# Trend and variability in a new, reconstructed streamflow dataset for West and Central Africa, and climatic interactions, 1950–2005

Moussa Sidibe, Bastien Dieppois, Gil Mahe, Jean-Emmanuel Paturel, Ernest Amoussou, Babatunde Anifowose, Damian Lawler

# Trend and variability in a new, reconstructed streamflow dataset for West and Central Africa, and climatic interactions, 1950 – 2005

Sidibe, M, Dieppois, B, Mahé, G, Paturel, J-E, Amoussou, E, Anifowose, B & Lawler, D

# Accepted Manuscript

Moussa Sidibe, Bastien Dieppois, Gil Mahé, Jean-Emmanuel Paturel, Ernest Amoussou, Babatunde Anifowose, Damian Lawler

Please cite this article as: Sidibe, M., Dieppois, B., Mahé, G., Paturel, J-E., Amoussou, E., Anifowose, B., Lawler, D., Trend and variability in a new, reconstructed streamflow dataset for West and Central Africa, and climatic interactions, 1950 – 2005, *Journal of Hydrology* (2018), doi: https://doi.org/10.1016/j.jhydrol.2018.04.024
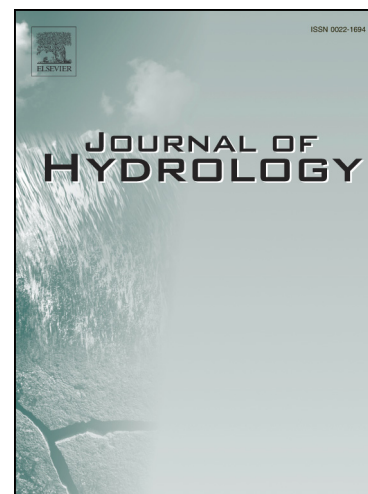
**Trend and variability in a new, reconstructed streamflow dataset for West and Central Africa, and climatic interactions, 1950 – 2005**

Moussa Sidibe[1], Bastien Dieppois[1, 2, 3], Gil Mahé[4], Jean-Emmanuel Paturel[4],
Ernest Amoussou[5], Babatunde Anifowose[6], Damian Lawler[1]

[1] Centre for Agroecology, Water and Resilience (CAWR), Coventry University, Coventry, UK

[2] Department of Oceanography, University of Cape Town, Cape Town, RSA

[3] School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK

[4] HydroSciences (HSM), IRD, Montpellier, France

[5] Département de Géographie et Aménagement du Territoire (DGAT), Université de Parakou, Parakou, Benin

[6] School of Energy, Construction & Environment, Coventry University, UK

*Abstract*

Over recent decades, regions of West and Central Africa have experienced different and significant changes in climatic patterns, which have significantly impacted hydrological regimes. Such impacts, however, are not fully understood at the regional scale, largely because of scarce hydroclimatic data. Therefore, the aim of this study is to (a) assemble a new, robust, reconstructed streamflow dataset of 152 gauging stations; (b) quantify changes in streamflow over 1950 – 2005 period, using these newly reconstructed datasets; (c) significantly reveal trends and variability in streamflow over West and Central Africa based on new reconstructions; and (d) assess the robustness of this dataset by comparing the results with those identified in key climatic drivers (*e.g.* precipitation and temperature) over the region. Gap filling methods applied to monthly time series (1950-2005) yielded robust results (median Kling-Gupta Efficiency >0.75). The study underlines a good agreement between precipitation and streamflow trends and reveals contrasts between western Africa (negative trends) and Central Africa (positive trends) in the 1950s and 1960s. Homogenous dry conditions of the 1970s and 1980s, characterized by reduced significant negative trends resulting from quasi-decadal modulations of the trend, are replaced by wetter conditions in the recent period (1993-2005). The effect of this rainfall recovery (which extends to West and Central Africa) on increased river flows are further amplified by land use change in some Sahelian basins. This is partially offset, however, by higher potential evapotranspiration rates over parts of Niger and Nigeria. Crucially, the new reconstructed

29  streamflow datasets presented here will be available for both the scientific community and water resource

30  managers.

