

Assessing the skill of precipitation forecasts on seasonal time scales over East Africa from a Climate Forecast System model

Emily Bosire, Franklin Opijah,
Wilson Gitau

Department of Meteorology, University
of Nairobi, Kenya

Abstract

It is becoming increasingly important to be able to verify the skill of precipitation forecasts, especially with the advent of high-resolution numerical weather prediction models. This study focused on assessing the skill of climate forecast system (CFS) model in predicting rainfall on seasonal time scales over East Africa region for the period January 1981 to December 2009. The rainfall seasons considered were March to May (MAM) and October to December (OND). The data used in the study included the observed seasonal rainfall totals from January 1981 to December 2009 and CFS model forecast data for the same period. The model had 15 Runs. The measure of skill employed was the categorical skill scores and included Heidke skill scores, bias, probability of detection and false alarm ratio. The results from the categorical skill scores confirmed relatively higher skills during OND season as compared to MAM. When compared with individual Runs, the mean of all the 15 Runs depicted relatively higher accuracy during OND season. Some individual Runs – 1, 7, 9 and 10 – also performed better during OND season. During MAM season, the mean of all the 15 Runs showed relatively lower accuracy in predicting rainfall. Some individual Runs – 5, 10, 12 and 14 – performed better than the mean of all the 15 Runs. The prediction of seasonal rainfall over East Africa region using CFS model depends on the season considered. During MAM, the prediction of seasonal rainfall is better as Runs are fewer, which showed relatively higher averaged skills; on the other hand, during OND the prediction of seasonal rainfall is better when using the mean of all the 15 Runs.

Introduction

The economies of countries in East Africa largely depend on agriculture, which is highly vulnerable to the amounts and distribution of

rainfall.

Numerical weather prediction models consist of ensembles of integrations. In principle, numerical models that represent the dynamics of the atmosphere and ocean should be able to give better seasonal forecasts than purely statistical approaches, because of their ability to handle a wide range of linear and non-linear interactions and their potential resilience against a changing climate.^{1,2} If the models were perfect, dynamical seasonal prediction would yield better forecasts. In practice model errors limit the skill of seasonal forecast, and it remains unclear to what extent the present generation of numerical forecast models is able to challenge existing empirical methods for seasonal forecasting.^{1,3} The skill in predicting precipitation depends very strongly on the region and season considered. During an El Niño southern oscillation (ENSO) event, the skill of the dynamical model in predicting precipitation is much higher than during neutral conditions.⁴ Accurate seasonal to inter-annual climate monitoring and forecasting could therefore contribute to improve planning and the management of climate-sensitive activities involving agricultural, water resources, hydro-electric power supply and tourism, among others. There is no single *ideal way* to characterize and compare model performances. Most previous validation studies used measures of accuracy and skill to determine the similarity between observed and modeled data. In this paper, the skill of the climate forecast system (CFS) model in predicting rainfall on seasonal time scale over East Africa is evaluated.

Materials and Methods

In the analysis monthly observed rainfall data for 41 stations across East Africa obtained from IGAD Climate Predictions and Applications Center for the period 1981 to 2009 were used. Spatial distribution of stations used in the study is shown in Figure 1 and the details of the stations are found in Table 1. Also used in the study were CFS model forecast datasets obtained from the International Research Institute for Climate and society for the period 1981 to 2009. The spatial resolution of the model data is 1.875° longitude by approximately 1.9° latitude.

Delineated homogeneous rainfall zones over East Africa

Figure 2 shows the East Africa climatological rainfall homogeneous zones for both March to May (MAM) and October to December (OND) rainfall seasons. The zonings were used to discuss the results obtained.

Correspondence: Emily Bosire, Department of Meteorology, Chiromo Campus, University of Nairobi, P.O. Box 30197-00100, Nairobi, Kenya. Tel. +254725174548 E-mail: ebosire@uonbi.ac.ke

Key words: Season; Runs; Forecast; Observed data.

Contributions: EB, data collecting and analyzing, manuscript writing and references search; FO and WG, manuscript reviewing.

Conflict of interest: the authors declare no potential conflict of interest.

Note: part of this paper was presented in the Department of Meteorology, University of Nairobi, for the award of Master of Science Degree in Meteorology.

Acknowledgments: the authors thank the entire administration of the Department of Meteorology, University of Nairobi and IGAD Climate Prediction and Applications Centre for the provision of data, encouragement and understanding at our respective places of work.

Received for publication: 6 August 2013.

Revision received: 19 May 2014.

Accepted for publication: 22 May 2014.

This work is licensed under a Creative Commons Attribution 3.0 License (by-nc 3.0).

©Copyright E. Bosire et al., 2014
Licensee PAGEPress, Italy
Global Meteorology 2014; 3:5020
doi:10.4081/gm.2014.5020

Methodology

The methods used in this study included standardization of observed and predicted rainfall, and categorical statistics analysis. Standardization was used to make observed and predicted data consistent, *i.e.* each data type having the same kind of content and format.

Categorical statistics is needed to evaluate binary forecasts (yes/no) of the type of statements that an event will or will not happen. Categorical statistics is not only limited to binary forecast but also multi-category contingency tables can be used. In this study verification of forecasts was compiled using a 3 by 3 contingency table showing the frequency of below normal, normal and above normal forecasts and corresponding observations (Table 2). The marginal distributions of the observations and forecasts are the totals that are provided in the right columns and lower rows of the contingency tables, respectively. A perfect forecast system would have all the entries along the diagonal (A, E, I) only, with the other cells being equal to zero.

Categorical skill scores were used to analyze

the relationship between the model outputs and the observed rainfall values. There are a number of categorical skill scores used as verification measures, some of which are defined in the next subsection highlighting some of their properties. No single score, measure or index has been found to satisfy this verification controversy.

