datasets
    - Figures 3 and 4: illustration of the method using two different types of data (transient and stationary)
    - Figure 5: illustration of the synthesis method including intermediate synthesis results
        - Supplement showing all other intermediate synthesis results
    - Figure 6: summary of all synthesis results
- Except for Figure 1 we do not show maps, as we are analysing time series of area-averaged quantities rather than spatial fields.
- As outlined in the text, the results in the "synthesis results" section are for all six regions. The synthesis figure 6 also shows these results for all six regions. It is just Figure 5 that focuses on the SS region, to illustrate the synthesis method that, for all regions and all variables, leads to the final synthesis statements. As the final synthesis statements are much more important than the step in between, that is shown in Figure 5, similar figures with intermediate results for the other five regions are only shown in the supplement. We direct the reader to the supplement to see these intermediate results.
- We have now made it clearer in the text that *intermediate* synthesis figures for all six regions can be found in the supplement, but intermediate synthesis figures are presented for one region (SS) in the main text to illustrate the synthesis method. As we already discuss findings per variable, we also added a couple of sentences outlining the structure of this section.

Changes:
The first paragraph of the synthesis section now reads: "In this section, to illustrate the synthesis method, intermediate synthesis figures are presented for the region SS for each of the four variables. See the caption of Fig. 5 for more information. The intermediate synthesis figures for all six regions can be found in the Supplementary Information. Table 3 and Fig.~6 summarize final synthesized findings for all regions. Using both the intermediate and final synthesis results we first draw conclusions based on different GCMs and hydrological models and then present conclusions per variable. "

Discussion: rather a list of shortcomings (which does not build trust in the analysis)than a discussion of the main findings and their relevance. Surely every

analysis has caveats, but in this paper the robust patterns need to be highlighted and then discussed in terms of their plausibility and potential further studies to be done as follow-up.

Response: The intended purpose of the discussion here (and in many other papers) is to discuss the main concerns and thus to what extent the reported results are sensitive to the choices and assumptions we have made, and to put the results in context of related studies. These choices and assumptions limit the study in the sense that they define what has been studied. They are not intended to be portrayed as shortcomings but rather as choices necessary (as in any study) to make it achievable, useful and appropriate, given the resources available.

We recognise, however, that the discussion lacks structure and is fragmented. As pointed out in this review process, other sections contain information more suited to the discussion. These paragraphs have been merged into the discussion and the structure will be sharpened up, with topics dealt with in a more logical order and long paragraphs condensed. With the revised structure the discussion is no longer a list of shortcomings but a more general overview of the context and the influence of our choices and assumptions on the results. Discussed topics include:

- The choice and definition of annual averaging scale: is the January-December definition appropriate? Would a different conclusion be reached using a sub-annual time scale?
- The potential influence of bias-correction on trends
- Our choice of model evaluation techniques in the light of recommendations from literature
- Our chosen approach towards communicating uncertainty of results
- The influence of the PET scheme on PET trends and the interpretation of PET trends in a water-limited regime, considering related studies
- The influence of (dynamic) vegetation schemes on drought trends, considering related studies and recommendations
- Factors beyond the scope of this study that may impact drought severity and food security

Conclusions: too long and not really conclusions but an extension of the Discussion.

Response: While restructuring the discussion section we moved some text from the conclusions to the discussion (i.e.. the discussion of food security and the use of

different PET schemes). In our opinion, the conclusions are now better structured and contain appropriate information.

Detailed/technical comments:

Abstract: Study period needs to be mentioned.
Response: Thank you for noticing.
Changes: To p.1 L3 we added "In the current study we focus on **trends in long-term** agricultural drought".

Line 12, "Nevertheless...", this info is not needed here.
Changes: "Using a combination of models and observational datasets, we studied trends, *spanning the period from 1900 (to represent the pre-industrial era) to 2018, …*"

Line 14/15, this is self-evident and no novel conclusion of this study I'd say.

Response: Whilst this conclusion might not be surprising in the light of other studies for different regions, it remains an important conclusion for Eastern Africa. This study was requested because the question of whether increasing temperatures are exacerbating drought keeps recurring.

Introduction: Line 30, GCM is the abbrevation for General Circulation Models.

