Major revision 1 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 have added 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 have edited the paragraph at the top of page 3 - see the changes made in 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 (i) that known recent droughts, particularly those that local readers will be familiar with, still stand out as extreme (have multiple year return periods) following the annual averaging and (ii) that the starting month of the annual averaging has no influence on the resulting trend detected. In addition to changing the subsection title we also now 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 investigate **trends in long-term** agricultural drought
- P3 L2-3 changed to: “Whilst short term single-season drought episodes can be severe, we choose to analyse changes in drought on annual rather than sub-annual time scales because the worst crises in food security in this region have occurred with multiple season droughts (Funk et al. 2015).”
- Title of subsection 3.2 changed to: Illustrative examples
- The explanation “in subsection 3.3 we provide an example of how the method is applied to real data.” is now inserted at the beginning of section 3,

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 have reduced the length of the introduction and made it more focussed on the study’s objectives.
Changes: We removed distracting lines from this paragraph on p.3. The remainder of paragraph has also been re-written. The purpose of the paragraph is now 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 have now removed that paragraph. The leading questions have been edited, and now read as below.

Changes:

- The paragraph is now used to highlight the use of PET and soil moisture in previous drought attribution/trend studies, with less detail on the outcomes of individual studies. p.4 L6-15 have been removed. Some of the preceding text has been moved to the discussion and also condensed.
- “... the objectives of this study are to (i) consider the attribution question ``do increasing global temperatures contribute to drier soils and thus exacerbate the risk of agricultural drought (low soil moisture) in eastern Africa?'' and (ii) to investigate if global-warming driven trends in precipitation or local temperature via PET explain any emerging trend in agricultural drought.”

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 from the National Meteorological Agency of Ethiopia. Experts from the Kenya Meteorological Department 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 area, averaging over smaller regions would result in too much uncertainty in the results and climate models would be less able to capture the small regions.
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 have edited the text to make this clearer.

Changes: We selected six regions based on precipitation zones, in which the annual mean precipitation and seasonal cycle are 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 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 the text about data projects 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 2-part 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: Our approach to attribution comprises the following steps: (1) Definition of the study variables 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, spatial resolution and references.

Observational dataset	**Full name**	**Time period used**	**Spatial resolution (°lat x °lon)**	**Reference(s)**
Observatational/reanalysis data set				
CenTrends (prcp)	Centennial Trends data set	1900–2014	0.1x0.1	Funk et al. (2015)
CRU TS4 (temp)	CRU TS4.01	1901–2019	0.5x0.5	Harris et al. (2014)
Berkeley (temp)	Berkeley Earth	1750–2019	1.0x1.0	Rohde et al. (2013b, a)
ERA-I	ERA-Interim	1979–2019	0.5x0.5	Dee et al. (2011)
Observation-driven hydro/impact model				
LPJmL-WFDEI (soil moisture)	Lund-Potsdam-Jena managed Land - WATCH-Forcing-Data-ERA-Interim	1971–2010	0.5 x 0.5	Bondeau et al. (2007); Rost et al. (2008); Schaphoff et al. (2013); Weedon et al. (2014)
PCRGLOB-WFDEI (soil moisture)	PCRaster GLOBal Water Balance model - WATCH-Forcing-Data-ERA-Interim	1971–2010	0.5 x 0.5	Sutanudjaja et al. (2018); Weedon et al. (2014)
CLM-ERA-I (soil moisture, PET)	Community Land Model version 4 - ERA-Interim	1979–2016	0.5 x 0.5	Oleson et al. (2010)
CLM-WFDEI (soil moisture, PET)	Community Land Model version 4 - WATCH-Forcing-Data-ERA-Interim	1979–2013	0.5 x 0.5	Lawrence et al. (2011); Weedon et al. (2014)
FLDAS (soil moisture)	Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System	1981–2018	0.1 x 0.1	McNally et al. (2017)
MERRA Ref-ET (PET)	Modern-Era Retrospective analysis for Research and Applications Reference Evapotranspiration	1980–2018	0.125 x 0.125	Hobbins et al. (2018)

Model dataset	Full name	Time period used	Spatial resolution (°lat x °lon)	Reference(s)
GCM/RCM				
GFDL	GFDL-ESM2M, Geophysical Fluid Dynamics Laboratory - Earth System Model 2M	1861–2018	2.02x2.5	Dunne et al. (2012, 2013)
HadGEM	HadGEM2-ES, Hadley Centre Global Environmental Model version 2-ES	1859–2018	1.25x1.88	Collins et al. (2011); Jones et al. (2011)
IPSL	IPSL-CM5A-LR, Institut Pierre Simon Laplace - CM5A-LR	1850–2018	1.89x3.75	Dufresne et al. (2013)
MIROC	MIROC5, Model for Interdisciplinary Research on Climate - version 5	1850–2018	1.4x1.4	Watanabe et al. (2010)
EC-Earth	EC-Earth 2.3	1850–2018	1.12x1.125	Hazeleger et al. (2012)
w@h (temp, prcp, soil moisture)	Weather@home	2005–2016 and counterfactual climate	0.11x0.11	Massey et al. (2015); Guillod et al. (2017)
Hydro/impact models				
H08 (soil moisture, PET)	H08	1861–2018	0.5x0.5	Hanasaki et al. (2008a, b)
LPJmL (soil moisture, PET)	Lund-Potsdam-Jena managed Land model	1861–2018	0.5x0.5	Bondeau et al. (2007); Rost et al. (2008); Schaphoff et al. (2013)
PCRGLOB (soil moisture, PET)	PCRGLOB-WB, PCRaster GLOBal Water Balance model	1861–2018	0.5x0.5	Sutanudjaja et al. (2018)
WaterGAP2 (soil moisture, PET)	Water Global Analysis and Progress Model version 2	1861–2018	0.5x0.5	Müller Schmied et al. (2016)

- To the first paragraph of the methods section we added “Furthermore, in subsection 3.2 we describe the assumptions and decisions that are made concerning the data/model setup and in subsection 3.3 we provide 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 and express changes in local variables with respect to GMST rather than just calculating a trend over time. This is a common approach in attribution science. We added this to the text. It is used for all transient model runs and observational time series.

- Concerning validation tests, the manuscript 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 of the observational data set(s). As there are sometimes also differences between observational data sets and, unless we have very good reasons otherwise, 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 as 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 mixing different time periods would not lead to more robust results, it would rather lead to an incomplete attribution result. We added this to item 4.

Added: However, we consider the use of all available observational and reanalysis data and 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: Section 3.2 was included in the manuscript to illustrate the method and show an example. All other data sets are analysed following similar steps. In case this was not clear enough in this section, we added a sentence to explain 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: Section 3.2 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 all intermediate results (an overwhelming number of figures). 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 the two different types of data used (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, which not only show the overall synthesis but also the results for individual models, 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 in our opinion 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
- 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) and reworded some paragraphs to make it clear that they concern concluding recommendations based on results. In our opinion, the conclusions now 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".

To p.1 L5 we added "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 12, "Nevertheless...", this info is not needed here.

Response and changes: This line has been left in, as it is an implication of our results. However we removed "as evaporation is water limited" from the sentence "However, the influence of these on soil moisture annual trends appears limited as evaporation is water limited", as this was not a direct result of our study.

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 similar studies for different regions, it remains an important conclusion for Eastern Africa. This study was requested (by CIFF, the Children's Investment Fund Foundation) 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), which is the outline of the paper, 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 have added the spatial resolution of each data set used to the data set table (Table 2 and 3 in section 2.2 of the revised manuscript). In spatial aggregation, land grid box values are aggregated by a simple average over the grid boxes. Over the coast the grid box values are included in the simple average if more than 50% of the grid box covers land, and are weighted by half if the center of the grid box lies on the coast.

Changed: the spatial resolution of each data set has been added to Table 2 and 3.

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

Thank you for the comment. Indeed the result from w@h is not a pure trend from a transient simulation but rather the difference in the variables between the present day climate and the counterfactual climate ensemble scenarios divided by the difference in GMST in the model in these two ensembles. We approach this study from an attribution perspective in which both transient simulations and factual/counterfactual simulations are commonly used. In essence we perform part of an attribution analysis i.e. the detection of trends in both observations and climate models but do not specifically link the study to a particular event severity. In the paper we now make clear how we calculate the trend from factual and counterfactual runs.

Added: P8. L10, now moved to supplement: "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 that the findings of Trambauer et al. (2014) also apply to our data. This is therefore not a novel idea nor a new finding, and besides we cannot 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. Note that this paragraph has since been moved to the discussion section.

Line33: what is refET?

Response: refET is daily reference evapotranspiration as mentioned in the data section and expanded on in the supplement.

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. Dynamical vegetation models indeed show significant changes in the future, but could also be responsible for some uncertainty in the modelled response of drought to climate change over recent years.

