



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/101997/>

Version: Accepted Version

Article:

Nkiaka, E, Nawaz, NR and Lovett, JC (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>

© 2016 Royal Meteorological Society. 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. doi: 10.1002/met.1600, 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 Self-Archiving. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

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

E. Nkiaka¹, N.R. Nawaz¹ and J.C. Lovett¹

¹School 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).

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

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

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

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

73

74 **2) Materials and methods**

75 **2.1) Study area**

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

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

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

97

98 **2.2) Data Sources**

99 **2.2.1) Rain gauge data**

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

110

111 **2.2.2) Reanalysis datasets**

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

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

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

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

137

138 **2.3) Method for comparison**

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

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

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

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

165

166 **3) Results**

167 **3.1) Monthly rainfall variation**

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

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

181

182 **3.2) Annual rainfall**

183 Annual gauge rainfall varies between 500-1500mm/year, ERA Interim varies between 500 –
184 1300mm/year while CFSR rainfall varies between 600 – 2600mm/year and consequently has the
185 highest spread among the three datasets (Figure 4a and Table 1). Koutsouris et al. (2015)
186 observed a similar spread in CFSR precipitation estimates in the Kilombero Valley in Tanzania.
187 Both reanalysis products overestimate annual rainfall in 68% of the stations in the catchment but
188 overestimation is greater for CFSR (27%) compared to 8% for ERA Interim.

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

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

205 The analysis also show that both products captured the spatial distribution of rainfall in the
206 Sudano and semi-arid areas of catchment fairly well. Annual rainfall generally ranges between
207 600-900mm/year and 900-1400mm/year in the semi-arid and Sudano areas respectively (Loth et
208 al., 2004); Figures 4(b) & (c). Fu et al. (2015) reported that CFSR and ERA Interim were able to
209 reproduce the observed spatial patterns of annual rainfall in Australia. Results also show that
210 annual rainfall estimates for CFSR grid box (grid point) located in Ngaoundere is significantly
211 higher than gauge data by more than 1000mm/year while ERA Interim underestimated rainfall in
212 that station by more than 150mm/year. Haensler, et al. (2013), reported that data from the
213 National Climatic Data Center (now National Centers for Environmental Information) strongly
214 overestimated precipitation in this station.

215 Furthermore, CFSR systematically underestimated annual rainfall during the period 1998 –
216 2000 across most stations in the catchment. Monteiro et al. (in press) observed a similar issue
217 during the same period while modelling the Tocantins catchment in Brazil and attributed it to an
218 artefact. ERA-Interim also systematically underestimated mean annual rainfall in the catchment
219 during the period 1997-2002 with annual average dropping below 300mm/year and 600mm/year
220 in the semi-arid and Sudano zones respectively in the year 2000 (Figures 4b&c).

221

222

223

224 3.3) Decadal monthly rainfall variability

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

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

240

241 3.4) Variation of rainfall with latitude

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

255

256 3.5) Statistical Analysis of monthly variation in rainfall

257 Although both reanalysis products generally show good correlation with monthly rain gauge
258 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.
259 (2014) in the Lake Tana Basin in Ethiopia. However, few large deviations still exist between the
260 reanalysis datasets and gauge data with variations in gauge rainfall measurements within the same
261 grid box (Table 2). For example: Mouvouday, Kalfou, Yagoua and Dana; Bailli and Bouso;
262 Pandzangue, Bekao and Moundou located inside the same ERA Interim grid boxes all produced
263 different values after statistical correlation with rainfall from the same grid points. A similar
264 situation was observed for Moundou and Delli; Dana and Bongor; Doukoula and Kalfou; Bekao
265 and Pandzangue located inside the same CFSR grid boxes.

266 The monthly bias values shown in Table 2 indicate that, both reanalysis products generally
267 overestimated rainfall in 68% of the stations with the highest overestimation recorded in
268 Ngaoundere for CFSR. Analysis also showed that ERA Interim had the lowest bias and MAE

269 with average values of 2% and 6.5mm/month respectively compared to 15% and 34mm/month for
270 CFSR.

271 Results of NSE as shown in Table 2 are in the range ($-1.15 \leq \text{NSE} \leq 0.56$) for CFSR and
272 ($0.11 \leq \text{NSE} \leq 0.63$) for ERA Interim, indicating that both forecasting models produced modest
273 results even though ERA Interim outperformed CFSR.

274

275 4) Discussion

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

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

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

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

313 rainfall estimates which are less than 1mm and these estimates when accumulated over a longer
314 time scale could have a significant influence on reanalysis estimates compared to gauge data.