Response: Thank you for noticing.
Changes: The correct expansion has been added.

Page 4 lines 20-27: can be deleted

Response: We are not sure whether the reviewer is really referring to p4 or not. Assuming he/she is: although it is not essential, for clarity we prefer to keep the paragraph (p4. L20-25) outlining the paper here, although L20-22 have been edited following Reviewer 2 specific point 4.

Datasets: what is the original spatial resolution of the different data, and how were they aggregated?

Response: We will refer the reader to references for the spatial resolution of the different data. In spatial aggregation, land grid points are weighted proportionally by the area represented.

W@home data: using the counterfactual climate dataset seems to make no sense here?

Added: Trends are calculated by dividing the difference in the variables between the present day climate and the counterfactual climate by the difference in GMST in the model in these two ensembles.

Page 8 line 27: why not shown, what sort of analysis is this?

Response: We analysed these datasets with different schemes to check the findings of Trambauer et al. (2014) also applies to our data. This is therefore not a novel idea nor a new finding, and besides we can not draw strong conclusions based on this that are relevant for the current analysis. Showing all details will distract the reader from the main findings. We therefore only mention that we checked this, but do not intend to include results.

Line33: what is refET?
Response: refET is daily reference evapotranspiration as mentioned in the data section.

Page 9 line 3-13: belongs to Discussion, as not studied here and probably not relevant for the historical time period.
Response: Thank you.
Changes: We indeed moved this paragraph to the discussion and shortened it.

Line 17: what is the relevance of the RCPs here, as you do not analyse future periods.

Response: It is not totally clear to us to which page this refers. However, differences in RCPs can account for uncertainty in the results, also for the near past. Between 2006 and 2018, there was a substantial increase in GMST and some spread in RCPs. Of course the difference would be larger if the analysis had extended to future periods.

Discussion:

page 19 line 17: where is this subannual analysis presented, and why not part of the Results (same for the analysis of PET differences, page 20 line 22)?

Response: We produced many more figures than those shown in this paper, e.g., the subannual analysis, the influence of different PET schemes on trends, the influence of different PET schemes compared to input datasets, the influence of using Jul-Jun instead of Jan-Dec etc. We think that presenting these extra analyses would add too much detail. We therefore only present the main findings and report the most important conclusions from additional analyses in only a few sentences. In doing so we keep the focus on the main findings.

Page2o line 21: uncertainties and origin are given: not so clear, and this is also in contrast to what is presented in Table 3.

Changed:

- We added: Rather, the uncertainties (confidence intervals) and their origin (e.g. natural variability or model spread) are given.
- we added “The table gives a concluding interpretation of the synthesized results shown in Fig. 6.”
- We added to the caption of Table 3 “The uncertainties associated with each result is depicted in Fig. 6.”

Reviewer Comment 2

In this paper, the authors obtained the sensitivities of soilmoisture, precipitation, potential evapotranspiration (PET) and local temperature to global mean surface temperature (GMST) from numerous datasets using statistical tools, and tried to explain the trend in soil moisture as a combination of trends in precipitation and potential evapotranspiration in eastern Africa. I believe that the authors did a lot of work to quantify the synthesized values of sensitivities, which may be helpful for the drought analysis of eastern Africa. However, as far as I am concerned, the writing of this manuscript need significant improvement, for example, the logical chain of the paper is poor; some expressions are not appropriate (i.e., temperatures?) and can be confusing to understand. Thus, as scientific research, it does need substantial improvements to presents a sufficiently significant advance to meet the ESD standards.

Response: Thank you for your thorough review. We will check the appropriateness of expressions used. We hope that with the responses given and the changes proposed will alleviate the main concerns and that the resulting revised manuscript will satisfy ESD standards.