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 to p20. L21: Rather, the uncertainties (confidence intervals) and their origin (e.g. natural variability or model spread) are given.
- we added to p16. L10: “The table gives a concluding interpretation of the synthesized results shown in Fig. 6.”
- We added to the caption of Table 3 on p19 “The uncertainties associated with each result are 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 have checked the appropriateness of expressions used (our manuscript has been internally reviewed by an American native speaker) and have we made changes where we think necessary or where confusion may have resulted. Note that the expression “temperatures” is commonly used when referring to temperature in a general sense, e.g. “the effects of high temperatures during ...” is fine. 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 the 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 goal. We also added this to the abstract. We will clarify intend 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 if global-warming driven trends in precipitation or local temperature via PET 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 at least one of these 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).

Changes:

- "The method is extensively explained in Oldenborgh et al. (2019), Philip et al. (2019), Oldenborgh et al. (2018) and van der Wiel et al. (2017)."
- The second mention of references Oldenborgh et al. (2019) and Philip et al. (2019) on p9 L30 has been removed.

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 manuscript. The effect of climate change is not straightforward to detect in such a complex region with complex climate dynamics and very large natural variability. The same holds for time as a covariate. No trend is therefore yet emerging over noise. We realised that we didn't add anything in the manuscript about the bootstrapping procedure, so we will add a sentence on that.

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 have adjusted the text to explain this. 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:

- p5 L6-7: After "we use as many datasets as readily available, provided that (i) the data are sufficiently complete over a time period long enough to be used for trend calculations", we add "and (ii) the model data pass the validation tests".
- To p6 L5 we added: Many model simulations stem from the ISIMIP project, which provides output of the variables under investigation for four different impact models. These simulations are complemented by other readily available model runs with different (but compatible) framings.
- Information on model projects has been moved to the supplement.

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.

Changes: We added to the conclusions “In this first multi-model, multi-method attribution study using several drought estimates in eastern Africa, we address the recurring question on whether increasing global 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 has been 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:
   - To the introduction: '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.'

   - At the beginning of section 2: 'In this section, we present the chosen study variables and study regions in eastern Africa and the datasets used to provide the variables to be analysed. Brief descriptions of the projects from which the datasets originate are provided in the supplement'.

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, so we have modified the last column of the table. 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. Concerning the precipitation zones, there is non-overlapping information: the figure shows the spatial distribution of CenTrends annual average precipitation and in the table we now provide information on the seasonal cycle - single or dual peak month(s).
Changes: In table 1, we made the land-use column conform to nomenclature in Fig1b, and we rename land-use type as livelihood zone, and we added a column to summarise climatological seasonal precipitation cycle (titled seasonal precipitation peak(s)) 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 have added 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: This remark 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 global impact models (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 as proposed below.

    Changes: “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 have changed instances of Figure to Fig., except where sentences begin with Figure(s).

12. **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 Distribution fits the observational or reanalysis data best, we use the following distributions:

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

Response: This is our second listed 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 new list is: 2, 5, 3, 8, 11, 10, 9, 4, 1, 6, 7.

Impact of precipitation and increasing temperatures on drought trends in eastern Africa

Sarah F. Kew[1], Sjoukje Y. Philip[1], Mathias Hauser[2], Mike Hobbins[3,4], Niko Wanders[5], Geert Jan van Oldenborgh[6], Karin van der Wiel[6], Ted I.E. Veldkamp[1], Joyce Kimutai[7], Chris Funk[8,9], and Friederike E.L. Otto[10]

[1]Institute for Environmental Studies, Vrije Universiteit, Amsterdam, The Netherlands
[2]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[3]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado
[4]Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado
[5]Department of Physical Geography, Utrecht University, Utrecht, the Netherlands
[6]Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands
[7]Kenya Meteorological Department, Nairobi, Kenya
[8]U.S. Geological Survey Center for Earth Resources Observation and Science, Sioux Falls, South Dakota
[9]University of California, Santa Barbara, Santa Barbara, California
[10]School of Geography and the Environment, University of Oxford, Oxford, UK

Correspondence: Sarah F. Kew (sarah.kew@knmi.nl)

Abstract. In eastern Africa droughts can cause crop failure and lead to food insecurity. With increasing temperatures, there is an a priori assumption that droughts are becoming more severe, however, the link between droughts and climate change is not sufficiently understood. In the current study we ~~focus on~~ investigate trends in long-term agricultural drought and the influence of ~~high~~ increasing temperatures and precipitation deficits~~on this~~.

Using a combination of models and observational datasets, we studied trends~~in~~, spanning the period from 1900 (to represent the pre-industrial era) to 2018, for six regions in eastern Africa in four drought-related annually averaged variables — soil moisture, precipitation, temperature and, as a measure of evaporative demand, potential evapotranspiration (PET). In standardized soil moisture data, we ~~find~~ found no discernible trends. Precipitation was found to have a stronger influence on soil moisture variability than temperature or PET, especially in the drier, or water-limited, study regions. The error margins on precipitation-trend estimates are however large and no clear trend is evident. We find significant positive trends in local temperatures. However, the influence of these on soil moisture annual trends appears limited~~as evaporation is water limited~~. The trends in PET are predominantly positive, but we do not find strong relations between PET and soil moisture trends. Nevertheless, the PET-trend results can still be of interest for irrigation purposes ~~as~~ because it is PET that determines the maximum evaporation rate.

We conclude that, until now, the impact of increasing local temperatures on agricultural drought in eastern Africa is limited and we recommend that any soil moisture analysis be supplemented by an analysis of precipitation deficit.

1 Introduction

In eastern Africa, drought has occurred throughout known history ~~,~~ and the phenomenon has incurred significant impacts on the agricultural sector ~~, and hence on the regional population and~~ and the economy, particularly thorough threats to food security. It is therefore important to examine the role of anthropogenic climate change in drought, particularly in the face of the large-scale droughts of 2010/11, 2014 ~~,~~ and 2015 in Ethiopia, and the 2016/17 drought in Somalia, Kenya, parts of Ethiopia and surrounding countries, which have recently raised the spectre of climate change as a risk multiplier in the region.

~~The complexity of droughts poses particular challenges for their attribution.~~ Droughts are triggered and maintained by a number of factors and their interactions, including meteorological forcings and variability, soil and vegetation feedbacks ~~,~~ and human factors such as agricultural practices and management choices, including irrigation and grazing density (**?**). ~~There are four main~~ Accordingly there are several definitions of drought in common use (**?**): meteorological drought (precipitation deficit), hydrological drought (low streamflow), agricultural drought (low soil moisture) ~~,~~ and socioeconomic drought (including supply and demand). ~~In the current study we focus on agricultural drought and how soil moisture and evapotranspiration is influenced by precipitation deficits and high temperatures~~This complexity of droughts poses challenges for their attribution. It is not straightforward to disentangle these interacting factors, but over a long time period it may be possible that a signal can be detected.

~~With respect to~~ Previous attribution studies for eastern Africa have mainly focussed on meteorological drought drivers ~~, the link between drought and climate change is not sufficiently understood~~(precipitation deficit), with recent studies finding little or no change in the risk of low-precipitation periods due to anthropogenic climate change (e.g., **??**) . Some weather stations in eastern Africa have recorded a decrease in precipitation in recent years~~as regional temperatures have increased in line with the global mean. While the regional temperature increase can be attributed to anthropogenic climate change, the effect of a changing climate on precipitation is generally much less straightforward (?) , particularly in eastern Africa. Climate~~ , however climate models generally project an increase in mean precipitation but give conflicting results for the probability of very dry rainy seasons (e.g. **?**). ~~Recent studies have found little or no change in the risk of low-precipitation periods due to anthropogenic climate change (e.g., ??) .~~ The reasons for the recent observed decrease in precipitation thus remains unclear, but the trend is within the large observed natural variability in the region, at least for the historical and current climate.

However, precipitation only covers one aspect of drought — that of the supply side of the water balance. The demand side is represented by actual evapotranspiration (ET), which is a function of moisture availability and evaporative demand~~(also referred to as "potential evapotranspiration" PET). PET is itself a function of temperature, humidity, solar radiation, and wind speed.~~ . With increasing temperatures, there is an a priori assumption that rising ~~PET~~ evaporative demand will increase the demand side of the water balance and, all else equal, droughts will become more severe. However, this assumption is not based on analyses, which motivates an objective study.

~~Aside from meteorological drought studies, the majority of drought studies on eastern Africa have focused on hydrological droughts. ? analysed high and low river flows including those on the Blue Nile. They found no significant change in low flows, but point out that uncertainties are large for low flows due to uncertainties inherent in Global Climate Models (GCMs), climate~~

scenarios, and hydrological models. This was confirmed by **?** , who identify no clear trends in low flows in the Blue Nile basin. **?** studied low flows on the global scale using a combination of models, comparing 2070–2099 to 1976–2005, and found little change in or reduced occurrence of drought projected across the Horn of Africa.