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

323

324 **5) Conclusions**

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

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

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

347

348 **6) Acknowledgement**

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

354

355

356

357

358 7) References

- 359 Ardoin-Bardin S. 2004. Hydroclimate variability and impacts on water resources of large
360 hydrological catchments in the Sudanese–Sahelian area (in French). *Thesis (Ph.D).*
361 *University of Montpellier II.*
- 362 Agusti-Panareda, A, Balsamo G, Beljaars A. 2010. Impact of improved soil moisture on the
363 ECMWF precipitation forecast in West Africa. *Geophys. Res. Lett.*, **37**: L20808,
364 doi:10.1029/2010GL044748.
- 365 Baisch J. 2010. Data Shortage in Africa. *Desalination*. **252**: 101-112.
- 366 Blacutt LA, Herdies DL, de Gonçalves LGG, Vila DA, Andrade M. 2015. Precipitation
367 comparison for the CFSR, MERRA, TRMM3B42 and combined scheme datasets in Bolivia.
368 *Atmos. Res.* **163**:117 – 131.
- 369 Boyer J, Dieulin C, Rouche N, Cres A, Servat E, Paturel J. 2006. SIEREM: An Environmental
370 Information System for Water Resources, Vol. IAHS Publication 308. IAHS Press:
371 Wallingford, United Kingdom
- 372 Buytaert W, Friesen J, Liebe J, Ludwig R. 2012. Assessment and Management of Water
373 Resources in Developing, Semi-arid and Arid Regions. *Water Resour. Manag.* **26**: 841-844.
- 374 Candela L, Elorza FJ, Tamoh K, Jiménez-Martínez J, Aureli A. (2014) Groundwater modelling
375 with limited data sets: the Chari– Logone area (Lake Chad Basin, Chad). *Hydrol. Processes*,
376 **28**: 3714-3727.
- 377 Dee DP, Uppala SM, Simmons AJ, Berrisford P, et al. 2011. The ERA-Interim ‘ reanalysis:
378 configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**:
379 553–597, doi: 10.1002/qj.828.
- 380 Di Giuseppe F, Molteni F, Dutra,E. 2013. Real-time correction of ERA-Interim monthly rainfall,
381 *Geophys. Res. Lett.* **40**: 3750– 3755. doi:10.1002/grl.50670.
- 382 Dile, Y.T. and Srinivasan, R. (2014). Evaluation of CFSR Climate Data for Hydrologic Prediction
383 in Data-Scarce Watersheds: An Application in the Blue Nile River Basin. *J. Am. Water*
384 *Resour. Assoc.* doi: 10.1111/jawr.12182.
- 385 Diro GT, Grimes DIF, Black E, O’Neill A, Pardo-Iguzquiza E. 2009. Evaluation of reanalysis
386 rainfall estimates over Ethiopia. *Int. J. Climatol.* **29**(1): 67–78.
- 387 Fu G, Charles SP, Timbal B, Jovanovic B, Ouyang F. 2015. Comparison of NCEP-NCAR and
388 ERA-Interim over Australia. *Int. J. Climatol.* doi:10.1002/joc.4499
- 389 Fuka DR, MacAllister CA, Degaetano AT, Easton ZM. 2013. Using the Climate Forecast System
390 Reanalysis dataset to improve weather input data for watershed models. *Hydrol. Processes*,
391 **28**: 5613-5623.
- 392 Haensler A, Saeed F, Jacob D. 2013. Assessing the robustness of projected precipitation changes
393 over central Africa on the basis of a multitude of global and regional climate projections.
394 *Clim. Change.* **121**(2): 349-363.
- 395 Koutsouris AJ, Chen D, and Lyon SW. 2015. Comparing global precipitation data sets in eastern
396 Africa: a case study of Kilombero Valley, Tanzania. *Int. J. Climatol.* doi: 10.1002/joc.4476.
- 397 Krogh SA, Pomeroy JW, Mcphee J. 2015. Physically based mountain hydrological modelling
398 using reanalysis data in Patagonia. *J. Hydrometeorol.* **16**: 172–193
- 399 Le Coz ML, Delclaux F, Genthon P, Favreau G. 2009. Assessment of Digital Elevation Model
400 (DEM) aggregation methods for hydrological modelling: Lake Chad Basin, Africa. *Comput.*
401 *Geosci.* **35**(8): 1661-1670.

