Journal of Climatology & Weather Forecasting
OPEN ACCESS Freely available online
Research Article

Assessing the Potential of Using ECMWF system-4 Seasonal Rainfall Forecasts over Central-West Ethiopia

Yate TA[1*], Hutjes RWA[2]

[1]Meteorology and Climate Chair, Water Technology Institute, Arba Minch University, Ethiopia; [2]Department of Environmental Sciences, Earth System Science Chair Group, Wageningen University and Research Center, the Netherlands

ABSTRACT

Generally, climate variability has imposed formidable uncertainties and risks in productivity in tropics and sub-tropics. More specifically, seasonal rainfall variability has posed serious challenges to management practices that are merely dependent on climatological probabilities. Moreover, climate change has imposed unequivocal uncertainties and risks in productivity. As Ethiopia relies on natural rainfall for most of its water supply, exacerbated by existing marginal capabilities, it is very sensitive to seasonal rainfall variability. Therefore, using seasonal rainfall forecasts provides a promising opportunity to deal with such risks in advance and to improve sectorial resilience. However, in order to be usable for such applications, the rainfall forecasts need to provide better information than the climatology-based predictions. In technical expression, the seasonal rainfall forecasts must have some skill. In this study, skill of seasonal rainfall forecasts from European Centre for Medium-Range Weather Forecasts (ECMWF) Systems-4 was assessed over central-west Ethiopia using 'WATCH Forcing Data methodology applied to ERA-Interim data' (WFDEI) meteorological forcing data for validation. We used mainly one deterministic metrics and two probabilistic metrics to assess forecast skills. The area under the ROC curve for precipitation forecast below the lower tercile was computed to be 0.53, 0.54 and 0.6 at 0-, 1- and 2-months lead-times respectively. The area under the ROC curve for the upper tercile category was 0.54, 0.44 and 0.38 at 0-, 1- and 2-month lead-times respectively. Whereas, the Ranked Probability Skill Scores (RPSS) were computed to be -0.108, -0.1713 and -0.1226 at 0-, 1- and 2-month forecast lead-times respectively. The predictive skill in ECMWF System-4 precipitation forecasts over the study area was generally poor, and nearly no benefit gained at any lead-time compared to the climatology.

Keywords: Skill; Seasonal forecasting; Forecast verification; Hindcast

INTRODUCTION

Ethiopia relies on natural rainfall for most of its water supply [1]. This makes the country very sensitive to seasonal rainfall variability. The existing marginal capabilities amplify the vulnerability even more. Any significant deviation from the normal seasonal rainfall may lead to a disaster that can severely affect many crucial sectors of the country. The agriculture sector, that forms the basis of the country's economy [2], obtains its water principally from natural rainfall making crop production susceptible to the risks of climate variability. Most severe droughts that Ethiopia experiences are often related to negative anomalies of seasonal rainfall, particularly of June-September (JJAS) rainfall [3]. This has been demonstrated with a number of droughts in Ethiopia in the last few decades which left the country with wide spread food crises [4-8]. Hence, monitoring seasons using seasonal rainfall forecasts is very crucial. Application of seasonal rainfall forecasts, at usable lead time, provides a promising opportunity to deal with climate variability risks in advance and to improve sectorial resilience.

As a result of advances in dynamical seasonal forecasting and its promise to avoid risks in advance, also due to a growing threat of climate change, seasonal rainfall forecasts are being growingly used to benefit decision-making in climate-sensitive sectors [9]. The recent improvement in the global seasonal forecast system has even intensified the prospect of forecast applications. However, in order to be usable for such applications, seasonal rainfall forecasts from dynamical physical models need to provide better information than the climatology-based predictions. In technical expression, the seasonal rainfall forecasts must have some skill [10]. Otherwise, substituting climatology-based predictions with seasonal rainfall forecasts becomes worthless. A skilful forecast is

Correspondence to: Tewodros A. Yate, Meteorology and Climate Chair, Water Technology Institute, Arba Minch University, Ethiopia; E-mail: tewodros.addisu@amu.edu.et

Received: February 25, 2021; **Accepted:** March 22, 2021; **Published:** March314, 2021

Citation: Yate TA, Hutjes RWA (2021) Assessing the Potential of Using ECMWF system-4 Seasonal Rainfall Forecasts over Central-West Ethiopia. J Climatol Weather Forecast. 9:279.

the one that has better correspondence to ensuing observations than simple climatological means have. It measures the accuracy of forecasts relative to climatological means or other standard reference forecasts.

Hence, assessing the skill of seasonal rainfall forecasts is very crucial by carrying out forecast verification if their application is to benefit decision-making. Murphy and Winkler [10] explained forecast verification as a process that investigate properties of "joint distribution of forecasts and observations" to determine quality of forecasts. The process aims at improving forecasting methodologies, and then ultimately the forecast products, basing on results from investigation of the joint distribution of forecasts and observations [10]. Some of the commonly used metrics for verification of a probabilistic forecast of an event include Ranked Probability Score, or Brie Score for binary event forecasts, Probability score, Relative Operating Characteristics (ROC) curve, reliability diagram and logarithmic score [11].

Forecasts need to score well above the "no skill" thresholds of verification metrics in order to be good enough to benefit appropriate users [10]. Skillful forecasts are forecasts that we can use with sufficient confidence to force agricultural and hydrological models. When droughts are predicted with sufficient confidence, disaster preparedness may be initiated to prevent the worst episodes of wide spread famine. When a highly productive season is forecasted with sufficient confidence, farmers may want to invest more in high yield crop variety seeds and fertilizer.

Scientific researches focusing on assessing skill of seasonal climate forecasts has dramatically increased [12-14]. There were also few related studies specific to seasonal climate of Ethiopia. Gissila, Black [15] developed seasonal forecast scheme for June to September rainfall of Ethiopia using a simple regression method and assessed its skill. Whereas, Korecha and Barnston [3] examined predictability of Kiremt rainfall of Ethiopia "based on the climate state prior to the onset of the rainy season using statistical techniques". Also, Diro, Grimes [16] evaluated the correspondence between gauge data and reanalysis rainfall estimates from ERA-40 and NCEP, numerical weather prediction models, for June to September and February to May rainfall of Ethiopia.

