International Journal of Science Academic Research
Vol. 02, Issue 07, pp.1898-1906, July, 2021
Available online at http://www.scienceijsar.com
ISSN: 2582-6425

Research Article

VERIFICATION OF FORECASTS GIVEN TO CIARE PROJECT INTERVENTION AREAS

*Melesse Lemma

National Meteorological Agency of Ethiopia, Addis Ababa Ethiopia

Received 10th May 2021; **Accepted** 18th June 2021; **Published online** 30th July 2021

Abstract

Downscaled forecasts or outlooks at timescale of 10 daily, monthly and seasonal have been prepared and provided to the CAIRE project areas since the third dekad of October 2016. In forecasting weather and climate, testing the value and quality of the forecasts must be taken as part of the whole forecasting process. Unless we test our forecast quality and measure where we stand we cannot see our progress or failures in forecast generation. So, verifying the issued forecasts and outlooks taken as one of the activities of the CIARE project. This study is done to test the performance of issued forecasts and outlooks from October 2016 third dekad to August 2017 first dekad for CIARE project areas. Methods for dichotomous (yes/no) forecasts and methods for forecasts of continuous variables used to test the performance of forecasts and outlooks given to the project areas. As we noticed, there are no meteorological stations in some of the CIARE interventions Woredas. For such Woredas, we cannot find normal meteorological data even for a single location in the Woreda. This means that our climatological knowledge of the dekal and monthly normal values can only be deduced from blended data or purely interpolation. Presenting the forecast for such Woredas in terms of normal rain could be very difficult for the assigned forecaster or meteorologist. In addition, if the season is not a rainfall season and dominantly sunny and dry, it does not matter if the forecast is presented in the form of rain, no rain or light to moderates rain. But for other Woredas where we have station and can get normal values, the dekadal and monthly forecast can be presented in relation to normal values. We can issue a forecast that clearly states whether the expected rainfall is above, near or below normal. The writer of this report finds very few forecasts of this kind. If possible, it would have value for the user as well as for one who verifies the forecast if the forecasts are presented in relation to normal values as it is done in the regular forecasts issued for the whole country in coarse resolution. The NWP disseminated forecasts are from the outputs of NOAA. We have more localized weather and climate information than NOAA has. It would be much better if we were able to run the WRF model we have at home and issue ten daily forecasts. The period we took for verification is too short to arrive at concrete conclusion, but with the sample we used we have found out that, in relative terms, the accuracy of the dekadal worded forecast is much better than the worded monthly forecasts. Other statistics can be inferred from the summary tables given in the document. When we come to the NWP outputs, in all aspects, we do not get promising results.

Keywords: Forecasts, Dekad, Woredas.

1. Introduction

1.1 Background

The Christian Aid Ethiopia Climate Information and Assets for Resilience in Ethiopia (CIARE) project is a DFID-funded project being implemented in Ethiopia under the global Building Resilience and Adaptation to Climate Extremes and Disasters (BRACED) programme. It is a multidisciplinary initiative aimed at building the resilience of 791,530 vulnerable people to climate extremes and disasters in seven high intensity and five medium intensity Woredas (districts) in Oromia and SNNPR of Ethiopia (Fig. 1). Such aim is being realized through undertaking packages of activities. These are climate information services, tangible resilience building schemes, capacity building support to local actors in disaster risk reduction and resilience programme planning and action, learning through various knowledge generation and sharing approaches and monitoring and evaluation activities. This project has been framed into four outputs and there are ten actors including the owner institute, Christian Aid. The first output focuses on generation and dissemination of climate information. The second output is all about supporting vulnerable peoples in seven districts. The third one deals with increased capacity of local actors for better resilience programming and action. The last one tries to understand what works in building resilience to climate extremes and disasters and what constitutes incremental and transformational changes including the factors affecting them. Each output has more than one activity. Christian Aid with partner organizations has been running this project for the last three years. Climate products have been generated and disseminated to users for quite a long time. Information providers, intermediaries and users have been involved in the generation, transmission and uptake of the climate information respectively.

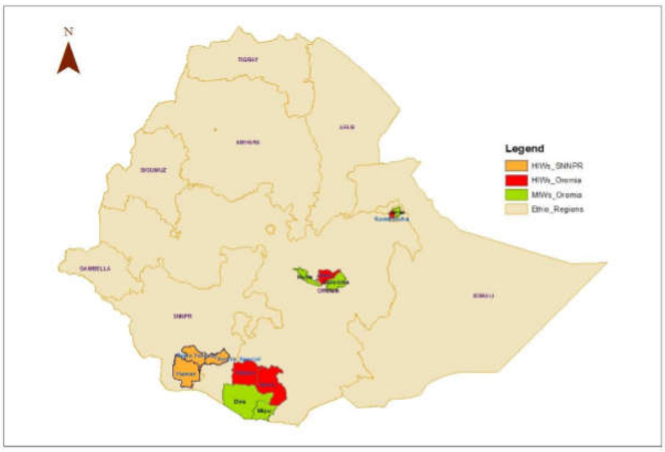


Figure 1.CIARE intervention Woredas in Ethiopia

In line with this, downscaled forecasts or outlooks at timescale of 10 daily, monthly and seasonal of project intervention Woredas have been prepared and provided to the project areas beneficiaries since the third dekad of October 2016 as far as