- 402 De Leeuw J, Methven J, Blackburn M. 2015. Evaluation of ERA-Interim reanalysis precipitation
 403 products using England and Wales observations. *Q. J. R. Meteorol. Soc.* **141**: 798–806,
 404 doi:10.1002/qj.2395.
- 405 Loth P. (Editor). 2004. The Return of the Water: Restoring the Waza Logone Floodplain in
 406 Cameroon. IUCN, Gland, Switzerland and Cambridge, UK.
- 407 Molua EL, Lambi CM. 2006. Climate, hydrology and water resources in Cameroon. In: CEEPA
 408 discussion paper no. 33, CEEPA, Pretoria.
- 409 Maidment RI, Grimes DIF, Allan RP, Greatrex H, Rojas O, Leo O. 2013. Evaluation of satellite-
 410 based and model re-analysis rainfall estimates for Uganda. *Meteorol. Appl.* **20**: 308 - 317.
- 411 Maidment RI, Allan IRP, Black E. 2015. Recent observed and simulated changes in precipitation
 412 over Africa, *Geophys. Res. Lett.* **42**: 8155–8164, doi: 10.1002/2015GL065765.
- 413 Monteiro JAF, Strauch M, Srinivasan R, Abbaspour K, Gücker B. Accuracy of grid precipitation
 414 data for Brazil: Application in river discharge modelling of the Tocantins catchment. *Hydrol.*
 415 *Processes*. doi: 10.1002/hyp.10708.
- 416 Morse A, Caminade C, Tompkins A, McIntyre KM, 2013. *The QWeCI project (Quantifying*
 417 *Weather and Climate Impacts on Health in Developing Countries) - Final report.* p. 60.
 418 (Unpublished).
- 419 Nkiaka E, Rizwan NR, Lovett JC. 2016. Using Self-Organizing Maps to infill missing data in
 420 hydro-meteorological time series from the Logone catchment, Lake Chad basin. *Environ*
 421 *Monit Assess.* **188**. doi: 10.1007/s10661-016-5385-1.
- 422 Rienecker MM, Suarez MJ, Gelaro R, Todling R, et al. 2011. MERRA: NASA’s Modern-Era
 423 Retrospective Analysis for Research and Applications. *J. Clim.* **24**: 3624–3648
- 424 Saha S, Moorthi S, Wu X, Wang J, Nadiga S, Tripp P, Behringer D, Hou YT, Chuang HY, Iredell
 425 M. 2014. The NCEP climate forecast system version 2. *J. Clim.* **27**: 2185 – 2208
- 426 Sharifi E, Steinacker R, and Saghafian B. 2016. Assessment of GPM-IMERG and Other
 427 Precipitation Products against Gauge Data under Different Topographic and Climatic
 428 Conditions in Iran: Preliminary Results. *Remote Sens.* **8**: 135. doi:10.3390/rs8020135
- 429 Van de Giesen N, Hut R, Selker J. 2014. The Trans-African Hydro-Meteorological Observatory
 430 (TAHMO). *WIREs Water.* **1**: 341-348.
- 431 Wang W, Xie P, Yoo S.-H, Xue Y, Kumar A, Wu X. 2011. An assessment of the surface climate
 432 in the NCEP climate forecast system reanalysis. *Clim. Dyn.* **37**:1601–1620.
- 433 Wang Y, Zhang Y, Fu Y F, Li R, Yang Y J. 2015. A climatological comparison of column-
 434 integrated water vapor for the third-generation reanalysis datasets. *Science China: Earth*
 435 *Sciences*. doi: 10.1007/s11430-015-5183-6
- 436 Washington R, Kay G, Harrison M, Conway D, Black E, Challinor A, Grimes D, Jones R, Morse
 437 A, Todd M. 2006. African Climate Change: Taking the Shorter Route. *Bull. Am. Meteorol.*
 438 *Soc.* **87**: 1355-1366.
- 439 Worqlul, A.W., Maathuis, B., Adem, A.A., Demissie, S.S., Langan, S., & Steenhuis, T.S. (2014).
 440 Comparison of rainfall estimations by TRMM 3B42, MPEM and CFSR with ground-
 441 observed data for the Lake Tana basin in Ethiopia. *Hydrol. Earth Syst. Sci.* **18**: 4871–4881.
- 442 Zengyun H, Zhang C, Hu Q, Tian H. 2014. Temperature Changes in Central Asia from 1979 to
 443 2011 Based on Multiple Datasets. *J. Clim.* **27**: 1143–1167.
- 444 Zhang Q, Kornich H, Holmgren K. 2013. How well do reanalyses represent the southern African
 445 precipitation? *Clim. Dynamics.* **40**: 951- 962. doi 10.1007/s00382-012-1423-z.
- 446

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
Bouso	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
Bouso	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; R²: coefficient of determination; MAE: mean absolute error; NSE: Nash Sutcliff Efficiency