In the current study we wish to align our drought definition as closely as possible with the major impact of drought — the threat to food security. Across eastern Africa, agricultural drought is intimately linked to the quality and quantity of food production for domestic consumption . In our analysis we focus on annual drought, as the worst crises in food security in this region have occurred with multiple season droughts (**?**) . Drivers of food security include precipitation and precipitation-related factors like reduced food production, low livestock prices, high food prices, and reduction in food access. Non-climatic factors are also important; these include population change, land-use change and change in water demand. **?** , and **?** show that vulnerability has a large influence on food security, e.g., socio-economic factors, stability, and the ability to trade with other regions. Food-security outcomes differ dramatically between different regions and between pastoralist and non-pastoralist livelihood zones: pastoral and agro-pastoral populations are more often food insecure than non-pastoral populations. Not only are these non-climate-related factors often more significant than climate (change) in determining food security, but they are often unpredictable. Thus, attributing crop growth to climate change is intimately linked to agricultural conditions. We therefore use the agricultural definition of drought — i.e., isolating the influence of climate change on crop growth — can be challenging and does not tell the full story. **?** argue that, if a good observational basis of factors influencing food security exists, the detection and attribution of impact-related variables to climate change is possible. If not, it is more informative to perform the climate change attribution on a variable less directly related to food security but for which the best observational basis exists. In this paper we aim for an intermediate step: the choice of variables is a compromise between the best-observed variable of relevance (precipitation) and the variables best related to food security/crop growth.

We study agricultural drought – defined as low soil moisture – as — because soil moisture is a better indicator of crop health than precipitation and is an important indicator alone and embodies the net effect of the supply and demand side of the water balance, in regions without irrigation. We discuss Whilst short term single-season drought episodes can be severe, we choose to analyse changes in drought on annual rather than sub-annual time scales because the worst crises in food security in this region have occurred with multiple season droughts (**?**) . We will also investigate the influence of meteorological drivers on soil moisture the main meteorological drivers of soil moisture trends, i.e. , precipitation and temperature.

— Ideally, we would study the influence of temperature on soil moisture via evapotranspiration (ET), however observational records are very limited in time and space and as the spatial correlation decorrelation lengths of evapotranspiration are short their informational value is limited. To more directly explore the influence of temperature on soil moisture we We therefore analyse evaporative demand, choosing as a measure PET(which is also referred to as “potential evapotranspiration” PET. PET is the amount of evaporation that would occur if an unlimited supply of water were available), which is calculable / or available for both observations and model simulations . Hydrological models driven by GCM or reanalysis data usually either take PET directly as an input or derive PET from the different components required by the chosen PET scheme. The theoretically maximum evaporation, as initially given by PET, is then limited by available soil moisture to simulate actual ET.

Evaporative demand can be regarded as even more important than soil moisturefor regions with irrigation and is a function of temperature, humidity, solar radiation and wind speed.

We investigate evaporative demand as a means to study the influence of temperature on soil moisture, however, for regions that are irrigated or where irrigation is ~~considered. Therefore, a step between attribution of food security and attribution of precipitation is the attribution of soil moisture and additionally PET .~~ being considered, evaporative demand itself can be regarded as more relevant than soil moisture as a measure of drought tendency.

Whilst attribution studies specifically for the east African region have not previously used soil moisture or PET to explore drought, PET has been used in various attribution or trend studies outside our region of study, to explore ~~specific events or causes of trends or changes in variability. For China, PET was used to study~~ for example, the influence of climate change on ~~the~~ hydrological cycle in China (e.g. **???**)~~. **?** study~~ , trends and variability ~~in PET at different~~ at sites in West Africa ~~, showing high variability of PET when using Penman-Monteith PET and station data for six sites in Benin. **?** used sites in Europe to study~~ **?** and compound events of low precipitation and high PET ~~. One of their findings is that while precipitation has the most influence on soil moisture for both wet and dry regions, the influence of PET on soil moisture differs between regions: in wet regions, high values of PET, signifying atmospheric conditions conducive to evaporation, can amplify low soil moisture anomalies during drought, whereas in dry regions where very little surface moisture is available for evaporation, the magnitude of PET has little influence on soil moisture.~~

? choose a drought index that does not include PET when assessing the utility of precipitation observations to anticipate food insecurity in eastern Africa, as including PET made little difference to the results and most observational data (e.g. temperature data) required to calculate PET is likely to be less accurate than observed precipitation data in this region. in Europe **?** .

~~**?** examined the efficacy of evaporative demand (E_0) in assessing drought projections across the generally water-limited US Northern Great Plains, comparing the variability in projected drought risk from temperature-, radiation- and Penman-Monteith-based E_0 parameterizations, two drought indices, and two CMIP5 climate driver sets. They found much larger E_0 trends in temperature- and radiation-based E_0 estimates, and a greater uncertainty due to the choice between these than to the choice of GCM drivers. They also found significant difference in projected drought risk as estimated by E_0-based and a combined precipitation-E_0 index (i.e., the SPEI). Within eastern Africa, **?** decomposed the variability of E_0 (using the Penman-Monteith reference ET dataset used in this study), and found that the dominant driver of daily E_0 variability varied across time and space, with all four drivers (temperature, solar radiation, humidity, and wind speed) variously dominating across the continent, but particularly across eastern Africa, where the spatial picture strongly reflected the heterogeneous topography and hydroclimate and the external sources of variation in the drivers.~~

~~The~~ Summarizing, the objectives of this study are to (i) consider the attribution question "do increasing global temperatures contribute to drier soils and thus exacerbate the risk of agricultural drought (low soil moisture) in eastern Africa?" and (ii) ~~more fully understand the interplay between temperature and precipitation and their influence on~~ to investigate if global-warming driven trends in precipitation or local temperature via PET explain any emerging trend in agricultural drought. Our approach to attribution comprises the following steps: (1) Definition of the study variables 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. Assessments will be based on both observations and climate and hydrological ~~models~~model output on the annual time scale. We will ~~relate the findings to~~ illustrate the method using examples of recent droughts in eastern Africa.

~~In the following sections we first describe the different study regions in eastern Africa and~~ 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 ~~to provide the four different variables we analyse — soil moisture, precipitation, temperature, and PET. We include brief descriptions of the (modelling) projects from which the datasets originate.~~ in the study. In Section **??** we describe the ~~methods~~ stepwise approach to attribution used in this paper, including ~~validation of the data and calculation of trends. Sections **??** and **??** list the~~ assumptions and decisions ~~and an example of the method, respectively. Next~~made and illustrative examples. In Section **??**, the results are synthesized per region. Finally, ~~a~~ the discussion and conclusions are presented in Sections **??** and **??**.

2 Study variables, region and datasets

In this section~~we show the regions under analysis, list~~ , we present the chosen study variables and study regions in eastern Africa and the datasets used ~~and summarize their advantages and drawbacks.~~

2.1 ~~Study region~~

~~Trend~~ to provide the variables to be analysed. Brief descriptions of the projects from which the datasets originate are provided in the supplement.

2.1 Study variables and region

We analyse four different variables: soil moisture, precipitation, temperature, and PET. We average these variables over six regions, as trend analyses of time series of regionally averaged quantities are more robust than the same analyses for point locations. This is especially true for precipitation, which shows small-scale spatial variability if the time period is not long enough to sufficiently sample the distribution from multiple precipitation events. It is however necessary to select homogeneous zones, so that the signals present are not averaged out.

The focus of the study is on eastern Africa — Ethiopia, Kenya ~~,~~ and Somalia (including the Somaliland region). We selected six regions based on ~~livelihood zones, homogeneous precipitation zones and local expert judgement~~precipitation zones, in which the annual mean precipitation and seasonal cycle are homogeneous (Fig. **??**a), livelihood zones (see Fig. **??**b) 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). The regions are shown in Fig. **??** and listed in Table **??**. Data is annually and spatially averaged over the study regions.