Several textbooks cover widely forecast verification methodologies.^{6,7} A historical survey on verification methodology has been compiled.⁸

Probability of detection

Probability of detection (POD), also known as the hit rate, is the number of correct divided by the total number of observed in each category. It measures the proportion of observed events successfully forecasted by the model. It ranges from 0 to 1 (100%) with a perfect score of 1 (100%). Equation 1 gives the formula for computing the POD for below normal, above normal and normal rainfall. Probability of detection is sensitive to hits and does not take into account the false alarms. It can be artificially improved by increasing the number of hits. While increasing the number of hits and reducing the number of false alarms is desirable, it is recommended that POD be examined together with false alarm ratio (FAR):

$$POD = \begin{cases} 100 \cdot \frac{A}{M} (\text{Below Normal}) \\ 100 \cdot \frac{I}{O} (\text{Above Normal}) \\ 100 \cdot \frac{E}{N} (\text{Normal}) \end{cases} \quad (\text{eq. 1})$$

False alarm ratio

The FAR defines the proportion of non-events for which a warning was provided incorrectly. The ratio is given by subtracting post agreement from 1. Post agreement is the ratio of correct to the total number of forecast in each category. The score ranges from 0 to 1 (100%) with a perfect score of 0. False alarm ratio for below normal and above normal rainfall is given by equation 2. False alarm ratio is sensitive to climatological frequency of the event. False alarm ratio is also sensitive to false alarms but takes no account of misses. It can be artificially improved by reducing the number of false alarms. Increase of POD is achieved by increasing FAR and a decrease of FAR by a decrease in POD, POD and FAR must therefore be examined together.

$$FAR = \begin{cases} 100 - 100 \cdot \frac{A}{I} (\text{Below Normal}) \\ 100 - 100 \cdot \frac{I}{E} (\text{Above Normal}) \end{cases} \quad (\text{eq. 2})$$

Bias score

Bias score, also known as the frequency bias index, compares the frequency of forecast rainfall events to the frequency of actual (observed) rainfall events. It measures the ability to forecast events at the same frequency as found in the sample without taking into account the accuracy of the forecast. Bias for

the below normal, normal and above normal rainfall is represented by the ratio in equation 3. Bias ranges from zero to infinity with a perfect score (unbiased score) of 1 (100%). It will indicate whether the forecast model is over forecasting (Bias>1) or under forecasting (Bias<1):

Table 1. Details of the rainfall stations used in the study.

| | Station | Country | Longitude | Latitude | Altitude asl (m) |
|-----|-------------|----------|-----------|----------|------------------|
| 1. | Gulu | Uganda | 32.33°E | 2.75°N | 1106 |
| 2. | Lira | Uganda | 32.93°E | 2.28°N | 1091 |
| 3. | Masindi | Uganda | 31.72°E | 1.68°N | 1146 |
| 4. | Soroti | Uganda | 33.62°E | 1.72°N | 1127 |
| 5. | Kasese | Uganda | 30.10°E | 0.18°N | 691 |
| 6. | Jinja | Uganda | 33.18°E | 0.45°N | 1173 |
| 7. | Tororo | Uganda | 34.17°E | 0.68°N | 1171 |
| 8. | Mbarara | Uganda | 30.65°E | 0.62°S | 1412 |
| 9. | Entebbe | Uganda | 32.45°E | 0.05°N | 1183 |
| 10. | Kabale | Uganda | 29.98°E | 1.25°S | 1867 |
| 11. | Lodwar | Kenya | 35.62°E | 3.12°N | 566 |
| 12. | Mandera | Kenya | 41.87°E | 3.93°N | 230 |
| 13. | Marsabit | Kenya | 37.90°E | 2.30°N | 1219 |
| 14. | Wajir | Kenya | 40.07°E | 1.75°N | 244 |
| 15. | Kakamega | Kenya | 34.78°E | 0.28°N | 1555 |
| 16. | Kisumu | Kenya | 34.75°E | 0.10°S | 1146 |
| 17. | Kisii | Kenya | 34.78°E | 0.67°S | 1171 |
| 18. | Kericho | Kenya | 35.55°E | 0.37°S | 2096 |
| 19. | Nakuru | Kenya | 36.10°E | 0.27°S | 1859 |
| 20. | Narok | Kenya | 35.83°E | 1.13°S | 1890 |
| 21. | Dagoretti | Kenya | 36.75°E | 1.30°S | 1798 |
| 22. | Makindu | Kenya | 37.83°E | 2.28°S | 1000 |
| 23. | Lamu | Kenya | 40.83°E | 2.27°S | 7 |
| 24. | Garissa | Kenya | 39.63°E | 0.48°S | 128 |
| 25. | Colchecchio | Kenya | 36.80°E | 0.63°N | 145 |
| 26. | Eldoret | Kenya | 35.28°E | 0.52°N | 2104 |
| 27. | Voi | Kenya | 38.57°E | 3.40°S | 579 |
| 28. | Malindi | Kenya | 40.10°E | 3.23°S | 3 |
| 29. | Mombasa | Kenya | 39.62°E | 4.03°S | 57 |
| 30. | Bukoba | Tanzania | 31.82°E | 1.33°S | 1143 |
| 31. | Musoma | Tanzania | 33.80°E | 1.50°S | 1147 |
| 32. | Mwanza | Tanzania | 32.92°E | 2.47°S | 1139 |
| 33. | Kigoma | Tanzania | 29.63°E | 4.88°S | 999 |
| 34. | Dodoma | Tanzania | 35.77°E | 6.17°S | 1120 |
| 35. | Dar Airport | Tanzania | 39.20°E | 6.87°S | 53 |
| 36. | Arusha | Tanzania | 36.38°E | 3.22°S | 1387 |
| 37. | Same | Tanzania | 37.40°E | 10.4°S | 872 |
| 38. | Tabora | Tanzania | 32.50°E | 5.05°S | 1182 |
| 39. | Mbeya | Tanzania | 33.47°E | 8.93°S | 1707 |
| 40. | Songea | Tanzania | 35.58°E | 10.68°S | 1067 |
| 41. | Mtwara | Tanzania | 40.18°E | 10.27°S | 113 |

$$BIAS = \begin{cases} \frac{J}{M} & (\text{Below Normal}) \\ \frac{L}{O} & (\text{Above Normal}) \\ \frac{K}{N} & (\text{Normal}) \end{cases} \quad (\text{eq. 3})$$

Heidke skill score

The Heidke skill score (HSS) measures the fraction of correct forecasts after eliminating those forecasts which would be correct due purely to random chance. It is the ratio of the sum of all entries in the diagonal minus the rows and columns sums divided by the sum total to sum total minus the rows and columns sums divided by the sum total to sum total. Heidke skill score is a measure of potential improvement in the number of correct forecasts over random forecasts. The score ranges from $-\infty$ to 1, with a perfect score of 1 and no skill forecast equal to zero. Any score that is less than zero means that the forecast model is worse off than climatology. Heidke skill score was computed using equation 4:

$$\text{Heidke Skill Score} = \frac{(A + E + I) - \frac{(JM + KN + LO)}{T}}{T - \frac{(JM + KN + LO)}{T}} \quad (\text{eq. 4})$$

Results and Discussion

The results of HSS, bias, POD and FAR calculated from a three by three contingency tables for some selected Runs for the East Africa region are shown in Figures 3 to 11.