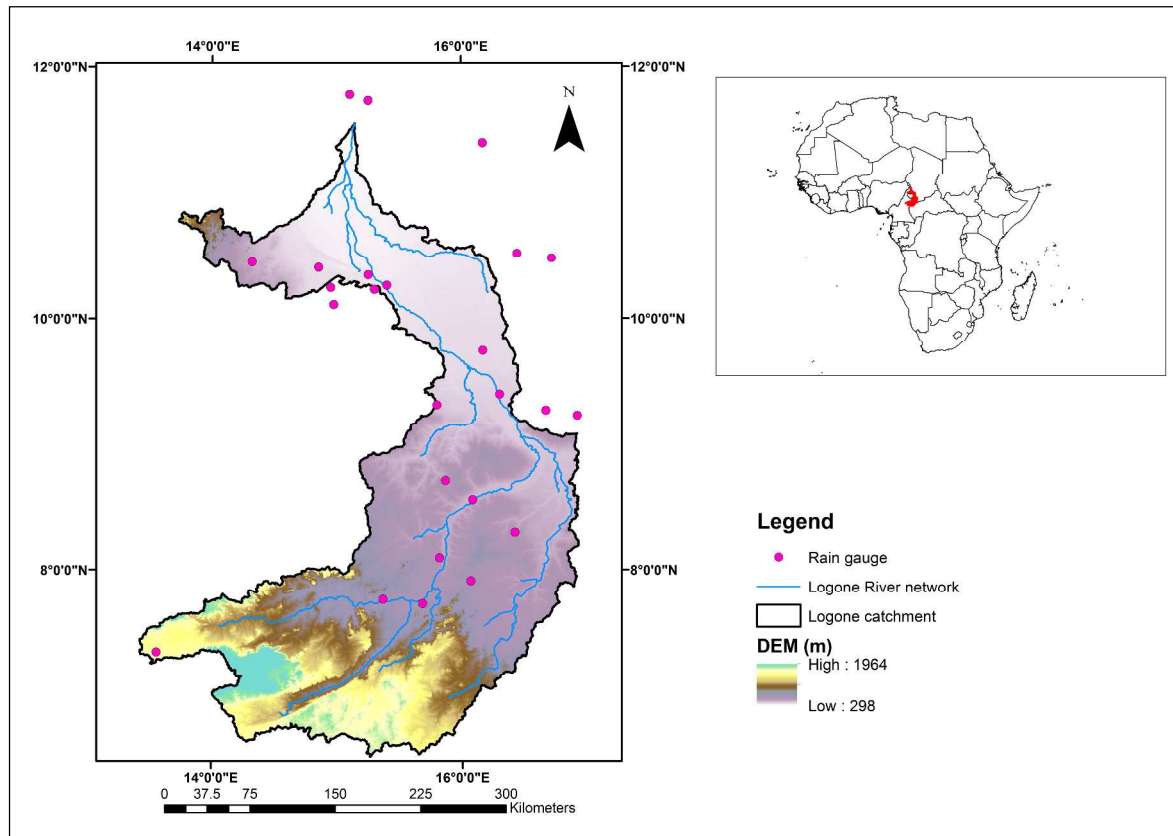


Figure 1: Map of study area (DEM: Digital Elevation Model)

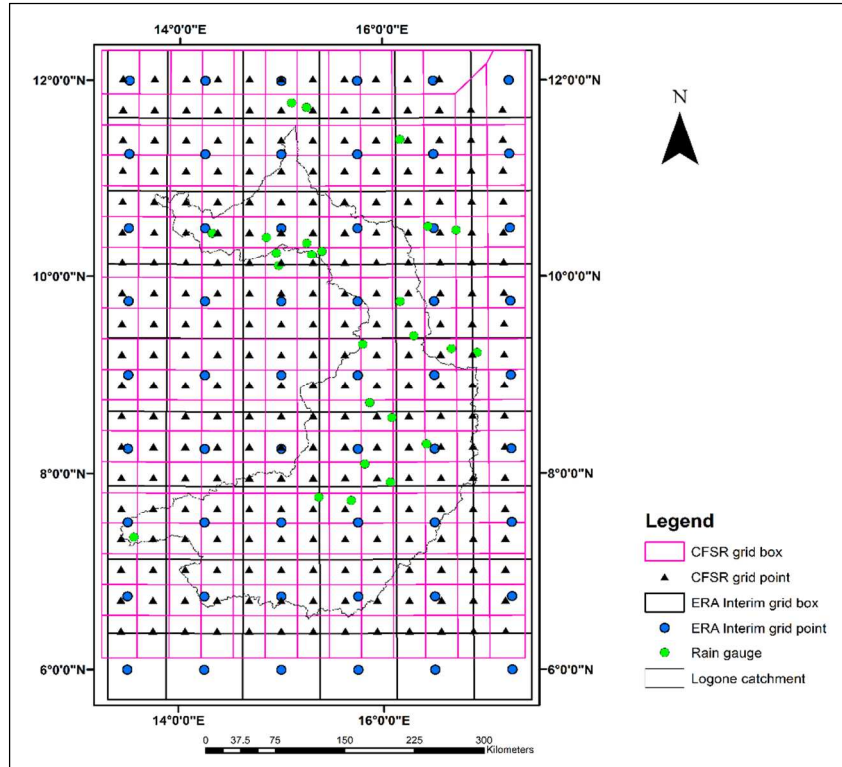


Figure2: Illustration of the grid boxes used in the assignment of a reanalysis grid point for comparison with a given rain gauge station. CFSR and ERA Interim grid boxes are represented by pink and black colours respectively, grid points are represented by black triangular and blue points and a rain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