Corresponding Author:** ***Melesse Lemma,
National Meteorological Agency of Ethiopia, Addis Ababa Ethiopia.

the writer of this study is aware of. Ten daily forecasts issue three times per month, monthly forecasts issue one time per month and seasonal forecast three times per year. The forecasts and outlooks were prepared by experienced staffs from NMA, mainly staffs from Meteorological Forecast and Early Warning Directorate (MFEWD). The forecasts and outlooks are sent to BBC Media Action. BBC Media Action prepares radio programmes based on the forecast or outlook it received and sent them to Ormia Broadcasting Network, SNNP Radio Agency, Arbaminch FM and Jinka FM for broadcasting. Then these FM Radio stations were airing them in local languages. In forecasting weather and climate, testing the value and quality of the forecasts must be taken as part of the whole forecasting process. Unless we test our forecast quality and measure where we stand we cannot see our progress or failures in forecast generation. So, verifying the issued forecasts and outlooks taken as one of the activities of the CIARE project. This study is done to test the performance of issued forecasts and outlooks from October 2016 third dekad to August 2017 first dekad for CIARE project areas. In the sections that follow we will see the data used for verification, the methodology we applied to verify the forecasts and outlooks against actual observations, the result we obtained from the analysis, and finally we give conclusion and recommendations.

1.2 Objective

The National Meteorological Agency of Ethiopia as a mandated meteorological Agency in the country has the responsibility of collecting, processing, analysing and interpreting meteorological data and issuing weather forecasts, climate outlooks and advices. Its services also encompass the delivery of early warnings for the public to reduce loss of life and damage to property. The forecasts, climate outlooks and early warnings disseminated to the public can refer different spatial and time scales. Short lead time forecast can have better success than long lead time forecasts. Similarly forecast for large area can see less failure than forecast for a smaller area. The performance of the Agency as well can be tested with the accuracy of the forecasts and early warnings it is giving. Recently, NMA in collaboration with Christian Aid Ethiopia and other organizations has started disseminating downscaled dekadal forecast, monthly and seasonal outlooks. Specifically, NMA is disseminating above mentioned forecasts and outlook to CAIRE intervention Woredas in collaboration Christian Aid and CIARE project partners. Users of the forecasts and outlooks give their views on the value of the forecast and how useful they are. The objective of this study is to scientifically check the quality of forecasts and outlooks issued to CIARE project intervention areas using standard forecast verification methodology and if possible indicate ways to improve accuracy of downscaled forecasts and outlooks.

2. Data and methodology

2.1 Data

The National Meteorological Agency of Ethiopia regularly issue daily, three daily, dekadal (ten daily) forecasts, monthly and seasonal outlooks and post them on the Agency's website. One can download and use these products. They are available freely. But these forecasts and outlooks are of a general type. They refer to large areas. The spatial scale of dekadal forecasts, monthly and seasonal outlooks disseminated on the NMA's website regularly do not provide detailed description of the likely occurrence of expected weather and climate conditions at smaller spatial resolution such as Woreda (district). They limit giving the details up to the zonal[1] level. Christian Aid Ethiopia understood that this coarse resolution forecasts and outlooks could not lead climate information users to reach to better informed decision at Woreda level. Therefore, through CIARE project, Christian Aid came into mutual agreement with NMA on the preparation of Woreda level forecasts specially for BRACED project interventions areas. The forecasts and outlooks were issued for the following Woredas.

- Gololcha
- Robe
- Seru
- Kombolcha
- Jarso
- Areo
- Bentsemay
- Dire
- Yabello
- Hamer
- Konso

Miyo Woreda is one of the project intervention Woreda but we do not get forecasts and outlooks except few. Hence, for the purpose of this study, forecasts and outlooks issued over the period October 2016 third dekad to August 2017 1st dekadof eleven Woredas were collected from NMA. During this period, a total of 28 dekadal forecasts and 9 monthly outlooks for each Woreda were issued. Out of the 28 dekadal forecasts two were missing. The missed forecasts are January 2017 2nd and April 2017 1st dekadal forecasts. From the monthly outlooks April 2017 forecast was missed. So, we are obligated to use 26 dekadal forecasts and 8 monthly outlooks for our verification analysis. In addition, ten daily Numerical Weather Prediction (NWP) outputs extracted from NOAA website for each project Woredas separately have been disseminated to the users. The NWP outputs that have been provided to the users along with the other forecasts and outlooks by NMA for the period mentioned above has also been obtained from the providers. The next action in data collection is obtaining observed data. Corresponding periods observed daily, ten daily, monthly and normal data obtained from the same institution for Woredas having observing sites. Among the Woredas tabled for the verification exercise we get stations data from five of them only. These are Golelcha, Robe, Seru, Yabello, and Konso. In the other Woredas there are no meteorological stations. The normal data collected for this analysis is the normal calculated based on 1981-2010 data. This is the latest available normal data.