The results of HSS are presented in Figure 3. For a forecast to be termed as perfect, the HSS should be 1 (100%). None of the Runs presented has values close to this perfect score. In Figure 3a, relatively higher HSS are within stations in Zones 2 and 8 representing the central and western Tanzania, central and western Uganda, respectively. In Figure 3b, relatively higher HSS are seen within stations in Zone 9, especially northern Uganda. In Figure 3e, relatively higher HSS are observed within stations in Zones 2 and 6 representing the east of Lake Victoria and northern Uganda and north west Kenya, respectively. Figure 3f, relatively higher HSS are observed over some stations in Zones 3 and 5 representing the southern and coastal Kenya respectively. Figure 3g, relatively higher HSS are seen within Zone 3, representing the central and northern Tanzania.

The results for bias score are presented in Figures 4, 5 and 6 for the above normal, below normal and normal rainfall categories, respectively.

In Figure 4, it is observed that during MAM season, perfect scores are noted in Runs 2, 8 and 10. These are presented in Figure 4a, b

and c. In Figure 4a, perfect scores are evident within some stations in Zones 4, 5, 8 and 9 representing the central highlands of Kenya, east of Lake Victoria, central and northern Uganda and north west Kenya, respectively. In Figure 4b, some stations within Zones 1 and 7 representing the southern Tanzania and north eastern Kenya have perfect scores. Chances of under forecasting exceed those of over forecasting for the Runs presented during MAM

season. For the OND season, perfect scores are evident in Runs 1, 9 and 14. These are shown in Figure 4e, f and g. In Figure 4f, east of Lake Victoria within Zone 2 has perfect scores. In Figure 4g, some stations within Zones 1 and 6 representing the southern Tanzania and northern Uganda, respectively have perfect scores. Chances of over forecasting exceed those of under forecasting for all the Runs presented during the OND season.

Table 2. A 3 by 3 contingency table showing the frequency of below normal, normal and above normal forecasts and corresponding observations.

| Observed | Forecast | | | Total |
|--------------|--------------|--------|--------------|-------|
| | Below normal | Normal | Above normal | |
| Below normal | A | B | C | M |
| Normal | D | E | F | N |
| Above normal | G | H | I | O |
| Total | J | K | L | T |

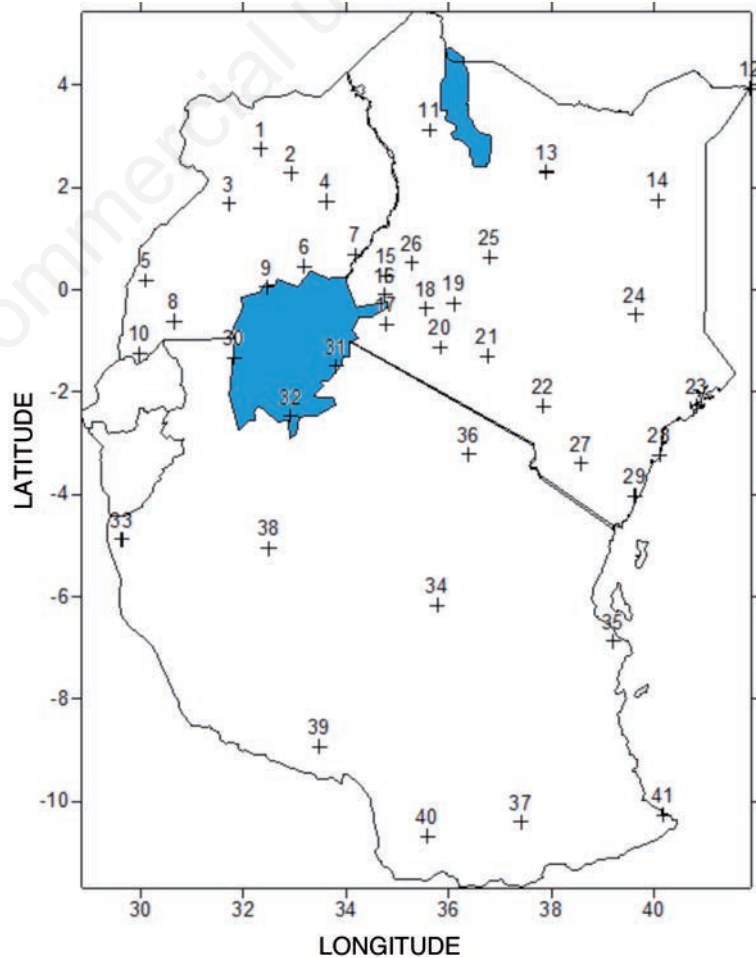


Figure 1. Study domain showing the spatial distribution of the rainfall stations numbered as in Table 1.

In Figure 5a, b, c, e, f, g and h it is observed that for the MAM season more stations with perfect scores were obtained in Runs 1, 2 and 7. Chances of over forecasting exceed those of under forecasting. For the OND season, Runs 8, 9, 15 and the mean of all the 15 Runs shows more stations with perfect scores as compared to the other Runs. In Figure 5c, perfect scores are seen within stations in Zones 3, 5 and 9 representing the coastal areas of Kenya, east of Lake Victoria and north west of Kenya, respectively. In Figure 5e, perfect scores are observed within stations in Zones 1 and 3 representing the southern and the northern Tanzania, respectively.

In Figure 6a, b, c, e and g it is observed that for the MAM season, more stations with perfect scores are seen in Runs 7, 9 and 11. Chances of over forecasting seem to be equal to those of under forecasting. For the OND season chances of under forecasting exceed those for over forecasting in most of the Runs but Run 5 and 13 have perfect scores at 10 stations. In Figure 6b, perfect scores are observed in Zones 3 and 5