31  **Keywords:** *West and Central Africa, streamflow trend and variability, hydroclimate variability, multi-*

32  *temporal trend identification, gap filling methods.*

33  **1.  INTRODUCTION**

34  Rainfall in Africa during the 20th century was characterized by decreasing annual trends over the continent

35  except regions in Cameroon, Sierra Leone and southern equatorial Africa (Hulme *et al*., 2001). Since 1950,

36  most of the extreme climatic conditions occurred in the Sahel region, which has experienced several

37  drought events from the end of the 1960s to the 1990s (Dai *et al.,* 2004; Lebel*,* 2003; Nicholson, 2013). For

38  the 1968−1997 period, rainfall in August in the West African Sahel showed a decrease of up to 37%

39  compared to the 1931−1960 period (Nicholson *et al.,* 2000). Rainfall patterns in the post-1990 period are,

40  however, less well documented, given data scarcity: this has led to controversial findings regarding the end

41  of Sahel drought. For example, Ozer *et al*. (2003) claimed that the Sahel drought ended in the 1990s,

42  whereas L´Hôte *et al.* (2002) suggested that the drought continued beyond the 1990s. These contradictions

43  partly reflect the significant changes in the spatial distribution of precipitation, which make findings highly

44  dependent on the specific region, and the years and even months considered. This underlines the need for

45  studies covering larger spatial scales. However, there is agreement on an increase in annual rainfall over the

46  West African Sahel since the 1990s (*e.g.* Ali and Lebel, 2009; Jury, 2013; Lebel and Ali, 2009; Mahé and

47  Paturel, 2009). See also Maidment *et al.* (2015), who described rainfall trends over Africa during the period

48  1983-2010, using different observational datasets and simulations from the current state-of-the-art global

49  climate models.

50  Interestingly, while rainfall variability has been investigated at the continental scale in Africa, its effects on

51  runoff regimes have mostly been investigated at catchment scales, using different statistical approaches and

52  hydrological models (*e.g.* Ibrahim *et al.,* 2015). This is mainly due to restricted data, and several factors

2

53 (*e.g.* Gyau-Boakye and Schultz, 1994) resulted in missing values in streamflow records. Such data

54 restrictions have limited attempts to systematically assess streamflow trend, variability and changes at the

55 regional scales. Descroix *et al.* (2009) reported a negative trend (more than 15% decrease) in streamflow

56 for the 1960-2010 period in Sudanian regions (receiving approximately 700 - 1300 mm yr-1 annual

57 rainfall) as a response to changes in rainfall regimes. Also, Mahé *et al.* (2013) found that a decrease in

58 annual rainfall of around 20% since 1970 has resulted in a streamflow decrease of up to 60% for some

59 rivers in West Africa (*e.g.* Niger and Senegal rivers). Amogu *et al.* (2010), in their attempt to regionalize

60 runoff evolution over western Africa (1950-2010), found a clear difference between the Sahelian zone

61 (where, curiously, runoff increases despite reduction in rainfall) and Sudanian and Guinean areas (where

62 runoff decreases logically with rainfall). While major rivers of West Africa show a distinct runoff decrease

63 since 1970, river flows in Central Africa do not show any long-term trend (Mahé *et al.,* 2013). However,

64 these results are restricted to a few long and gap-free time series, making it difficult to describe regional

65 streamflow variability.

66 Changes in the observational networks (*e.g.* station locations, gauge maintenance and data collection

67 methods) have limited attempts to study streamflow trends and variability at regional scales. Different gap

68 filling methods have been used (*e.g.* regression analysis, time series analysis, artificial neural network and

69 interpolation). Multiple imputations approaches, such as proposed by Rubin (1987), were recently

70 implemented to construct a complete rainfall-runoff database in Iran (Kalteh and Hjorth, 2009). More

71 complex methods such as artificial neural networks (Kim and Pachepsky, 2010) and random forest-based

72 approaches (Stekhoven and Bühlmann, 2012) have also been implemented for gap filling problems with

73 satisfactory results. Despite many hydrological data gap filling studies, few African examples exist. Most

74 African studies focus on gap-free stations (*e.g.* Nka *et al.,*2015) or reconstructions using linear interpolation

75 techniques. A decision support system based on length of data-gaps, seasons, climatic zones, model

