Article:

Nkiaka, E., Nawaz, N.R. and Lovett, J.C. (2017) Evaluating global reanalysis precipitation datasets with rain gauge measurements in the Sudano-Sahel region: case study of the Logone catchment, Lake Chad Basin. Meteorological Applications, 24 (1). pp. 9-18. ISSN 1350-4827

https://doi.org/10.1002/met.1600

This is the peer reviewed version of the following article: Nkiaka, E., Nawaz, N.R. and Lovett, J.C. (2017), Evaluating global reanalysis precipitation datasets with rain gauge measurements in the Sudano-Sahel region: case study of the Logone catchment, Lake Chad Basin. Met. Apps, 24: 9-18, which has been published in final form at https://doi.org/10.1002/met.1600. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Evaluating global reanalysis precipitation datasets with rain gauge measurements in the Sudano-Sahel region: case study of the Logone catchment, Lake Chad Basin

E. Nkiaka[1], N.R. Nawaz[1] and J.C. Lovett[1]
[1]School of Geography, University of Leeds:
Correspondence to: E. Nkiaka (gyenan@leeds.ac.uk)

Abstract

Africa has a paucity of long-term reliable meteorological ground station data and reanalysis products are used to provide the climate estimations that are important for climate change projections. This paper uses monthly observed precipitation records in the Logone catchment of the Lake Chad Basin (LCB) to evaluate the performance of two global reanalysis products: the Climate Forecasting System Reanalysis (CFSR) and ERA Interim datasets.

The two reanalysis products reproduced the monthly, annual and decadal cycle of precipitation and variability relatively accurately albeit with some discrepancies. The catchment rainfall gradient was also well captured by the two products. There are good correlations between the reanalysis and rain gauge datasets though significant deviations exist, especially for CFSR. Both reanalysis products overestimated rainfall in 68% of the rain gauge stations. ERA Interim produced the lowest bias and mean absolute error (MAE) with average values of 2% and 6.5mm/month respectively compared to 15% and 34mm/month for the CFSR. However, both reanalysis products systematically underestimated annual rainfall in the catchment during the period 1997-2002 for ERA-Interim and 1998-2000 for CFSR. This research demonstrates that evaluating reanalysis products in remote areas like the Logone catchment enables users to identify artefacts inherent in reanalysis datasets. This will facilitate improvements in certain aspects of the reanalysis forecast model physics and parametrisation to improve reanalysis dataset quality.

Our study concludes that the application of each reanalysis product in the catchment will depend on the purpose for which it is to be used and the spatial scale required.

Key words: CFSR, ERA Interim, rain gauge, reanalysis, Sudano-Sahel region, Logone catchment, Lake Chad basin.

1) Introduction

Scarcity of meteorological data is a major bottleneck that retards advancement of knowledge on water management and climate change in many parts of the world, especially in developing regions (Buytaert et al., 2012). Reliable, long-term, and well distributed climate information is essential to informing policies that aim to address the consequences of climate variability and change (Baisch, 2010; van de Giesen et al., 2014) and enhance water resource management.

In Sub Sahara Africa there is uneven distribution of hydro-meteorological stations and many of these are in decline, with the result that most areas of Africa, particularly those in Central Africa, are unmonitored (Washington et al., 2006). Another challenge in these regions is that, even when data is collected and archived, accessing it requires much effort and money as the data are not digitised or readily available (Fuka et al., 2013).

Data scarcity in Central Africa in particular has been identified by many researchers as a constraint to modelling and validation e.g. (Haensler et al., 2013; Candela et al., 2014; Maidment et al., 2015).

To overcome this challenge, multiyear global gridded representation of weather known as reanalysis datasets are now available. The large number of variables makes reanalyses datasets ideal for investigating climate variability and to enhance management of water resources. Examples of reanalysis datasets currently in use include: National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) Climate Forecasting System Reanalysis (CFSR) (Saha et al., 2014); European Center for Medium-Range Weather Forecasts (ECMWF) ERA-Interim (Dee et al., 2011); and Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al. 2011). These reanalysis datasets have spatial resolutions of 0.3125° (~38km), 0.703° (~82km) and 0.50° (~50km) for CFSR, ERA Interim and MERRA respectively.

The reanalysis products have been used for many different applications around the world (Fuka et al., 2013; Blacutt et al., 2015; Krogh et al., 2015; Sharifi et al., 2016). Many studies have also been carried out in Africa to evaluate the accuracy of precipitation estimates from reanalysis datasets at a monthly time scale or more e.g. (Maidment et al., 2013; Zhang et al., 2013; Worqlul et al., 2014; Koutsouris et al., 2015). However, in the course of modelling climate change impacts on the Logone catchment of the Lake Chad Basin (LCB), no evaluation studies were found for Central Africa. Before reanalysis datasets are used in this region, their accuracy needs to be tested against in situ measurements.

The main objectives of this study are to: (i) evaluate the accuracy of precipitation estimates from two reanalysis datasets, CFSR and ERA Interim against rain gauge data in the Logone catchment of the LCB, and (ii) evaluate how data from these reanalysis datasets are able to reproduce the monthly, annual and decadal rainfall cycle. The results will identify which of the reanalyses products better reproduces precipitation and variability estimates for the catchment and so validate their use in hydrological and climate models in this data scarce region.

The paper is structured as follows: Section 2 describes the data and methodology used in the study; Section 3 presents the results obtained; Section 4 provides a general discussion on the results and Section 5 gives a general summary and conclusion.

2) Materials and methods

2.1) Study area

The Logone catchment is a transboundary catchment shared by Cameroon, Chad and the Central Africa Republic with an estimated catchment area of 86500km^2 at Logone Gana discharge station (Figure 1). It lies between latitude 6˚ - 12˚ N and longitude 13˚ - 16˚ E. The Logone River forms part of the international boundary between Cameroon and Chad. The Logone floodplains are the most extensive and among the richest ecological wetlands in the African Sahel covering an estimated area of 6000 km^2 (Loth et al., 2004). There is a high concentration of wildlife, which is protected in two National Parks (Waza and Kalamaloue). The Waza National Park is a Ramsar site and a Biosphere Reserve of international importance. Many migratory birds make use of the seasonally abundant food resources (Loth et al., 2004).

The Logone has its source in Cameroon through the Mbere and Vina Rivers which flow from the north eastern slopes of the Adamawa Plateau (Molua and Lambi, 2006). In Lai, the Logone is joined by the Pende River from the Central Africa Republic and flows from south to north. In this region elevation ranges from 300 masl around Kousseri to about 1200 masl in the Adamawa Plateau. Apart from some local mountains in the south the basin topography is quite flat with an average slope of about 1.3% in a south to north gradient (Le Coz et al., 2009).

The catchment has both a Sudano climate in the south and semi-arid climate in the north. Estimated average annual rainfall varies between 600 mm/year in the north to about 1200 mm/year in the south (Molua and Lambi, 2006). The climate in the region is characterized by high spatial variability and is dominated by the tropical continental air mass (the Harmattan) and the marine equatorial air mass (monsoon) (Candela et al., 2014). Almost all rain falls during the rainy season from April/May/ to September/October and mean annual temperature is 28°C.

2.2) Data Sources

2.2.1) Rain gauge data

Monthly gauge rainfall was obtained from “Système d'Informations Environnementales sur les Ressources en Eau et leur Modélisation” (SIEREM) (Boyer et al., 2006). Quality control of the gauge data was done in three steps: (i) selecting only stations that had monthly data dating back to 1979 to match the period of the reanalysis data; (ii) selecting stations that had data for a minimum of 15 years and (iii) eliminating stations that had extended gaps of more than six months in each year. Gaps in the monthly rainfall time series were filled using the Artificial Neural Network (ANN) Self-Organizing Map (SOM) technique (Nkiaka et al., 2016). Using these criteria, out of 55 rain gauge stations located inside the catchment, only 19 stations had consistent data spanning the period 1979 - 2002. To increase the number of rain gauges, six additional stations located outside the catchment but with the same climate conditions were selected.

2.2.2) Reanalysis datasets

A reanalysis project involves the reprocessing of observational data spanning an extended historical period: “It makes use of a consistent modern analysis system, to produce a dataset that to a certain extent can be regarded as a "proxy" for observation with the advantage of providing coverage and time resolution often unobtainable with normal observational network” (Morse et al., 2013). It is generated by a data assimilation system combining observations with a numerical weather prediction model. For the entire reanalysis period, the model physics remain unchanged in the forecast model for consistency of the output data. The reanalysis consequently provides a picture of the global climate over a period during which observational data are available. Reanalysis data can provide a multivariate, spatially complete, and coherent record of the global atmospheric circulation (Dee et al., 2011).