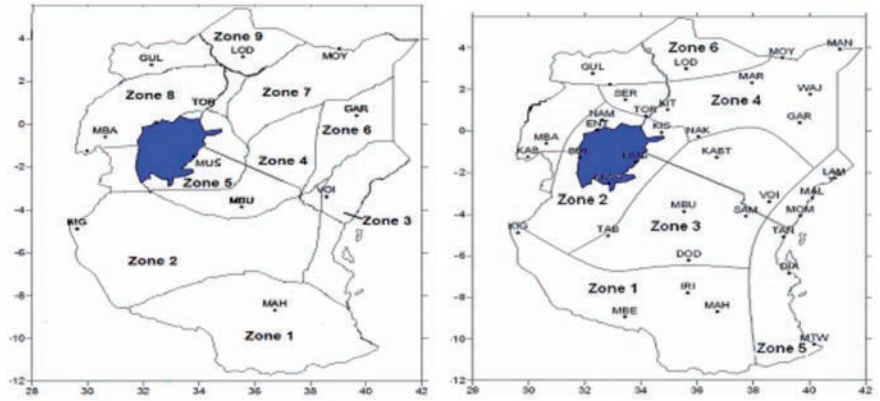
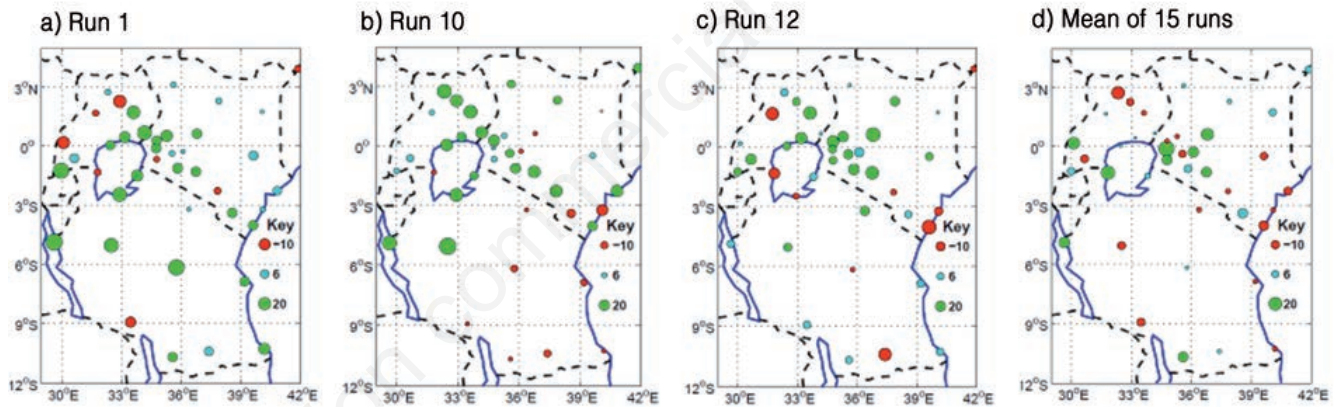


Figure 2. Climatological zones in East Africa for March to May season (left panel) and October to December season (right panel).⁵

March to May season



October to December season

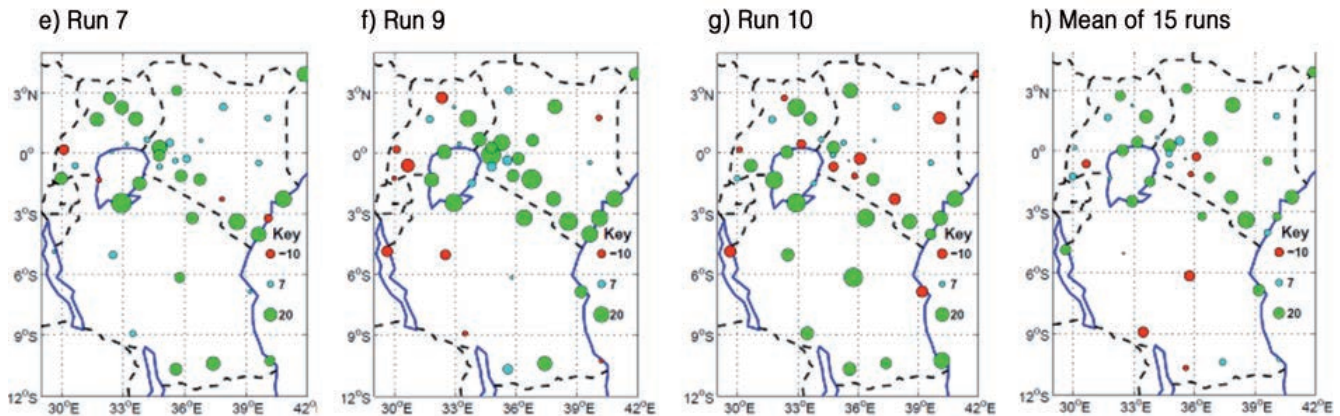


Figure 3. Spatial distribution of Heidke skill scores (%) for the period 1981 to 2009. The red and green colors indicate that the model performs worse off and better than climatology, respectively.