2.2 Datasets

Table 1. The six study regions. See also Fig. **??**

Region	Long name	Latitude	Longitude	~~Primary land-use type~~Seasonal precipitation peak(s)	Primary liveliho
WE	West Ethiopia	7°N-14°N	34°E-38°E	Aug	agropastoral/mix
EE	East Ethiopia	8°N-13°N	38°E-43°E	Apr, Jul/Aug	pastoral
NS	North Somalia/Somaliland region and East Ethiopia	5°N-12°N	43°E-52°E	Apr/May, Oct	pastoral
NK	North Kenya	2°N-4.5°N	34°E-41°E	Apr, Oct/Nov	pastoral
CK	Central Kenya	1.5°S-1.5°N	35°E-38.5°E	Apr, Nov	agropastoral/mix
SS	South Somalia	2°S-5°N	41°E-48°E	Apr/May, Oct/Nov	pastoral/agropast

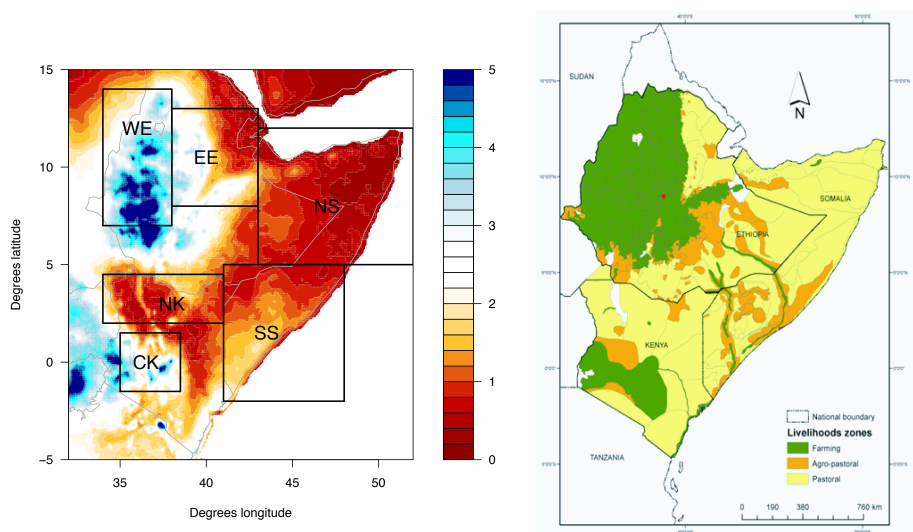

Figure 1. Left: annual mean precipitation [mm/day] and the six study regions. Note that only land values are used. Right: livelihood zones after **?**, which were also used to define the study regions.

We analyse four different variables: soil moisture, precipitation, temperature, and PET. Soil moisture and PET are estimated using data-model chains. For the four study variables, we use as many datasets as readily available over the study area, provided that (i) the data are sufficiently complete over a long-enough time period time period long enough to be used for trend calculations , and (ii) the model data pass the validation tests (see Sect. **??**). For this purpose, we decided to use time series of 35 years and longer. As the focus of this paper is on annual time scales, using monthly data is sufficient. The observational and model datasets used in this study are shown in Fig. **??** and listed belowin tables **??** and **??** below. For brief descriptions of the projects from which these data originate, please see the Supplement. Brief project and (model) data descriptions are given below the list. Note that we use the data as it is available without applying any additional bias correction. Some of the data has undergone bias correction in the projects, as listed belowwithin project of origin, as described in the Supplement.

Datasets used in this paper. Top: observational precipitation (prcp) and near-surface temperature (temp) datasets, bottom: models. Listed under PET is the PET scheme (T: Priestley-Taylor, M: Penman-Monteith, H: Hamon, B: Bulk formula) and, under SM, is the depth of the top soil moisture layer available (RD: depends on rooting depth (0.1-1.5m for WaterGAP2); IL: integrated over all layers). Shading indicates an experiment with either multiple input datasets or multiple hydrological models. The number of resulting hydrological model simulations are indicated by horizontal lines on the right side of the figure.

For observations and reanalyses we use (see below for more detailed explanations): For soil moisture: CLM-ERA-I, CLM-WFDEI, FLDAS (?) and because soil moisture is only calculated with models we additionally used the two available ISIMIP runs with WFDEI reanalysis input (PCRGLOB-WFDEI **?** and LPJmL-WFDEI). For PET: CLM-ERA-I, CLM-WFDEI, MERRA RefET (**?**) , ERA-I (**?**) For precipitation : CenTrends (**?**) For near-surface temperature: CRU TS4 (**?**) , Berkeley (anomalies), ERA-Interim (**?**) For model data the following simulations are used: 16 (4 GCMs x 4 hydrological models) transient runs from the ISIMIP ensemble that have data available for both soil moisture and PET. For observations of precipitation and daily mean near-surface temperatureand precipitation series from the 4 GCMs that are used in these ISIMIP runs. The combination EC-Earth - PCR-GLOBWB for all four variables The weather@home ensemble for near-surface temperature, precipitation and soil moisture

The reanalysis and model data are obtained from different model projects. A brief description of these projects and models is given below, see the references herein for more details, we use gridded observational data sets and reanalyses.

The Community Land Model version 4 (CLM, **??**) was forced with 6-hourly atmospheric forcing from two datasets. Once with raw ERA-Interim data (CLM-ERA-I), and once with bias-corrected ERA-Interim data from the Water and Global Change (WATCH) project; i.e., the WATCH-Forcing-Data-ERA-Interim (WFDEI) (**?**) (CLM-WFDEI). CLM is the land component of the Community Earth System Model and has 10 hydrologically active soil layers with exponentially increasing thickness, as well as a groundwater module.

The daily reference ET (RefET; **?**) used in this study was generated by the US National Oceanic and Atmospheric Administration from hourly drivers provided by the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) of the US National Aeronautics and Space Administration (NASA) Global Modeling and Assimilation Office (GMAO). Hourly MERRA-2 drivers for 2-m air temperature, 2-m specific humidity, surface pressure, 2-m wind speeds, and downwelling shortwave radiation at the surface were aggregated to daily input to the American Society of Civil Engineers Standardized

Reference ET Equation ? ; identical to the international standard FAO-56 report (?) at the daily timescalefor a 0.12-m short grass reference crop. The reference ET data are available daily from January 1, 1980 to within a few weeks of the present at a resolution of 0.125, which was downscaled from the native MERRA2 resolution of 0.5latitude × 0.625longitude ().

Within the second phase of the Inter-Sectoral Impact Model Intercomparison project (ISIMIP2 ?) , global hydrological models are used for intercomparison of climate impacts. The model simulations use input from GCMs or observations, e.g., precipitation and PET. For more details on the ISIMIP project, please refer to ? and ? . We used modelled soil moisture (0.5× 0.5spatial resolution) for the period 1850/1861–2018 from four global hydrological models: H08 (??) , LPJmL (???) , PCR-GLOBWB (???) and WaterGAP2 (?) . The GCMs used are GFDL-ESM2M (GFDL, ??) , HadGEM2-ES (HadGEM, ?) , MIROC5 (MIROC, ?) , IPSL-CM5A-LR (IPSL, ?) . All simulations were carried out under the modelling framework of phase 2b of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b: ?, https://www.isimip.org/protocol/#isimip2b) . For the ISIMIP models we also have access to the adjusted (i.e., bias corrected for the ISIMIP project) GCM data that is used as input for the hydrological For soil moisture and PET, no direct observations meeting the above criteria exist. Instead, we use observational estimates of soil moisture and PET resulting from various combinations of observational forcing data and models (see also Section ??).For precipitation and temperature, the original GCM data is analysed for trends.

For weather@home we use the large-ensemble regional modelling approach as in (?) employing the distributed computing framework climateprediction.net (??) . For this study, two large ensembles of simulations of temperature, precipitation and soil-moisture are available for the present day climate (2005–2016) and acounterfactual climate representing how conditions might have been without anthropogenic greenhouse gas and aerosol emissions for the same time-frameFig. ??a).

EC-Earth - PCR-GLOBWB has been developed using large-ensemble simulations of EC-Earth (?) in combination with the PCR-GLOBWB model (?) . The EC-Earth - PCR-GLOBWB (EC-PCRGLOB)ensemble was originally developed by ? to study changes in hydrological extremes.

There is little available observational data for Concerning soil moisture, with most series being observational series are few and generally too short to use for trend analysis . ? used CCI satellite-derived microwave soil moisture data to show that observational, reanalysis and model datasets for soil moisture are not always highly correlated with each other and they do not correlate well with reanalysis or model data over eastern Africa . Using a set of models and observations or assimilated data for soil moisture is therefore always necessary to (?) . It is therefore important to use multiple observationally forced model estimates to span the large uncertainties from inter-dataset differences. There being no a priori reason to favour one soil moisture dataset over another, we treated treat all resulting soil moisture datasets equally. For soil moisture For all soil moisture data sets, observed and modelled, we use the topmost layer (see Fig. ?? for the depth of the topmost layer) provided by each dataset and scale each time series to have a standard deviation of 1 in order to make comparisons in trends possible. An exception to this is weather@home where the available soil moisture variable is an integrated measure of all four layers of soil moisture in the model, including the deep soil.

