Comments - Referee #2 B. F. Zaitchik (Referee)

zaitchik@jhu.edu

Received and published: 5 June 2014

Summary:
This Discussion Paper builds on a series of important papers that the authors have written on patterns, drivers, and change in East African precipitation. In this paper the authors use this foundation to show that variability in March-May East African precipitation can largely be explained by variability in two key SST indices. They then propose that these indices can be used as tools for seasonal prediction. The paper is well motivated and clearly written, and as far as I can judge it is methodologically sound. I do have one general comment about the purpose of the paper and a number of specific comments on methodology and presentation that I suggest the authors consider before the paper is accepted for final publication. These are listed below.

General Comment:
The core motivation for the paper is to identify indices that can be used in seasonal prediction. But the authors stop short of proposing a complete, vetted prediction system: there are no out-of-sample tests of predictive skill, predictors are not combined to generate a prediction of total precipitation anomalies, and despite the fact that the authors note nonlinearities and nonstationarities in their correlations they do not propose any way to incorporate this information into operational predictions. Given these limitations, it seems that this paper is intended as an intermediate step rather than an end in itself. I have no objection to this, but I do think that the paper would stand on its own more effectively if the authors could frame its contributions in more specific terms. As it stands, Section 1 is mostly general background on the region and Section 4 is mostly general conclusions on statistical approaches to prediction.

I would be interested to see these sections address questions like: What is the advantage of predicting PCs of precipitation rather than total precipitation? Is there any limitation to using PCs that explain only 38% of total variance as the prediction target? What do the proposed predictor indices offer that other predictors do not? What do these results suggest about the baseline period or the model structure of existing operational seasonal prediction systems in East Africa? Does the analysis offer any specific insights on how to deal with nonstationarity in this region, beyond just watching it as a general point of concern?

These are just examples. I would defer to the authors regarding the best emphasis, but I do think it could be more specific than the framing that is currently offered.

Responses: We appreciate greatly appreciate Reviewer 2's thoughtful comments, in both their general sense and in their specific applications. We thank the reviewer for these suggestions, which should substantially improve the manuscript, and address each of these suggestions below.

1. Concerns about paper scope and lack of cross-validation. As mentioned by reviewer 2, this paper seems like an intermediate step; which in fact it is. This paper will hopefully be accompanied by a companion article *"A seasonal agricultural drought forecast system for food-insecure regions of East Africa"* by Shraddhanand Shukla, that describes in a much more rigorous way (including cross-validation) how we integrate climate forecasts to bootstrap CHIRPS data to drive a land surface model to make hydrologic forecasts.

   That said, we concede that the paper needs to have its mission made more transparent, and its text strengthened to effectively convey that message. The main objective of this article is to communicate to a broad audience the west Pacific teleconnection pattern and to describe how it can be used as a basis for more rigorous prediction. Regionally, there is still a great need to share this basic understanding. For example, members of UNEP are currently mobilizing the early warning to prepare for drought impacts in Kenya related to the current El Niño; when the latest science indicates above normal rainfall would be expected. So, as an adjunct to the paper by Shukla, this discussion was crafted to communicate the new sensitivity to a broad audience in a way that could be replicated by region climate experts. We hope to provide training along these lines this summer at the upcoming pre-GHACOF training.

   Nonetheless, reviewer 2's comments indicate that we have failed to adequately explain our objectives. We will amend the paper appropriately, and also explicitly reference and discuss Shukla's HESSD paper.

   Technically, there are several improvements that we can make to the paper. While simple, the proposed system is quite close to a 'vetted prediction system', and we have modified the paper to move it closer to that goal, while also clarifying the fact the article's main purpose is illustrative and heuristic. Specifically, we suggest making the following modifications:

   1. Adding cross-validation to our estimation process, to provide estimate of out of sample error. The prediction skills for the WPG/PC1 estimates hold up well (cross-validated correlation ~0.63), while the Indian Ocean/PC2 values do not hold up well (cross-validated correlation ~0.2).
   2. Expressing our predictions and observations in Fig. 4 as anomalies in mm (as opposed to SPI values).
   3. More specific discussion of how we can link these indices to quantitative rainfall forecasts (both in reference to Shukla et al., and to a sample application that was made this winter (to successfully predict the poor 2014 long rains). A section will be added addressing this specific issue.