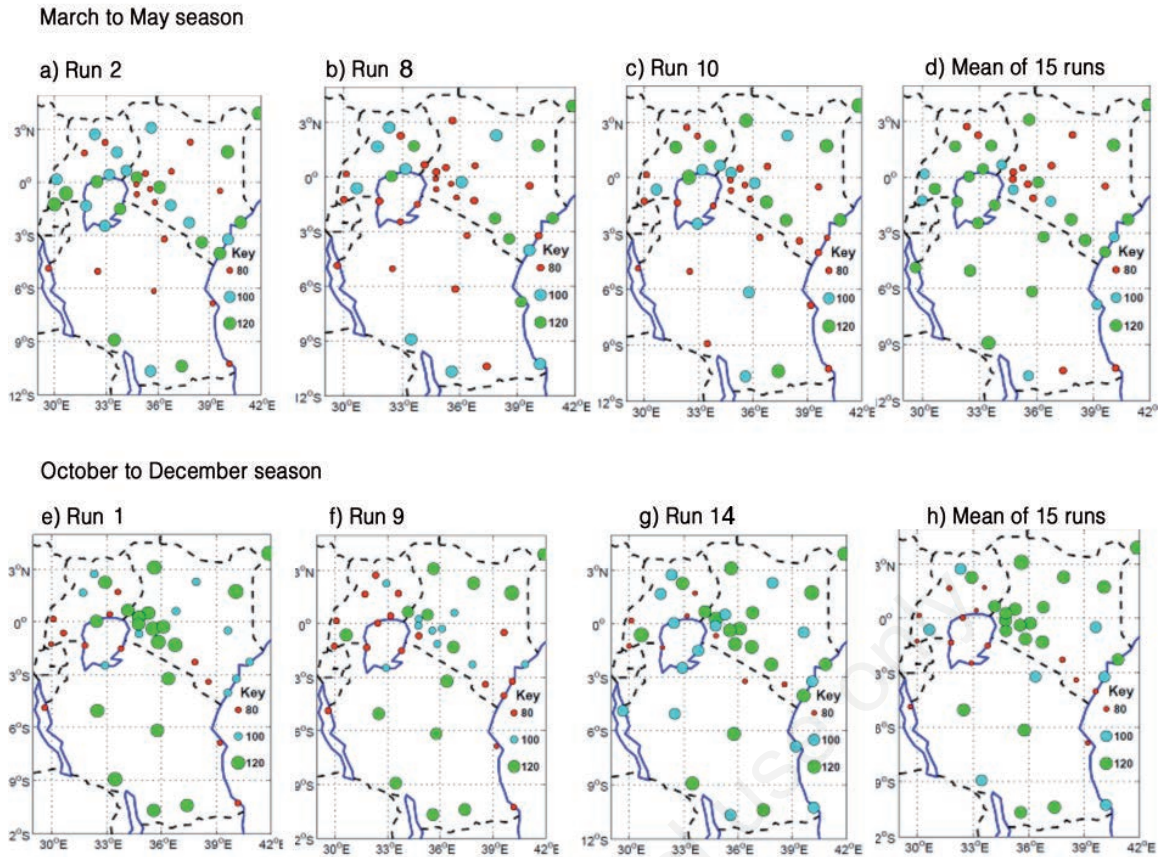


Figure 4. Results for bias score for the above normal rainfall categories.

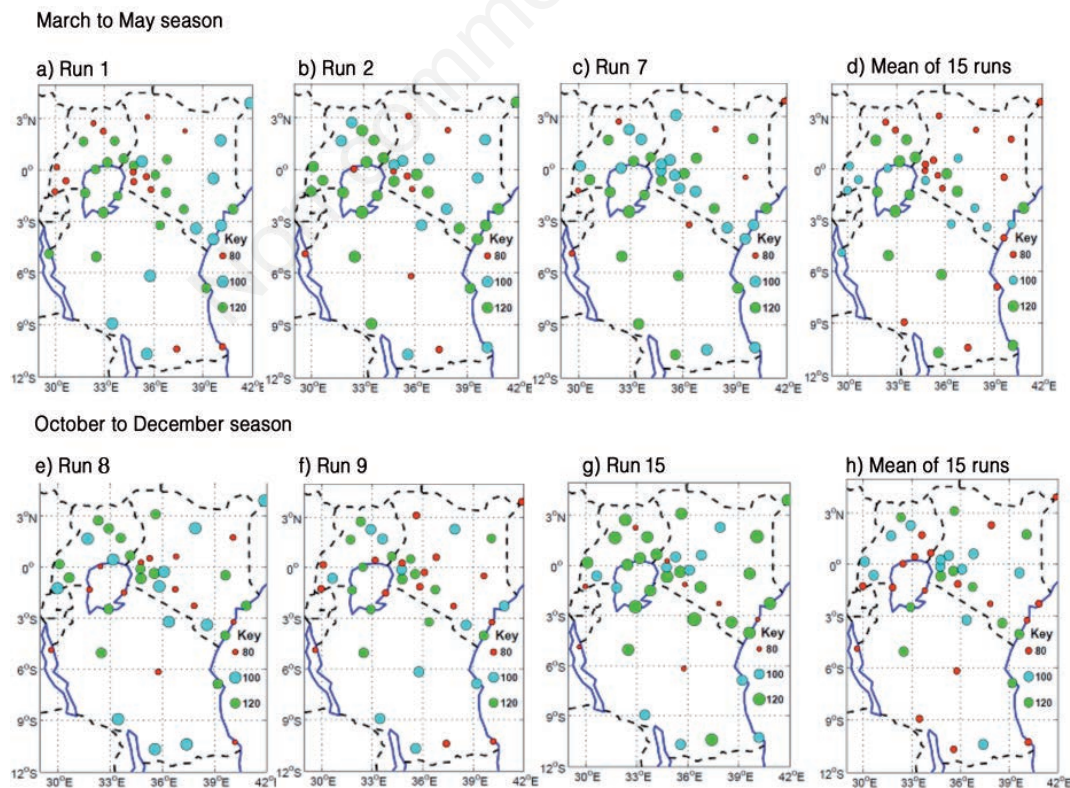


Figure 5. Results for bias score for the below normal rainfall categories.

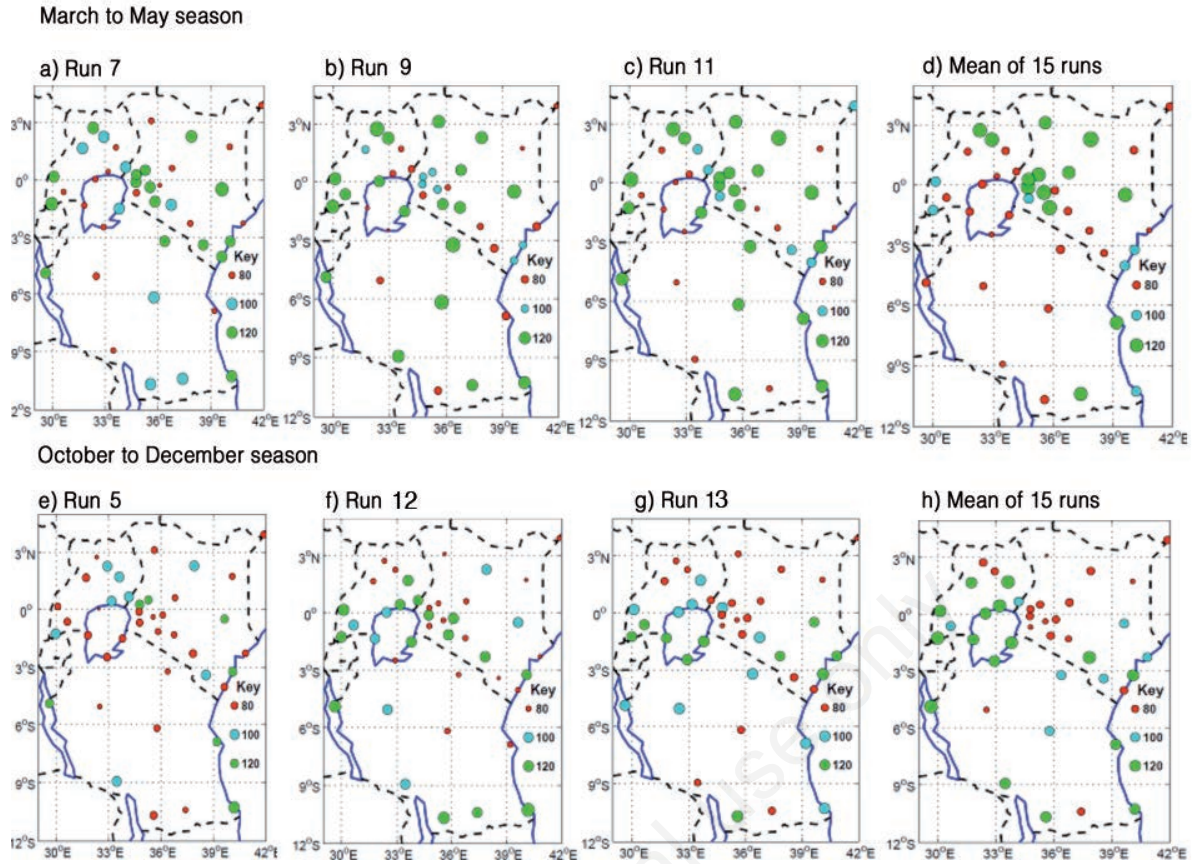


Figure 6. Results for bias score for the normal rainfall categories.