76 performances and data availability has been provided by Gyau-Boakye and Schultz (1994), but such a

77 system would be difficult to implement at the regional scale due to substantial input data requirement and it

78 may result in: i) spatially non-homogenous reconstructions, and ii) non-statistically independent

79 reconstructions from climate variables. The development of regional climate models (RCM) also open new

80 prospects for hydrological data reconstruction. For instance, Moalafhi *et al.* (2017), testing such approaches

81 over the Limpopo basin, in southern Africa, found that dynamical downscaling of reanalysis products can

82 be very useful for semi-arid, data-scarce environments. However, important biases in RCM physics

83 combined with uncertainties in hydrological modeling could substantially impact the quality of streamflow

84 estimates. The present study aims at (1) providing a new, robust reconstructed streamflow dataset, using

85 only streamflow records as predictands, over West and Central Africa, and (2) using the new dataset,

86 together with gridded climatic data, to examine and assess flow changes and variability over the region

87 between 1950 and 2005. This paper is organized as follows. After introducing the study area and the

88 different datasets in section 2, we present the methods in section 3. In section 4.1, two gap filling methods

89 are used to generate a robust and complete streamflow dataset for West and Central Africa. Then, we

90 examine changes (abrupt and gradual) and variability in streamflow, and we compare these results to those

91 observed in climatic variables from section 4.2 to section 4.4. Results are interpreted, and their wider

92 implications discussed in Section 5.

93 **2. STUDY AREA AND DATASETS**

94 **2.1 Research Area**

95 The study area covers West and Central Africa, from -10°N to 25°N and -20°E to 30°E. The selection of

96 the research area was motivated by the key importance of climate change and variability in this region,

97 which also has a dense streamflow gauging network (Figure 1). Some records were short or incomplete,

98 mainly due to equipment failure, instrumentation maintenance issues and sometimes political instability.

99 Most hydrological studies in the region primarily refer to four eco-climatic zones, which are based on both

100 annual rainfall amounts and agricultural properties (FAO, 2004; Roudier *et al.,* 2014): (a) the Sahelian zone

101 (mean annual precipitation ranges between 250 and 500 mm yr-1), (b) the Sudano-sahelian zone (mean

4

102 annual precipitation ranges between 500 to 900 mm yr-1), (c) the Sudanian zone (mean annual precipitation

103 ranges from 900 to 1100 mm yr-1) and (d) the Guinean zone (mean annual precipitation exceeds 1100 mm

104 yr-1). However, more complex classifications based on seasonal to interannual variability of rainfall can be

105 found (Badr *et al.,* 2016; Hermann and Mohr, 2011; Janicot, 1992; L´Hôte *et al.,* 1996; Mahé *et al.,* 2001).

106 West African rainfall is primarily related to the displacement of the Intertropical Convergence Zone

107 (ITCZ), which results in two major seasonal cycles. Regions with less than ~1100 mm yr-1 annual rainfall

108 are characterized by a single rainy season with a maximum in August, while, further south, the rainfall

109 seasonal cycle is characterized by two rainy seasons (September-November and March-July) (*e.g.* L´Hôte

110 *et al.,* 1996; Roudier *et al.,* 2014). The boundary between these two zones is however not very clear, with

111 areas experiencing both cycles from year to year because of high rainfall variability (*e.g.* Le Barbé *et al.,*

112 2002).

113 These different rainfall patterns result in different streamflow regimes. If the characteristics of the flow

114 hydrographs coincide with the rainfall seasonal cycle, aspects such as the timing of the peak and the shape

115 of hydrographs are mainly related to the size and physical properties of drainage basins (Roudier *et al.,*

116 2014). For example, headwater catchments in the Niger river basin (*e.g.* Mopti, Koulikoro), are

117 characterized by hydrographs with shorter lag times compared to their downstream counterparts (*e.g.*

118 Niamey, Malanville).

119 **Figure 1: Study area with locations of the main catchments (grey shaded), the river network (blue) and streamflow**

120 **gauges collected from the SIEREM database (light blue dots).**

121 Furthermore, water related issues have led to the construction of several hydraulic structures, which can

122 have significant impacts on hydrological regimes in some basins. According to the Global Reservoir and