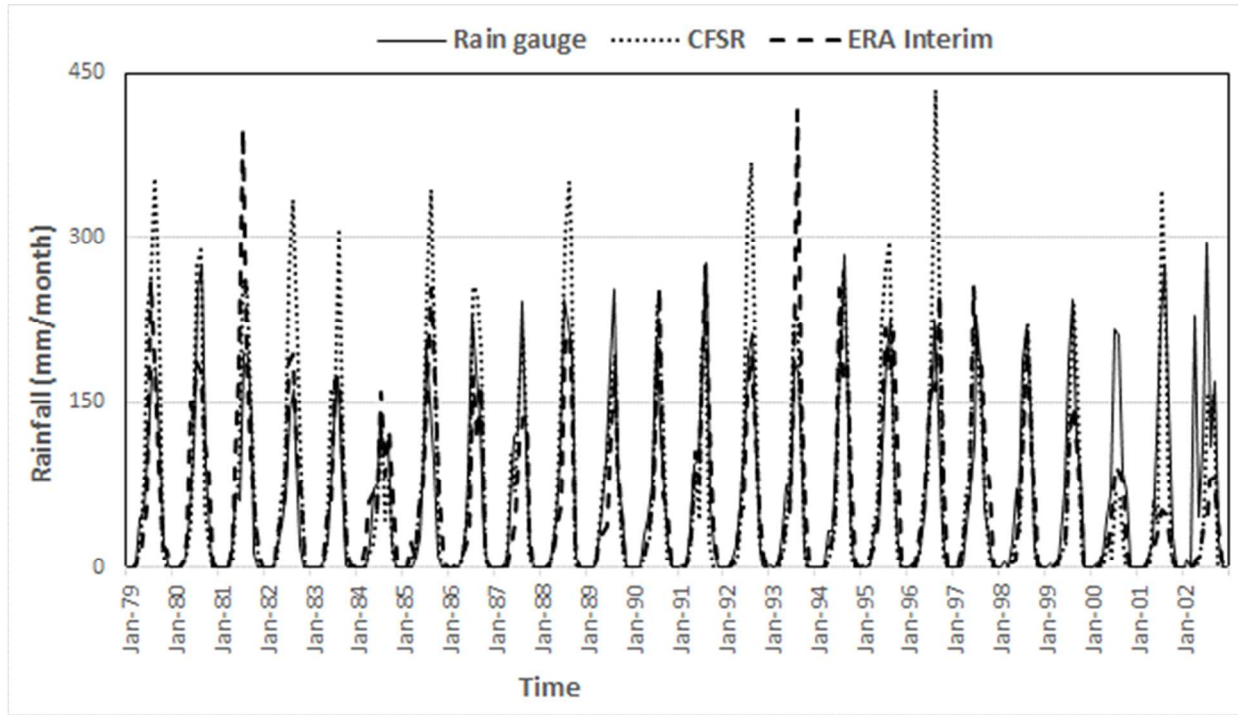


Figure 3(a): Average monthly rainfall time series for semi-arid area obtained by averaging monthly rainfall for all stations located in the semi-arid area.

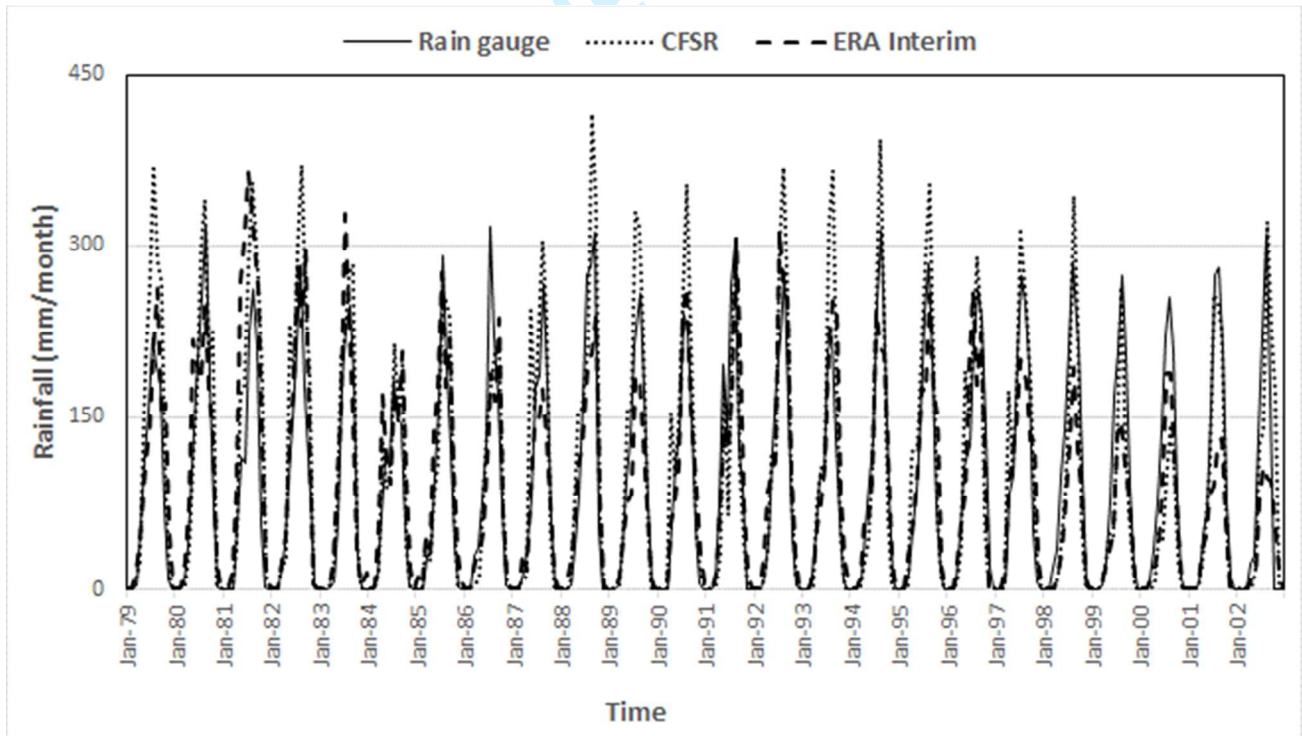


Figure 3(b): Average monthly rainfall time series for Sudano obtained by averaging monthly rainfall for all stations located in the Sudano area.