2.2 Methodology

Before we proceed to present and interpret the results of this work, it is of interest to provide some detail points on the methods we have used. Forecast verification involves exploring and summarising the relationship between sets of forecasts and observed data and making comparisons between the performance of forecasting systems and that of reference forecasts. Verification is therefore a statistical problem (Jolliffe and D. B. Stephenson, Eds., 2003). Hence, we have employed

[1]Zones are the second administrative hierarchy in the Ethiopian government administration structure next to regional states.

statistical verification techniques to compare observations and forecasts. The nature of the forecasts guided the method we used to verify the forecasts and outlooks. The dekadal forecasts and monthly outlooks are worded forecasts and outlooks. These forecasts tell the occurrence or non-occurrence of rainfall in terms of rain or no rain, below normal rain or no below normal rain, normal rain or no-normal rain, above normal rain or no above normal rain. The forecasts and outlooks does not stop here, they further give the expected spatial coverage of the forecasts and outlooks. The nature of these forecast and outlooks are discussed in detail further in the next section. These are yes/no or dichotomous forecasts or outlooks. The logical way of verifying these types of forecasts or outlooks is to use the contingency table method. The details are as follows as taken from Forecast Verification - Issues, Methods and FAQ (WCRP, 2009) which is available online from http://www.cawcr.gov.au/projects/verification/and accessed on 02 August 2017.

Methods for dichotomous (yes/no) forecasts

A dichotomous forecast says, "yes, an event will happen", or "no, the event will not happen".

To verify this type of forecast we start with a contingency table that shows the frequency of "yes" and "no" forecasts and occurrences. The four combinations of forecasts (yes or no) and observations (yes or no), called the joint distribution, are:

hit - event forecast to occur, and did occur
miss - event forecast not to occur, but did occur
false alarm - event forecast to occur, but did not occur
correct negative - event forecast not to occur, and did not occur
The total numbers of observed and forecast occurrences and non-occurrences are given on the lower and right sides of the contingency table, and are called the marginal distribution.

Contingency Table				
		Observed		
		yes	no	Total
Forecast	yes	hits	false alarms	forecast yes
	no	misses	correct negatives	forecast no
Total		observed yes	observed no	total

The contingency table is a useful way to see what types of errors are being made. A perfect forecast system would produce only hits and correct negatives, and no misses or false alarms. A large variety of categorical statistics are computed from the elements in the contingency table to describe particular aspects of forecast performance. Categorical statistics that are computed from the yes/no contingency table are.

1) **Accuracy (fraction correct)** -

$$Accuracy = \frac{hits + correct\ negatives}{total}$$

Answers the question: Overall, what fraction of the forecasts were correct?

2) **Bias score (frequency bias)** -

$$BIAS = \frac{hits + false\ alarms}{hits + misses}$$

Answers the question: How did the forecast frequency of "yes" events compare to the observed frequency of "yes" events?

3) **Probability of detection (hit rate)** -

$$POD = \frac{hits}{hits + misses}$$

Answers the question: What fraction of the observed "yes" events were correctly forecast?

4) **False alarm ratio –**

$$FAR = \frac{false\ alarms}{hits + false\ alarms}$$

Answers the question: What fraction of the predicted "yes" events actually did not occur (i.e., were false alarms)?

5) **Probability of false detection (false alarm rate)** -

$$POFD = \frac{false\ alarms}{correct\ negatives + false\ alarms}$$

Answers the question: What fraction of the observed "no" events were incorrectly forecast as "yes"?

6) **Threat score (critical success index)** -

$$TS = \frac{hits}{hits + misses + false\ alarms}$$

Answers the question: How well did the forecast "yes" events correspond to the observed "yes" events?

7) **Equitable threat score (Gilbert skill score)**-

$$ETS = \frac{hits - hits_{random}}{hits + misses + false\ alarms - hits_{random}}$$

where

$$hits_{random} = \frac{(hits + misses)(hits + false\ alarms)}{total}$$

Answers the question: How well did the forecast "yes" events correspond to the observed "yes" events (accounting for hits due to chance)?

8) **Hanssen and Kuipers discriminant (true skill statistic, Peirces's skill score)** -

$$HK = \frac{hits}{hits + misses} - \frac{false\ alarms}{false\ alarms + correct\ negatives}$$

Answers the question: How well did the forecast separate the "yes" events from the "no" events?

9) **Heidke skill score (Cohen's k)** -

$$HSS = \frac{(hits + correct\ negatives) - (expected\ correct)_{random}}{N - (expected\ correct)_{random}}$$

where

$$(expected\ correct)_{random} = \frac{1}{N}\left[\begin{array}{l}(hits + misses)(hits + false\ alarms) \ + \\ (correct\ negatives + misses)(correct\ negatives + false\ alarms)\end{array}\right]$$

Answers the question: What was the accuracy of the forecast relative to that of random chance?

10) **Odds ratio** –

$$OR = \frac{hits * correct\ negatives}{misses * false\ alarms} = \frac{\left(\frac{POD}{1-POD}\right)}{\left(\frac{POFD}{1-POFD}\right)}$$

Answers the question: What is the ratio of the odds of a "yes" forecast being correct, to the odds of a "yes" forecast being wrong?

11) **Odds ratio skill score (Yule's Q)**-

$$ORSS = \frac{hits * correct\ negatives - misses * false\ alarms}{hits * correct\ negatives + misses * false\ alarms}$$

Answers the question: What was the improvement of the forecast over random chance?

The NWP presented in ranges of rainfall amount. The forecast areas shaded with different colours that indicate the expected amount of rainfall. So, we have used the methods for continuous variables to verify these forecasts. These methods as described in Verification - Issues, Methods and FAQ (WCRP, 2009) which is available online from http://www.cawcr.gov.au/projects/verification/ and accessed on 02 August 2017 are given below.