In this study, skill of seasonal rainfall forecasts from European Centre for Medium-Range Weather Forecasts (ECMWF), one of world's most renowned prediction centers with global coverage, is evaluated over central-west Ethiopia. This study specifically uses forecast products from ECMWF System4 forecast system. ECMWF System-4 is a latest seasonal state-of-the-art forecasting system providing high quality forecast products [17]. More on the forecasting system and its forecast products is found in the materials and method section of this report. The study used 'WATCH Forcing Data methodology applied to ERA-Interim data' (WFDEI) meteorological forcing data to validate System-4 reforecasts.

MATERIALS AND METHODS

Study area

The study area, manually designed for the purpose of this work, is located in the central-west Ethiopia in 7.5°N-10.5°N and 34°E-37.5°E latitudinal and longitudinal ranges respectively. It encompasses 68 weredas, a national term to 3rd-level administrative divisions of Ethiopia; 59 from Oromia, 6 from Benishangul-Gumuz, 2 from SNNPR and 1 from Gambela regions. The areal altitude ranges from 550 m.a.s.l to more than 2,500 m.a.s.l, as shown in Figure 1. The study area falls within rainfall regime "A" of the country where mono-modal rainfall pattern is observed [16,18]. The main rainfall period runs from June to September, with seasonal rainfall climate ranging from 800 mm to 1400 mm [3]. As a result, these four months constituted the growing season of the area [18].

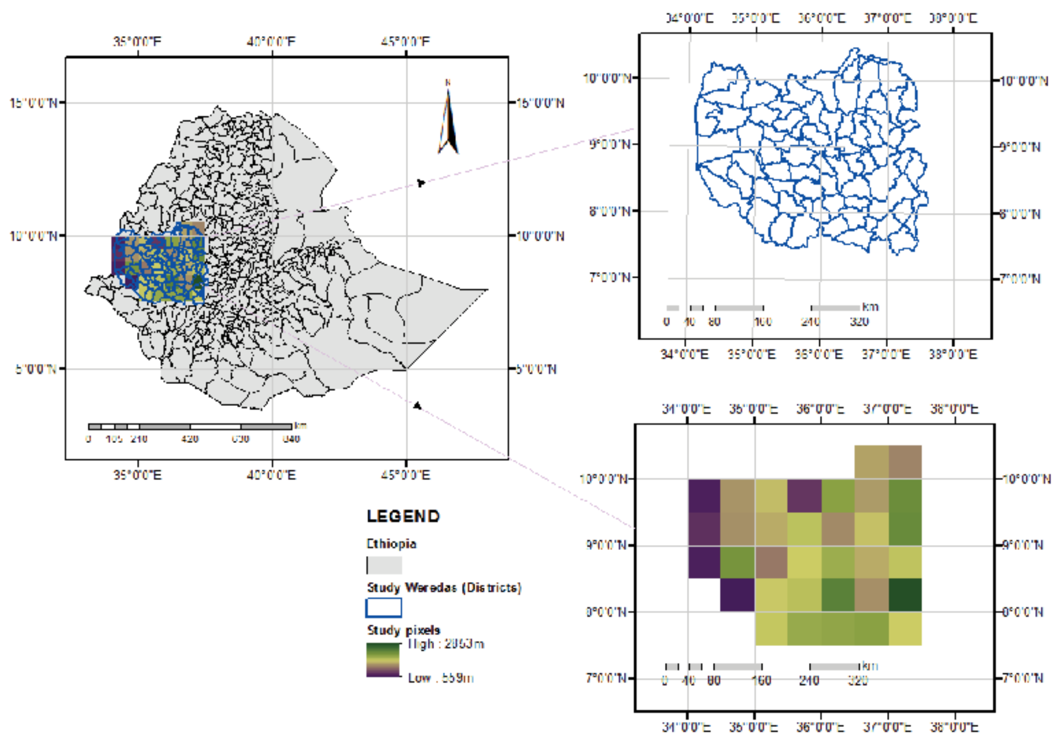


Figure 1: The study area is depicted in the central-west Ethiopia; (a): The upper figure on the right hand side shows administrative boundaries of the weredas of the study area; (b): the lower figure on the right handside shows the elevation of the study area on 0.5° × 0.5° pixels.

Data

In this study, skill of seasonal rainfall forecasts from European Centre for Medium-Range Weather Forecasts (ECMWF) is evaluated over central-west Ethiopia. The European Centre for Medium-Range Weather Forecasts (ECMWF) generates and provides numerical weather predictions. ECMWF's forecasting system gathers an atmospheric general circulation model, an ocean wave model, an ocean circulation model, a land surface model and perturbation models. It produces medium-range (up to 15 days), extended-range (up to 30 days) and long-range (up to 7 months) forecasts [19]. Since the aim of this study was to assess the skill in seasonal forecasts, the long range forecast data were used. Long-range retrospective forecasts were produced using seasonal forecast System 4, a system of fully coupled atmospheric and oceanic models, for the seven months ahead of every month since 1981 [17,20]. The forecast product includes seasonal forecasts of temperature, precipitation and means sea level pressure and sea surface temperature variables. In this study, ECMWF system4 long-range forecasts for 30 years with starting dates from first of January, 1981 to first of December, 2010 were used. Thus, for example, a data file of January contains forecasts for the months January (lead month 0), February (lead month 1) and so on up to June (lead month 6) for all the 15 ensemble members.

'WATCH Forcing Data methodology applied to ERA-Interim data' (WFDEI) meteorological forcing data was produced by European Water and Global Change Program (WATCH). WATCH program was a research project, ran between 2007 and 2011, bringing together several hydrological and climate experts from outstanding European research centres and institutes (http://www.eu-watch.org/). WATCH produced a large global datasets that can be used to force hydrological and land surface models. WFDEI dataset is one of these forcing datasets comprising of eight meteorological variables, for the global land surface at 0.5° × 0.5° grid resolution, at 3 hourly time step and also as daily averages for the period 1979 to 2012 [21]. Monthly bias corrections were made using CRU TS3.1/TS3.101/TS3.21 and GPCCv5/v6 data resulting in two WFDEI dataset catalogues to choose from. In this study, the WFDEI forcing data with monthly bias correction based on CRU was used for forecast verification.