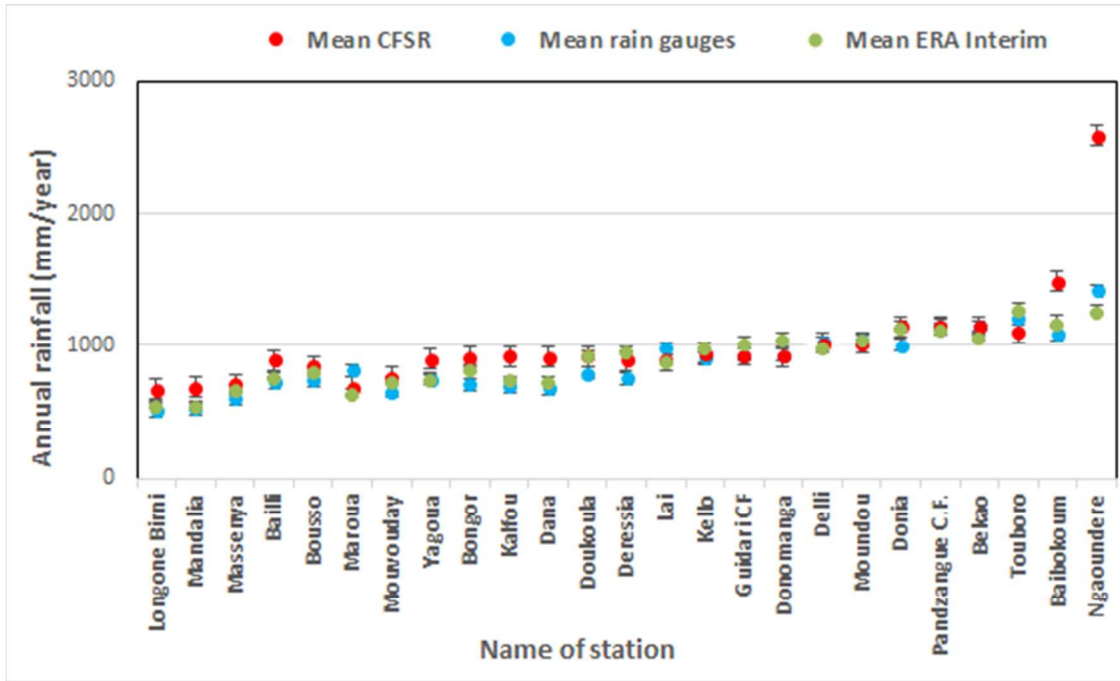


Figure 4a: Mean annual rainfall with error bars for all stations (1979-2002). (Stations from left to right are in descending order of latitude from north to south)

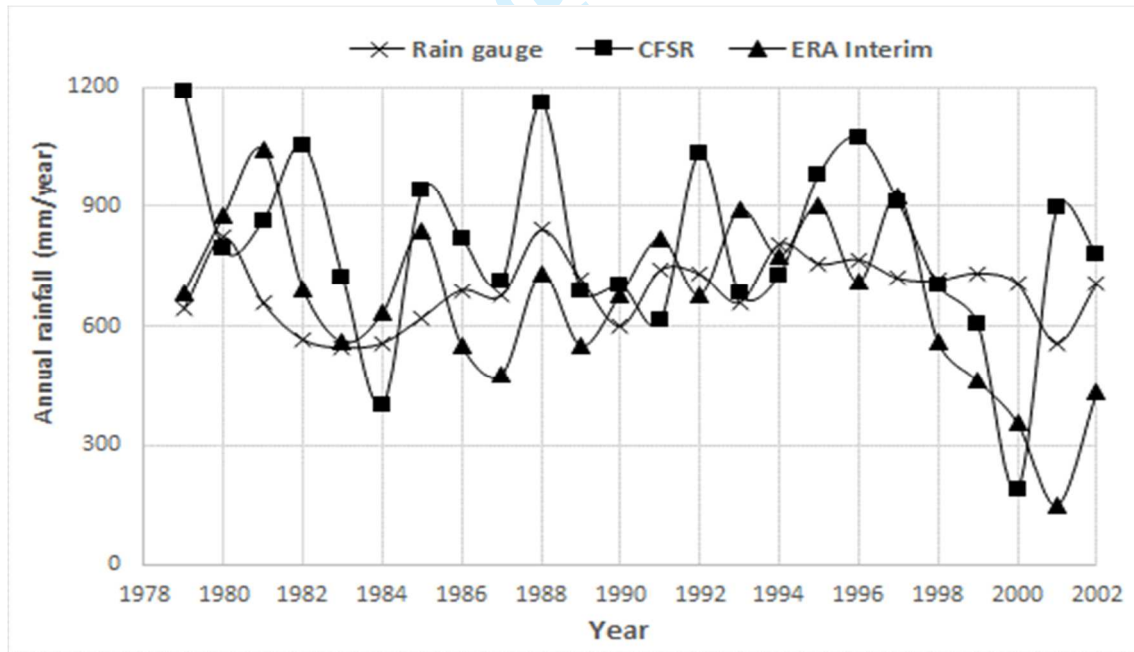


Figure 4b: Mean annual rainfall for the semi-arid (northern) part of the catchment obtained by averaging mean annual rainfall for all the stations located in the semi-arid area