Methods for forecasts of continuous variables

Verification of continuous forecasts often includes some exploratory plots such as scatter plots and box plots, as well as various summary scores. For our verification exercise we have done scatter plot and box plots.

1) **Scatter plot** – is plots of the forecast values against the observed values. It shows how well did the forecast values correspond to the observed values?
2) **Box plot -** Plot boxes to show the range of data falling between the 25^{th} and 75^{th} percentiles, horizontal line inside the box showing the median value, and the whiskers showing the complete range of the data. Box plots answer the question how well did the distribution of forecast values correspond to the distribution of observed values?

The summary score taken into consideration for this study are:

1) Mean error –

$$Mean\ Error = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)$$

Where F_i is the i_{th} forecast and O_i is the i_{th} corresponding observation

Answers the question: What is the average forecast error?

2) (Multiplicative) bias(Bias)-

$$Bias = \frac{\frac{1}{N}\sum_{i=1}^{N}F_i}{\frac{1}{N}\sum_{i=1}^{N}O_i}$$

Where F_i is the i_{th} forecast and O_i is the i_{th} corresponding observation

Answers the question: How does the average forecast magnitude compare to the average observed magnitude?

3) Mean absolute error (MAE) –

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|F_i - O_i|$$

Where F_i is the i_{th} forecast and O_i is the i_{th} corresponding observation

Answers the question: What is the average magnitude of the forecast errors?

4) Root mean square error (RMSE) –

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2}$$

Where F_i is the i_{th} forecast and O_i is the i_{th} corresponding observation

Answers the question: What is the average magnitude of the forecast errors?

The root mean square factor is similar to RMSE, but gives a multiplicative error instead of an additive error.

5) Mean squared error (MSE)-

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2$$

Measures the mean squared difference between the forecasts and observations.

Where F_i is the i_{th} forecast and O_i is the i_{th} corresponding observation

6) Correlation coefficient (r) –

$$r = \frac{\sum(F-\bar{F})(O-\bar{O})}{\sqrt{\sum(F-\bar{F})^2}\sqrt{\sum(O-\bar{O})^2}}$$

Addresses the question: How well did the forecast values correspond to the observed values?

Where F is the forecast, F and O are the average forecast and observations respectively.

7) Anomaly correlation (AC)

$$AC = \frac{\sum(F-C)(O-C)}{\sqrt{\sum(F-C)^2}\sqrt{\sum(O-C)^2}}(\text{uncentered}) \quad \text{or} \quad \frac{\sum\left[(F-C)-\overline{(F-C)}\right]\left[(O-C)-\overline{(O-C)}\right]}{\sqrt{\sum(F-C)^2}\sqrt{\sum(O-C)^2}}(\text{centered})$$

Addresses the question: How well did the forecast anomalies correspond to the observed anomalies?

To have better insight and better understand the strong and weak points of the methods used in verifying dichotomous forecast and methods for continuous variables the reader can consult well cited literatures on forecast verification. Wilks

(2006) and Jolliffe and Stephenson (2003) give a detail description on the methods.

3. Results and discussion

3.1 Dekadal and Monthly Forecast Analysis

As stated above forecast verification is testing the performance of the forecast by comparing it with observed values. Twenty-six dekadal forecasts, twenty-six NWP products and eight monthly outlooks are taken for the verification. As observed data are available for only five stations in five Woredas the verification is done for these Woredas only. The Woredas for which the verification is done are Golelcha, Robe, Seru, Yabello and Konso. In fact, the forecasts referred to whole Woreda but the observed data we obtained or had is station or point data. Here we face the problem of verifying forecasts or outlooks given for the whole Woreda against point observed data. Even if we wish to produce Woreda rainfall analysis map, there are no enough meteorological data observation sites in all the Woredas, as many as possible to do map analysis. Due to this problem we are not able to bring observations and forecasts to the same spatial extent. In case of a forecast having one interpretation there was no problem. But if the forecast has more than one interpretation we are forced to do some assumptions. Let's explain this using some examples. For example, a forecast expects dry condition all over the Woreda. This forecast has only one meaning. That is dry all over. It is dry at the point where we have observation site and elsewhere. The problem comes when the forecast has more than one meaning. For example, a given ten daily forecast may say there will be light to moderate rain over few places in Jarso Woreda in the coming ten days. This forecast can as well mean no rain over the other areas. In such cases we need to do some assumptions. The observation sites are used as input when issuing the forecast. The forecasters are using these observational data for monitoring their forecasts. Hence it is logical to conclude that the forecast at the station point will align with that of a better rainfall activity. According to our assumption the forecast at the station is light to moderate rain. With this understanding, the verification was undertaken. The dekadal forecasts consist of two types of outputs. The first one is worded forecast issued by NMA staffs and the second one is Numerical Weather Prediction (NWP) extracted for each project Woredas separately from forecasts posted on the NOAA website. Ten daily forecasts and monthly worded outlooks are expressed in terms of occurrence of rain or non-occurrence of rain, existence of light, light to moderate, heavy, in excess or lower in amount to the areas normal rainfall along with the spatial extent of the forecast. The spatial extents are described as rain over few places, rain over scattered areas, rain over most places (wide spread rainfall activity) and the like.