Method

Most of statistical and graphical analysis in this study were done using R program (http://www.r-project.org/). ArcGIS program was used to produce maps including the study area. Bilinear method was applied to interpolate data from ECMWF System4 0.75° grids to 0.5° grids in order to match the horizontal resolution of WFDEI. We used mainly one deterministic metrics and two probabilistic metrics to assess forecast skills. These were Pearson product moment correlation coefficient, Relative Operating Characteristics (ROC) curve (including ROC score) and Ranked Probability Skill Score (RPSS).

Correlation coefficient is a measure of a linear correspondence between forecasts and the corresponding observations of the predicted. Pearson product moment correlation coefficient can be computed as [22]:

$$r_{f^\circ} = \left[\frac{\sum_{i=1}^{n} (f_i - \overline{f_i})(o_i - \overline{o_i})}{(n-1)(s_{f*} s_o)} \right] \quad \text{(Equation 1)}$$

Where r_{f° is Pearson product moment correlation coefficient between *f*, forecast, and 0, observation; *i* is the index of each elements of the data; *n* is total number of observation or forecast; *S* is standard deviation. Correlation values range between -1 and 1, with -1 indicating perfect negative linear association and 1 indicating perfect positive linear association. Whereas 0 score show absence of linear correspondences between the forecasts and observations [22]. However, it is noteworthy to mention that correlation coefficient can misleadingly be low for distributions with high non-linear association, as this metric often measure linear association [22]. In order to particularly examine forecast performances in anomalous seasons, correlation coefficients were calculated between seasonal precipitations forecast anomaly and observation anomaly of the period 1981 to 2010. The climatology mean obtained from the observation datasets were subtracted from the ensemble mean forecasts and from the observation datasets to get the corresponding anomalies, which were then used to compute anomaly correlation coefficients. The ability of forecasts to predict anomalies is suggested as one of the key factors determining their usefulness [17]. Murphy and Epstein [23] recommended squared values of anomaly correlation coefficients to consider as a measure of potential forecast skills. We also computed correlation coefficient between forecast and observation considering four anomalous years over the study area: 1985, 1991, 2004 and 2009.

Relative Operating Characteristics (ROC) curve is a signal detection curve for dichotomous categorical forecasts plotted with hit rate value on the y-axis and false alarm rate value on the x-axis [12]. The area under the ROC curve was computed using a method proposed by Mason and Graham [24]. The area measures the ability of a forecast system to correctly predict the occurrence or non-occurrence of predefined events [24]. When the observation occurs, it is referred as event, when not it is referred as no-event. Thus, ROC area weighs the forecast probability of successfully discriminating an event from a no-event. ROC curve is commonly used method to explore if forecast is well discriminated given different outcomes [11]. When the forecasts have some skills the

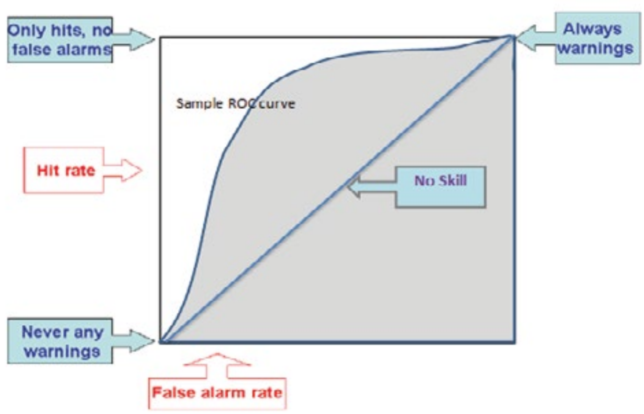


Figure 2: ROC diagram regions, adopted and modified from Persson and Grazzini [25]. The grey shaded region is area under ROC curve. The diagonal line, where area is 0.5, represents no skill in the forecast system, whereas the maximum area of 1 represents perfect forecast.

area under ROC curve exceeds 0.5, with the maximum value being 1 [24].

Since this method requires the categorical forecasts to be of binary outcome, separate results were calculated for events of the lower tercile and events of the upper tercile in this study. In order to do this, first, the ensemble forecast was transformed to forecast frequency distribution using tercile probability threshold, resulting in 30 × 3 matrixes with each column containing the number of forecast ensembles falling in the lower, middle and upper tercile categories of the corresponding years. Then, forecast probabilities for each of the lower and upper tercile categories was computed separately. Similarly, the observed dataset was also transformed to frequency distribution using tercile probability threshold, which again resulted in 30 × 3 matrixes, but in this case the rows of the tercile columns were populated with 1 when the observation falls in and 0 when the observation falls in the other terciles, and so on. Finally, the ROC diagram was produced separately for the lower and upper tercile categories. The top left corner of the ROC diagram, as seen in the figure 2, characterizes a forecast system with high skill scores, whereas, the top right corner characterizes a forecast system where the event is always warned for [25].

Skill score for the area under the ROC curve (ROCS) can then be computed as equation (2). ROCS range between 1, for perfect forecast, and -1, for perfectly wrong forecast, with 0 indicating no skill in the forecast [14].

$$ROCS = 2(ROC\ Area - 0.5) \quad \text{(Equation 2)}$$

Ranked Probability Score (RPS) measures inaccuracy in discriminating among categories. It integrates a concept of distance to give forecasts higher credits when their probabilities are close to the verifying category, a category which the observation fell in. Assume, we have probability forecasts of three rainfall categories: R ≤ 5 mm, 5 mm<R ≤ 10 mm and 10<R ≤ 15. Suppose, two forecasts A=(0.1, 0.5, 0.4) and B=(0.5, 0.1, 0.4) were made, while the observation fell in the third category. If the two forecasts are compared, both predicted that there was 40% probability for the observation to fall in the third category. However, forecast A was a better forecast because categories 2 and 3 are closer to one another than categories 1 and 3 [26]. Different definitions of RPS have been proposed, but for this study Ranked Probability Score (RPS) definition given by Mason and Stephenson [11] was considered:

$$RPS = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}(P_{ij} - O_{ij})^2 \quad \text{(Equation 3)}$$

Where RPS is Ranked Probability Score, n is number of forecasts, m is number of forecast categories, O_{ij} is the cumulative probability upto category j in the i^{th} forecast, O_{ij} is 1 if the i^{th} observation falls in any of the category less or equal to j, else O_{ij} is zero. It is noteworthy to mention that Ranked Probability Score (RPS) takes the value of 0 for perfect forecast, with less accurate forecasts taking higher RPS values ranging up to 1 [27]. For dichotomous categorical forecasts having only two possible outcomes, i.e when m=2, Ranked Probability Score (RPS) is the same as Brier Score. The Ranked Probability Score (RPS) measure the accuracy of the forecast in predicting the category the observation fell into, whereas Ranked Probability Skill Scores (RPSS) determines how good the forecast is over the climatology in predicating the category the observation fell in. Thus, in this study, Ranked Probability Skill Score was applied