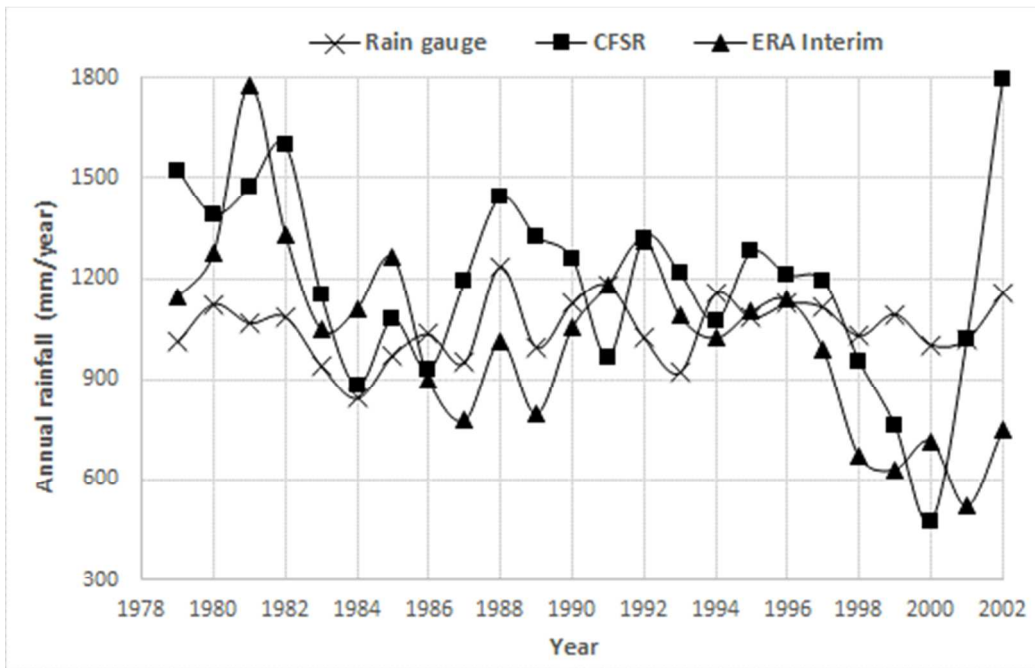


Figure 4c: Mean annual rainfall variation for the Sudano (southern) part of the catchment obtained by averaging mean annual rainfall for all the stations located in the Sudano area