The Climate Forecast System, NCEP version 2 (CFSv2) is an upgraded version of CFS version 1 (CFSv1). It is a reanalysis product first developed as part of the Climate Forecast System by NCEP in 2004 with quasi-global coverage and is a fully coupled atmosphere-ocean-land model used by NCEP for seasonal prediction (Saha et al., 2014). CFSR has a 3D-variational analysis scheme of the upper-air atmospheric state with 64 vertical levels and a horizontal resolution of 38km spanning the period 1st January 1979 to present day (Saha et al., 2014).

ERA-Interim is the latest global atmospheric reanalysis produced by ECMWF and covers the period from 1st January 1979 to present day (Dee et al., 2011). The core component of the ERA-Interim data assimilation system is the 12-h 4D-variational analysis scheme of the upper-air atmospheric state, which is on a spectral grid with triangular truncation of 255 waves (corresponding to approximately 80 km) and a hybrid vertical coordinate system with 60 vertical levels. Details concerning the two reanalysis products can be found in Dee et al. (2011) and Saha et al. (2014) for ERA Interim and CFSR respectively.

Reanalysis data for the study area was obtained for an area bounded by latitude 6°N-12.0°N and longitude 13°E-17.25°E.

2.3) Method for comparison

To identify which reanalysis grid point to compare with which rain gauge station(s), grid boxes were created with the same native resolution for all reanalysis grid points (Figure 2). Where two or more rain gauges were located inside the same grid box, their precipitation estimates were compared with the precipitation estimate of that grid box. This method has been used to evaluate reanalysis datasets with in situ measurements by Diro et al. (2009); Worqlul et al. (2014) and Zengyun et al. (2014). Pairwise statistical analyses were carried out between reanalysis grid point precipitation estimates and rain gauge data located within the grid box. This was done assuming that the reanalysis grid point precipitation estimate within each grid is the average for the whole of that grid box.

Five different statistical measures were used to evaluate the results: correlation coefficient (R), coefficient of determination (R^2), mean absolute error (MAE), Bias and the Nash Sutcliff Efficiency (NSE) (Maidment et al., 2013; Worqlul et al., 2014; Koutsouris et al., 2015; Sharifi et al., 2016).

Graphical plots were used to compare monthly, annual and decadal rainfall located inside the grid box with the rainfall estimate of that grid box. Reanalysis precipitation estimates were aggregated to monthly and annual totals to match the available rain gauge data. Mean annual rainfall of the reanalysis products and gauge data were also calculated for each station with their respective error bars. In addition, a graphical plot of mean annual rainfall was made to show variation of rainfall with latitude.

Annual and monthly rainfall data for each station was averaged over two different climatic zones in the catchment: Sudano and semi-arid. The distinction between the climatic zones was based on rainfall gradients in the catchment. Stations located between latitude 6°-10°N were grouped together in the Sudano area while stations located above latitude 10°N were grouped together as the semi-arid area. Following these criteria, 12 stations were located in the semi-arid area and 13 were located in the Sudano area. Diro et al. (2009) and Maidment et al. (2013) also used this approach to evaluate reanalysis products in Ethiopia and Uganda respectively.

3) Results

3.1) Monthly rainfall variation

Results shown in Figures 3(a) & (b) reveal that, despite the smoothing effect over large areas, rainfall estimates from the two reanalysis datasets follow the same monthly cycle shown by gauged rainfall in the two spatial zones. The general pattern is that rainfall in the area begins in April/May and lasts until September/October (Loth et al., 2004; Molua and Lambi, 2006). CFSR overestimated monthly rainfall in most stations while ERA-Interim underestimated in some stations and over-estimated in others. The ability of CFSR and ERA Interim to efficiently capture monthly precipitation cycles has been reported from Ethiopia, Australia, Tanzania and South America (Worqlul et al., 2014; Fu et al., 2015; Koutsouris et al., 2015; Krogh et al., 2015). Blacutt et al. (2015) reported that; CFSR consistently overestimated precipitation estimates in three basins in South America (La Plata, Altiplano and Amazon) while Koutsouris et al. (2015) reported that CFSR underestimated seasonal precipitation in the Kilombero Valley in Tanzania.

Maidment et al. (2013) also reported that ERA Interim overestimated seasonal precipitation in Uganda.

3.2) Annual rainfall

Annual gauge rainfall varies between 500-1500mm/year, ERA Interim varies between 500 – 1300mm/year while CFSR rainfall varies between 600 – 2600mm/year and consequently has the highest spread among the three datasets (Figure 4a and Table 1). Koutsouris et al. (2015) observed a similar spread in CFSR precipitation estimates in the Kilombero Valley in Tanzania. Both reanalysis products overestimate annual rainfall in 68% of the stations in the catchment but overestimation is greater for CFSR (27%) compared to 8% for ERA Interim.

Figures 4(b) & (c) show the average annual rainfall from stations located in the semi-arid (latitude 10° - 12°) and Sudano (latitude 6° -10°) areas respectively. The figures show that both reanalysis products were able to capture inter-annual rainfall variability in the catchment, though with some differences. Figures 4(b) & (c) also show that the reanalyses products were able to capture the droughts that affected the region especially in 1984 which is reported as the driest year in the region during the period under study (Molua and Lambi, 2006). However, for this extreme drought year (1984) CFSR slightly under estimated rainfall in the semi-arid area and over-estimated in the Sudano area, while ERA Interim overestimated in both.

Generally, CFSR overestimated annual rainfall by an average of 19% in the semi-arid area compared to 11% in the Sudano area of the catchment. ERA Interim demonstrated almost perfect performance in the semi-arid area and overestimated by 3% in the Sudano area. Dile and Srinivasan (2014) and Worqlul et al. (2014) reported that CFSR over/underestimates rainfall in some stations within the same catchment in their respective studies, De Leeuw et al. (2015) and Sharifi et al. (2016) reported that ERA Interim generally underestimated rainfall in their respective study areas. Fu et al. (2015) reported that ERA Interim underestimated mean annual rainfall in Australia while CFSR overestimated it in some regions and underestimated in others.

The analysis also show that both products captured the spatial distribution of rainfall in the Sudano and semi-arid areas of catchment fairly well. Annual rainfall generally ranges between 600-900mm/year and 900-1400mm/year in the semi-arid and Sudano areas respectively (Loth et al., 2004); Figures 4(b) & (c). Fu et al. (2015) reported that CFSR and ERA Interim were able to reproduce the observed spatial patterns of annual rainfall in Australia. Results also show that annual rainfall estimates for CFSR grid box (grid point) located in Ngaoundere is significantly higher than gauge data by more than 1000mm/year while ERA Interim underestimated rainfall in that station by more than 150mm/year. Haensler, et al. (2013), reported that data from the National Climatic Data Center (now National Centers for Environmental Information) strongly overestimated precipitation in this station.

Furthermore, CFSR systematically underestimated annual rainfall during the period 1998 – 2000 across most stations in the catchment. Monteiro et al. (in press) observed a similar issue during the same period while modelling the Tocantins catchment in Brazil and attributed it to an artefact. ERA-Interim also systematically underestimated mean annual rainfall in the catchment during the period 1997-2002 with annual average dropping below 300mm/year and 600mm/year in the semi-arid and Sudano zones respectively in the year 2000 (Figures 4b&c).

3.3) Decadal monthly rainfall variability

To compare the monthly rainfall cycle between gauge and reanalysis grid points a 10-year mean was calculated for each month starting from the first year of the data period in each decade i.e. for the 1980 decade (1981-1990) and for the 1990 decade (1991-2000). Figures 6(a) & (b) show that at a decadal time scale both reanalysis datasets follow the same monthly cycle as the gauge data for stations located in the semi-arid (Bailli) and Sudano (Bekao).