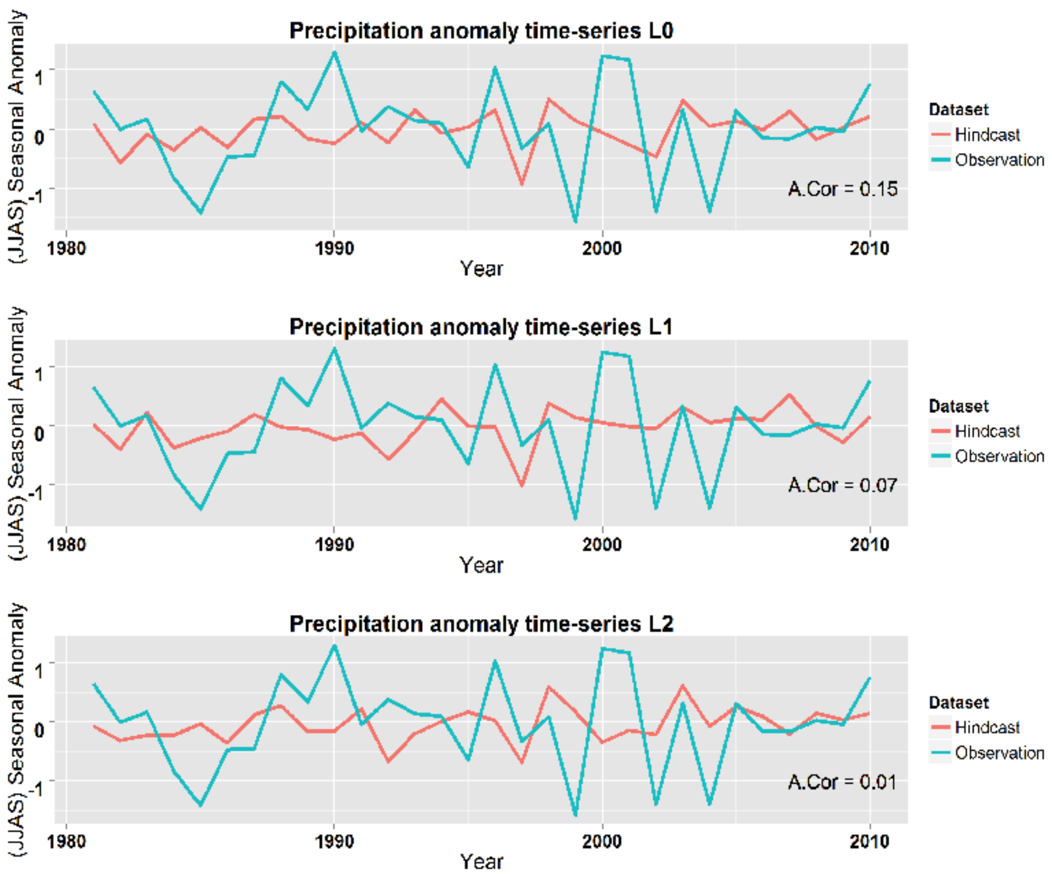


Figure 3: Time-series for precipitation anomaly over the study area for the period 1981 to 2010. The red line shows (JJAS) forecast anomaly and the cyan line shows (JJAS) observation anomaly; (a): at lead 0; (b): at lead 1; (c): at lead 2. A.Cor denotes anomaly correlation coefficient.

to evaluate the forecast benefit gained over the climatological probability forecast. RPSS can be computed from RPS as [11]:

$$RPSS = 1 - \frac{RPS}{RPS_{ref}}, \quad \text{(Equation 4)}$$

Where *RPSS* is Ranked Probability Skill Score, *RPS* is Ranked Probability Score for the forecast; *RPSref* is RPS for the long-term climatology forecasts. In this case, a perfect forecast takes RPSS value of 1 while the worst score is 0 [26]. The skill score defines whether having forecast information is better or not than merely having historical climatological probabilities. Negative values refer to forecasts worse than climatology.

RESULTS

Correlation between hindcast and observation

The anomaly correlation coefficients between the precipitation forecast and observation were found to be very low; 0.15, 0.07 and 0.01 at zero, one and two months forecast lead-times respectively. Figure 3 shows precipitation anomaly times series for the period 1981 to 2010. From the figure, it could be seen that precipitation

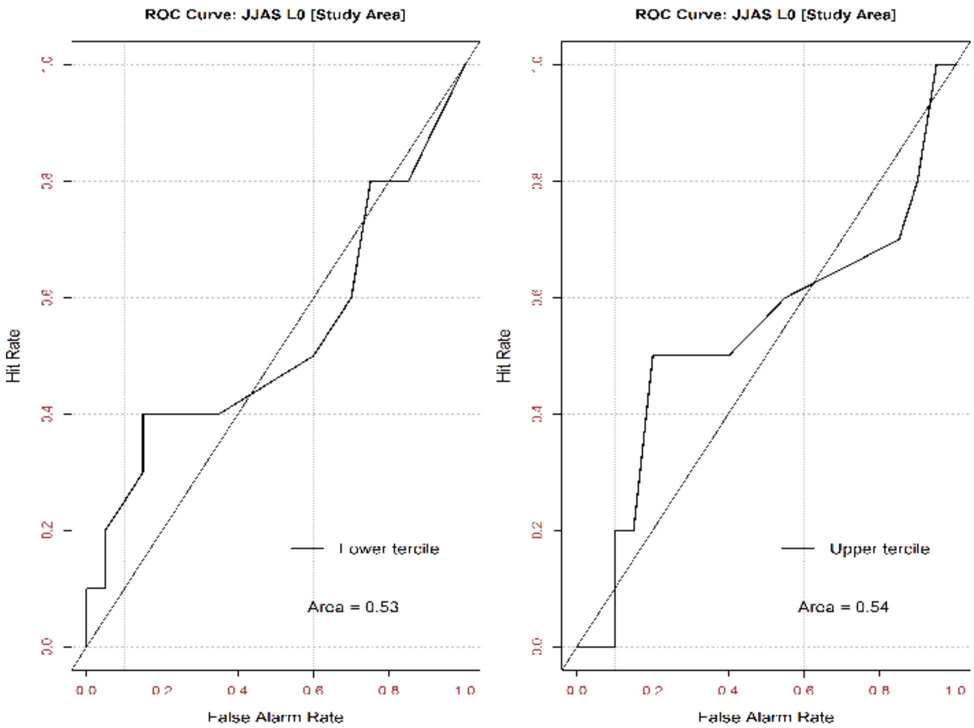


Figure 4: ROC diagram for JJAS precipitation hindcasts for 0-month lead time (L0) over the study area; (a): below lower tercile and (b): above upper tercile.