123 Dam database (GRanD; Lehner *et al.,*2011), large dams (capacity $>10^6.m^3$), as defined by the International

124 Commission on Large Dams (ICOLD; http://www.icold-cigb.net/GB/Dictionary/dictionary.asp), are

125 primarily located in the Volta basin (53.5%) and in the Niger River basin (35.2%; Figure 2). The other

126 large dams are distributed within the Lake Chad basin (9.4%), the Senegal River basin (1.2%) and the

127 Congo basin (<1%; Figure 2). This corroborates the study by Adeaga *et al.* (2012) who found that the Volta

128 River and the lower Niger River are the most impacted rivers in western Africa. A summary of the key

129 characteristics of the existing major water resource schemes (hydropower, irrigation) in the Volta basin is

130 provided by McCartney *et al.* (2012).

131 **Figure 2: Large dams (Capacity > $10^6$ .m³) in the study area and their start of operation (purple: 1920-1950; blue: 1950-**

132 **1970; green:1970-1990; red: 1990-2006). Data source: Global Reservoir and Dam database (GRanD; Lehner *et al.,* 2011).**

133 **2.2 Data**

134 **2.2.1 Streamflow data**

135 Mean daily streamflow data were collected from the SIEREM ("Système d'Informations

136 Environnementales sur les Ressources en Eaux et leur Modélisation") database, which initially consisted of

137 data collected by the "Institut de Recherche pour le Dévelopement" (IRD). Further developments include

138 data quality assessment and a coupling to gridded environmental data over West and Central Africa (Boyer

139 *et al.,* 2006). Station metadata and GIS format files (basin contours, hydrological network, soil water

140 holding capacity, vegetation, and geology) can freely be retrieved from

141 http://www.hydrosciences.org/sierem. Additional streamflow data for the Niger River (Idah, Lokoja,

142 Makurdi and Onitsha) were collected from the National Inland Waterways Authority of Nigeria.

143 Over the study area, 863 daily streamflow datasets were collected, and monthly time series were generated

144 but only for the complete months. The percentage of missing data was then calculated for the entire region

145 (Figure 3A), and only stations with less than 50% missing records were selected for analysis. This approach

146 covered most of the study area, and rigorously minimized reconstruction errors. (see Appendix A for the

147 list of reconstructed stations). Figure 3B shows that most gaps are in the 1950s and 2000s for the 152

148 selected stations. This is due, in some countries, to the absence of gauging stations (*e.g.* Burkina Faso) and

149 lack of updated records for the recent period (*e.g.* Central African Republic).

150 **Figure 3: A) River network (blue lines) and spatial distribution of stream gauges over the study area, and with their**

151 **percentage of missing data (purple=0-25%, blue=25-50%, green=50-75% and red= >75%. Major catchments are**

152  **displayed in grey shaded. B) Time-evolution of missing values for the 152 selected stream gauges. Missing values are in**

153  **red, while observations are in grey. Stations are ordered by country (BF: Burkina Faso, BJ: Benin, CF: Central African**

154  **Republic, CG: Democratic Republic of Congo, CI: Cote d´Ivoire, CM: Cameroon, GA: Gabon, GH: Ghana, GN: Guinea**

155  **Conakry, ML: Mali, NG: Nigeria, SN: Senegal, TD: Chad, TO: Togo). The black line represents the number of records**

156  **per month over the study area for the 1950–2005 period.**

157      **2.2.2    Gridded climate data**

158  To investigate climate variability and its impact on streamflow regimes over West and Central Africa,

159  gridded monthly climatic datasets (P, T, PET) from the Climatic Research Unit (Mitchel and Jones, 2005)

160  were used. The dataset consists of monthly climatic data for the entire world (generated with more than

161  4000 weather stations at the global scale) at half a degree resolution. The development of this database

162  required seven data sources, the most important being: the Global Historical Climatology Network (GHCN;

163  Peterson and Vose, 1997), Jones Surface Temperature Anomaly dataset (Jones, 1994; Jones and Moberg,

164  2003) and Hulme Historical Monthly Precipitation (Hulme *et al.,* 1998). The latest release (CRU TS