Monthly precipitation estimates from CFSR are similar to measured data with the peak occurring in the month of August in the 1980 decade in Bailli while peak rainfall during the same period in the ERA Interim occurred in July. For Bekao station, peak rainfall in the ERA Interim coincides with measured data and occurred in July while CFSR peak rainfall occurred in August during the 1980 decade. During the 1990 decade, peak rainfall in CFSR and ERA Interim occurred during the same month but ERA interim underestimated peak decadal monthly rainfall in both stations. Similar results on decadal monthly rainfall by both reanalysis products were reported by Zhang et al. (2013) in the Southern African region. Di Giuseppe et al. (2013) also reported that there were some discrepancies in decadal precipitation estimates from ERA Interim over Africa with 1979-1989 decade presenting a wet bias compared to other decades.

3.4) Variation of rainfall with latitude

Figure 5 shows the variation of mean annual rainfall with latitude in the study area. In general rainfall isohyets in the Lake Chad basin form parallel east to west lines. Both reanalysis datasets are able to reproduce the rainfall gradient in the catchment with rainfall increasing as latitude decreases southwards. Annual rainfall is highest in the south of the Logone catchment at the source of the Logone River in Cameroon on the Adamawa Plateau (Loth et al., 2004). Rainfall is lower in the north because of the Saharan anticyclone and dry continental trade winds that blow as far south as 5° N (Molua and Lambi, 2006). The Adamawa Plateau acts like a shield preventing progress of the Atlantic air masses northwards and forcing the Intertropical Convergence Zone (ITCZ) towards the western part of Cameroon. (Ardoin-Bardin, 2004; Molua and Lambi, 2006). During the rainy season, the monsoon winds blow from the south and push the continental trade winds northwards. Higher rainfall during this period is indicative of the strength of the monsoon winds (Molua and Lambi, 2006). Both reanalysis products could accurately capture the rainfall gradient in the catchment.

3.5) Statistical Analysis of monthly variation in rainfall

Although both reanalysis products generally show good correlation with monthly rain gauge data ($0.70 \leq r \leq 0.85$ and $0.60 \leq R^2 \leq 0.78$; Table 2). Similar values were obtained by Worqlul et al. (2014) in the Lake Tana Basin in Ethiopia. However, few large deviations still exist between the reanalysis datasets and gauge data with variations in gauge rainfall measurements within the same grid box (Table 2). For example: Mouvouday, Kalfou, Yagoua and Dana; Bailli and Bousso; Pandzangue, Bekao and Moundou located inside the same ERA Interim grid boxes all produced different values after statistical correlation with rainfall from the same grid points. A similar situation was observed for Moundou and Delli; Dana and Bongor; Doukoula and Kalfou; Bekao and Pandzangue located inside the same CFSR grid boxes.

The monthly bias values shown in Table 2 indicate that, both reanalysis products generally overestimated rainfall in 68% of the stations with the highest overestimation recorded in Ngaoundere for CFSR. Analysis also showed that ERA Interim had the lowest bias and MAE with average values of 2% and 6.5mm/month respectively compared to 15% and 34mm/month for CFSR.

Results of NSE as shown in Table 2 are in the range (-1.15≤NSE≤0.56) for CFSR and (0.11≤NSE≤0.63) for ERA Interim, indicating that both forecasting models produced modest results even though ERA Interim outperformed CFSR.

4) Discussion

Rainfall over/under estimation by both reanalysis products over the Logone catchment as observed in this study may be attributed to the fact that rainfall in the region is highly variable (Molua and Lambi, 21006). Furthermore, there are only few surface observation stations in Central Africa, leading to uncertainty in forecast model input. In a study of observed and simulated precipitation changes over Africa using different datasets, Maidment et al. (2015) attributed the large discrepancies in results observed over Central Africa to low rain gauge station density. Dee et al. (2011) reported that the large differences observed in precipitation estimates from ERA Interim over Central Africa were due to uncertainties as a result of sparse radiosonde coverage. The authors also attributed it to the possible presence of a substantial warm bias in the model associated with underestimated aerosol optical depth. Wang et al. (2015) also reported on the paucity of radiosonde observations for different reanalysis products over Central Africa. Meanwhile, Agusti-Panareda et al. (2010) stated that, the biases in ERA Interim precipitation estimates over the African continent could be attributed to “scarcity of observations to constrain the assimilation cycle and the limitation of the convection and land surface parameterizations over the region”.

In fact, a variable such as precipitation is not directly assimilated but constrained by observations used to initialize the forecast model, therefore the accuracy of model-generated estimates depends on the quality of the model physics as well as the observations. The quality of precipitation estimates from reanalysis products also depends on sea surface temperature boundary conditions, other assimilated observations and on the physical parameterization of the model (Zhang et al., 2013).

The large discrepancy between CFSR and gauge data observed in Ngaoundere could be attributed to the complex topography of the region, as suggested by Zhang et al. (2013) in southern Africa. In addition, The Intertropical Convergence Zone (ITCZ) creates a complex and unpredictable movement of air masses in the Sudano-Sahelian region making it difficult for the CFSR forecast model to produce accurate precipitation estimates. Furthermore, given that rainfall estimates from the two reanalysis products for the station are opposite, with CFSR producing high and ERA Interim producing low estimates, the ITCZ could be located in different positions in the two forecast models.

There are also potential errors due to the comparison being made between the rain gauge point measurement and grid point, which is the average of a grid box measuring 38 km x 38 km and 80 km x 80 km for CFSR and ERA Interim respectively. For example, grid boxes with two or more rain gauges could have significant differences in gauge measurements that can be attributed to the generally high spatio-temporal precipitation variability in the region. It could also be due to elevation differences between reanalysis grid average and the individual rain gauge station site(s). Furthermore, rain gauge measurements could also be subject to under catch; measurement errors; or may not register rain showers less than 1mm. While reanalysis forecast models can produce rainfall estimates which are less than 1mm and these estimates when accumulated over a longer time scale could have a significant influence on reanalysis estimates compared to gauge data.

Wang et al. (2011) attributed the artefact identified in CFSR precipitation estimates during the period 1998-2000 as observed in this study to be related to possible changes in the assimilation of solar radiation and surface wind data by CFSR. It could also be due to changes in instrument(s) used for obtaining the data, faulty instrument(s), and/or recalibration of the data acquisition instrument after replacement of defective part(s). Meanwhile Di Giuseppe et al. (2013) attributed the decadal discrepancy in ERA Interim precipitation estimates over Africa to a model artifact that generated an unrealistic strengthening and northward displacement of the monsoon cycle in the first decade of the data set.

5) Conclusions

The main objectives of this study were to evaluate the accuracy of precipitation estimates from two reanalysis datasets; CFSR and ERA Interim with rain gauge measurements, and compare how these products reproduce the monthly, annual and decadal rainfall cycle in the Logone catchment. Results obtained show that;

- Both reanalyses products could reproduce the precipitation cycle in the catchment at monthly, annual and decadal time scale and the inter-annual variability is well captured.
- Both products were also able to reproduce the rainfall gradient in the catchment, although they overestimated rainfall in 68% of the stations across the catchment.
- At the monthly time scale both reanalysis products show good correlation with rain gauge data although differences still exist between the reanalyses datasets and rain gauge data especially for CFSR.

Results from in this study are comparable to those obtained from Africa by other researchers (Maidment et al., 2013; Zhang et al., 2013; Dile, and Srinivasan, 2014; Worqlul et al., 2014; Koutsouris et al., 2015) and globally (Blacutt et al., 2015; Fu et al., 2015; Krogh et al., 2015; de Leeuw et al., 2015; Sharifi et al., 2016). From these results, the application of each reanalysis product in the catchment will depend on the purpose for which it is to be used and on the spatial scale required, given that both products have the same temporal resolution. However users may need to exclude the period during which rainfall is systematically underestimated in their analysis. The research also shows that evaluating reanalysis products in remote locations like the Logone catchment may enable users to identify artefacts inherent in reanalysis datasets and so may enable the model developers to improve on certain aspects of the model physics and parametrisation scheme to improve the reanalysis datasets quality.

6) Acknowledgement

This research was supported by a Commonwealth Scholarship award to the first author. The authors are grateful to NCEP/NCAR, ECMWF and SIEREM for providing the data used in this study and to the two anonymous reviewers for helping to improve on the manuscript. We are indebted to the two anonymous reviewers whose valuable contributions helped to greatly improve the manuscript.