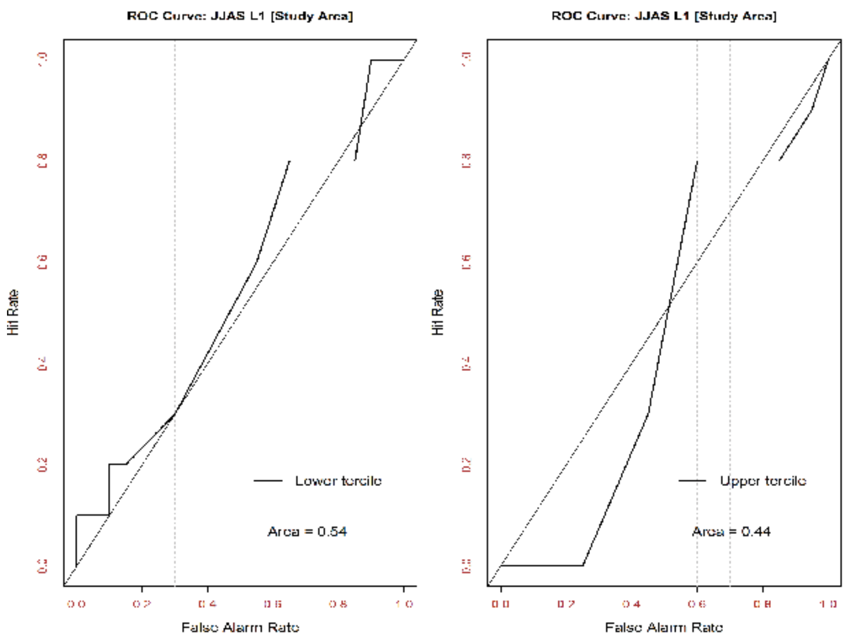


Figure 5: ROC diagram for JJAS precipitation hindcasts for 1-month lead time (L1) over the study area; (a): below lower tercile and (b): above upper tercile.

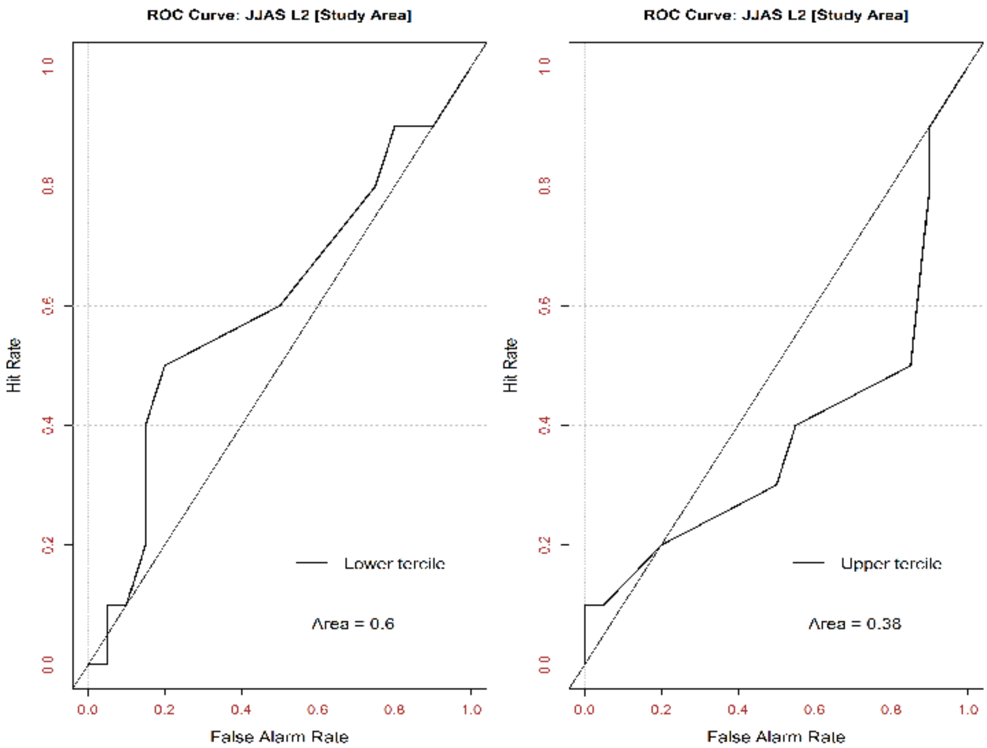


Figure 6: ROC diagram for JJAS precipitation hindcasts for 2-months lead time (L2) over the study area; (a): below lower tercile and (b): above upper tercile.

hindcasts anomaly poorly corresponded to the reference observation anomaly over the study area with no considerable sensitivity to the forecast lead-times.

Good correlation between forecast and observation was found for the four anomalous years over the study area. Again no trend was observed between the correlation values and lead times; however, from table 1, it can be seen that correlation values are relatively higher for the 2-month lead time forecasts.

Table 1: Areal Pearson product-moment correlation coefficients between mean forecast ensembles of JJAS precipitation hindcasts and observation, for each of the four anomalous years at 0-, 1- and 2-month lead times.

	0-month	1-month	2-month
1985	0.73	0.56	0.76
1991	0.9	0.92	0.91
2004	0.83	0.86	0.85
2009	0.78	0.75	0.83

Correlation between hindcast and altitude

In this study, bias corrected seasonal precipitation obtained from WFDEI global data, considered as reference observation, were used to validate ECMWF System-4 seasonal reforecasts. As the study area demonstrates complex terrain, we also computed correlation between simulated seasonal climate and local altitude on the four anomalous years. The correlation between reference seasonal precipitation and local altitudes was calculated to be 0.42, 0.37, 0.62 and 0.31 for 1985, 1991, 2004 and 2009 respectively. The correlation between precipitation hindcast and local altitudes, at all the three lead-times, is shown in Appendix B. The result showed that there was no sufficient correspondence between the simulated seasonal climate and altitude.