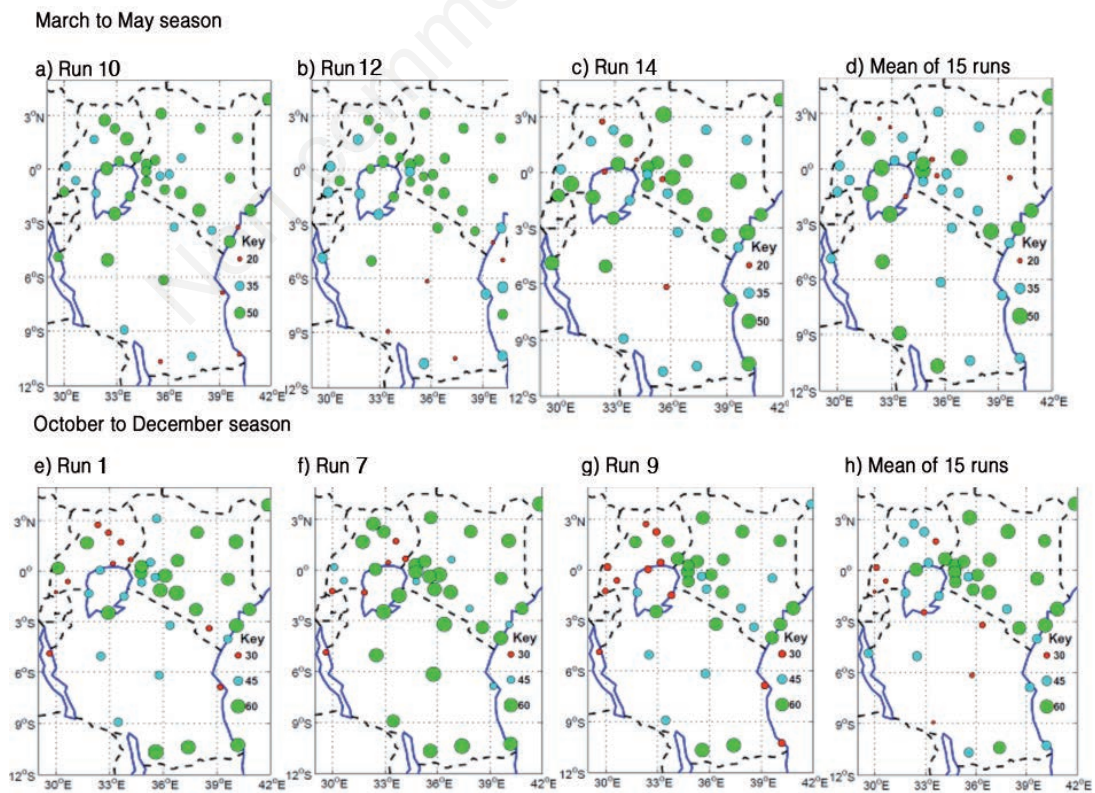


Figure 7. Results for probability of detection for the above normal rainfall categories.

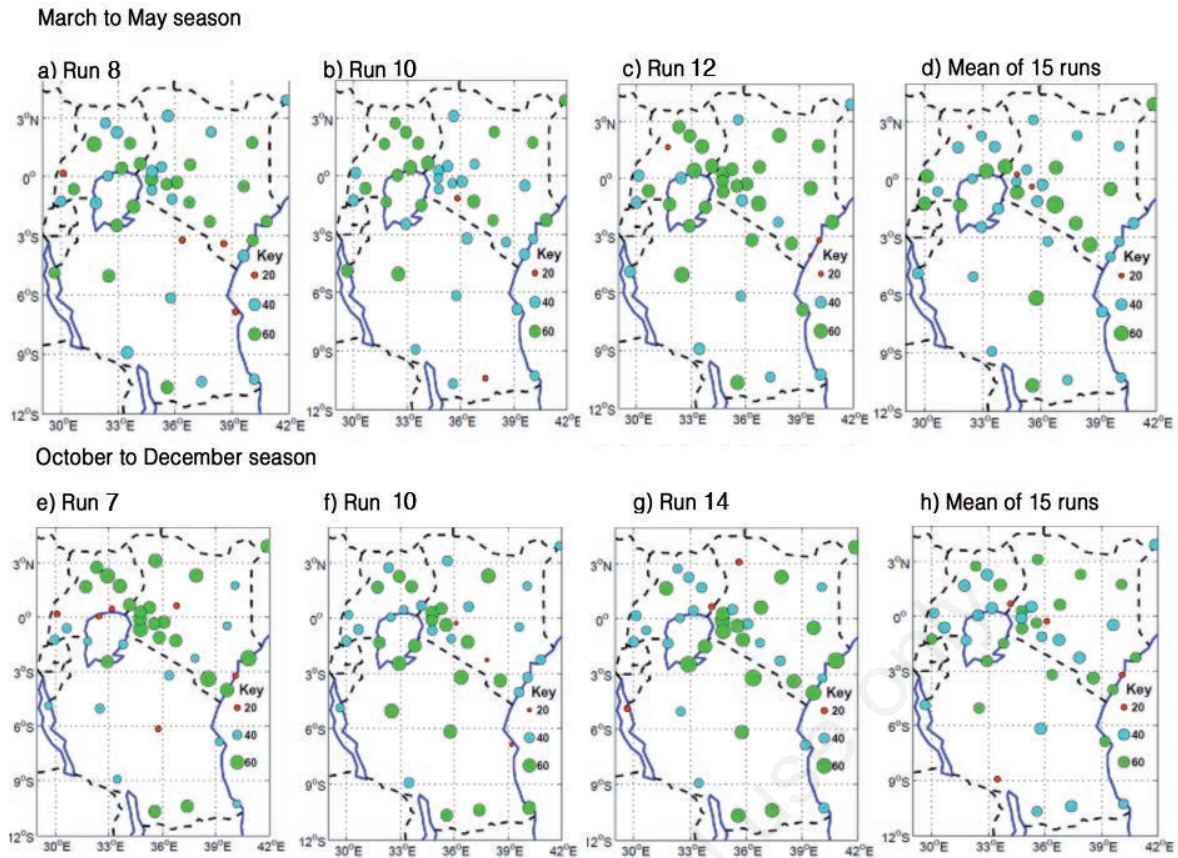


Figure 8. Results for probability of detection for the below normal rainfall categories.