For example, a typical worded forecast of rain or not rain is like this: -

"In the coming ten days, Jarso and Combolcha Woredas are expected to be dry and sunny but there will be some cloud developments over few areas". This forecast is issued on 21 October 2017 for Jarso and Combolcha Woredas and the forecast is valid for the third dekad of October 2017. This is a forecast of first type. It states non-occurrence of rain in the coming ten days over Jarso.

A typical forecast presented in reference to normal values is like this: -

"In the coming ten days Jarso and CombolchaWoredas are expected to receive near normal rainfall over most places". This forecast is issued on 01 July 2017 for Jarso and Combolcha Woredas and the forecast dekad is first dekad of July 2017. This is a forecast of second type. It seems that the forecasters are sure of the occurrence of rain and expected the amount to be close to the areas normal rainfall.

The third type of worded forecast makes it basis on the rainfall intensity. A typical example for a forecast of this type is like this: -

In the coming month, Benatsemy, Hamer and Konso Woredas will be dominantly sunny but there will be light rains after the middle of the month. This forecast is issued on 01 Sep 2017. It is valid for Sep 2017. Worded dekadal and monthly forecasts are verified using the statistical dichotomous (yes/no) contingency table method. Dekadal NWP forecasts are verified using the methods for continuous variables. Here it is necessary to give the definition of yes and no we adopted in preparing contingency tables. In dichotomous method, for the forecast of type one, yes is occurrence of rain and no is non-occurrence of rain. For the same method and forecast of the second type, yes is taken when the rain is normal or above normal and no is otherwise. In the forecast of third type yes is when there is light, light to moderate rain and no is when there is no rain. With these yes and no definitions in mind the joint distribution of forecast and observation are counted and contingency tables are prepared for the five Woredas separately. Based on the entries in the contingency tables different statistics are computed.

The four-combinations or joint distribution of observed (yes or no) and forecast (yes or no), hits, misses, false alarms and correct negatives, frequencies for each station is shown in the contingency tables below. A contingency table is a useful way to see what types of errors are being made. A large variety of categorical statistics can be computed from the elements in the contingency table to describe particular aspects of forecast performance (WCRP, 2009). The calculated statistics are shown in Table 2. The accuracy (fraction correct), bias score (frequency bias), probability of detection (hit rate), false alarm ratio (FAR), probability of false detection (false alarm rate), threat score (critical success index), equitable threat score (Gilbert skill score), Hanssen and Kuipers discriminant (true skill statistic, Peirce's skill score), Heidke skill score (Cohen's K), Odds ratio, and Odds ratio skill score have been computed for all the stations. The results are shown in Table 2. Over all, in the dekadal forecast, the accuracy is higher for Konso and Golelcha. The other three stations, Robe, Seru and Yabello, are at the same level of accuracy. Eighty-five percent of the forecast were correct of the forecasts given for Konso while seventy seven percent were correct out of the forecasts issued for Robe, Seru and Yabello. The bias score in Table 2a shows that rains are forecasted more than they occur in Golelcha, Seru and Robe. While the opposite happened in Yabello. A perfect score observed for Konso. The probability of detection ranges 79% to 94% meaning that more than 78% of the rains that occurred were detected. The highest probability of detection of rain is seen in Golelcha. The false alarm ratios are lowest at Konso and Yabello, 11% and 12% respectively. The highest is observed at Seru. It is 26%.