Major points:

1. **The logical chain of the paper seems to be incorrect. The target of this paper is to investigate the impact of precipitation and temperature on drought, which, however, was not quantified in the paper. In fact, the authors only showed the sensitivity of soil moisture, precipitation, PET and local temperature to GMST without any details of the physical mechanism.**

Response: Thank you for drawing our attention to this. We now see that the title does not reflect the paper content well enough, and that our approach to assessing the link between trends in precipitation, temperature and drought is poorly expressed. Indeed we do not intend to examine the *mechanism* by which precipitation affects drought, but rather (i) to investigate if there is a signal of change in agricultural drought indicators and (ii) to investigate which global-warming driven trends in precipitation or local temperature explain any emerging trend in agricultural drought. We propose to add the word 'trends' to the title as indeed the study concerns trends in drought rather than drought itself or specific episodes. We hope this clarifies the confusion about the target. We also added this to the abstract. We will clarify that we do not intend to examine mechanisms but primarily to detect (necessarily using both observations and models) whether there are GMST-driven trends in drought indicators, including temperature and precipitation, and if trends in precipitation and/or temperature are related to trends in agricultural drought. To make the procedure in the method more evident we added a paragraph in the introduction explaining the steps in the method.

Changes:
Manuscript Title: Impact of precipitation and increasing temperatures on drought *trends* in eastern Africa
In the abstract, p.1 L3: In the current study we focus on *trends in long-term* agricultural drought
Introduction: we changed the sentence on the second objective to (ii) to investigate which global-warming driven trends in precipitation or local temperature explain any emerging trend in agricultural drought.

Introduction: added before the last paragraph:
We have a stepwise approach to assess the link between trends in precipitation, temperature and drought:

1. Definition of the event and explanation of the study regions
2. Description of observational data and detection of trends in observations
3. Model evaluation including description of the models
4. Attribution of trends in models
5. Synthesis of the results

2. **The method to quantify the sensitivity of different variables to GMST is unclear in the paper. As shown in line 27, page 9, "The method is extensively explained in van Oldenborgh et al. (2019) and Philip et al. (2019)", however, van Oldenborgh et al. (2019) is in review (line 29, page 28) and Philip et al. (2019) is in preparation (line 22, page 27). Therefore, I believe it's better to illustrate some necessary mechanism of the method in the paper.**

Response: the revision of both papers is nearly finalized but neither has yet been published, so we decided to additionally refer to two other published papers in which the method is also described well. These are van Oldenborgh et al. (2018) and van der Wiel et al. (2017).

van der Wiel, K., Kapnick, S. B., van Oldenborgh, G. J., Whan, K., Philip, S. Y., Vecchi, G. A., Singh, R. K., Arrighi, J., and Cullen, H.: Rapid attribution of the August 2016 flood-inducing extreme precipitation in south Louisiana to climate change, Hydrol. Earth Syst. Sci., 21, 897–921, https://doi.org/10.5194/hess-21-897-2017, 2017.

3. **As shown in Figure 3b, the 95% confidence interval for fitted location parameter of precipitation to GMST is quite large. I wonder how precise the sensitivity of precipitation to GMST in the paper is since even the paper itself referred to the fact that "the effect of a changing climate on precipitation is generally much less straightforward" in line 17, page 2.**

Response: the 95% confidence intervals are calculated using a non-parametric bootstrapping procedure, i.e., we repeat the fit a large number of times (1000) with samples of (covariate,observation) pairs drawn from the original series with replacement. This is discussed in the papers we refer to, but we now also added this to the papers. The effect of climate change is not straightforward as natural variability is simply very high. No trend is therefore yet emerging over noise.

Changes: in the methods section we added: “Confidence intervals (CI) are estimated using a non-parametric bootstrapping procedure.”

4. **The authors claim to use as many datasets as readily available, provided that the data are sufficiently complete over a long-enough time period. Moreover, there are different hydro/impact models being applied to simulate PET and SM. Two questions are raised here, first, since the accuracy of different datasets may vary spatially, is it reasonable to use as many datasets as readily available, particularly, without applying any additional bias correction (as suggested in Page 6 line 2); second, a very long paragraph is organized here to describe different projects and models, however, differences among these models are not highlighted and the reasons why these projects and models were selected are not clear. Section 2.2 needs serious revisions.**

Response:

1. Generally, our approach to attribution studies is to use and synthesize data that could have been produced by different teams separately, and to arrive at a conclusion based on a range of models and different (but compatible) framings of the research (attribution) question. However, we do reject models that are not fit for purpose in the validation step. Generally, we take the data as it comes and ideally as it would have been used in individual method studies, therefore including any corrections already applied to the data but not applying any more. Furthermore, in this case, we do not need a bias correction on the mean, as we are only looking at trends.
2. We use ISIMIP data because the ISIMIP project provides readily available model output of the variables under investigation. This is complemented by other readily available model runs with different (but compatible) framings. We will explain this in the text. The aim is however not to show differences between models, but rather to get a more complete answer on the attribution question. Different (types of) models could lead to different conclusions. With a multi-method and multi-model set-up study we make the attribution result more robust and thus gain confidence in the result.

