**Interactive comment on** “Comparison of simulated and reconstructed variations in East African hydroclimate over the last millennium” by F. Klein et al.

**Anonymous Referee #2**

Received and published: 25 March 2016

Summary: This study uses a paleoclimate model data comparison framework to analyze East African lake levels over the last millennium. GCMs struggle to represent the seasonal cycle of precipitation and teleconnections over East Africa. Nevertheless, the teleconnections appear to be variable over the last millennium, and between fixed forcing and variable forcing simulations. For the Masoko/Malawi region, in particular, anthropogenic forcing appears to influence the teleconnections. On centennial timescales the variation in teleconnections are large for both regions and this is explained by changes to natural forcing. Despite a clear link between forcing and teleconnection changes in the models over the last millennium, there is no relationship between forcing and hydroclimate changes. By contrast, internal atmosphere-ocean variability is shown to be
the dominant driver of simulated hydroclimate changes over East Africa, even on centennial timescales (although anthropogenic has driven consistent simulated changes in the Masoko/Malawi region over the most recent \(\sim 150\) years). The dominant role for internal atmosphere-ocean variability in driving hydroclimate changes can explain the mismatch between the time histories of hydroclimate over East Africa simulated by models and that reconstructed for the four lakes.

General Remarks: This represents an interesting and important contribution to our understanding of low-frequency hydroclimate variability over East Africa as simulated by models. My comments are largely minor although I do have three major concerns that I hope that the authors will consider addressing.

Major Concerns:

1. The manuscript, in general, is clear although I would strongly suggest revisiting the manuscript and editing for grammar and sentence structure.

2. (Page 23, Line 6) “An interesting question is whether the forcing actually alters the dynamical link between East African rainfall and SSTs, or if it only masks it because of a different impact on continental rainfall and SSTs. Answering this question is out of the scope of this study, but it is of interest for the interpretation of records used for reconstructing phenomena like the IOD. Indeed, if dynamical relationships are not stable when considering different time scales, a record calibrated in observations of the recent period may not be representative of the studied phenomena over longertime scales.” This is very important and likely the most important conclusion of the paper, is there not some relatively simple way to approach this (specifically determining if the forcing actually alters the dynamical link between East African rainfall and SSTs, or if it only masks it because of a different impact on continental rainfall and SSTs)? I think it would really improve the contribution that this paper will make to our understanding of simulated and real-world East African rainfall. In particular, the CESM last millennium ensemble would seem to be well suited to answering this question given that the ten
ensemble members can be used to robustly determine the impacts of forcing relative to internal variability.

3. When analyzing the reconstructions, Challa and Naivasha look very different, as do Malawi and Masoko. I found that the descriptions of common changes here were not consistent with what can be seen by eye in the figure. Perhaps there is a more quantitative way to approach this? Generally, I suggest that this section be revisited. If the reconstructions do not line up why might this be and what does this suggest for our interpretation of the model simulations?

Specific Comments:

Abstract: “The bimodal seasonal cycle characterizing the Challa/Naivasha region, except that in the latter the relative magnitude of the two rainy seasons is less well captured.” This language doesn’t seem fully consistent with the results, it appears that the models generally struggle to reproduce the characteristics of the seasonal cycle at both locations.

Introduction: Southeast and equatorial east African lakes versus east African rainfall. Make as clear as possible at the outset that east African is covering all four lakes but that the distinction between southeast and equatorial east African is how you will describe the two sets of two lakes.

(Page 2, Line 8) “through atmospheric adjustments to the Walker circulation”. This statement is unclear to me, perhaps you mean oceanic driven changes to the atmospheric Walker circulation?

(Page 2, Line 23) “Enhanced pattern” perhaps better to say an increase in because you are not talking about a spatial pattern.

(Page 2, Line 33) “poor ability” suggest using the inability.

(Page 3, Line 1) “reached contrasting conclusions depending on the region or spatial scale, or on the variables and models considered.” Given this, basing any conclusions
regarding the role of internal variability on the fact that the models do not match the reconstructions seems problematic given that only two model grid points are being analyzed. A larger spatial scale might provide more confidence here, however the results using the CESM ensemble do provide strong evidence for the role of internal variability.