7) References

Ardoin-Bardin S. 2004. Hydroclimate variability and impacts on water resources of large hydrological catchments in the Sudanese–Sahelian area (in French). *Thesis (Ph.D). University of Montpellier II.*

Agusti-Panareda, A, Balsamo G, Beljaars A. 2010. Impact of improved soil moisture on the ECMWF precipitation forecast in West Africa. *Geophys. Res. Lett.*, **37**: L20808, doi:10.1029/2010GL044748.

Baisch J. 2010. Data Shortage in Africa. *Desalination.* **252**: 101-112.

Blacutt LA, Herdies DL, de Gonçalves LGG, Vila DA, Andrade M. 2015. Precipitation comparison for the CFSR, MERRA, TRMM3B42 and combined scheme datasets in Bolivia. *Atmos. Res.* **163**:117 – 131.

Boyer J, Dieulin C, Rouche N, Cres A, Servat E, Paturel J. 2006. SIEREM: An Environmental Information System for Water Resources, Vol. IAHS Publication 308. IAHS Press: Wallingford, United Kingdom

Buytaert W, Friesen J, Liebe J, Ludwig R. 2012. Assessment and Management of Water Resources in Developing, Semi-arid and Arid Regions. *Water Resour. Manag.* **26**: 841-844.

Candela L, Elorza FJ, Tamoh K, Jiménez-Martínez J, Aureli A. (2014) Groundwater modelling with limited data sets: the Chari– Logone area (Lake Chad Basin, Chad). *Hydrol. Processes*, **28**: 3714-3727.

Dee DP, Uppala SM, Simmons AJ, Berrisford P, et al. 2011. The ERA-Interim ′ reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**: 553–597, doi: 10.1002/qj.828.

Di Giuseppe F, Molteni F, Dutra,E. 2013. Real-time correction of ERA-Interim monthly rainfall, *Geophys. Res. Lett.* **40**: 3750– 3755. doi:10.1002/grl.50670.

Dile, Y.T. and Srinivasan, R. (2014). Evaluation of CFSR Climate Data for Hydrologic Prediction in Data-Scarce Watersheds: An Application in the Blue Nile River Basin. *J. Am. Water Resour. Assoc.* doi: 10.1111/jawr.12182.

Diro GT, Grimes DIF, Black E, O'Neill A, Pardo-Iguzquiza E. 2009. Evaluation of reanalysis rainfall estimates over Ethiopia. *Int. J. Climatol.* **29**(1): 67–78.

Fu G, Charles SP, Timbal B, Jovanovic B, Ouyang F. 2015. Comparison of NCEP-NCAR and ERA-Interim over Australia. *Int. J. Climatol.* doi:10.1002/joc.4499

Fuka DR, MacAllister CA, Degaetano AT, Easton ZM. 2013. Using the Climate Forecast System Reanalysis dataset to improve weather input data for watershed models. *Hydrol. Processes,* **28**: 5613-5623.

Haensler A, Saeed F, Jacob D. 2013. Assessing the robustness of projected precipitation changes over central Africa on the basis of a multitude of global and regional climate projections. *Clim. Change.* **121**(2): 349-363.

Koutsouris AJ, Chen D, and Lyon SW. 2015. Comparing global precipitation data sets in eastern Africa: a case study of Kilombero Valley, Tanzania. *Int. J. Climatol.* doi: 10.1002/joc.4476.

Krogh SA, Pomeroy JW, Mcphee J. 2015. Physically based mountain hydrological modelling using reanalysis data in Patagonia. *J. Hydrometeorol.* **16**: 172–193

Le Coz ML, Delclaux F, Genthon P, Favreau G. 2009. Assessment of Digital Elevation Model (DEM) aggregation methods for hydrological modelling: Lake Chad Basin, Africa. *Comput. Geosci.* **35**(8): 1661-1670.

De Leeuw J, Methven J, Blackburn M. 2015. Evaluation of ERA-Interim reanalysis precipitation products using England andWales observations. *Q. J. R. Meteorol. Soc.* **141**: 798–806, doi:10.1002/qj.2395.

Loth P. (Editor). 2004. The Return of the Water: Restoring the Waza Logone Floodplain in Cameroon. IUCN, Gland, Switzerland and Cambridge, UK.

Molua EL, Lambi CM. 2006. Climate, hydrology and water resources in Cameroon. In: CEEPA discussion paper no. 33, CEEPA, Pretoria.

Maidment RI, Grimes DIF, Allan RP, Greatrex H, Rojas O, Leo O. 2013. Evaluation of satellite-based and model re-analysis rainfall estimates for Uganda. *Meteorol. Appl.* **20**: 308 - 317.

Maidment RI, Allan IRP, Black E. 2015. Recent observed and simulated changes in precipitation over Africa, *Geophys. Res. Lett*. **42**: 8155–8164, doi: 10.1002/2015GL065765.

Monteiro JAF, Strauch M, Srinivasan R, Abbaspour K, Gücker B. Accuracy of grid precipitation data for Brazil: Application in river discharge modelling of the Tocantins catchment. *Hydrol. Processes.* doi: 10.1002/hyp.10708.

Morse A, Caminade C, Tompkins A, McIntyre KM, 2013. *The QWeCI* project (Quantifying Weather and Climate Impacts on Health in Developing Countries) - Final report. p. 60. (Unpublished).

Nkiaka E, Rizwan NR, Lovett JC. 2016. Using Self-Organizing Maps to infill missing data in hydro-meteorological time series from the Logone catchment, Lake Chad basin. *Environ Monit Assess*. **188**. doi: 10.1007/s10661-016-5385-1.

Rienecker MM, Suarez MJ, Gelaro R, Todling R, et al. 2011. MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Clim.* **24**: 3624–3648

Saha S, Moorthi S, Wu X, Wang J, Nadiga S, Tripp P, Behringer D, Hou YT, Chuang HY, Iredell M. 2014. The NCEP climate forecast system version 2. *J. Clim.* **27**: 2185 – 2208

Sharifi E, Steinacker R, and Saghafian B. 2016. Assessment of GPM-IMERG and Other Precipitation Products against Gauge Data under Different Topographic and Climatic Conditions in Iran: Preliminary Results. *Remote Sens.* **8**: 135. doi:10.3390/rs8020135

Van de Giesen N, Hut R, Selker J. 2014. The Trans-African Hydro-Meteorological Observatory (TAHMO). *WIREs Water*. **1**: 341-348.

Wang W, Xie P, Yoo S.-H, Xue Y, Kumar A, Wu X. 2011. An assessment of the surface climate in the NCEP climate forecast system reanalysis. *Clim. Dyn*. **37**:1601–1620.

Wang Y, Zhang Y, Fu Y F, Li R, Yang Y J. 2015. A climatological comparison of column-integrated water vapor for the third-generation reanalysis datasets. Science China: Earth Sciences. doi: 10.1007/s11430-015-5183-6

Washington R, Kay G, Harrison M, Conway D, Black E, Challinor A, Grimes D, Jones R, Morse A, Todd M. 2006. African Climate Change: Taking the Shorter Route. *Bull. Am. Meteorol. Soc.* **87**: 1355-1366.

Worqlul, A.W., Maathuis, B., Adem, A.A., Demissie, S.S., Langan, S., & Steenhuis, T.S. (2014). Comparison of rainfall estimations by TRMM 3B42, MPEG and CFSR with ground-observed data for the Lake Tana basin in Ethiopia. *Hydrol. Earth Syst. Sci*. **18**: 4871–4881.

Zengyun H, Zhang C, Hu Q, Tian H. 2014. Temperature Changes in Central Asia from 1979 to 2011 Based on Multiple Datasets. *J.Clim.* **27**: 1143–1167.

Zhang Q, Kornich H, Holmgren K. 2013. How well do reanalyses represent the southern African precipitation? *Clim. Dynamics.* **40**: 951- 962. doi 10.1007/s00382-012-1423-z.