165  v.4.00) was preferred as it was built using a new gridding technique (Angular Distance Weighting), which

166  provides more robust results due to a better selection of observation stations for gridding (Harris and Jones,

167  2017). Unlike precipitation and temperature, which are observed variables, potential evapotranspiration

168  was derived based on a variant of the Penman-Monteith formula, i.e. the FAO (Food and Agricultural

169  Organization) grass reference evapotranspiration equation (Ekström *et al.,*2007), which assumes a

170  homogenous grass surface (0.12 m height) with no moisture stress, surface albedo of 0.23 and bulk surface

171  resistance of 0.70 s/m. All climate variables are measured at 2m AGL (Above Ground Level), except for

172  wind speed (commonly recorded at 10m AGL) which has been reduced to 2m AGL using a conversion

173  coefficient. Absolute values of the different variables were computed using the baseline values (*i.e.* 1961–

174  1990 long-term average) (see Harris *et al.,* 2014; Appendix 1).

175  Even though the high spatial resolution of the dataset makes it very convenient for investigating local

176  processes, limited number of operational stations over West and Central Africa before 1940 could have

177  resulted in inconsistencies in the CRU dataset (Mitchell and Jones, 2005). Therefore, analyses in this study

178 will be performed from 1950 to 2005. In addition, Harris *et al.* (2014) compared the CRU dataset to

179 datasets developed by the University of Delaware (UDEL) and the Global Precipitation Climatology Centre

180 (GPCC), which both used different observation datasets, interpolation and quality control techniques than

181 the CRU dataset and found good agreement. For instance, for the period considered in this study, mean

182 annual precipitation values from the CRU dataset and the GPCC dataset have a correlation coefficient of

183 0.9885 significant at $p \leq 0.1$.

### 184 3. METHODOLOGY

185 Methods have been implemented using R, a free software environment for statistical computing and

186 graphics (https://www.R-project.org/).

### 187 3.1 Gap filling Methods

188 Although parametric gap-filling methods are more commonly used (*e.g.* Gyau-Boakye and Schultz, 1994;

189 Kalteh and Hjorth, 2009), non-parametric tests are more suitable for hydroclimate variables, as there is no

190 assumption regarding the distribution of the data. Both parametric and non-parametric gap filling methods

191 are therefore tested in this study to generate robust streamflow reconstructions.

### 192 3.1.1 Multiple Imputation by Chained Equations (MICE)

193 Based on a set of imputation models defined for individual variables with missing values, Multiple

194 Imputation by Chained Equations (MICE; Van Buuren and Oudshoorn, 1999) is a practical approach for

195 handling missing data. The method has been successfully tested for both continuous and categorical

196 variables in hydrology (*e.g.* Kalteh and Hjorth, 2009).

197 In this study, for each incomplete streamflow record, the first step consists of imputing missing values by

198 randomly sampling with replacement from the distribution of observed values. The observed values of each

199 streamflow station are then regressed to other streamflow stations, and missing values are completed by

200 simulated draws from the corresponding posterior predictive distribution of the considered variable (*e.g.*

201 observed values of $x_1$ are regressed on all other variables $x_2 ... xk$, and the missing values of $x_1$ are sampled

202   from its posterior predictive distribution). Several simulations are required to generate a stable single

203   reconstructed streamflow dataset, and the process is repeated several times to generate multiple complete

204   streamflow datasets. In most applications, linear regression models are used for imputing normally

205   distributed continuous variables. The different steps are summarized below:

206   Considering an incomplete variable z (with $n_{obs}$ observed values) to be reconstructed using other complete

207   variables $X = (x_1 ... x_2)$ the following linear model is used:

208   $$z|x; \beta \sim N(\beta x, \sigma^2) \qquad \text{(eq. 1)}$$

209   Let $\hat{\beta}$ be a row vector of length $k$, a realization of the estimated parameters from fitting the model with the

210   observed $z$. $V$ represents the covariance matrix of $\widehat{\beta}$, and $\hat{\sigma}$ the estimated root mean-squared error.