We also add the following:

a) A specific discussion of why EOFs were used, and contrast the variance explained by the EOFs with the variance explained by using a regional mean. One objective of the EOF analysis is to evaluate the appropriateness of bootstrap-based forecasts in Shukla et al. (i.e. using one set of probabilities for the whole region); EOF1 suggests this is generally acceptable, within reason. The 2nd value of the EOF results is that they so nicely partition the Indian and Pacific signals; we discuss this more explicitly. There clearly are limitations to the EOF approach, since the climate is so complex. These limitations will be more explicitly enumerated.

b) In the final discussion, we add more comparisons between these predictors and other potential candidates. Our point here is not the proposed predictors are ultimately the 'best' but that the emergent equatorial SST gradients provide a new powerful foundation for early warning. This does have implications for baseline periods and current operational seasonal prediction systems. These may use too long of a baseline period and focus too much on the Indian Ocean. This is nicely underscored by our new cross-validation results.

c) We better address the question of non-stationarity and global warming in the conclusion. Honestly, such changes can wreak havoc with statistical forecast methods, and we suggest that perhaps mesoscale model may ultimately be the way forward.

Specific Comments:

Section 2:

1. How was the geographic area of analysis selected? Is there evidence that the region is coherent when it comes to large scale drivers of March-May precipitation? And can the authors demonstrate that their results are not overly sensitive to small changes in the region used to calculate PCs? If the region includes significant heterogeneity in Mar-May teleconnections and/or if the analysis is sensitive to small changes in the borders of the analysis region then the authors might want to target a more specific region defined on the basis of correlation strength with proposed SST drivers or on the basis of homogeneity in March-May precipitation variability.

These are great questions. The geographic region was chosen based on 1) the region receiving at least some March-May rainfall, 2) prior research linking rainfall variations to SST anomalies in the Indian and Pacific Oceans, and 3) climatological moisture transports associated with the onshore flow from the Indian Ocean and Somali jet. Hence western Kenya and Ethiopia were not included. While our prior work had suggested that eastern Kenya, the Belg growing regions of Ethiopia and Southern Somalia would covary, the covariance structure in north Ethiopia, Djiboutia, Eritrea and Yemen was unknown. There positive covariance with eastern Kenya, the Belg growing regions of Ethiopia and Southern Somalia is a new, and quite interesting result that may be related to changes in the Somali Jet (all these regions are on the Jets western flank).

Experiments were made with the choice of region, and it was surprisingly robust to variations in latitude, and only slightly more sensitive to changes in longitude. We believe that this is due to the fact that EOF is based on the covariance matrix, and the most coherent part of this is driven by changes in the Somali Jet, precipitation across the eastern portion of the Horn, and Indo-Pacific SST variations. These changes in domain will be detailed and their effects described. Several specific experiments (e.g. for Tanzania) will be used to show that while the results are basically robust, there are areas in the study region that are not well described by the EOFs.

2. The authors note that CHIRPS agrees well with other state-of-the-art precipitation datasets, but it is not clear whether they mean that the products agree well in this region or just that they agree well in general. Given the importance of CHIRPS to this analysis I think that it would make sense to include a time series of CHIRPS and other common precipitation datasets for the study region and/or for key precipitation zones within the study region.

Okay. As discussed we will add a forecast application to the 2014 MAM rainfall season, along with a time series of GPCC and CHIRPS (see below for April-May, we will show MAM in the paper revision). We will also add some other sources of precipitation estimates, i.e. at least some or all of the following: TRMM v7, CPC-unified, RFE2, ECMWF, CFS.