? show that the influence of different choices in the calculation of PET and soil moisture can be quite large, with two of the choices being the input dataset for the hydrological model and PET scheme. ? also showed that both the scheme used to calculate PET and the input data used for calculation of PET have a large influence on PET values. We confirm this using the

CLM-ERA-PT (Priestley-Taylor), CLM-WFDEI-PT and CLM-ERA-PM (PET is a function of temperature, humidity, solar radiation and wind speed, and as such is not a directly observable variable. Observational estimates of PET used here originate from reanalysis data sets or reanalysis-driven impact models. For both observed and modelled PET, there are various ways of parametrizing PET, ranging from simple temperature or radiation-based schemes to sophisticated schemes based on all the aforementioned components. Whilst the Penman-Monteith) datasets (not shown). In our study regions, PET values are consistently higher when using PM then when using PT. The differences in trends in PET using ERA or WFDEI input or using PT or PM input are sometimes significant. However, comparing study regions, there is no consistency in the difference; in four out of the six regions the PM data shows a higher trend than the PT data and in four out of the six regions WFDEI data shows a higher trend than the ERA data. We thus chose to use a variety of PETparameterizations and input datasets in order to cover the range of possible PET values and trends in PET. If several schemes are available there is an a priori preference for Penman-Monteith PET scheme , as demonstrated by ? , who showed that the dominant driver of RefET variability is often neither temperature nor radiation. Howeverscheme is often considered superior (e.g. ?) , one is often constrained from using a Penman-Monteith parameterization due either to the lack of accurate or reliable input data or because the choice of PET parameterization within a given hydrological model setting is already prescribed, as in the ISIMIP ensemble. We thus chose to use a variety of PET parameterizations and input datasets in order to cover the range of possible PET values and trends in PET. The PET scheme used by each data set is noted in Fig. ??.

It is still unknown how vegetation will respond to substantial increases of CO_2 concentration. Two counteracting effects — physiological and structural responses — are expected, but their net effect is unknown (e.g. ?) . The physiological response is that increased CO_2 results in increased water use efficiency (WUE), that is reduced water loss through evapotranspiration by reducing stomatal openings.The structural response is increased growth, which may lead to increased leaf area and thereby increased water loss through evapotranspiration (?) .

So-called 'dynamic vegetation models' include these CO_2 effects, whereas others do not. ? studied uncertainties of irrigation water demand under climate change and found much lower demand is simulated by global impact models that model Concerning model data sets, most simulations stem from the ISIMIP project, which provides output of the CO_2 effect than those that do not. ? found that models that include the CO_2 and vegetation effects show a weaker response of global hydrological drought to climate change. They therefore advise users to ensure that the impact (hydrological) models they select must capture the diverse ways that models represent the response of plants to CO_2 and climate — the largest source of uncertainty in their study. In this study we use different combinations of observational input datasets, GCM input and hydrological models. Our selection of hydrological models is restricted by the variables we require, however, out of the four ISIMIP hydrological models that match our criteria, one (LPJmL) uses dynamic vegetation modelingvariables under investigation for four different impact models driven by four different GCMs. These simulations are complemented by other readily available model runs with different (but compatible) framings.

With the datasets we use we cover a wide range of different factors that influence PET and soil moisture. The different factors include meteorological forcing, model choice, RCP scenario for the greenhouse gas concentration trajectory, PET

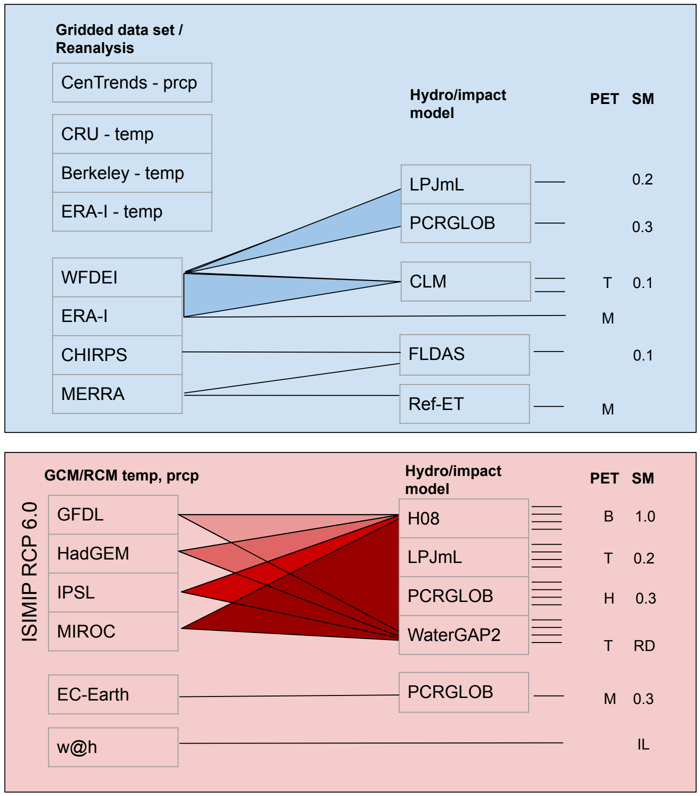

Figure 2. Datasets used in this paper. Top: observational precipitation (prcp) and near-surface temperature (temp) datasets, bottom: models. Listed under PET is the PET scheme (T: Priestley-Taylor, M: Penman-Monteith, H: Hamon, B: Bulk formula) and, under SM, is the depth of the top soil moisture layer available (RD: depends on rooting depth (0.1-1.5m for WaterGAP2); IL: integrated over all layers). Shading indicates an experiment with either multiple input datasets or multiple hydrological models. The number of resulting hydrological model simulations are indicated by horizontal lines on the right side of the figure.

scheme, number of soil layers and depth of topsoil layer, dynamic vegetation modelling (LPJmL only) and transient versus time slice runs (see next section on 'Methods').

3 Synthesis results

In this section~~synthesis figures~~ , to illustrate the synthesis method, intermediate synthesis figures, which not only show the overall synthesis but also the results for individual models, are presented for the region SS for each of the four variables. See the caption of ~~Figure~~Fig. **??** for more information. The intermediate synthesis figures of all six regions can be found

Table 2. Observational data used in this study.

Observational dataset	Full name	Time period used	Spatial resolution (°lat x °lon)	Reference(s)
Observatational/reanalysis data set				
CenTrends (prcp)	Centennial Trends data set	1900–2014	0.1x0.1	?
CRU TS4 (temp)	CRU TS4.01	1901–2019	0.5x0.5	?
Berkeley (temp)	Berkeley Earth	1750–2019	1.0x1.0	??
ERA-I	ERA-Interim	1979–2019	0.5x0.5	?
Observation-driven hydro/impact model				
LPJmL-WFDEI (soil moisture)	Lund-Potsdam-Jena managed Land - WATCH-Forcing-Data-ERA-Interim	1971–2010	0.5 x 0.5	????
PCRGLOB-WFDEI (soil moisture)	PCRaster GLOBal Water Balance model - WATCH-Forcing-Data-ERA-Interim	1971–2010	0.5 x 0.5	??
CLM-ERA-I (soil moisture, PET)	Community Land Model version 4 - ERA-Interim	1979–2016	0.5 x 0.5	?
CLM-WFDEI (soil moisture, PET)	Community Land Model version 4 - WATCH-Forcing-Data-ERA-Interim	1979–2013	0.5 x 0.5	??
FLDAS (soil moisture)	Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System	1981–2018	0.1 x 0.1	?
MERRA Ref-ET (PET)	Modern-Era Retrospective analysis for Research and Applications Reference Evapotranspiration	1980–2018	0.125 x 0.125	?

Table 3. Model data used in this study.

Model dataset	**Full name**	**Time period used**	**Spatial resolution (°lat x °lon)**	**Reference(s)**
GCM/RCM				
GFDL	GFDL-ESM2M, Geophysical Fluid Dynamics Laboratory - Earth System Model 2M	1861–2018	2.02x2.5	??
HadGEM	HadGEM2-ES, Hadley Centre Global Environmental Model version 2-ES	1859–2018	1.25x1.88	??
IPSL	IPSL-CM5A-LR, Institut Pierre Simon Laplace - CM5A-LR	1850–2018	1.89x3.75	?
MIROC	MIROC5, Model for Interdisciplinary Research on Climate - version 5	1850–2018	1.4x1.4	?
EC-Earth	EC-Earth 2.3	1850–2018	1.12x1.125	?
w@h (temp, prcp, soil moisture)	Weather@home	2005–2016 and counterfactual climate	0.11x0.11	??
Hydro/impact models				
H08 (soil moisture, PET)	H08	1861–2018	0.5x0.5	??
LPJmL (soil moisture, PET)	Lund-Potsdam-Jena managed Land model	1861–2018	0.5x0.5	???
PCRGLOB (soil moisture, PET)	PCRGLOB-WB, PCRaster GLOBal Water Balance model	1861–2018	0.5x0.5	?
WaterGAP2 (soil moisture, PET)	Water Global Analysis and Progress Model version 2	1861–2018	0.5x0.5	?