Table 1: Overview of rain gauge stations and corresponding grid points with their annual rainfall totals

Rain gauge station	Latitude	Longitude	Elevation (m)	Period of rainfall record	Mean annual rainfall rain gauge	Mean annual rainfall CFSR	Mean annual rainfall ERA Interim
Longone Birni	11.78	15.10	300	1979-1996	517.06	672.96	536.61
Mandalia	11.73	15.25	300	1979-1996	529.76	686.78	536.61
Massenya	11.40	16.17	328	1979-2001	607.45	711.59	659.59
Bailli	10.52	16.44	330	1979-2003	726.25	891.74	750.06
Bousso	10.48	16.72	336	1979-2001	743.59	845.07	789.36
Maroua	10.45	14.32	384	1979-2003	811.46	688.86	632.23
Mouvouday	10.41	14.85	336	1979-1994	657.93	759.99	721.68
Yagoua	10.35	15.25	325	1979-1997	747.78	897.69	742.78
Bongor	10.27	15.40	328	1979-1999	707.94	914.67	810.27
Kalfou	10.25	14.95	340	1979-1996	692.06	919.83	732.59
Dana	10.23	15.30	310	1979-1995	680.00	914.67	724.18
Doukoula	10.12	14.98	340	1979-1996	789.18	919.83	909.75
Deressia	9.75	16.17	344	1979-1998	757.84	891.77	943.26
Lai	9.40	16.30	358	1979-2002	980.13	890.98	873.26
Kello	9.32	15.80	378	1979-2003	909.00	945.64	971.14
Guidari CF	9.27	16.67	369	1979-2001	929.18	927.79	1007.78
Donomanga	9.23	16.92	370	1979-2001	925.45	916.96	1035.84
Delli	8.72	15.87	427	1979-2002	1027.21	1015.72	971.14
Moundou	8.57	16.08	410	1979-2003	1043.57	1015.72	1033.11
Donia	8.30	16.42	414	1979-2001	1000.18	1149.91	1125.16
Pandzangue	8.10	15.82	345	1979-1999	1154.65	1150.89	1111.67
Bekao	7.92	16.07	528	1979-2002	1150.40	1150.89	1052.18
Touboro	7.77	15.37	1430	1979-1995	1206.59	1102.62	1256.88
Baibokoum	7.73	15.68	1323	1979-1999	1090.50	1489.02	1163.57
Ngaoundere	7.35	13.56	1113	1979-2003	1420.61	2584.17	1246.67

Lat. and Long.: Latitude and Longitude in degrees respectively, annual rainfall in mm. (Stations from top to bottom are in descending order of latitude from north to south)

Table 2: Statistical performance of reanalysis datasets

Rain gauge station	CFSR					ERA Interim				
	R	R^2	Bias	MAE	NSE	R	R^2	Bias	MAE	NSE
Longone Birni	0.76	0.57	1.30	12.99	0.18	0.80	0.64	1.04	1.63	0.60
Mandalia	0.72	0.52	1.26	11.58	-0.15	0.80	0.64	0.95	2.46	0.60
Massenya	0.75	0.57	1.14	7.36	0.11	0.78	0.62	1.04	1.92	0.56
Bailli	0.81	0.65	1.20	12.12	0.45	0.70	0.49	1.04	2.31	0.41
Bousso	0.82	0.67	1.14	8.46	0.51	0.78	0.61	1.06	3.81	0.55
Maroua	0.70	0.49	0.85	10.22	0.35	0.64	0.41	0.71	19.35	0.36
Mouvouday	0.85	0.72	1.13	7.84	0.58	0.78	0.60	1.03	1.98	0.58
Yagoua	0.81	0.66	1.20	12.49	0.35	0.79	0.62	0.99	0.42	0.60
Bongor	0.81	0.66	1.26	15.60	0.38	0.79	0.63	1.12	7.48	0.56
Kalfou	0.78	0.61	1.35	20.04	0.17	0.77	0.59	1.00	0.19	0.52
Dana	0.63	0.40	1.33	18.83	-0.20	0.55	0.30	1.07	3.82	0.11
Doukoula	0.83	0.69	1.17	10.89	0.48	0.77	0.59	0.99	10.05	0.51
Deressia	0.72	0.52	1.12	7.84	0.22	0.65	0.42	1.21	13.59	0.16
Lai	0.76	0.58	0.91	7.43	0.52	0.74	0.55	0.89	8.91	0.53
Kello	0.82	0.68	1.02	1.19	0.61	0.72	0.51	1.07	5.09	0.47
Guidari CF	0.81	0.65	1.00	0.12	0.56	0.81	0.66	1.08	6.55	0.63
Donomanga	0.79	0.63	0.99	0.71	0.52	0.82	0.67	1.12	9.20	0.61
Delli	0.68	0.46	0.97	2.28	0.29	0.71	0.50	0.93	6.04	0.48
Moundou	0.81	0.66	0.99	0.96	0.57	0.75	0.57	1.01	0.49	0.54
Donia	0.80	0.64	1.24	18.70	0.34	0.78	0.62	1.22	16.64	0.50
Pandzangue	0.81	0.66	1.07	6.69	0.59	0.79	0.62	0.97	3.34	0.62
Bekao	0.81	0.65	1.00	0.21	0.55	0.78	0.60	0.91	8.44	0.59
Touboro	0.78	0.61	1.37	33.21	0.20	0.73	0.53	1.15	13.86	0.47
Baibokoum	0.75	0.56	0.91	8.69	0.52	0.72	0.52	0.95	5.03	0.51
Ngaoundere	0.81	0.66	1.82	96.96	-1.15	0.76	0.58	0.89	13.24	0.52

CFSR: Climate Forecasting System Reanalysis; R: correlation coefficient; R^2: coefficient of determination; MAE: mean absolute error; NSE: Nash Sutcliff Efficiency

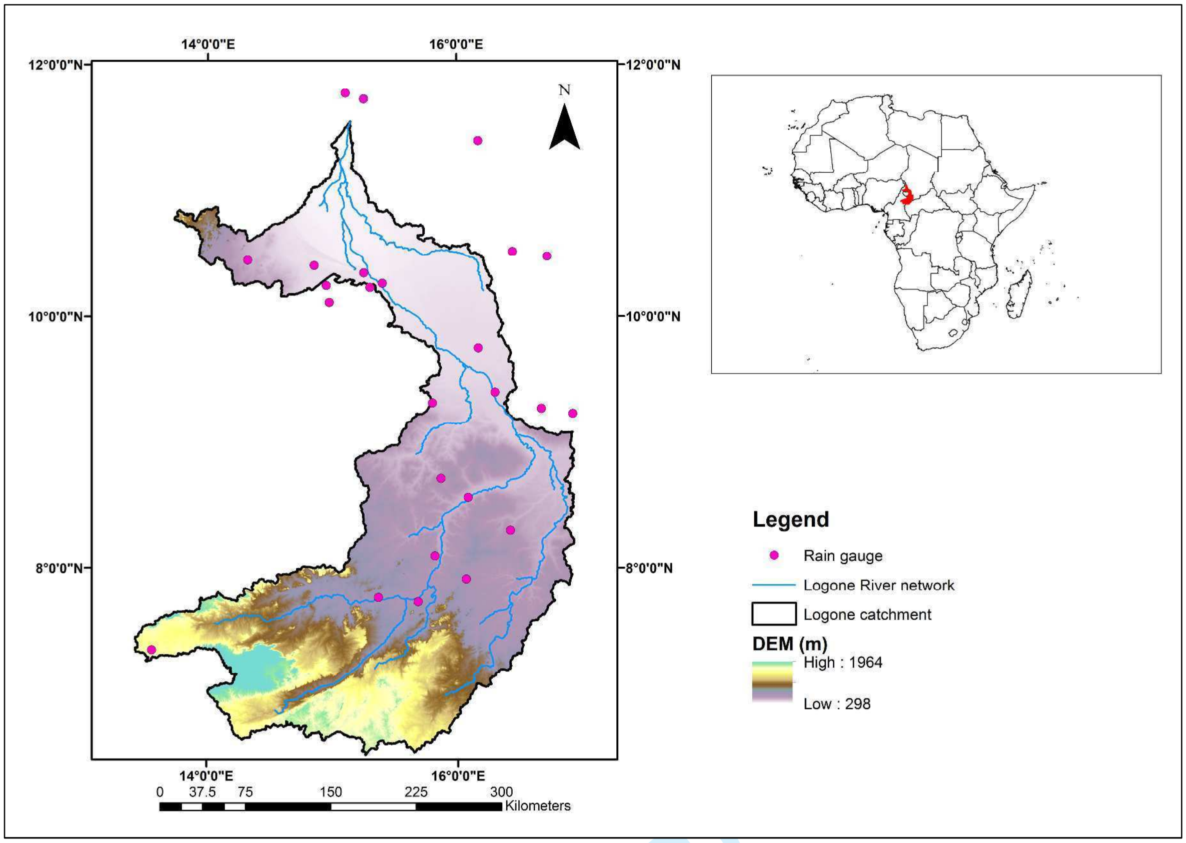

Figure 1: Map of study area (DEM: Digital Elevation Model)