211   Imputation parameters $\sigma^*$ and $\beta^*$ are drawn from the exact joint distribution of $\sigma$, $\beta$ such that:

212   $$\sigma* = \hat{\sigma}\sqrt{(nobs - k)/g} \qquad \text{(eq. 2)}$$

213   $$\beta* = \hat{\beta} + \frac{\sigma*}{\hat{\sigma}} u_1 V^{1/2} \qquad \text{(eq. 3)}$$

214   with $g$, a random draw from a $\chi^2$ distribution on $n_{obs} - k$ degrees of freedom, $u_1$ a row vector of $k$

215   independent random draws from a standard Normal distribution and $V^{1/2}$ the Cholesky decomposition of $V$.

216   For each missing observation $z_i$ estimates are calculated:

217   $$z_i{}^* = \beta^* x_i + u_{2i} \sigma^* \qquad \text{(eq. 4)}$$

218   where $u_{2i}$ is a random draw from a standard normal distribution.

219   As the normal assumption is not often valid for streamflow data (*e.g.* Kundzewicz and Radziejewski,

220   2006), missing values were estimated using the Predictive Mean Matching (PMM) approach, which

221   samples estimates from the observed values of the variable $z$. Instead of estimating missing values of $z$ as in

222   eq. 4, PMM identifies $\alpha$ elements with the smallest error $|\hat{\beta}x_h - \beta^*x_i|$ *(h=1,..., nobs).* One of these

223   elements is randomly selected and the imputed value of $z_i$ is $z_{i'}$. The method was implemented using 50

224   iterations and 100 multiple imputations, which produce a standard deviation only 0.25% wider than a case

9

225    of infinite multiple imputations according to Rubin (1987). The median was taken as the better estimate to

226    combine the multiple reconstructed datasets.

### 3.1.2   Random forest-based reconstruction

228    The method is based on the random forest (RF) technique (Breiman, 2001), and involves iteratively training

229    a RF on observed values for predicting the missing values. This method was chosen for its ability to

230    perform under high dimensions, complex interactions and non-linearity (Stekhoven and Bühlmann, 2012).

231    Furthermore, compared to other gap filling methods (*e.g.* KNNimpute: Troyanskaya *et al.,* 2001; MICE:

232    Van Buuren and Oudshoorn, 1999), it does not require tuning parameters and prior knowledge of the data,

233    making it computationally attractive. The main limitation, however, is the lack of understanding around the

234    construction of the different trees. The different steps are presented below:

235    Assuming $X = (X_1, X_2, \ldots X_p)$ a $n*p$-dimensional dataset (in our case $n$ observations and $p$ streamflow

236    gauges), the missing values are estimated based on a RF trained on the observed parts of the dataset. For a

237    given gauging station $X_s$ with missing values at the indices $i_{mis}^{(s)}$, the dataset is separated in four parts:

238    •   The observed streamflow values at the station $X_s$, denoted by $Y^{(s)}_{obs}$;

239    •   The missing values at the station $X_s$, denoted by $Y^{(s)}_{mis}$;

240    •   The other gauging stations with streamflow records at the indices $i^{(s)}_{obs} = \{1,\ldots n\}\setminus i^{(s)}_{mis}$ denoted $X^{(s)}_{obs}$

241    •   The other gauging stations with streamflow records at $i^{(s)}_{mis}$ denoted by $X^{(s)}_{mis}$.

242    The initial step consists of an initial guess of missing values using mean values. The data frame is then

243    sorted and gauging stations are placed in increasing order, based on the proportion of missing data. For

244    each gauging station $X_s$, the missing data is imputed by first fitting a RF taking $Y^{(s)}_{obs}$ as response variable

245    and $X^{(s)}_{obs}$ as predictors; then estimating missing values $Y^{(s)}_{miss}$ by applying the trained RF to $X^{(s)}_{mis}$. The

246    procedure is repeated until the difference between the newly filled data matrix and the previous one

247    increases for the first time. The stopping criteria is defined as follows:

248    $$\Delta = \frac{\sum_{j \in N}(X_{new}^{imp} - X_{old}^{imp})^2}{\sum_{j \in N}(X_{new}^{imp})^2}$$    (eq. 5)

10

249    The simulations were performed using 1000 trees with the maximum number of iterations set to 100.

250    ### 3.1.3    Validation of gap filling methods

251    The validation method used to assess the performance of the implemented reconstruction techniques

252    involves generating artificial gaps in the time series, performing the reconstructions on the new dataset and