Table 1. Dekadal and monthly Contingency tablesby stations

(a)Golelcha – Dekadal forecast

	Observed			
		Yes	No	Total
Forecast	Yes	15	4	19
	No	1	5	6
	Total	16	9	25

(b) Gololcha – Monthly forecast

	Observed			
		Yes	No	Total
Forecast	Yes	5	1	6
	No	1	1	2
	total	6	2	8

(c) Robe – Dekadal forecast

	Observed			
		Yes	No	Total
Forecast	Yes	16	4	20
	No	2	4	6
	Total	18	8	26

(d) Robe – Monthly

	Observed			
		Yes	No	total
Forecast	Yes	4	1	5
	No	2	1	3
	Total	6	2	8

(e) Seru – Dekadal forecast

	Observed			
		Yes	No	Total
Forecast	Yes	14	5	19
	No	1	6	7
	Total	15	11	26

(f) Seru– Monthly

	Observed			
		Yes	No	Total
Forecast	Yes	3	3	6
	No	1	1	2
	Total	4	4	8

(g) Konso– Dekadal forecast

	Observed			
		Yes	No	Total
Forecast	Yes	17	2	19
	No	2	5	7
	Total	19	7	26

(h) Konso – Monthly

	Observed			
		Yes	No	Total
Forecast	Yes	4	2	6
	No	2	0	2
	Total	6	2	8

(i)Yabello – Dekadal forecast

	Observed			
		Yes	No	Total
Forecast	Yes	15	2	17
	No	4	5	9
	Total	19	7	26

(j) Yabello – Monthly

	Observed			
		Yes	No	Total
Forecast	Yes	3	2	5
	No	3	0	3
	Total	6	2	8

Twenty six percent of the forecasts for rain turn out to be false for Seru. No rain occurred in the 26% of the cases though the forecast anticipated a rain. The fraction of no rain events that are incorrectly forecast as rain event is indicated by the probability of false detection (false alarm rate or POFD). POFD is highest at Robe and lowest at Konso. Fifty percent of the cases no rain event forecasted as rain event for Robe. The threat score for these stations varies between 71% and 81%. No much variation detected in the threat score across the stations. The values for the equitable threat score, Peirces's skill score, odds ratio and odds ratio skill score are shown in Table 2a. The accuracy of the forecast relative to that of random chance (HSS) ranges between 42% and 61% for the stations under investigation. The one for Konso is the highest. Over all, in the monthly forecast, the accuracy is higher for Golelcha compared to the other stations. The accuracy is lower than 65% in the other four stations. Seventy-five percent of the forecast were correct of the forecasts given for Golelcha. The accuracy is lower compared to the dekadal forecast in all the stations. Here bear in mind that the sample size very small. The bias score in Table 2b shows that rains are forecasted more than they occur in Seru. While the opposite happened in Robe and Yabello. A perfect score observed at Konso and Golelcha. The probability of detection ranges 50% to 83% meaning that more than 49% of the rains that occurred were detected. The highest probability of detection of rain is seen in Golelcha. The false alarm ratios are lowest at Golelcha and Robe, 17% and 20% respectively. The highest is observed at Seru. It is 50%. Fifty percent of the forecasts for rain turn out to be false for Seru. No rain occurred. The fraction of no rain events that are incorrectly forecast as rain event is indicated by the probability of false detection (false alarm rate or POFD). POFD is highest at Robe and lowest at Konso.

In all cases no rain event forecasted as rain event for Konso and Yabello. The threat score for these stations varies between 38% and 71%. Much variation detected in the threat score among the stations. The values for the equitable threat score, Peirces's skill score, odds ratio and odds ratio skill score are shown in Table 2b. The accuracy of the forecast relative to that of random chance (HSS) ranges between -43% and 33% for the stations under investigation. The one for Golelcha is the highest. Meteorological parameters taken as continuous variables within the range they operate. Verifying forecasts of continuous variables measure how the values of the forecast differ from the values of the observations (WCRP, 2009). As mentioned earlier in this document, twenty-six dekadal NWP outputs are available for the analysis. These NWP outputs cover the period from October 2017 third dekad to August 2017 first dekad. All the NWP outputs are downloaded from NOAA website. The NWP are presented in ranges of rainfall amount. The forecast areas shaded with different colours that indicate the expected amount of rainfall. See figure 2. According to the legend the forecast expects 30 to 250 mm rainfall over Arero Woreda. Observed data obtained for five stations in five Woredas. Hence, the verification is done for these stations only. The stations for which the verification is done are Golelcha, Robe, Seru, Yabello, and Konso. We extract the NWP value at the station location. The NWP value even at the station location is given in range not in single figure. For comparing with observed point data, we took the average over the forecast range. For example, if the indicated NWP at Robe station is 0-10 mm, we average it and get 5 mm. Using the observed data at the station location and the NWP values different statistics computed and graphs plotted that shows the relation between observed and predicted values.

Table 2. Categorical statistics computed from frequencies given in the contingency tables above. (a) Dekadal forecast (b) Monthly forecast. These values describe how well the forecast were performing

(a) Dekadal forecast

Categorical statistics	Station				
	Golelcha	Konso	Robe	Seru	Yabello
Accuracy (fraction correct)	0.80	0.85	0.77	0.77	0.77
Bias score (frequency bias) – BAIS	1.19	1.00	1.11	1.27	0.89
Probability of detection (hit rate) - POD	0.94	0.89	0.89	0.93	0.79
False alarm ratio – FAR	0.21	0.11	0.20	0.26	0.12
Probability of false detection (false alarm rate) – POFD	0.44	0.29	0.50	0.45	0.29
Threat score (critical success index) - TS	0.75	0.81	0.73	0.70	0.71
Equitable threat score (Gilbert skill score) – ETS	0.36	0.44	0.26	0.34	0.30
Hanssen and Kuipersdiscriminant (true skill statistic, Peirce's skill score) - HP	0.49	0.61	0.39	0.48	0.50
Heidke skill score (Cohen's K) - HSS	0.53	0.61	0.42	0.50	0.46
Odds ratio – OR	4.00	5.50	3.33	3.33	3.33
Odds ratio skill score – ORSS	0.90	0.91	0.78	0.89	0.81

(b) Monthly forecast

	Station				
Categorical statistics	Golelcha	Konso	Robe	Seru	Yabello
Accuracy (fraction correct)	0.75	0.50	0.63	0.50	0.38
Bias score (frequency bias) – BAIS	1.00	1.00	0.83	1.50	0.83
Probability of detection (hit rate) - POD	0.83	0.67	0.67	0.75	0.50
False alarm ratio – FAR	0.17	0.33	0.20	0.50	0.40
Probability of false detection (false alarm rate) - POFD	0.50	1.00	0.50	0.75	1.00
Threat score (critical success index) - TS	0.71	0.50	0.57	0.43	0.38
Equitable threat score (Gilbert skill score) – ETS	0.20	-0.14	0.08	0.00	-0.18
Hanssen and Kuipersdiscriminant (true skill statistic, Peirce's skill score) - HP	0.33	-0.33	0.17	0.00	-0.50
Heidke skill score (Cohen's K) - HSS	0.33	-0.33	0.14	0.00	-0.43
Odds ratio - OR	3.00	1.00	1.67	1.00	0.60
Odds ratio skill score – ORSS	0.67	-1.00	0.33	0.00	-1.00

Table 3. Statistics computed using observed data at station location and NWP extracted for the same location and period. These values describe how well the NWP were performing