Relative Operating Characteristics (ROC) curve

As described previously, the area under Relative Operating Characteristics (ROC) curve needs to exceed 0.5 in order for forecasts to have some skill [28]. An area of 0.8 or more is reported to be a measure of good forecast system; whereas, a minimum value of 0.7 should be obtained for the forecast system to be referred as useful [29]. According to the results of this study, the area under the ROC curve for precipitation hindcasts below the lower tercile demonstrated only marginal skill over the study area. The ROC area values for this tercile category were 0.53, 0.54 and 0.6 at 0-, 1- and 2-months lead-times respectively. The values showed slight increment from shorter to longer leads. Similarly, marginal predictive skill was detected for hindcasts above the upper tercile for 0-month lead forecasts, with ROC area of 0.54. However, no skill was found for this category at 1- and 2-months lead-times.

We presented the area under the ROC curve results for precipitation hindcasts below the lower tercile and above the upper tercile, for all the three lead times, in a bar plot. As can be observed from figure 7, the forecast skills were relatively better for the lower tercile hindcast categories for which the ROC area did not fall below the no-skill value of 0.5 at any of the leads. Whereas, hindcasts above the upper tercile showed some skill only at zero lead time.

Ranked Probability Skill Score (RPSS)

The Ranked Probability Skill Scores for precipitation hindcasts were found to be below the no skill value of 0 at any of the forecast lead-times. RPSS for precipitation hindcasts were computed to be -0.108, -0.1713 and -0.1226 at 0-, 1- and 2-month forecast lead-times respectively. Wilks [22] said, compared to other skill indicators, forecasts often gets low RPSS scores. Low skill scores could also be obtained when the ensemble sizes are small [30].

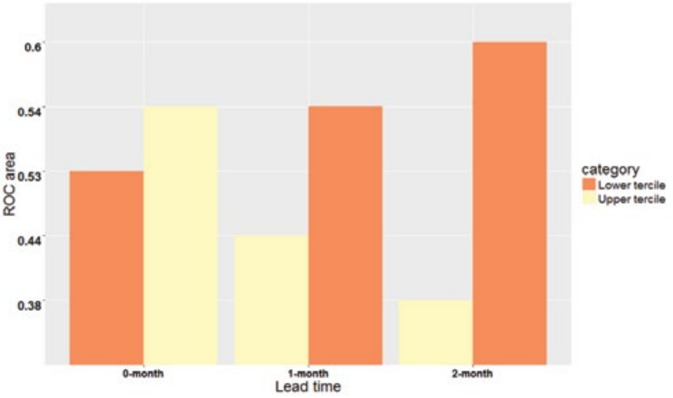


Figure 7: Bar plot of the values of the area under ROC curve for hindcasts below the lower tercile and above the upper tercile at lead 0, 1 and 2.

DISCUSSION

There was hardly any research works on seasonal prediction skill of ECMWF system4 precipitation reforecasts executed in analogous experimental setup to this study. Firstly, most studies considered forecast products of older ECMWF systems [13,14,31-33]. The fact that ECMWF system4 became available only in 2011, as marked by Molteni Stockdale [17], may have contributed to the limited existing studies on its forecasts. Secondly, all of them were performed at higher scales; at continental, sub-continental, or country level to the least. Thirdly, some of the existing published works on seasonal prediction skill of ECMWF system-4, such as [17,34], did not apply ROC and RPSS verification metrics in their researches. Consequently, comparing results of this study in order to check consistence with results of other related works was not straightforward.

Kim, Webster [34] assessed predictive skills in precipitation forecasts of ECMWF system 4, deploying precipitation validation dataset of Global Precipitation Climatology Project (GPCP), by computing anomaly correlation. According to their results, the precipitation forecast of ECMWF system4 overestimated the NH summer mean precipitation along the Inter-Tropical Convergence Zone (ITCZ), an area of intertropical front moving back and forth across equatorial Africa. They particularly found out low predictive skill in the retrospective forecasts for Asian monsoon regions. Similarly, in this study the ECMWF system4 precipitation forecasts overestimated June-September seasonal precipitation over the study area at all the three forecast lead-times, as witnessed from areal maps of the anomalous years. As the correlation coefficients were computed between the mean of the forecast ensembles and observation, the positive bias was more revealed by means of the forecast maps, see Appendix A, where pixel-wise forecast values could be visually compared to the observation. From a closer inspection of figure 4 of Molteni Stockdale [17], depicting their anomaly correlation result for the globe between ECMWF system 4 JJA precipitation hindcast and JJA ERA interim reanalysis at 1-month lead, it can be seen that the study area of this research, central-west Ethiopia, fell in -0.2 to 0.2 anomaly correlation coefficient range. Comparably, in this research at the same lead month, for JJAS season, anomaly correlation coefficient was calculated to be 0.07.

As explained in results section of this report, ROC and RPSS scores also indicated generally poor skills in ECMWF system4 precipitation reforecasts for JJAS season over the study area. According to Buizza, Hollingsworth [29] thresholds, ECMWF system4 demonstrated ability which was below par with respect to utility at this local scale. Buizza, Hollingsworth [29] recommended a ROC area value of 0.7 at the minimum to refer a forecast system as valuable. However, Wilson [35] advised to be cautious while applying the thresholds as these values were intended to be applied on condition that the observations used for verification were instrumental records. Nevertheless, the obtained peripheral skill upholds the recommendations by Troccoli [36] and Hansen, Mason [37] to apply seasonal climate forecasting over larger spatial resolutions for better accuracies.