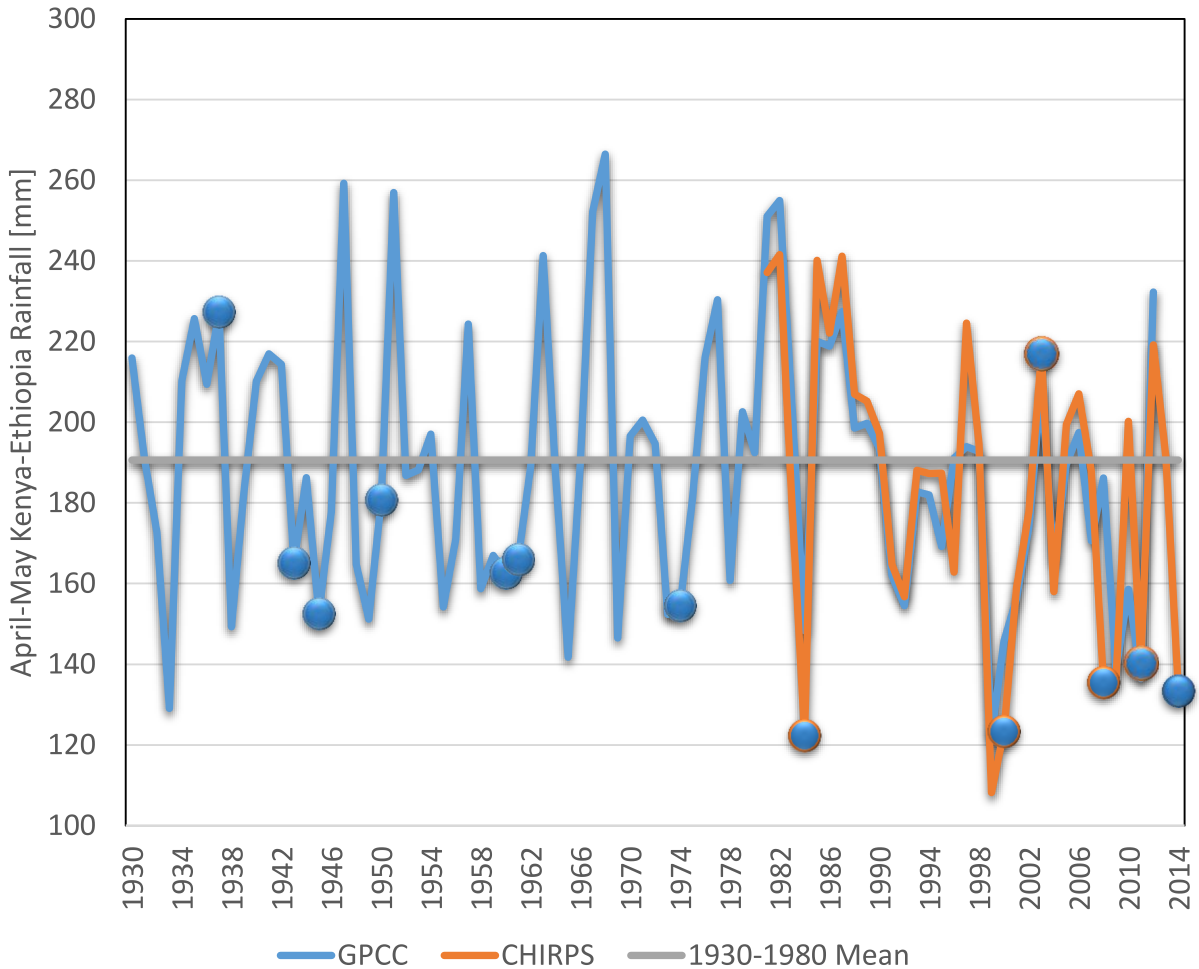
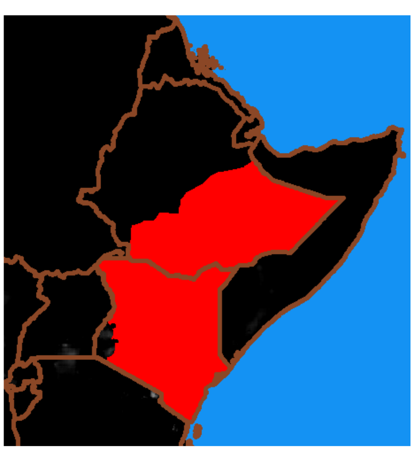


Fig. X. Time series of GPCC and CHIRPS data for Kenya and southern Ethiopia. Dots in the time series indicate strong Pacific SST gradient seasons. Correlation between CHIRPS and GPCC = 0.87.