Categorical statistics	Station				
	Golelcha	Konso	Robe	Seru	Yabello
Additive bias or mean error	0.28	3.92	-6.56	1.66	17.28
Multiplicative bias	0.98	1.29	0.75	1.15	2.96
Mean absolute error	16.71	20.49	12.71	18.64	17.80
Root Mean Square Error	23.57	33.77	19.24	28.44	25.48
Correlation	0.62	0.36	0.78	0.64	0.76
Anomaly correlation	0.61	0.38	0.79	0.64	0.76

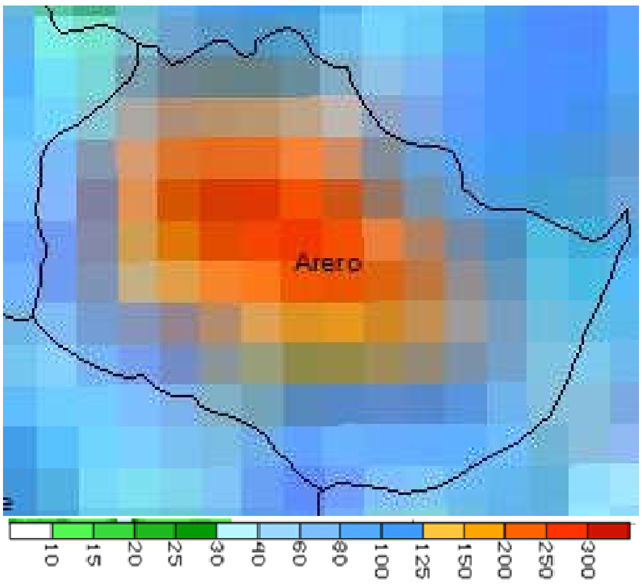


Figure 2. NWP forecast for AreroWoreda valid for 1-10 May 2017

The values for additive bias, multiplicative bias, mean absolute error, root mean square error, correlation and anomaly correlation are shown in Table 3 above. On the average, the forecast exceeds the observed value at Yabello, Konso and Seru as indicated by the additive bias. While under forecasting is the case over Robe. The average difference between forecast and observed is much higher at Yabello. The additive bias exceeded 15 in Yabello. Compare to the other stations the value is much higher. Even the multiplicative bias is the highest over the same station. The additive bias for Golelcha is close to zero slightly to the positive direction indicating very low over forecasting. The root mean square error and mean absolute error or in other words the Magnitude of the errors are high in Konso. In most respect, the statistics do not show promising results for Konso. We cannot much rely on the NWP outputs we get for Konso.

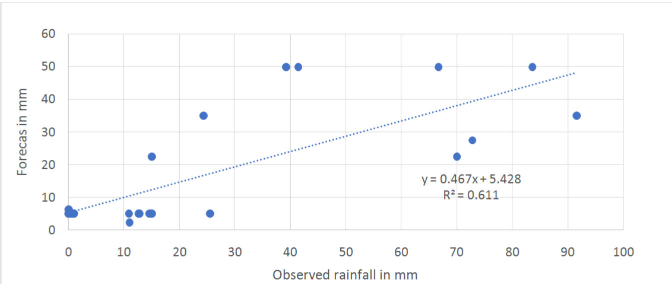


Fig.3. Scatter plot of NWP rainfall forecast versus observed over Robe

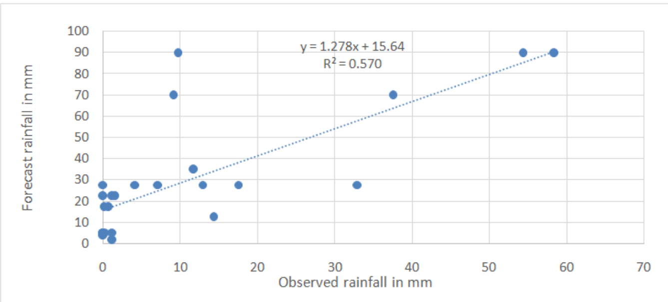


Fig.4. Scatter plot of NWP rainfall forecast versus observed over Yabello

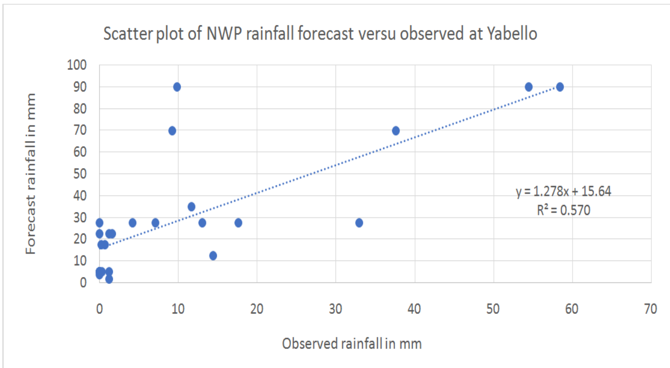
No much difference observed between correlation of basic data and anomalies. The correlation coefficient is very promising at Yabello and Robe while the correlation at Konso is very low.