253    estimating agreements between predictions and observations. Over the study area, the assumption of data

254    missing completely at random was considered. First, observed values (12, 24, 36, 48, 60 and 120 months)

255    over the entire period, 1950–2005, were randomly removed in each of the stations and later compared to

256    the predictions. Secondly, we randomly removed entire segments of observed data to assess the ability of

257    the gap filling methods to reconstruct contiguous missing data. The modified Kling-Gupta Efficiency

258    (KGE) was used as an indicator of agreement between observations and predictions. This efficiency

259    criterion ensures that the temporal dynamics (measured by the correlation coefficient) as well as the

260    distribution of flows (measured by both the bias and variability ratio) are well represented (Kling *et al.,*

261    2012).

262    **3.2 Step change detection and trend analysis**

263    Changes (natural or artificial) in hydro-climatic time series can occur abruptly (step change) or gradually

264    (trend) or in more complex forms (Machiwal and Jha, 2006). Step-like changes, induced by reservoir

265    construction and changes of gauging structures, for example, might also result from gradual changes since

266    nonlinear system dynamics may show cumulative effects and thresholds (Kundzewicz and Radziejewski,

267    2006). In this study, step-like changes in the mean are investigated in reconstructed mean annual

268    streamflow time series using a multiple change-points detection analysis (Killick and Eckley, 2014). This

269    technique, which is similar to the method proposed by Hubert *et al.* (1989), is based on the segment

270    neighborhood algorithm (Auger and Lawrence, 1989). The non-parametric cumulative sum test statistic

271    (Page, 1954) is used to assess the optimal position of change-points.

272    Linear trends are then investigated for periods defined based on the results of the multiple change-points

273    analysis at the regional scale. The significance of the Mann-Kendall (MK) test (Kendall, 1975; Mann,

274 1945) is highly sensitive to positive serial correlation (Von Storch, 1995), so its variant (Yue *et al.,* 2002)

275 was preferred for linear trend detection here. The Yue *et al.* (2002) method assumes trends are linear;

276 datasets are first detrended before extracting the lag-1 serial correlation from the detrended series (i.e. a

277 trend-free pre-whitening procedure (TFPW) which should generate independent residuals series). The

278 detected trend and the residuals are combined, before the MK test for significance is applied. The Theil Sen

279 Approach (TSA) is used to estimate the slope of the trend in a dataset. This approach is less sensitive to

280 outliers and therefore provides a better estimate of slope for skewed data, compared to regression methods.

281 In addition, as trend values are highly dependent on start and end dates, a multitemporal trend analysis

282 approach has been implemented here (Liebmann *et al.,* 2010; McCabe and Wolock, 2002). Trends here are

283 calculated for all possible segments (minimal length of 5 years) from 1950 to 2005 to explore and define

284 the time series internal variability. For each time series, the multitemporal trend analysis generates a

285 diagram in which each possible pair of start and end dates is associated with a trend value. To investigate

286 the spatial extent and zonal coherence of the different variability patterns in precipitation and streamflow,

287 the multi-temporal trend analysis results were grouped using hierarchical clustering, using the Euclidean

288 distance as the metric of dissimilarity. Different approaches exist to determine the optimal number of

289 clusters (Charrad *et al.,* 2014), but we used the multiscale bootstrapping approach of Suzuki and

290 Shimodaira (2006), which allows uncertainty estimation for each cluster. This is achieved through

291 thousands of bootstraps resampling and used to estimate the probability that a cluster appears in the

292 replicates.

293 ## 4. RESULTS AND DISCUSSIONS

294 ### 4.1 Reconstruction outputs

295 Two reconstruction methods were applied to the subset of streamflow stations with less than 50% missing

296 data *(i.e.* 152 streamflow gauges here). All 152 stations were reconstructed with satisfactory results as

297 illustrated in Figures 4 and 5.