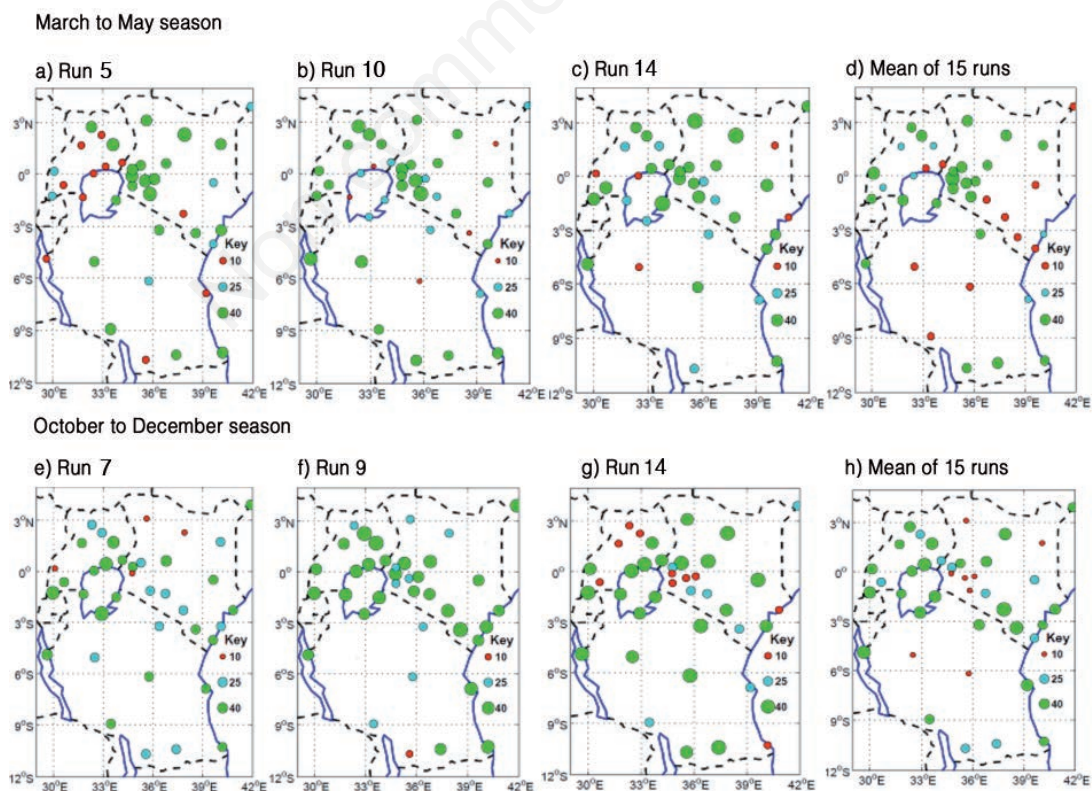
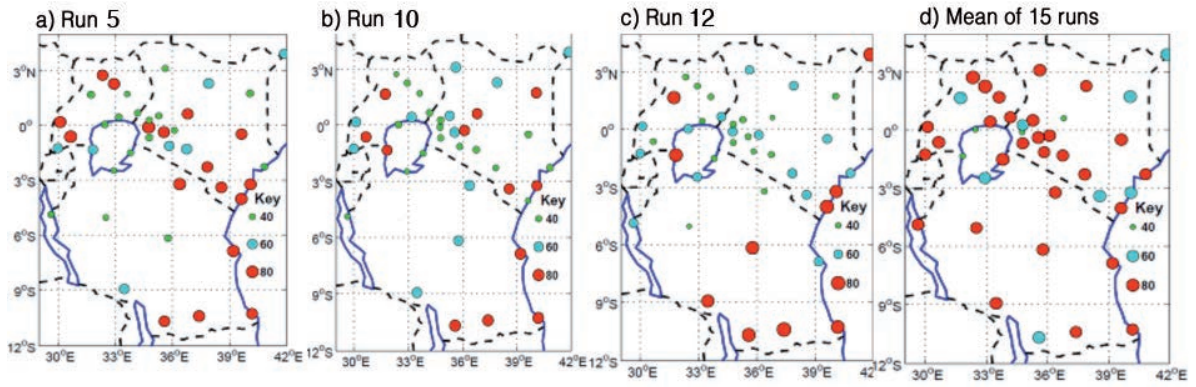


Figure 9. Results for probability of detection for the normal rainfall categories.

March to May season



October to December season

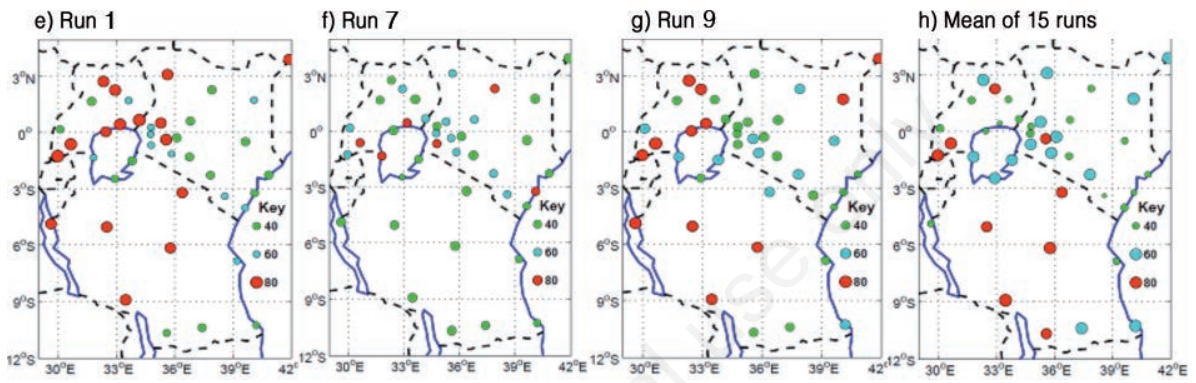
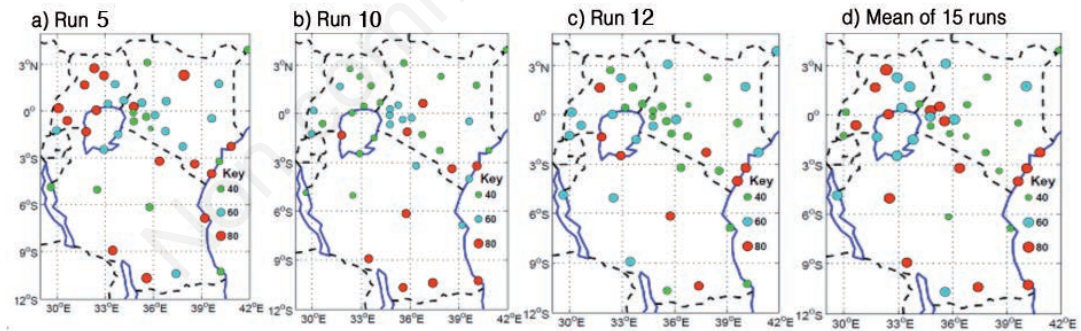


