Interactive comment on "Forecasting droughts in East Africa" by E. Mwangi et al

Anonymous Referee #1

Received and published: 9 October 2013

General Comments

The paper addresses the important topic of drought forecasting in Africa, analyzing the possibilities for the region of East Africa. The research presented is interesting and relevant and deserves publication in HESS. It needs minor revision.
In general the paper is well written and concise. The introductory section gives a short but clear introduction to the topic, including a description of the current practice in the region and highlighting possibilities for improvement.
The materials and methods section is well described with references for more detailed information on the indicators used. However, this section would benefit from more detailed explanations. Some statements are unclear and need improvement (see specific comments).

With the exception of a few statements and inaccuracies, the results section is clear and easily understandable. Figures support the conclusions.
The English is generally good. Errors are noted in the technical comments section. English spelling and grammar should be checked once more carefully.

We thank the reviewer for the comments on the paper. See our comments in red below.

Specific Comments

1. Introduction

In the Introduction the East African GHACOF is described as being based on (a) the forecasters subjective knowledge of the relationship between SST and rainfall patterns/amounts, (b) rain gauge data, and (c) dynamical forecasts by other international centers. I assume that "other" stands for non-ECMWF. It may be better to say so or else to delete "other". Later in the introduction, it is twice said that the forecast relies mostly on precipitation/station data. This is contradicting what was said before. To clarify you could indicate the weight of the individual knowledge components in the GHACOF or describe in more detail how decisions are taken (e.g., Which information has the most influence? How are the forecasts entering the decision process?, etc).

ECMWF is among the International centers whose dynamical forecasts are considered; hence we will replace other by "several international centers (including ECMWF). The following information about these centres has also been added to the paper "These are centres that in 2006 were designated by the World Meteorological organisation (WMO) to be making global seasonal forecasts as WMO Global Producing Centres (GPCs) for long Range forecasts. The centres maintain seasonal forecasts of upto four months lead time for the globe."

The following detailed information on the forecasting process has been added to the paper "Two weeks prior to the GHACOFs, meteorologists from the various countries National Meteorological and Hydrological Services (NMHSs) convene at ICPAC where

they produce a seasonal statistical rainfall forecast for each country. Historical rainfall data from synoptic stations is correlated with global Sea Surface Temperatures (SSTs) up to 6 months prior to the rainy season. SSTs from ocean regions that have correlation values > 0.5 are extracted and used to develop a forecast model using linear regression. However, to make computation easier Principal Component Analysis (PCA) was used to divide each country into homogeneous zones and only one representative rain gauge station is used for each zone.

Each country's forecast is compared to that of the neighbouring countries. If the forecasts don't agree then the forecasters look at historical years that had similar SST patterns to see which of the forecasts is similar to the events of that historical year. The forecasts are also compared with dynamical forecasts from several international centers. The country forecasts are consolidated into a consensus forecast for the Greater Horn of Africa which is manually smoothed to cater for all the information above. The forecast gives the tercile probabilistic forecast of rainfall being in a particular category (above normal, near normal or below normal).
"

2. Material and methods

2.1. Observations and model data

Is there any literature reference on how the sub-division in 34 homogeneous regions has been done? A short explanation of the methodology would be useful, as these regions are the principle spatial reference for the analysis. Is each region represented by only one "representative" station?

Principal Component Analysis (PCA) was used to divide each country into homogeneous zones and only one representative rain gauge station is used for each zone. Unfortunately there is no published literature to refer to nonetheless the subdivision of these areas is a well established practice used in GHACOF. Since this paper wants to propose a methodology that could provide added information to the forecaster involved in the issuance of the consensus forecasts it was ideal to use the same subdivisions.

2.2. Quantitative assessment of the forecast skill

The phrase ("If CRPSS <= 0, no .....") is unclear or incomplete. Please re-phrase to make clear what your statement is. If CRPSS = 0, then the value of the forecast is equal to the value of a climatology. If it is < 0, than it is actually worse

The phrasing was changed to "CRPSS values above (below) zero denote forecast skill better (worse) in comparison with the reference forecast."

2.3. Qualitative assessment of skill

Headline: add "the forecast" (to be in line with section 2.2)

Was added.

First paragraph: It is unclear what happens here. What does manually smoothed actually mean? How have the proxies been generated? These are important issues in order to evaluate the results of this whole exercise. Please be more precise.
The GHACOF forecast maps were obtained from ICPAC as images and during the forecasting process, as explained in section 1, a large amount of information goes into the process. Therefore, the forecast maps generated by

the statistical models are later manually smoothed to cater for forecasters' knowledge and dynamical model forecasts. This has been better explained in the introduction now. See also our comment to your first comment.