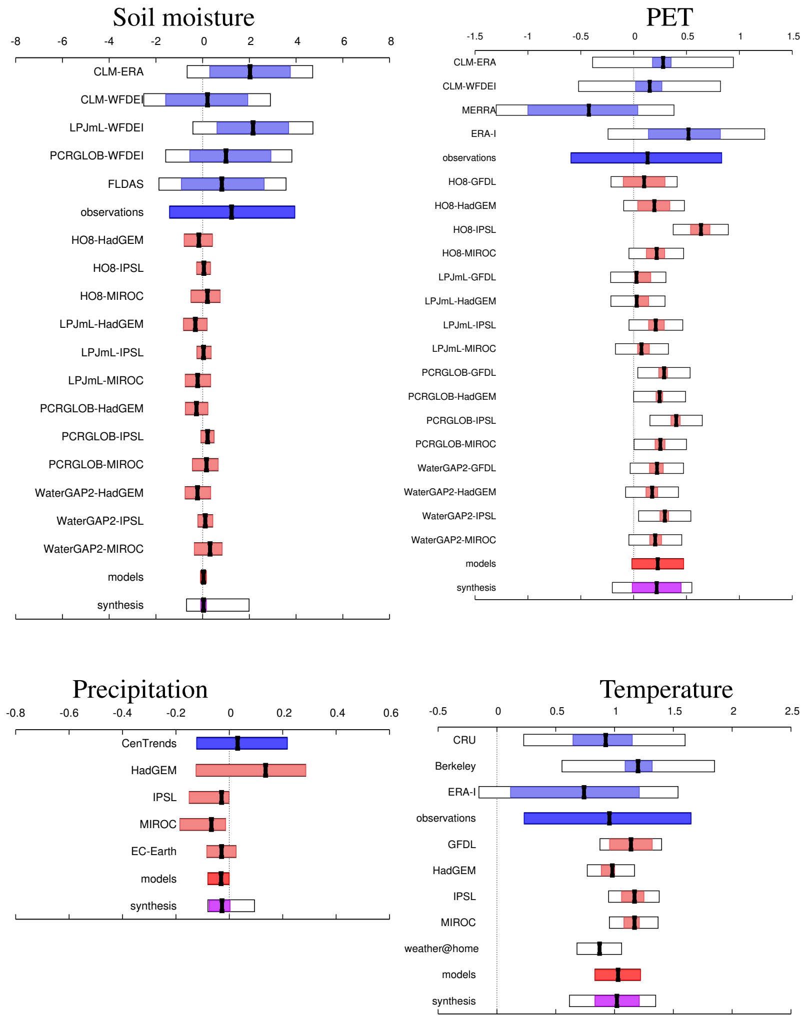

Figure 3. Illustrative examples of the synthesized values of trends per degree GMST rise for soil moisture (top left), PET (top right), precipitation (bottom left) and temperature (bottom right) for region SS. Black bars are the average trends, colored boxes denote the 95% CI. Blue represents observations and reanalyses, red represents models and magenta the weighted synthesis. Coloured bars denote natural variability, white boxes also take representativity / model errors into account if applicable (see Sect. **??**). In the synthesis, the magenta bar denotes the weighted average of observations and models and the white box denotes the unweighted average. Soil moisture trends are based on standardized data, the other trends are absolute trends.

the Supplementary Information. Table **??** and Fig. **??** summarize all ~~findings.~~ final synthesized findings. Using both the intermediate and final synthesis results we first draw conclusions based on different GCMs and hydrological models and then turn to conclusions per variable.

First we look for consistent behaviour in the trends from individual GCMs across the four variables. Some general conclusions ~~looking at~~ about the different GCMs are as follows: (i) for GCM-driven model runs with stronger positive trends in

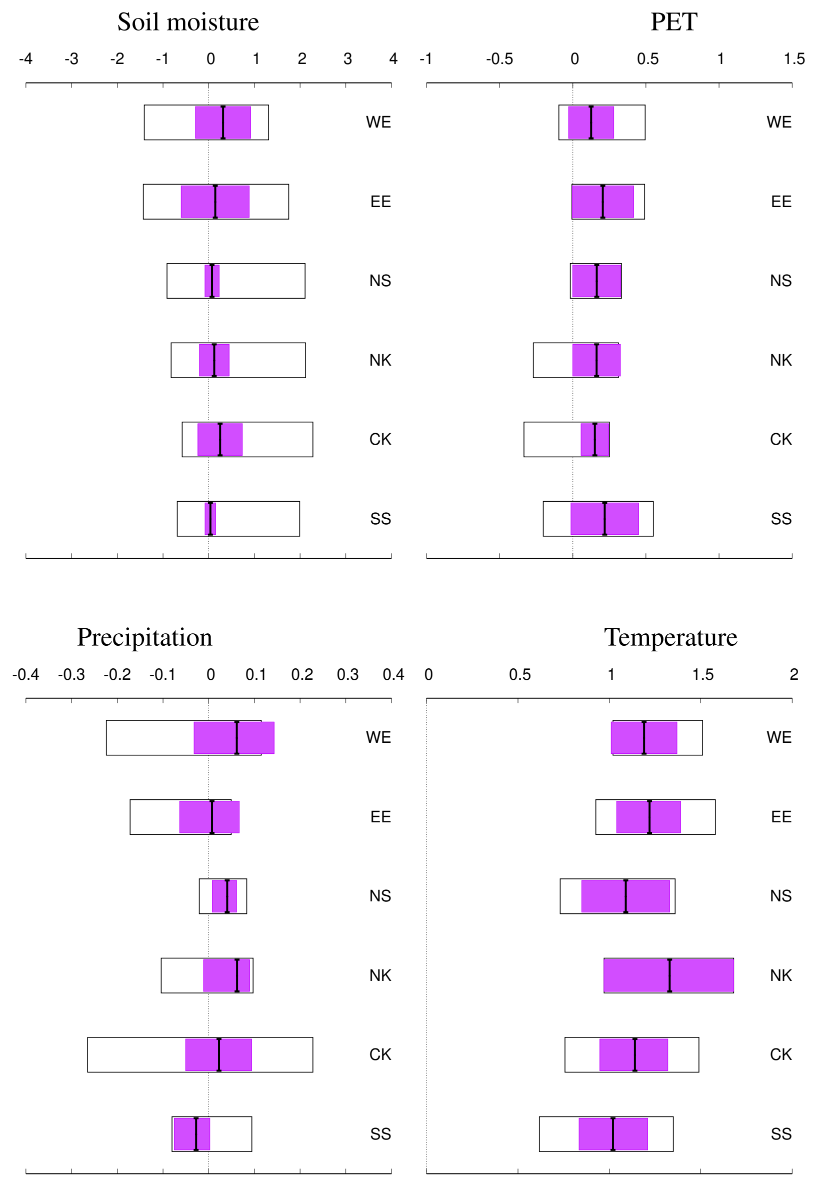

Figure 4. Summary of the synthesized values for soil moisture, PET, precipitation and temperature in the six regions. The magenta bars denote the weighted averages of observations and models and the white boxes denote the unweighted averages.

temperature, there is a tendency that the positive trends in PET are also stronger , and vice versa; (ii) the uncertainty in precipitation trends is high compared to the trend magnitudes. This is one of the reasons why a clear relation with soil moisture trends is not evident; (iii) no clear relation between local temperature trends and soil moisture trends is evident.

Looking at the different hydrological models we conclude that the trend in PCR-GLOBWB PET, which uses the Hamon PET scheme that depends only on temperature, is generally higher than the trend in in EC-Earth PET, which uses the more-complex

Table 4. Summary of synthesis results for each region and study variable. Note that '0' means no *significant* change and a '+' sign indicates a positive trend, where in soil moisture this means a change towards a *wetter* soil. The uncertainties associated with each result are depicted in Fig. **??**

Region	Soil moisture	Precipitation	Temperature	PET
WE	0/+	0/+	+	+
EE	0	0	+	+
NS	0/+	+	+	+
NK	0/+	0/+	+	0/+
CK	0/+	0	+	0/+
SS	0/+	0/-	+	+

Penman-Monteith PET scheme that additionally depends on humidity, wind and radiation. Using this more complex scheme can influence the trend in soil moisture, especially in wetter regions.

The analyses of the individual model runs, stratifying by GCM or hydrological model, do not lead to a clear conclusion on the relation between the trends in precipitation, temperature, PET and soil moisture. We therefore turn to the analysis of the synthesized values, see Table **??** and Fig. **??** for a summary of the outcome ~~,~~ and Fig. **??** and Figs. S1 to S6 in the Supplementary Information for synthesis diagrams. The table gives a concluding interpretation of the synthesized results shown in Fig. **??**.

For soil moisture we find no significant synthesized trends: there is practically no change in region EE ~~,~~ and no trend to a small positive non-significant trend in regions WE, NS, NK, CK and SS.

For precipitation, regions WE and NK show a positive but non-significant trend, in region NS there is a small positive trend, regions EE and CK show no trend and region SS a negative non-significant trend.