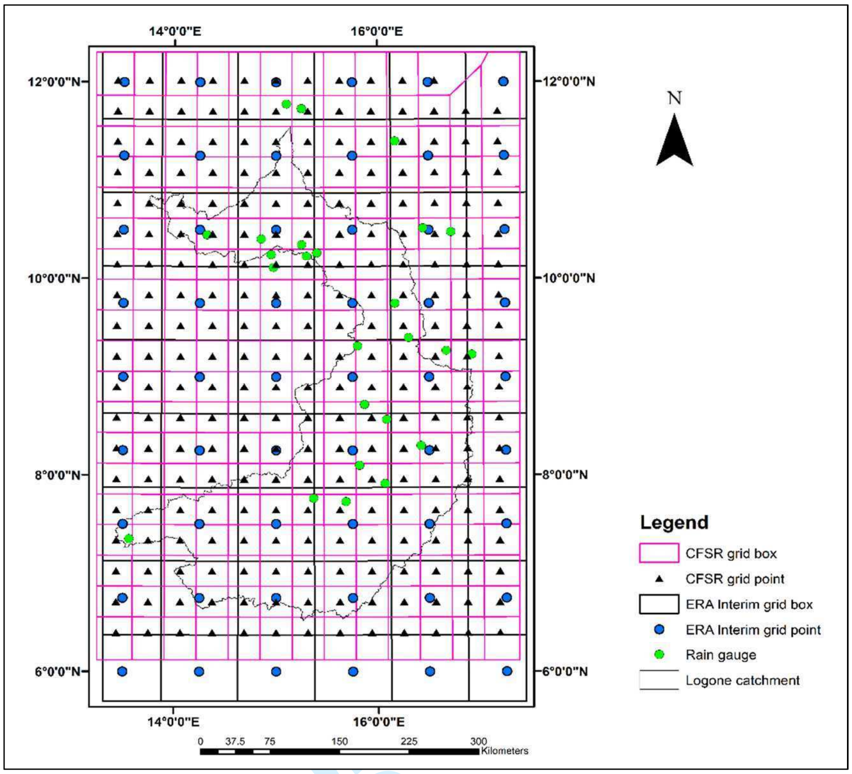

Figure2: Illustration of the grid boxes used in the assignment of a reanalysis grid point for comparison with a given rain gauge station. CFSR and ERA Interim grid boxes are represented by pink and black colours respectively, grid points are represented by black triangular and blue points and a rain gauge station is represented by a green point located inside either of the grid boxes.

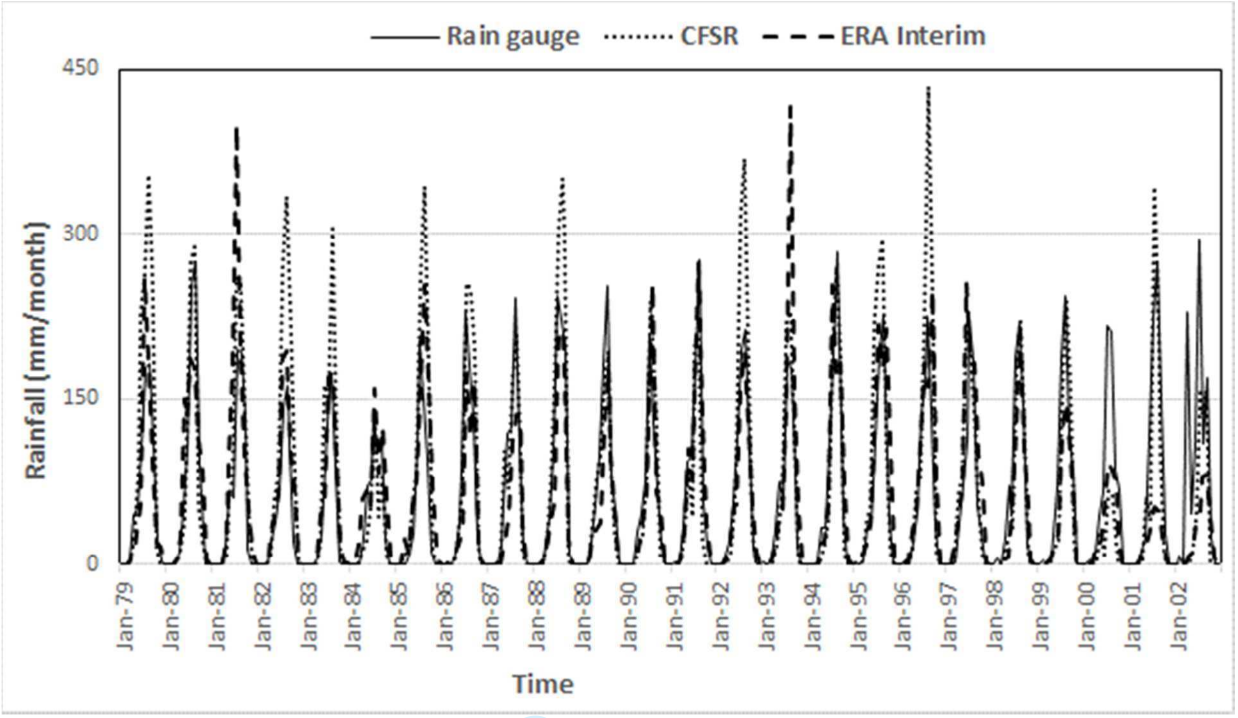

Figure 3(a): Average monthly rainfall time series for semi-arid area obtained by averaging monthly rainfall for all stations located in the semi-arid area.

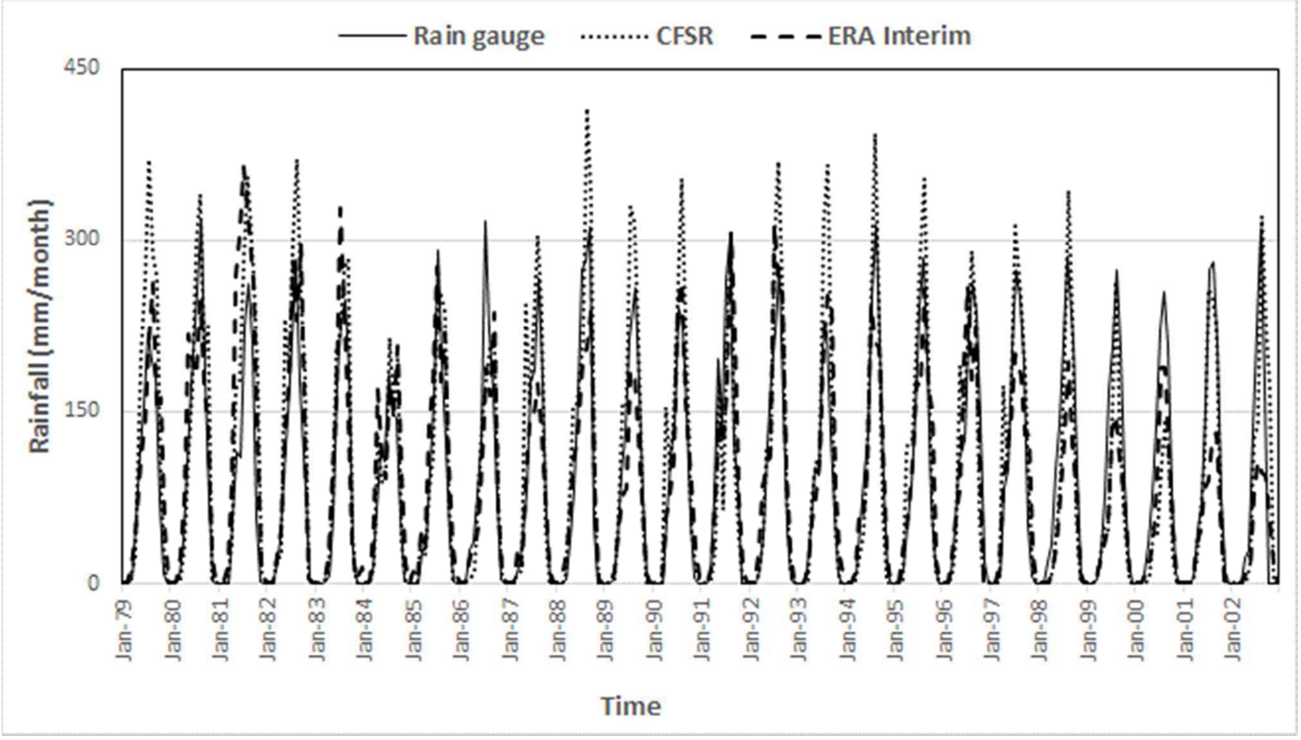

Figure 3(b): Average monthly rainfall time series Sudano obtained by averaging monthly rainfall for all stations located in the Sudano area.