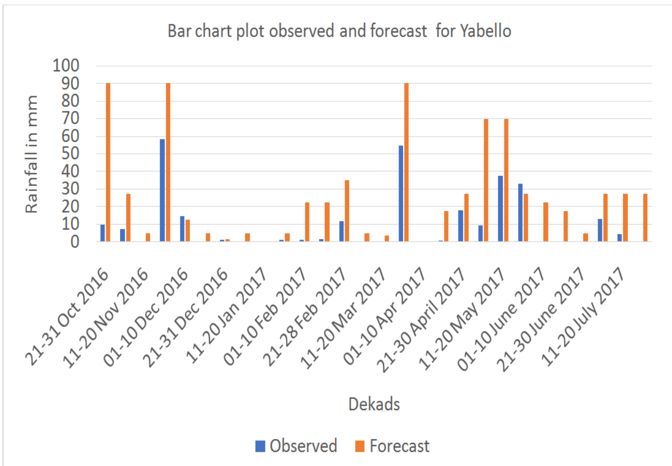
The next step in the NWP verification is generating scatter and box plots as well as bar graph (Annex I) to show how the observed and forecast values correspond each other. Two scatter plots are shown in Figures 3 and 4 for Robe and Yabello. Both observation and forecast are expressed in mm. Each dot is a particular observation and forecast pair. Two lines are shown on the plots. One is a 45-degree line that passes through the origin and the other is least square regression line fit. If the forecasts were perfect the least square regression line fit passes through the origin. In both cases the lines do not pass through the origin. On the other hand, the forecast seems ok as most points are not much far away from 45-degree line. Though not systematic there are occasions when low rainfalls are forecasted high and vice versa.

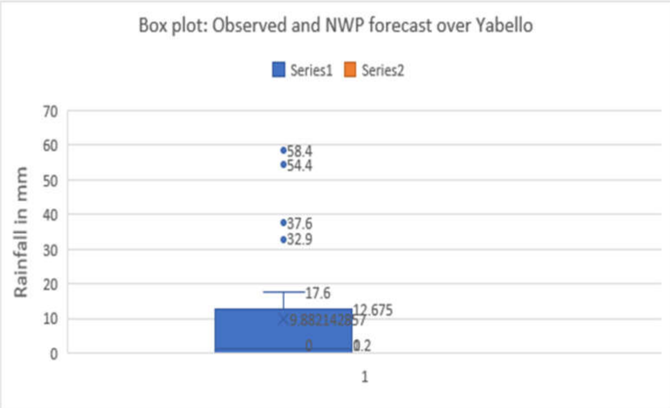
Conclusions and recommendation

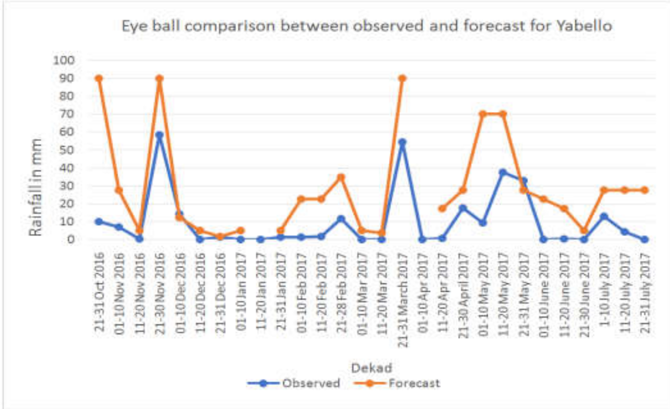
There are no meteorological stations in some of the CIARE interventions Woredas. Specifically, no meteorological stations in Jarso, Bentsemay, Dire and Hamer. For such Woredas, we cannot find normal data even for a single location in the Woreda. This means that our climatological knowledge of the dekal and monthly normal values can only be deduced from blended data or purely interpolation. Presenting the forecast for such Woredas in terms of normal rain could be very difficult for the assigned forecaster or meteorologist. In addition, if the season is not a rainfall season and dominantly sunny and dry, it does not matter if the forecast is presented in the form of rain, no rain or light to moderates rain. But for other Woredas where we have station and can get normal values, the dekadal and monthly forecast can be presented in relation to normal values. We can issue a forecast that clearly states whether the expected rainfall is above, near or below normal. The writer of this report finds very few forecasts of this kind. If possible, it would have value for the user as well as for one who verifies the forecast if the forecasts are presented in relation to normal values as we do in the regular forecast issued for the whole country in coarse resolution. The NWP disseminated are outputs of NOAA. As we know we have more localized weather and climate information than NOAA has. It would be much better if we were able to run the WRF model we have at home and issue ten daily forecasts. The period we took for verification is too short to arrive to concrete conclusion, but with the sample we used we have found out that, in relative terms, the accuracy of the dekadal worded forecast is much better than the worded monthly forecasts. Other statistics can be inferred from the summary tables given in the documents. When we come to the NWP output, in all aspects we do not get promising results when we look the verification statistics we found for Konso.

Annex I









References

WCRP, 2009: Forecast Verification - Issues, Methods and FAQ. Available online from http://www.cawcr.gov.au/projects/verification/ [Accessed on 02 August 2017].

Jolliffe, I. T., and D. B. Stephenson, Eds., 2003: Forecast Verification: A Practitioner's Guide in Atmospheric Science. John Wiley and Sons, 240 pp. Katz, R. W., and A. H. Murphy, Eds., 1997:

Wilks, D.S., 2006. Statistical Methods in Atmospheric Sciences. 2nd ed., Department of Earth and Atmospheric Sciences Cornell University, Academic Press is an imprint of Elsevier.

****