We do however agree that for this purpose the section on data is rather long. We therefore moved part of the model descriptions to the supplement.

Changes:
After "we use as many datasets as readily available, provided that the data are sufficiently complete over a long-enough time period to be used for trend calculations"

(p5 line 6-7) we add "and, for model data, provided that the model data pass the validation tests".

5. **The result that 'Precipitation has a stronger influence on soil moisture variability than temperature or PET in the drier or water-limited region' seems to be one of the major conclusions in this study. In fact, there are studies revealing the fact that precipitation is more influential on soil moisture over dry regions and temperature is more influential on soil moisture in wet regions. The authors may need to highlight the novelty of this study in different ways.**

Response: We will express that this is the first multi-model attribution study on several drought estimates in a highly vulnerable area, addressing a recurring question on whether increasing temperatures exacerbate drought.

Some specific points:

1. **Page 1 line 5, we studied trends in six regions or four drought-related variables? I suppose they refer trends in four drought-related variables, however, the statement is not appropriate.**

   Response: We are not sure what the reviewer misunderstands here, or why s/he thinks the statement is not appropriate. The text does not read 'in six regions **or** four … variables' but 'in six regions **in** four … variables'. In case it is the formulation which is confusing, we propose to change the text.

   Changes: Using a combination of models and observational datasets, for six regions in eastern Africa we studied trends in four drought-related annually averaged variables.

2. **Page 4 lines 6-15, this paragraph doesn't seem to be closely related to the topic of this manuscript?**

   Response: we agree; the detail of this paragraph is more distracting than helpful.
   Changes: This paragraph is deleted.

3. **Page 4 line 24, a discussion and conclusions are... this is suggested to be changed to discussions and conclusions are…**

   Response: we agree ‘a discussion’ sounds strange. We will change ‘a’ to ‘the’, rather than ‘discussion’ to ‘discussions’ as then we still use the exact words in the section titles.
   Changes: The text will now read ‘... the discussion and conclusions are presented in …’

4.

5. **Page 4 line 27, in this section we show... this sentence can be moved to Line 22 before in Section 3 to keep consistency and avoid a one-sentence paragraph.**

   Response: It would reduce consistency to remove the sentence as we have such an introductory sentence at the beginning of each section (note, this doesn’t apply to subsections). A one-sentence paragraph should not in itself be a problem, however, it would make sense to reduce the information on Section 2 in the paper outline (in lines 20-22) and transfer that information to the beginning of Section 2.
   Changes: ‘The outline of the remainder of the paper is as follows: In Section 2 the chosen study regions are presented followed by a description of the datasets used in the study.’

   ‘In this section, we present the chosen study regions in eastern Africa and the datasets used to provide the four different variables (soil moisture, precipitation, temperature and PET) to be analysed. Brief descriptions of the (modelling) projects from which the datasets originate are provided’.

6. **Table 1 and Figure 1 have basically the same information, no need to keep both. Suggest keeping only Figure 1. The authors mentioned that six regions are selected based on livelihood, precipitation zones and local expert judgment, suggest clarifying these criteria clearly in Table 1.**

   Response: The reviewer is correct that there is overlap of information, however, we think it is helpful to retain both means of presenting the information. A map shows very quickly the spatial relation between the study regions but it is easier to read off their coordinates in a collective table. We now notice it would be better

to use the same names of livelihood type as in the key of Fig.1b. Local expert opinion did modify our original study zone borders, for example, the Kenyan Meteorological Department suggested a westward extension of the original NK box and an increased separation between the original box NK and CK, according to their understanding of climatological and agricultural zones in Kenya. Also, acting on advice of Ethiopians in an earlier study, we shifted the northern boundary of box EE from 14degN to 13degN. We think these details, however, are too much for the manuscript.
Changes: In table 1, make the 5th column conform to nomenclature in Fig1b, land-use type → livelihood zone, add column to summarise climatological precipitation for each region.