In another study, Dutra, Di Giuseppe [38] found considerable anomaly correlation coefficients between JJA seasonal ECMWF system-4 precipitation reforecast and Global Precipitation Climatology Project over Blue Nile basin for the period 1981-2010. Unlike this study, they used Global Precipitation Climatology Project (GPCP) for verification. Moreover, the area of the Blue Nile basin is by far greater than the study area of this research. They carried out the skill assessment by merging ECMWF system4 with GPCP. Dutra, Di Giuseppe [38] found some skills in Standardized Precipitation Index (SPI) seasonal forecasts which also showed skills in underlying seasonal precipitation forecasts. Figure 7 of Dutra, Di Giuseppe [38] presented Continuous Ranked Probability Skill Scores (CRPSS) for SPI-3 of Blue Nile basin at different accumulation months (x-axis) and lead times (y-axis). From the 8th column of the figure, it could be seen that CRPSS for JJA SPI-3 ranges between 0.2 and 0.25 over the basin. In this study, RPSS values for JJAS were found to be slightly below 0. Obviously, the spatial coverage of the study areas, the predicted variables and the verification datasets used were not the same between these two studies. However, the CRPSS values from Dutra, Di Giuseppe [38] as well were not significantly higher than the no-skill value of 0.

CONCLUSIONS

The focus of this study was to assess predictive skills in seasonal precipitation retrospective forecasts of ECMWF system-4 over central-west Ethiopia. The results of such studies principally designate the experimental setups pursued, and hence require interpretations accordingly. Some of the factors that can determine results of skill assessments include the extent of verification metrics applied, the type of validation dataset used, target variables, resolution, time and regions.

The predictive skill in ECMWF System-4 precipitation forecasts over the study area was generally poor, and nearly no benefit gained at any lead-time compared to the climatology. The model showed a tendency to better discriminate the lower tercile categories than the upper tercile categories.

ACKNOWLEDGMENT

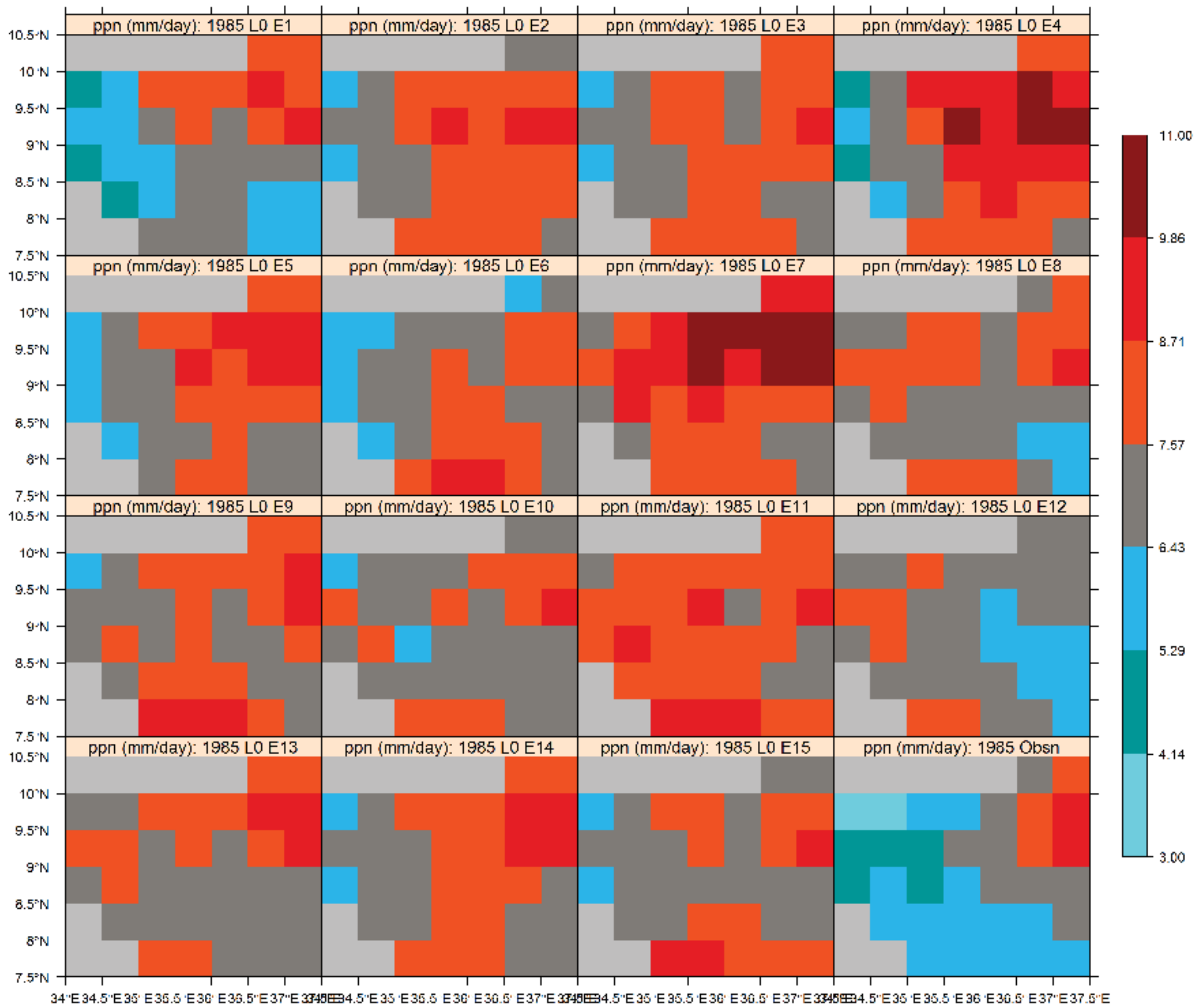
Tewodros A. Yate was supported by Netherlands Fellowship Program (NFP)-The Netherlands Organization for International Cooperation in Higher Education (NUFFIC) to do his M.Sc. study in Wageningen University and Research. We would like to thank Geoffrey Ogutu and Ir. Wietse Franssen for their overall valuable collaboration in the course of this research.