As expected from global climate change, the local annually averaged temperatures all have a significant positive trend, with best estimates between 1.0°and 1.3°per degree of GMST increase. Related to this, trends in PET are also positive in four of the six regions but lower than for temperature and generally with larger confidence intervals. The ~~exceptions are~~ regions NK and CK ~~, where, although the~~ are the exceptions. Although weighted averages show positive trends, ~~the models show opposite tendency to observationsand are thus incompatible with them, rendering the result~~ models show tendencies opposite to observations. This incompatibility renders the results uncertain.

We can identify the following relationships between different variables: (i) Precipitation trends have a (small) influence on soil moisture trends in regions WE, NS and NK; (ii) in regions WE, EE, NS, NK and CK, temperature and PET have no discernible influence on soil moisture trends; (iii) in region SS, the non-significant negative trend in precipitation does not lead to lower soil moisture ~~,~~ and neither do ~~temperature increase or the trend in~~ the trends in temperature or PET.

4 Discussion

In this section, we discuss ~~ways in which our chosen approach to studying drought in eastern Africa~~ the interpretation of our results in the light of how choices and assumptions made may have influenced the ~~results obtained~~outcome and we compare previous studies on similar topics.

~~We focus on drought on annual~~ We study drought trends on annual as opposed to sub-annual time scales, as long-term drought presents a greater risk for food security~~, with~~ . We define the annual period ~~defined~~ to be from January to December. This definition is a natural choice for each of our study regions, where the single or dual seasonal cycle peaks in precipitation (rainy seasons) and temperature do not extend beyond December into the next year. The Jan–Dec definition has the consequence that multi-season droughts out of phase with this period ~~, such as~~ do not appear extreme in the observational time series used here, whilst they would appear extreme in a Jul–Jun series. For example, in the well-documented 2010/2011 drought event in eastern Africa~~that affected~~ , only the second rainy season in 2010 and first rainy season of 2011 ~~, do not appear extreme in the observational time series used here, whilst they would appear extreme in a Jul–Jun series. This does however~~ were exceptionally dry. This choice however does not affect the ~~annual trends in the region~~ resulting annual trends, which are similar ~~independent of defining the year from~~ for both the Jan–Dec ~~or from~~ and Jul–Jun ~~. Hence our choice of annual definitiondoes not significantly influence the results~~annual definition.

On the annual time scale, we do not see strong explanatory relationships between the *trends* in the four studied variables. To gain insight in the relationships between the variables, we additionally looked at correlations on a sub-annual time scale. Simple correlations between monthly precipitation, temperature, PET ~~,~~ and soil moisture (not shown) support the conclusions of **?** on the influence of precipitation and PET on soil moisture at dry sites in Europe. They found that at water-limited sites the influence of precipitation on soil moisture is much larger than the influence of temperature, via PET, on soil moisture. In our study, we find the same for the driest regions and the driest months in the wetter regions, and for the more temperature-based PET schemes.

Looking at seasonal cycles — monthly means averaged over recent decades — a comparison between seasonal cycles of the different variables shows that the seasonal cycle of soil moisture is similar to that of precipitation in all six study regions. In contrast, the inverse seasonal cycle of temperature is not similar to that of soil moisture. Whether the PET seasonal cycle reflects elements of the soil moisture cycle or not depends on the PET scheme used: temperature- or radiation-based schemes show a seasonal cycle that is similar to that of temperature, whereas more advanced schemes reflect a mixture between the seasonal cycles of precipitation and temperature.

We thus conclude that the influence of precipitation on soil moisture is higher than that of temperature or PET. This is supported by the synthesized results that show negligible or no trends in soil moisture and precipitation ~~,~~ whereas the trends in temperature and PET are strongly positive.

If temperature has, via PET, an influence on trends in soil moisture, we expect to see that the positive trend in temperature is coupled to a drying trend in soil moisture. As we average over the annual scale, we may miss parts of the season when this effect is strongest. Therefore we selected a region and period outside the rainy season, in which the seasonal peak in temperature corresponds to a dip in soil moisture (region CK, months Feb–Mar), to inspect sub-annual trends (not shown). Even then, we find that there is no negative trend in soil moisture accompanying the positive temperature trends.

The approach taken in this paper towards uncertainty has been to Perform a multi-model and multi-observation analysis that summarises what we know at the present moment, using readily available data and methods. Apply simple evaluation techniques to readily available data, treating datasets that satisfy evaluation criteria equally and rejecting the others. Communicate uncertainties from synthesis. A simple 'yes' or 'no' is not appropriate in this analysis where there is no clear significant positive or negative trend. Rather, While improving the data with respect to some characteristics, an additional uncertainty arises from the bias correction of the GCM data prior to use in the hydrological model. The bias correction in ISIMIP was set up to preserve the long-term trend, but it also decreases the daily variability by truncating extreme high values (e.g., in precipitation) (**?**) . The most important element for our analysis is that it also increases the daily variability by removing excessive drizzle, which is often present in GCM precipitation data. **?** noted that such a statistical bias correction can influence the signal of runoff changes but that the uncertainties and their origin are given.effect generally remains smaller than the uncertainty from GCMs and global impact models. By far the largest difference we found in our analysis between trends in original and bias-corrected data was for temperature for IPSL in region NK: we found 1.9 K/K (95% CI 1.8 to 2.1 K/K) for the original trend and 1.4 K/K (95% CI 1.3 to 1.5 K/K) for the trend in bias-corrected data. All other differences were smaller and non-significant.

We find that the PET scheme used has a large influence on the PET seasonal cycle and trend (**?**) .In any case, in the long run, a trend in PET only has meaning for crop growth if there is water available for evaporation. Much of eastern Africa is in a water-limited evaporation regime. In the case that irrigation would be locally applied, more water would become available for evaporation, shifting the situation away from a water-limited regime and towards an energy limited regime. A trend in PET seen in our analyses (especially if the analysis using different schemes produces a robust PET trend) could then signify a trend in real evaporation and would therefore be accompanied by an increase in irrigation water demand.

Note that irrigation is not accounted for by the models or reanalysis datasets used here.

A study by **?** discussed the possibility that climate model trends precipitation trends in East Africa are influenced by inability of the models to represent key physical processes reliably, and flagged this issue as a topic for further study. In attribution studies on drought, especially for this region, it is therefore high priority to extend model evaluation techniques to assess models' representation of key physical processes. The approach taken in this paper has been to apply simple evaluation techniques to on the seasonal cycle and frequency distributions of readily available data , in order to advance and that results from models passing validation tests represent the status of our current knowledge. Precipitation Rainy seasons in this region are governed by large-scale processes, such as the shifting of the ITCZ , and ENSO dynamics. The ability of a model to capture the seasonal cycle in precipitation and temperature thus provides some assurance that large-scale physical processes are reasonably well described by the model. The frequency of extremes in precipitation is influenced by variability with respect to the general magnitude, therefore we also check that variability in analysed time series is similar to observations. We see these tests We see the tests we perform as a minimum requirement for model validation. However, to improve the performance of models and to understand the discrepancies between models and observations, a much more thorough investigation into the models' representation of physical processes and feedbacks is required, such as demonstrated by **?** and encouraged by the IMPALA (Improving Model Processes for African Climate) project (https://futureclimateafrica.org/project/impala/).

While improving the data with respect to some characteristics, an additional uncertainty arises from the bias correction of the GCM data prior to use in It is still unknown how vegetation will respond to substantial increases of CO_2 concentration. Two counteracting effects — physiological (restriction of stomatal openings leading to decreased evapotranspiration) and structural (increased leaf area leading to more stomata and increased evapotranspiration) responses — are expected, but their net effect is unknown (e.g. **?**) . So-called 'dynamic vegetation models' include these CO_2 effects and there are indications that these models show a weaker response of drought to climate change (**??**) . In this study our selection of hydrological models is restricted by the variables we require, however, out of the four ISIMIP hydrological models that match our criteria, one (LPJmL) uses dynamic vegetation modeling. The soil moisture response to increasing GMST in LPJmL simulations is mid-range amongst the ISIMIP results. The PET response for LPJmL simulations is, however, somewhat on the low side of the ISIMIP results. It has not been verified if this behaviour is linked to dynamic vegetation modelling, but with confidence intervals generally overlapping with the synthesized model outcome, there is no exceptional difference.