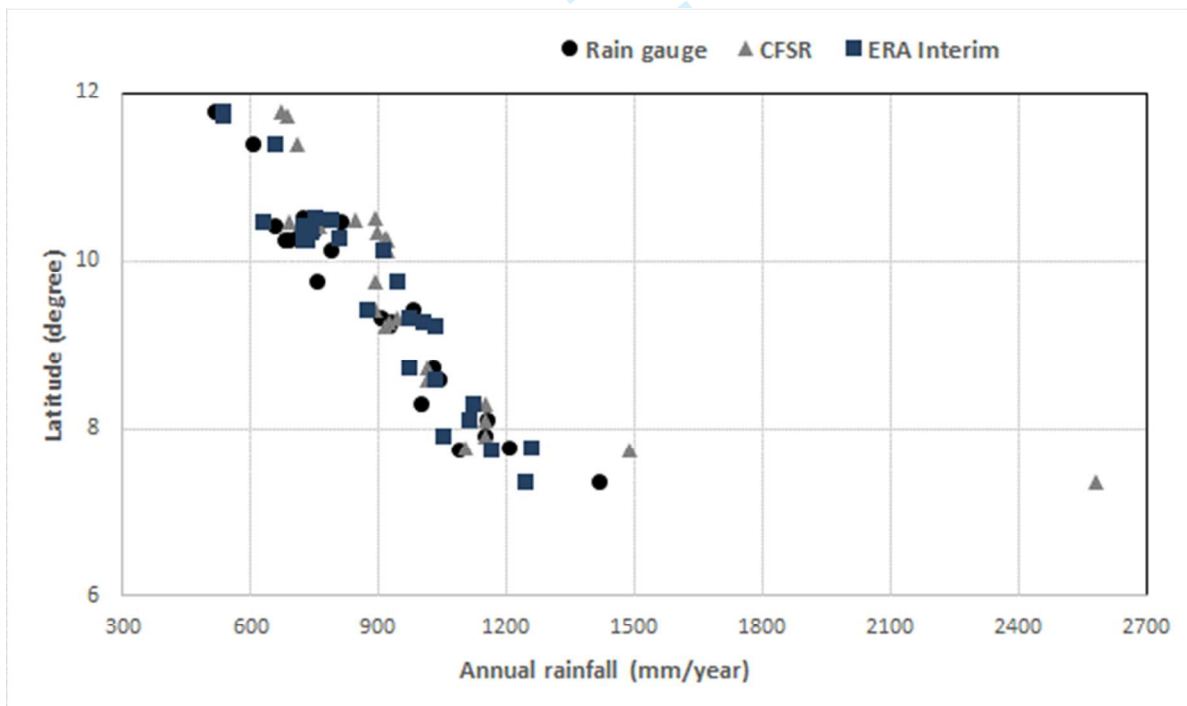


Figure 5: Variation of rainfall with latitude in the Logone catchment

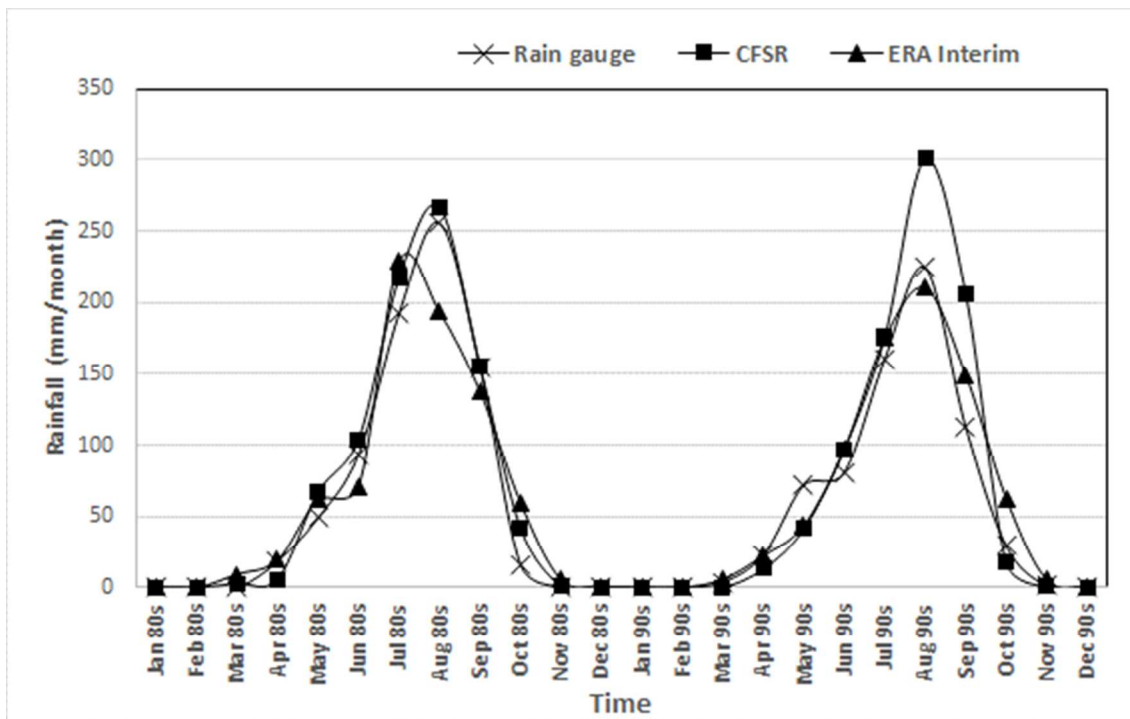


Figure 6(a): Monthly decadal rainfall variation Bailli (1980-1999) located in the northern part of the catchment.

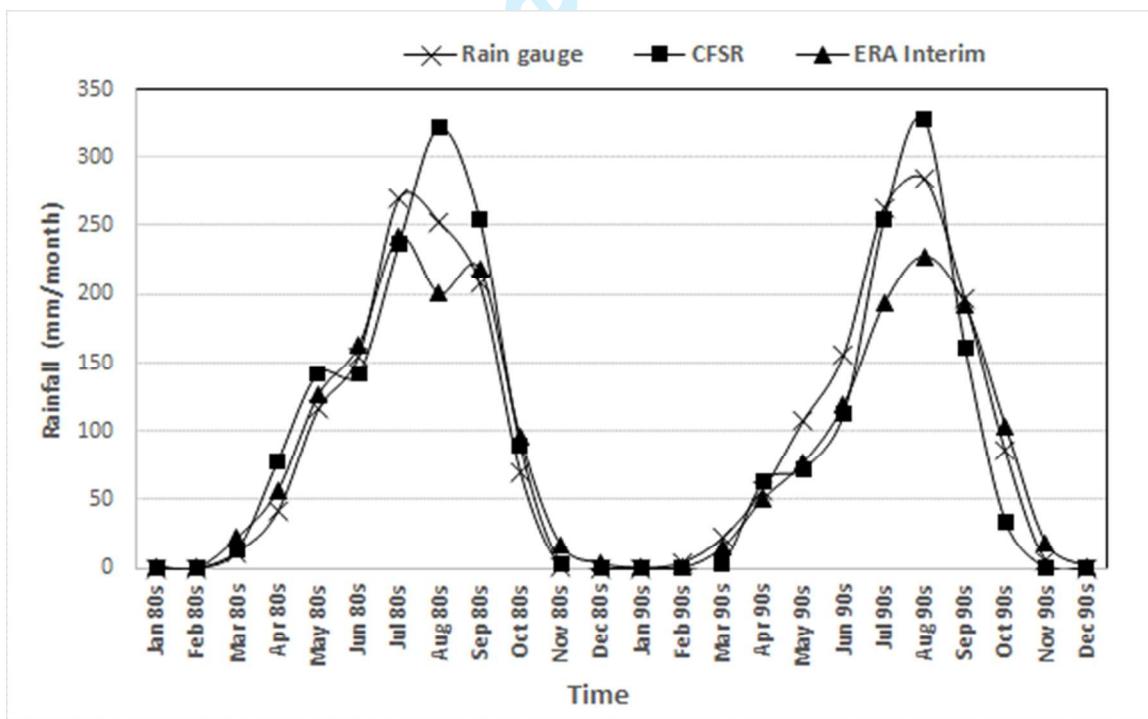


Figure 6(b): Decadal variation of rainfall for Bekao (1980-1999) located in the southern part of the catchment