(Page 3, Line 20) It is not clear how the annual means follow from the long-term changes. Maybe save the discussion regarding the use of annual means portion for later.

(Page 3, Line 22) Is there any reason to suggest that models should capture the reconstructed changes? Perhaps change to analyze GCM simulations of long-term change relative to reconstructions over the last millennium. I do not think this study is truly aimed at investigating GCM performance.

(Page 4) An aside, but I really appreciate the detail that you have gone into with regards to the set up of the models and the slight differences.

Section 2.2. I think it is important to note how comparable each of these records are. Are there reasons to expect systematic differences given that they each are reconstructing different things? How might this impact the interpretation of the results and conclusions?

(Figure 2) On the x-axis Fev should be Feb, the labels overlap so potentially make text smaller.

(Page 7, Line 15) To my eye CESM1 doesn’t look better than the other models, BCC is arguably more realistic, maybe just remove the second part of that sentence.

(Page 8, Line 1) Put a space in “Lake Naivashais”.

(Page 8, Line 4) I am not sure the split in how realistic the models are is that clear. When looking at CCSM4, CESM1 and BCC-CSM1-1, IPSL looks just as reasonable to me. Maybe be a bit more general.
I think this is a bit of an oversimplification in the second part of this sentence as the timing is also off (not just the magnitude).

"Since Lake Challa and Lake Naivasha on the one hand and Lake Masoko and Lake Malawi on the other have a very similar climatology and seasonal cycle both in models and observations (Fig. 2)". We are interested in long term changes, nothing that can really be done but the reconstructions for each lake look very different suggesting that on long timescales things might not be expected to be similar.

"However, using the larger grid boxes raises the issue of whether the proxy-based reconstructions employed to assess model performance through the last millennium are representative for these larger spatial domains". Be more exact here, don’t use assess model performanceâ˘A¶that’s not really what’s being done here.

"Regarding annual mean absolute values". What do you mean by this?

"Nevertheless, with values equal to 0.27 and 0.16 the significant correlation coefficients between, respectively." Are these spatial correlations? This section is a bit confusing; suggest revisiting with an eye towards clarifying the language.

"Making abstraction of chronological uncertainties, the Challa and Naivasha proxy records show discrepancy during the first four centuries of the last millennium". This sentence is confusing.

"humidity", suggest changing to wetting.

"besides", suggest removing.

"thus leads to larger river runoff towards the lakes". I am not sure that I understand this as the models do not have these actual lakes.

"This implies that most changes of P-E over time are due to changes"
in precipitation”. Is this necessarily true on longer timescales?

(Page 15, Line 8) Why not a more typical and interpretable standardization using the standard deviation?

(Page 15, Line 9) How does this or does this not match the temporal resolution of those reconstructions. I suppose I am just interested in a more thorough justification for this choice of smoothing.

(Page 15, Line 11) By eye in Figure 7 the model variability looks to be about the same magnitude as the reconstructions.

(Page 16, Line 8) “Spatial representivity”, should likely just be locations.

(Page 16, Line 12) “confronted”, should be compared.

(Page 17, Line 10) This is an important sentence. At least remove “that” to make it clear, however, restructuring the sentence would probably be good.

(Page 17, Line 22) “In this regard, it is of interest to note that for the one model for which multiple ensemble members were available (CESM1), there is also no correlation between the different ensemble members that differ only from slightly different air temperature at the start of the experiments (Otto-Bliesner et al., 2015)”. This is very important and gets lost a bit as cast.

(Section 5.2) I feel like this section would be clearer with the unfiltered pre-industrial control run teleconnections also shown. It becomes a bit confusing with the pre-industrial portion of the last millennium, pre-industrial control runs, and historical simulations all being compared simultaneously.

(Conclusion) This ends on a weak note, I might suggest finishing the paper with the last sentence of the previous paragraph.