REFERENCES

1. Wagesho N, Goel N, Jain M. Temporal and spatial variability of annual and seasonal rainfall over Ethiopia. Hydrol Sci J. 2013;58:354-373.
2. Dercon S, Hill RV. Growth from agriculture in Ethiopia: Identifying key constraints. IFPRI's ESSP-II policy conference, Accelerating agricultural development, economic growth and poverty reduction in Ethiopia. Addis Ababa. 2009.
3. Korecha D, Barnston AG. Predictability of June-September rainfall in Ethiopia. Monthly Weather Rev. 2007;135:628-650.
4. Zeleke TT, Giorgi F, Diro GT, Zaitchik BF. Trend and periodicity of drought over Ethiopia. Int J Climatol. 2017;37:4733-4748.
5. Suryabhagavan K. GIS-based climate variability and drought characterization in Ethiopia over three decades. Weather Clim Extreme. 2017;15:11-23.
6. Edossa DC, Babel MS, Gupta AD. Drought analysis in the Awash river basin, Ethiopia. Water Res Manage. 2010:24:1441-1460.
7. De Waal A, Taffesse A, Carruth L. Child survival during the 2002–2003 drought in Ethiopia. Glob Public Health. 2006;1:125-132.
8. Viste E, Korecha D, Sorteberg A. Recent drought and precipitation tendencies in Ethiopia. Theor Applied Climatol. 2013;112:535-551.
9. Hewitt C, Buontempo C, Newton P. Using climate predictions to better serve society's needs. Eos Transac Am Geophys Union. 2013;94:105-107.
10. Murphy AH, Winkler RL. A general framework for forecast verification. Monthly Weather Rev. 1987;115:1330-1338.
11. Mason SJ, Stephenson DB. How do we know whether seasonal climate forecasts are any good?. Seasonal Climate: Forecasting and Managing Risk. Springer. 2008.
12. Doblas-Reyes F, Pavan V, Stephenson D. The skill of multi-model seasonal forecasts of the wintertime North Atlantic Oscillation. Climate Dynamics. 2003;21:501-514.
13. Lavers D, Luo L, Wood EF. A multiple model assessment of seasonal climate forecast skill for applications. Geophys Res Letters. 2009;36:1-6.
14. Frías MD, Herrera S, Cofino AS, Gutierrez JM. Assessing the skill of precipitation and temperature seasonal forecasts in Spain: windows of opportunity related to ENSO events. J Climate. 2010:23:209-220.
15. Gissila T, Black E, Grimes DIF, Slingo JM. Seasonal forecasting of the Ethiopian summer rains. Int J Climatol. 2004;24:1345-1358.
16. Diro GT, Grimes DIF, Black E, Neill AO, Pardo-Iguzquiza E. Evaluation of reanalysis rainfall estimates over Ethiopia. Int J Climatol. 2009;29:67-78.
17. Molteni F, Stockdale T, Balmaseda MA, Balsamo G. The new ECMWF seasonal forecast system (System 4). European Centre for Medium-Range Weather Forecasts. 2011.
18. Segele ZT, Lamb PJ. Characterization and variability of Kiremt rainy season over Ethiopia. Meteorol Atmos Phys. 2005;89:153-180.
19. Persson A, Grazzini F. User guide to ECMWF forecast products. Meteorol Bull. 2005;3:1-161.
20. Weisheimer A, Corti S, Palmer T, Vitart F. Addressing model error through atmospheric stochastic physical parametrizations: impact on the coupled ECMWF seasonal forecasting system. Phil Transac R Soc A. 2014;372:1-21.
21. Weedon GP, Balsamo G, Bellouin N, Gomes S, Best MJ, Viterbo P. The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. Water Resourc Res. 2014;50:7505-7514.
22. Wilks DS. Statistical methods in the atmospheric sciences. Academic Press. 3rd edition. 2011.
23. Murphy AH, Epstein ES. Skill scores and correlation coefficients in model verification. Monthly Weather Rev. 1989;117:572-582.
24. Mason SJ, Graham NE. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. Q J R Meteorol Soc. 2002;128:2145-2166.
25. Persson A, Grazzini F. User guide to ECMWF forecast products. Meteorol Bull. 2007:3:1-153.
26. Epstein ES. A scoring system for probability forecasts of ranked categories. J Applied Meteorol. 1969;8:985-987.
27. Weigel AP, Liniger MA, Appenzeller C. The discrete Brier and ranked probability skill scores. Month Weather Rev. 2007;135:118-124.
28. Mason SJ, Graham N. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. Q J R Meteorol Soc. 2002;128:2145-2166.
29. Buizza R, Hollingsworth A, Lalaurette F, Ghelli A. Probabilistic predictions of precipitation using the ECMWF Ensemble Prediction System. Weather Forecast. 1999;14:168-189.
30. Müller WA, Appenzeller C, Doblas-Reyes FJ, Liniger MA. A debiased ranked probability skill score to evaluate probabilistic ensemble forecasts with small ensemble sizes. J Clim. 2005;18:1513-1523.
31. Diro GT, Tompkins AM, Bi X. Dynamical downscaling of ECMWF Ensemble seasonal forecasts over East Africa with RegCM3. J Geophys Res. 2012;117:1-20.
32. Jury MR. Evaluation of coupled model forecasts of Ethiopian highlands summer climate. Adv Meteorol. 2014;2014:1-10.
33. Palmer TN, Alessandri A, Andersen U, Cantelaube P, Davey MK, Delécluse P, et al. Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). Bull Am Meteorol Soc. 2004;85:853-872.
34. Kim HM, Webster PJ, Curry JA, Toma VE. Asian summer monsoon prediction in ECMWF System 4 and NCEP CFSv2 retrospective seasonal forecasts. Clim Dyn. 2012;39:2975-2991.
35. Wilson LJ. Probabilistic predictions of precipitation using the ECMWF ensemble prediction system. Weath Forecast. 2000;15:361-364.
36. Troccoli A. Seasonal climate forecasting. Meteorol App. 2010;17:251-268.
37. Hansen JW, Mason SJ, Sun L, Tall A. Review of seasonal climate forecasting for agriculture in sub-Saharan Africa. Exp Agric. 2011;47:205-240.
38. Dutra E, Di Giuseppe F, Wetterhall F, Pappenberger F. Seasonal forecasts of drought indices in African basins. Hydrol Earth Syst Sci. 2013;9:2359-2373.

APPENDIX-A

Areal maps for JJAS precipitation depicting the fifteen ensemble forecasts from ECMWF system-4 (from E1 to E15 as shown as the last substrings of the titles above each panels). The seasonal averages were computed from June, July, August and September monthly corresponding values. The bottom most right panel depicts the reference JJAS observation. The corresponding year and lead are indicated at the top of each panel: for example, '1985 L0 E1' refers area map of 1985 at 0-month lead for the case of ensemble member 1 and so on.



APPENDIX B

Bar plot of areal Pearson product-moment correlation coefficients between mean forecast ensembles of June – September seasonal precipitation and elevation, for each of the four anomalous years at each of the three leads (L0, L1 and L2 denote the forecast lead times).