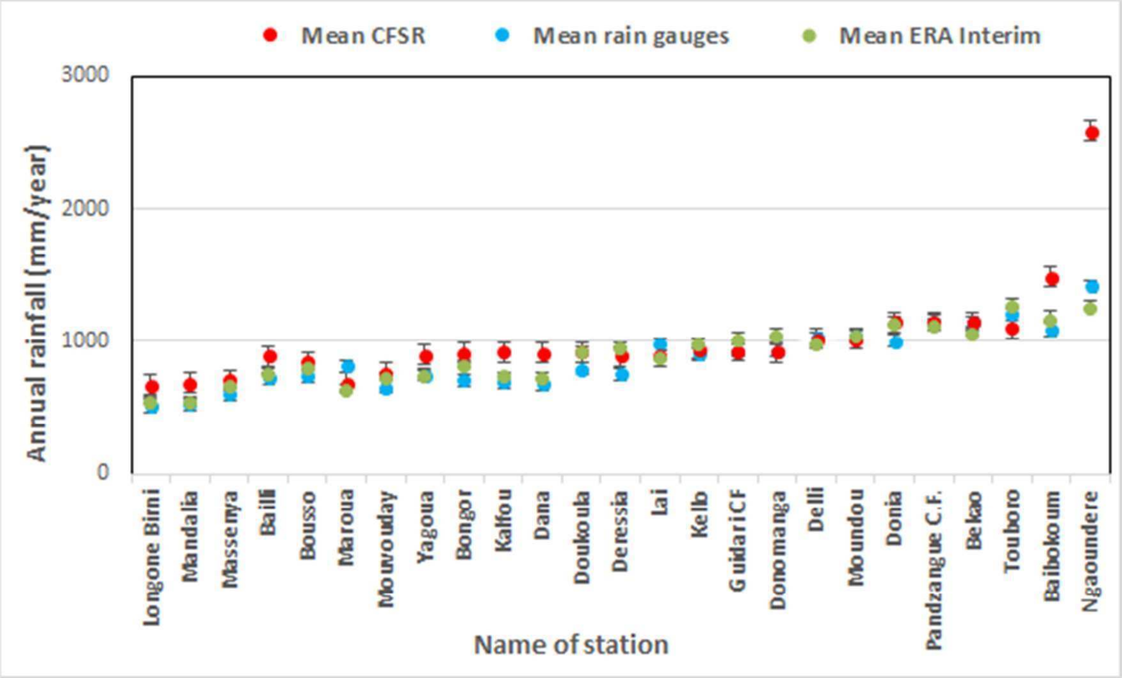

Figure 4a: Mean annual rainfall with error bars for all stations (1979-2002). (Stations from left to right are in descending order of latitude from north to south)

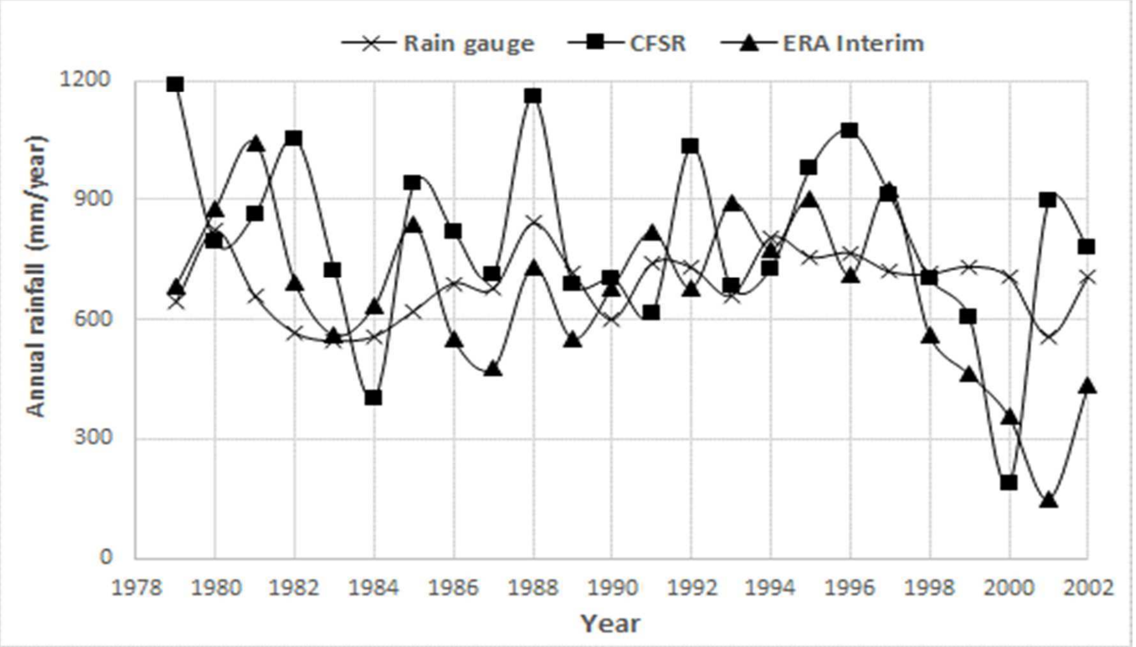

Figure 4b: Mean annual rainfall for the semi-arid (northern) part of the catchment obtained by averaging mean annual rainfall for all the stations located in the semi-arid area

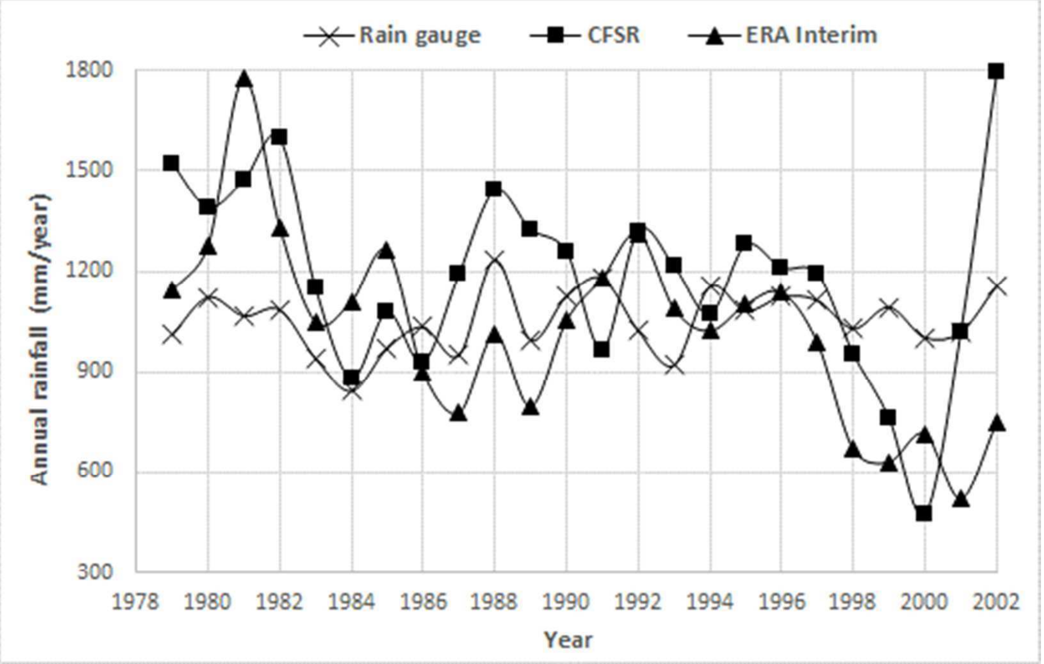

Figure 4c: Mean annual rainfall variation for the Sudano (southern) part of the catchment obtained by averaging mean annual rainfall for all the stations located in the Sudano area