3. Results and discussion

3.1. System-4 verification against in situ observations

Please re-phrase the first sentence, which is concise but a bit cryptic (e.g. Correlation coefficients between the precipitation anomalies derived from ECMWF system-4 forecasts and in situ measurements during the MAM and OND seasons . . . as well as CRPSSs are .....). This is also true for the headings of Figs 2 and 3.

We have decided that after all, this sentence is not needed and has been deleted accordingly figure headings for Fig 2 and 3 were rephrased

In Figures 4 and 5 I miss an explanation what the different bar-widths mean. Is the white line the mean or the median? Which percentiles are represented by the different part of the bars?

The boxplots extend from the minimum (whiskers), percentiles 10, 30, 50 (white line), 70, 90 and maximum. This was added to the plot

3.2. Use of system-4 in the consensus framework

This section conveys a very positive message. However, it would be interesting to discuss also the situation when the model fails. How should the decision maker manage the uncertainties? What are the consequences?

This is an interesting question, but beyond the scope of this paper. To assess this we would have to carry out an experiment in conjunction with the forecasters and this is a possible future application of the method.

It has to be noted that this paper is envisaged as a theoretical exercise of the possible added benefit of including in an operational framework the information provided by frequently updated model outputs.

The 5 authors independently did try to "quantify" the added benefit of having ECMWF products in a proxy consensus meeting simulations. Due to the subjective nature of the exercise it was nevertheless very difficult to extract a quantitative assessment. We have to leave this question open.

In Figs 7 to 9 the acronym ECFS4 is used without previous explanation in the text.

Sorry that was a typo. The acronym was changed to SYS-4, which is carefully explained in the text

4. Conclusions

In general well written. I have two remarks:

In the first paragraph you talk about statistical downscaling. This again refers to the GHACOF procedure and is explained nowhere in more detail. A said

before, it is necessary to clarify the procedure and the importance of the various information available to the forecasters involved.

Refer to our comments on your first comment.

I would further remark that (with reference to lines 6 to 9 on page 10219) the "reality check" performed is only valid for this particular region as it relies on specific teleconnections. The sentence as it stands now suggests a more general reliability, which is not proven by this analysis.

This is a fair comment. We have changed East Africa to "the study region".

Technical Corrections
Thanks for the notes. All corrections were done

Page 10210:

Line 5: insert "the" before "Standardized Precipitation Index (SPI)"

Line 10: insert "The forecast for" before "the October-December rain season
Line 11: "that" should read "than"
Line 11: insert "than the one" before "for the March-May season" Line 17: insert "and humanitarian" before "impacts since"
Line 21: "at-least" should read "at least"
Line 21: delete each in "one major drought per each decade" Page 10211:
Line 6: ";" should read ","

Line 16: "from other international centres" – do you refer to centres other than ECMWF? Which ones? Page 10212:
Line 16: "forecast" should read "forecasting . . ." Line 20: "gauges" should read "gauge"
Page10213:

Line 1: "precipitation experienced" would better read "precipitation climatology"
Line 22: "hindcast" should read "hindcasts"
Line 28: use capitals for "Analysis of the Correlation Coefficient" Page 10214:
We belive that Capitals should not be used for correlation coefficient apart from the acronym, neither for anomaly correlation coefficient nor standardized precipitation index.
Line 5: there should be a comma behind "members"

Line 11: "particularly effective" would better read "appropriate" Line 12: insert "a" before "probabilistic"
Line 20: "hind cast" should read "hindcast"

Line 22: "using the grid nearest neighbour being" Should read "using the nearest neighbour grid, being"

Page 10215:

Line 7: "output" should read "forecasts"

Line 19: "long-term precipitation record which is" should read "long-term precipitation records, which are"
Line 25: Insert "the" before "Standardized . . ." Line 26: No comma after "Index"

Page 10216:

Line 18: “forecast” should read “forecasts” Line 21: “horn” should be with capital H
Line 24: “season rain” should read “seasonal rains” Page 10217:

Line 4: Replace “For” by “Due to”

Was replaced by “because of”

Line 4: “qualitative” should read “quantitative” (please check!)

No, it is qualitative. However, qualitative should be quantitative in the previous sentence. The two sentences now read: “Because of the subjective nature of the consensus forecasts a purely quantitative assessment of its skill was not possible. We therefore resorted to perform a qualitative analysis based on subjective examinations of 11 years of forecast.”

Line 6: “yr” should read “years”
Line 18: “wet condition” should read “wet conditions” Page 10218:
Line 6: “or” should read “for” Line 14: delete “on”
Line 19: “Forums” should read “Forum” Page 10219:
Line 14: “informations” should read “information” Line 15: “exhisting” should read “existing”
Line 19: “Standardise” should read “Standardized”

Line 20: insert “data” after Precipitation and insert an “a” before “proxy”