The approach taken in this paper towards uncertainty has been to

- Perform a multi-model and multi-observation analysis that summarises what we know at the present moment, using readily available data and methods.
- Apply simple evaluation techniques to readily available data, treating datasets that satisfy evaluation criteria equally and rejecting the others.
- Communicate uncertainties from synthesis. A simple 'yes' or 'no' is not appropriate in this analysis where there is no clear significant positive or negative trend. Rather, the hydrological model . The bias correction in ISIMIP was set up to preserve the long-term trend, but it also decreases the daily variability by truncating extreme high values (uncertainties (confidence intervals) and their origin (e.g. , in precipitation) (**?**) . The most important element for our analysis is that it also increases the daily variability by removing excessive drizzle, which is often present in GCM precipitation data. **?** noted that such a statistical bias correction can influence the signal of runoff changes but that the effect generally remains smaller than the uncertainty from GCMs and global impact models. By far the largest difference we found in our analysis between trends in original and bias-corrected data was for temperature for IPSL in region NK: we found 1.9 K/K (95% CI 1.8 to 2.1 K/K)for the original trend and 1.4 K/K (95% CI 1.3 to 1.5 K/K) for the trend in bias-corrected data. All other differences were smaller and non-significant. natural variability or model spread) are given.

It is also important to note the significant scientific uncertainty relating to the effects of increased atmospheric CO_2 concentrations on the relationship between PET and plant growth and so on drought

In the long term, a trend in PET only has meaning for crop growth if there is water available for evaporation. Much of eastern Africa is in a water-limited evaporation regime. In the case that irrigation would be locally applied, more water would become available for evaporation, shifting the situation away from a water-limited regime and towards an energy limited regime. A trend in PET seen in our analyses (especially if the analysis using different schemes produces a robust PET trend) could then signify a trend in real evaporation and would therefore be accompanied by an increase in irrigation water demand. Note that irrigation is not accounted for by the models or reanalysis datasets used here.

Previous studies have shown that both the PET scheme and the input data used for calculation of PET can have a large influence on PET values (**??**) . We confirm this using the CLM-ERA-PT (Priestley-Taylor), CLM-WFDEI-PT and CLM-ERA-PM (Penman-Monteith) datasets (not shown). In our study regions, PET values are consistently higher when using PM then when using PT. The differences in trends in PET using ERA or WFDEI input or using PT or PM input are sometimes significant. However, comparing study regions, there is no consistency in the difference; in four out of the six regions the PM data shows a higher trend than the PT data and in four out of the six regions WFDEI data shows a higher trend than the ERA data.

There is some evidence that warm spells are increasing in length, particularly in Ethiopia and northern Somalia/Somaliland region (**?**), as is the number of consecutive dry days in some parts of eastern Africa, which may have an impact on drought length and increase the ~~onset rapidity and~~ rapidity of onset and the intensity of drought (**?**). However, the overall impact on crops and food security during long-duration droughts on annual timescales is probably insensitive to this.

It is possible that increasing temperatures have a negative impact on food security during droughts in ways that are beyond the scope of this study, e.g., decreased immunity of livestock, or increased water demand for cooling and water supply (**?**, and references therein). In addition, in regions suffering from recent meteorological drought, non-meteorological factors such as increasing population and land-use changes also play a role in worsening the declining vegetation conditions, even after precipitation returns to normal (**?**).

5 Conclusions

In this first multi-model, multi-method attribution study using several drought estimates in eastern Africa, we address the recurring question on whether increasing global temperatures exacerbate drought. Previous attribution studies for the eastern Africa region have examined drought from a meteorological perspective (precipitation deficit) ~~,~~ and have found no clear trends above the noise of natural variability. In this study, we examined trends in eastern African drought from an agricultural perspective (soil moisture) as well as the meteorological perspective ~~, and additionally investigated whether increasing global and local temperatures and trends in evaporative demand~~ (precipitation, temperature and PET) for six regions in eastern Africa. We also investigate whether global-warming driven trends in these meteorological variables can be seen to contribute to trends towards drier soils.

~~Using a combination of models and observational datasets we studied trends in four drought related annual variables — soil moisture, precipitation, PET, and temperature — for six regions in eastern Africa.~~ In this section, we draw conclusions for each variable in turn ~~.~~ and make recommendations.

Out of the four studied variables, ~~food security~~ soil moisture is most closely related to ~~soil moisture anomalies~~food security via crop health. In standardized soil moisture data, we ~~find~~ found no discernible trends. The uncertainties in trends from model runs ~~are large ,~~ were found to be large and there are no long observational runs available. This ~~makes~~ emphasizes that the use of an ensemble of models is imperative. Due to the large uncertainties in both soil moisture observations and simulations, we conclude that soil moisture cannot be relied upon on its own as a drought indicator and it is therefore important to examine other drought indicators as well. Besides, soil moisture also has a physical lower limit: once the soil is dry it will remain dry. In water limited regions an analysis of precipitation is thus a helpful addition.

Precipitation was found to have a stronger influence than temperature or PET on soil moisture variability, especially in the drier study regions (the significant positive trend in temperature is not reflected by a decrease in soil moisture). However, the ~~error margins~~ confidence intervals on precipitation trend estimations are large and no clear trend is evident.

As expected from the increase in global temperatures, we find significant positive trends in local temperatures in all six regions. The synthesized trend is between 1.0 and 1.3 times the trend in GMST, which corresponds to a local temperature rise of 1.1 to 1.4 degrees from pre-industrial times to 2018. However, the influence of this on annual soil moisture trends appears limited.

PET has a more direct link via evaporation to soil moisture than temperature. The trends in PET are predominantly positive, although in the regions NK and CK the ~~value is uncertain~~uncertainty in this trend is large. This generally agrees with the positive trends in temperature. Similar to the results for temperature, we do not find strong relations between PET and soil moisture trends. Nevertheless, the results can still be of interest, especially in irrigated regions. ~~Note, however, that there are generally large differences between~~ Due to large differences in results from different ~~model runs due to the different PET schemes and input datasets used. An analysis of PET should therefore~~ hydrological model runs, we recommend that PET attribution analyses be carried out using ~~a model ensemble that (i) covers various PET schemes, especially if the additional drivers required for the more advanced PET schemes are not monitored well, and (ii) uses multiple input datasets to span the uncertainty from driving GCMs . Conclusions on resulting trends should be drawn with caution, giving more weight to the confidence interval around the best estimate rather than the best estimate alone.~~ an ensemble of hydrological models. These should use various (observational) input datasets and driving GCMs and cover various PET schemes, in order to be representative of the uncertainty surrounding all valid approaches and not bias results towards a particular method.

~~Considering food security, a drought becomes more problematic if it endures for more than one season.~~

So although the trends in seasonally averaged variables may be larger than those in annually averaged variables, the conclusions drawn for annual droughts are of higher relevance for food security.

Whilst it may be preferable to use soil moisture as a drought indicator, observations and simulations of precipitation are more reliable in this region (?) . Precipitation has a large influence on agricultural droughts and is therefore appropriate to use in attribution studies in eastern Africa, supplementing the analysis of soil moisture. The outcome of previous studies that have focussed on precipitation deficits only (e.g., **??**) are thus still relevant and compare well with our results~~here~~, that no consistent significant trends on droughts are found.

Finally, communication of the uncertainties in the analyses of soil moisture, precipitation, temperature ~~,~~ and PET (and any drought indicators) to policy makers, the media and other stakeholders is crucial. Without insight into the uncertainties in synthesized trends in the different drought indicators, conclusions become meaningless and results can easily be misinterpreted.

Data availability. Almost all time series used in the analysis are available for download under https://climexp.knmi.nl/EastAfrica_timeseries.cgi (last access: 29 April 2019).

Author contributions. Sarah Kew and Sjoukje Philip designed and coordinated the study, analysed all data and led the writing of the manuscript. Mathias Hauser contributed the CLM datasets including PET calculations and substantially contributed to writing. Mike Hobbins produced the RefET dataset and substantially contributed to writing. Niko Wanders and Karin van der Wiel collaborated to create the EC-Earth - PCR-GLOBWB data, including PET calculations and bias correction. Niko Wanders additionally advised on the use and validation of PET and soil moisture data for the analysis of drought. Geert Jan van Oldenborgh contributed analysis tools, monitored progress and contributed to writing. Ted I.E. Veldkamp provided guidance on the use of ISIMIP data and contributed to discussions. Joyce Kimutai and Chris Funk provided local information. Friederike E.L. Otto conceived the idea for the study, monitored development, provided weather@home results and contributed to writing.

Competing interests. We declare that there are no competing interests.

Acknowledgements. We would like to acknowledge funding for this work from the Children's Investment Fund Foundation (CIFF) grant #1805-02753. For their roles in producing, coordinating ,and making available the ISI-MIP model output, we acknowledge the modelling groups and the ISI-MIP coordination team. We would also like to thank the volunteers running the weather@home models as well as the technical team in OeRC for their support. MERRA-2 data were developed by the Global Modeling and Assimilation Office (GMAO) at NASA Goddard Space Flight Center under funding from the NASA Modeling, Analysis, and Prediction (MAP) program; the data are disseminated through the Goddard Earth Science Data and Information Services Center (GES DISC), preliminary data have been made available thanks to GMAO.