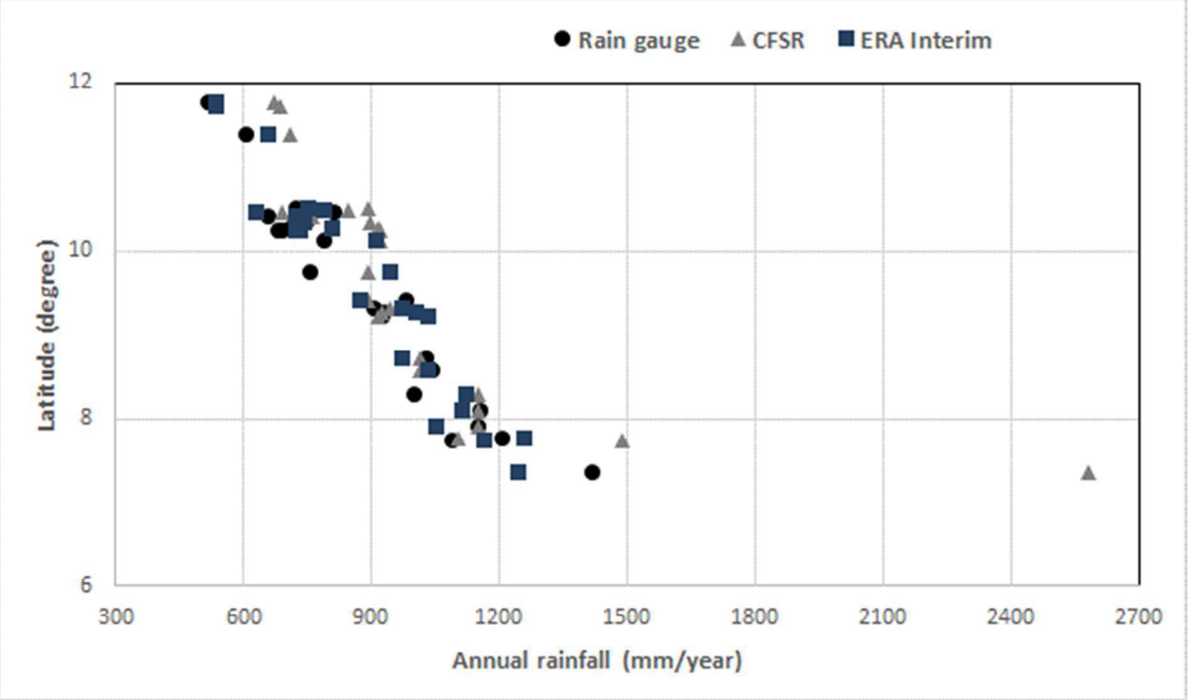

Figure 5: Variation of rainfall with latitude in the Logone catchment

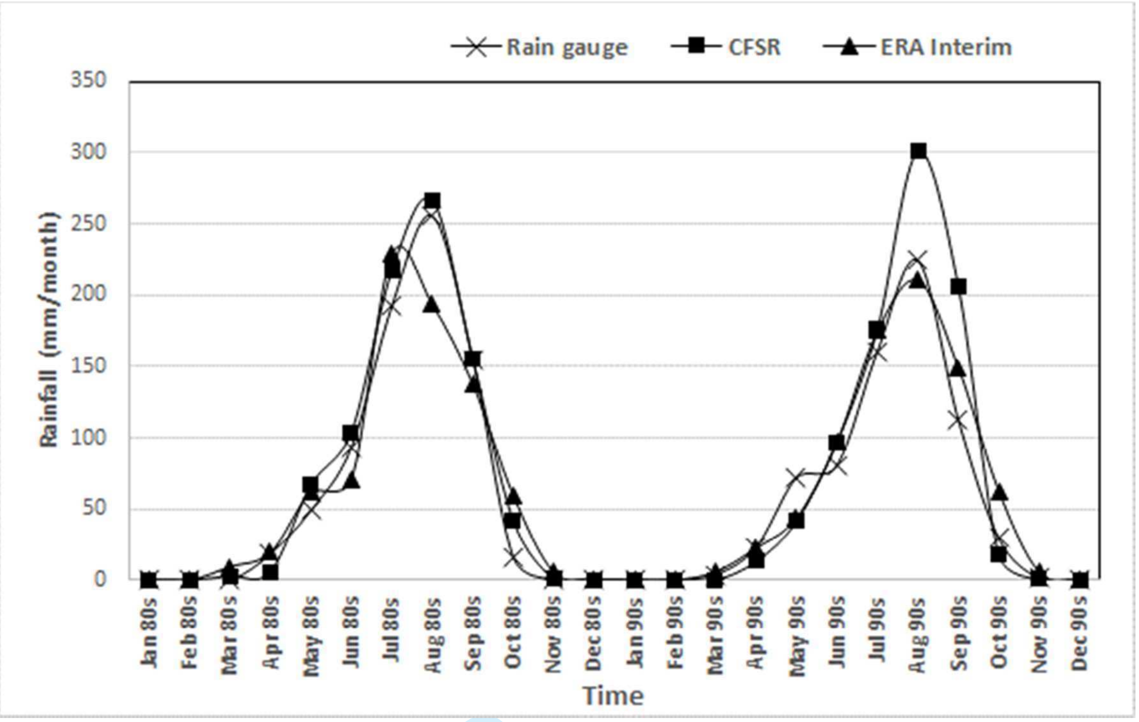

Figure 6(a): Monthly decadal rainfall variation Bailli (1980-1999) located in the northern part of the catchment.

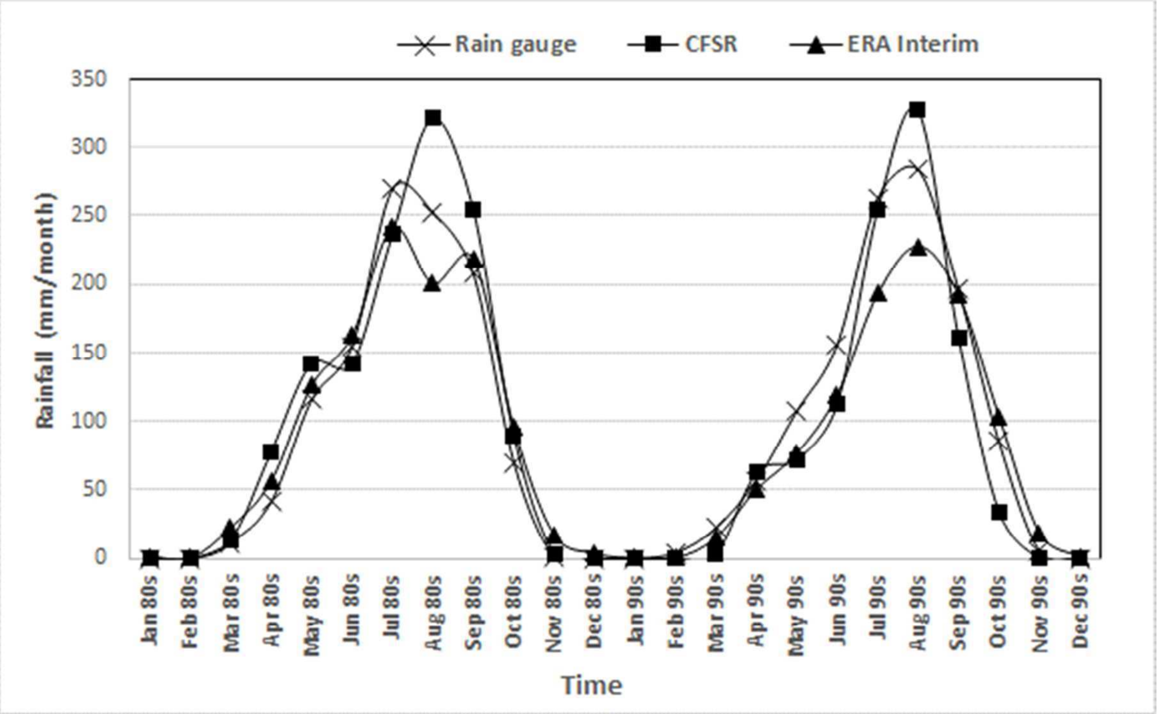

Figure 6(b): Decadal variation of rainfall for Bekao (1980-1999) located in the southern part of the catchment