7. **In Table 3, the description of '-' (a negative trend) is missing. A small comment, I think table 3 may not be necessary here since similar information has been conveyed in Fig. 6.**

   Response: True, there is a negative sign in the table, and only the positive sign is explained. Whilst similar information is conveyed in table 3 and Fig. 6., we would argue for keeping both. The table summarises our interpretation of the numerical results in Fig. 6. The table is our conclusion on whether or not there are significant changes in the four variables in each region. Fig. 6 shows the numbers behind these conclusions and more importantly illustrates the uncertainties associated with each.

   Changes: We will explain the above, i.e. the purpose of table 3 and Fig. 6 in the text, as well as adding a description of the '-' sign in the caption of Table 3.
   To the caption of table 3 we will add that 'the uncertainties associated with each result are depicted in Fig.6'.

8. **Lines 6-7 in page 19, not clear**

   Response: Concerns the words "In this section, we discuss ways in which our chosen approach to studying drought in eastern Africa may have influenced the results obtained." In our opinion, the purpose of a discussion is to interpret the results in the light of (i) how choices that have been made impact the outcome, and (ii) how the results relate to those previous studies on similar topics. It is not obvious what the reviewer finds unclear in the lines mentioned. Either s/he does not think the sentence describes what we do in the discussion, or perhaps s/he doesn't understand what is meant by "chosen approach".

Changes: Incase the latter is true, we will change “chosen approach” to “choices and assumptions”.

9. **Page 20 line 22, we find that ... (Prudhomme et al., 2014). It is not clear whether the conclusion comes from the author or from other’s work.**

   Response: We see that the reason for including the reference is not clear at all. In a global study, Prudhomme et al., found that GIMs contribute more than GCMs to the uncertainty in projected changes in drought and the uncertainty associated with GIMs has been attributed to differences in the number and type of processes represented in the GIMs (e.g., water balance, energy balance) and to differences in the details of their implementations. They do not specifically talk about PET schemes, however, so the reference will be removed.

   Changes: Prudhomme reference removed.

10. **Page 20 line 31, it is therefore (of) high priority? And line 32.....has been to apply simple, these sentences seem problematic to me.**

    Response: “therefore high” or “therefore of high” are both fine so we can add “of”. We cannot see anything grammatically incorrect with line 32 “The approach taken in this paper has been to apply simple evaluation techniques to readily available data, in order to advance our current knowledge.”, however we can change it to “In order to advance our current knowledge, in this paper we applied simple evaluation techniques to readily available data”.

11. **The inconsistency between Figure and Fig. in the manuscript (e.g. Page15 lines 27-28).**

    Response: thank you for noticing.
    Changes: we will change instances of Figure to Fig., except where sentences begin with Figure(s).

12.

13. **For soil moisture and precipitation, both low extremes are targeted. Why the distribution functions are different?**

Response: We alluded to this in point 7 in the list of assumptions (section 3.1, page 14, line 7-10) but we could be more explicit in the main text as to why we use a specific distribution for a specific variable.
Changes: After inspection of whether a Gaussian or a General Pareto Distributions fits the observational or reanalysis data best, we use the following distributions:

14. **Are there any proofs suggesting that the CenTrends precipitation dataset is better than others?**

Response: This is our second assumption, but we will change that sentence slightly.
Changes: As was shown by Funk et al. (2015), the CenTrends precipitation dataset includes many different sources of precipitation data and more stations than most other datasets. We therefore assume for precipitation that the CenTrends dataset is superior to other datasets over our region of study. We therefore only use the CenTrends dataset for observations of precipitation.

Content in Section 3.1 is hard to follow due to the poor logic. Suggest reorganizing.

Response: Thank you for your suggestion to re-order this list into an order that makes more sense. We reorganised the assumptions, starting with all assumptions related to data issues, observational data and model data, continuing with assumptions that could impact the trend and finishing with the assumptions made on the fits. In the old numbering, the order of the list is: 2, 5, 3, 8, 11, 10, 9, 4, 1, 6, 7.