Figure 10. Results for false alarm ratio for the above normal rainfall categories.

March to May season



October to December season

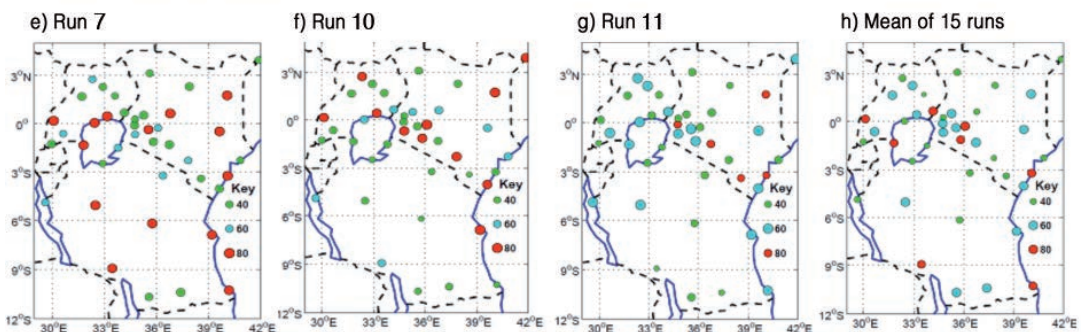


Figure 11. Results for false alarm ratio for the below normal rainfall categories.

representing the south coast of Kenya and east of Lake Victoria, respectively.

Conclusions

Some individual Runs (for example Runs 5, 10, 12 and 14 for the MAM season, and Runs 1, 7, 9 and 10 for the OND season) performed better than the mean of all the 15 Runs in most cases. During the OND season the mean of all the 15 Runs showed relatively higher skill when compared to the individual Runs. The study has shown that the skill of the CFS model is highly dependent on the season considered. The skill is relatively higher during OND season as compared to the MAM season. The skill of the CFS model also depends on the ENSO events: it is relatively higher during the start of the ENSO events (OND season) and it becomes relatively lower towards the end of the ENSO events (MAM season).

In conclusion, the study has shown that the CFS model is able to predict rainfall on seasonal time scales in the East Africa region during OND season. Reliable predictions of seasonal rainfall would allow more efficient planning in various sectors like agriculture, hydrology, energy, health, insurance, finance among others therefore, improving the quality of life, health, and safety. Following the results obtained, the use of the mean of all the 15 Runs will improve seasonal rainfall prediction over the East Africa region during OND season and hence improve early warning. Early warning will help in reducing the impacts of climate extremes on agricultural production, hydropower generation, water use, and water availability at household level. The results of

this research can be used as a guide by all forecasters within East Africa to identify which CFS Runs skillfully predict rainfall in the region so as to reduce high computing resources and time.

There is need to improve the initial conditions of the CFS model especially the oceanic ones since in the CFS model the pentad Runs use the same oceanic conditions. This will help in producing a good forecast. Matching observations and forecast is the most difficult part in verification process. Matching approach considered can have an impact on the results of verification. There are different ways of matching observations and forecast; these include point to grid or grid to point. In this study observed rainfall data was matched to grid. Further studies can be done by interpolating gridded model data to point data to check for similarities or differences. There are different ways of assess the quality of a forecast too. The method employed in this study is looking at the skill of the forecast. Further studies can be done by using other measures of forecast such as bias, resolution and sharpness. This is because focusing on only one measure forecast quality may be misleading.

Further studies of the systems that influence MAM rainfall are however still required to improve the skill of MAM season. A similar study should be done for other weather parameters such as temperature and wind among others in order to understand the overall performance of the model.

References

1. Latif M, Sperber K, Arblaster J, et al. ENSIP:

the El Niño simulation Intercomparison project. *Clim Dynam* 2001;18:255-76.

2. Palmer T, Alessandri A, Andersen U, et al. Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull Amer Meteor Soc* 2004;85:853-72. Available from: <http://journals.ametsoc.org/doi/abs/10.1175/5/BAMS-85-6-853>
3. Barnston AG, Glantz MH, He Y. Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997/98 El Niño Episode and the 1998 La Niña Onset. *Bull Amer Meteor Soc* 1999;80:217-43. Available from: <http://journals.ametsoc.org/doi/abs/10.1175/1520-0477%281999%29080%3C0217%3APSOSA-D%3E2.0.CO%3B2>
4. Geert JVO, Magdalena AB, Ferranti L, et al. Did the ECMWF seasonal forecast model outperform statistical ENSO forecast models over the last 15 years? *J Climatol* 2005; 18:3240-9.
5. Mutua FM, Mutemi JN, Oludhe C. Homogeneous climatological zoning. *DMC Lecture Notes* 1999;3:29-43.
6. Jolliffe IT, Stephenson DB. *Forecast verification: a practitioner's guide in atmospheric science*. Hoboken, NJ, USA: Wiley; 2003.
7. Wilks DS. *Statistical methods in the atmospheric sciences: an introduction*. Forecast verification. Waltham, MA: Academic Press; 1995.
8. Stanski HR, Wilson LJ, Burrows WR. Survey of common verification methods in meteorology 1989. WMO Research Report No. 89-5. Available from: http://www.cawcr.gov.au/projects/verification/Stanski_et_al/Stanski_et_al.html