Section 3:

3. Section 3.1: Were the PCs calculated from standardized precipitation fields or from the raw precipitation data? The decision to plot standardized precipitation in Figure 1 suggests the former, but reading the first paragraph of 3.1 I see no mention of standardizing precipitation anomalies prior to deriving the PCs, and the description of the calculation makes me think the data were not standardized. If this is the case then it would be useful to see versions of Fig 1a-c that use raw precipitation. As it stands there are some areas that receive relatively little Mar-May precipitation that feature very strongly in these maps. This would be an imperfect representation of the PC loadings if the PCs were in fact calculated without standardization.

Okay. They were calculated from the covariance matrix; standardized anomalies were chosen to emphasize the relative impacts in places like Yemen; but we agree this may be misleading. We will show these as changes in mm, which will highlight Kenya and southern Ethiopia. This probably also provides more useful information about where the EOF/forecasts are targeting.

4. Section 3.3: The term "skill" is used when presenting regression results, but I don't see any quantitative evaluation of the predictive skill of these models presented in the paper. The authors offer descriptive plots and in-sample correlations, but if the goal is to inform operational prediction systems then it really is necessary to present some kind of evaluation of out-of-sample predictive skill. This is particularly the case considering that the authors have chosen to use a relatively short time period to fit their models and because trends over the time period of analysis contribute to observed model fit. I recognize that this might be asking for more of a prediction system than the authors want to attempt in this paper, but given the challenges of developing a statistical prediction model on a short baseline–which the authors seem to be recommending–some discussion of predictive skill evaluation would seem to be important.

Totally agreed. We will add cross-validated regressions to analysis, and really should have done so from the outset. This will really strengthen the paper by a) making the case the emergent Pacific sensitivity is quite stable, and holds up well to cross-validation while b) highlighting that more predictive skill seems to be coming from the Pacific (as opposed to the Indian) Ocean.

Section 4:

5. This section offers an excellent summary of challenges facing seasonal prediction in a changing climate, but it offers relatively little in the way of conclusions and interpretation specific to this study. I would encourage the authors to use the Summary section to integrate the analyses presented in the paper–e.g., the relevance of the CMIP5 analysis to seasonal prediction using their proposed indices–and to tie their results more specifically to a way forward on seasonal climate prediction for East Africa. This could include questions I listed in my General Comment, but it doesn't have to.

Agreed; this section will be strengthened and made more 'proactive', and linked better with the introduction. We will make the case that 1) strong Pacific SST gradient events appear to be happening more frequently, but 2) these events offer opportunities for prediction, and 3) there are probably multiple ways to use this information: simple indices, more sophisticated statistical climate models, COF forecasts, hybrid coupled model predictions and mesoscale model simulations. This may allow East Africa to adapt through better predictions.

Tables and Figures:

6. Table 1: It would be useful to provide the name of the model in addition to the name of the modeling group (e.g., which NCAR model was used)?

Good suggestion. This will be added.

7. Figure 3d: The spatial variability in correlations suggests that the defined study region does not respond homogeneously to the selected large-scale forcers. This argues for a more systematic method for defining the region of analysis or, at the very least, a discussion of how and why correlation values differ within the region (e.g., why correlation in northern Tanzania is opposite that in surrounding regions).

Agreed, there certainly are regions (like northern Tanzania) that appear to be very poorly represented by the model. Conversely, however, given the incredible complexity of East African precipitation, and the size of the domain (~3,850 km by ~2,220 km) it is quite shocking that the that this simple forecast model works anywhere near as well as it does. We certainly agree, however, that the region does not respond homogeneously to large scale climate forcing. We will report tests with the spatial domain (as described above) and emphasize the heuristic emphasis of our study, and the need to tailor forecasts regionally.

We appreciate very much these thoughtful comments, which should make the revised manuscript much better.