12

298 The validation shows that gap filling methods perform well for both cases of randomly removed

299 observations and contiguous missing segments. Figure 4 shows that the median of the KGE is always

300 greater than 0.75, which indicates that for half of the stations, the worst component (*e.g.* correlation, bias

301 ratio or variability ratio) is higher or equal to 0.75: this suggests good reconstruction performance. Very

302 similar results were achieved using the Nash-Sutcliffe Efficiency and the normalized Root Mean Squared

303 Error (not shown). Also, both methods are reasonably stable when artificially increasing the number of

304 missing observations and when artificially increasing the length of missing segments, despite an artefact

305 suggesting better performances with increasing missing data, which is in fact caused by the sensitivity of

306 efficiency criteria to sample size (*e.g.* Schönbrodt and Perugini, 2013). However, MICE seem to perform

307 better than RF when increasing the number and the length of missing data (Figure 4).

308 **Figure 4: Validation of gap filling methods: boxplot of validation efficiencies for all the reconstructed stations; upper**

309 **panels for randomly removed values and lower ones for cases of randomly missing data segments. A red line is drawn at**

310 **KGE=0.75. Outliers are represented in blue dots.**

311 To compare both gap filling methods, results from five stations from different climatic zones and

312 hydrological regimes are presented in Figure 5. While both methods show similar results overall,

313 significant dissimilarity appears in some cases, such as in the Niger River at Niamey, where MICE show an

314 abrupt increase in minimum flow, and decrease in peak flow from 1999 (Figure 5). This pattern, which is

315 similar to those induced by large dams (higher low flows and lower peak flows in downstream reaches), is

316 not consistent with recent studies in the region (*e.g.* Amogu *et al.,* 2010; Mahé *et al.,* 2013), highlighting

317 increased runoff coefficients at Niamey from the 1990s. MICE generate estimates of missing values by

318 sampling from the observed values and might therefore fail at reconstructing flows beyond observed

319 ranges. Thus, even though MICE often seem to provide better estimates than the RF based method, the

320 latter appears to be more appropriate in the context of changing hydrological regimes, because of its ability

321 to capture complex nonlinear relations between predictors and predictands.

322 **Figure 5: Reconstructed time series for five streamflow stations representative of different climatic conditions: Wayen**

323 **(Sahelian), Bonou (Tropical humid), Mbasso (Tropical humid), Niamey (Tropical humid, Sudanian and Sahelian),**

324 **Bangui (Tropical humid). Blue lines represent observations; black dotted lines represent MICE estimates and Red dotted**

325 **lines represent Random Forest estimates. Red Boxes highlight time windows of interest.**

326 ### 4.1. Streamflow changes between 1950 and 2005

327 With the assumption that two major break points occurred in the streamflow time series, the step change

328 analysis detected changes in 147 stations over the study area. Both reconstruction methods presented

329 similar results and only those of random-forest based reconstructions are presented. At the regional scale,

330 the first discontinuity in mean annual streamflow occurred in 1970 (Figure 6), with a marked negative shift

331 in the mean (up to -60%). Similar results were found by Hubert *et al.* (1989), for the Niger and Senegal

332 rivers. The second discontinuity at the regional scale occurred around 1993 and is characterized by a

333 positive shift for more than 70% of the stations (with an average increase of about +23%, Figure 6).

334 Despite this positive shift in mean streamflow, recent conditions are still below the 1950s and 1960s wet

335 periods.

336 Some sub-regional differences, however, emerge along the Gulf of Guinea and regions in Central Africa,

337 where a discontinuity in the mean annual streamflow occurred in the 1950s and early 1960s, with an

338 average positive shift of around 18% (Figure 6). These results are consistent with the findings of Mahé *et*

339 *al.* (2001), underlining differences in rainfall variability between West and Central Africa from 1951 to

340 1989. Also, some discontinuities are revealed before the 1990s in some stations (Figure 6), probably

341 induced by the wet episodes observed at the end of the 1980s. Based on the data collected from the Global

342 Reservoir and Dam database (GRanD; Lehner *et al.*,2011), presented in Figure 2, regional scale

343 discontinuities in streamflow were more likely induced by climate variability and land use change rather

344 than reservoirs as only 4% of the large dams in the region were completed between 1968 and 1970 and

345 14% between 1985 and 1993.

346 **Figure 6: Locations of step changes in random-forest based streamflow reconstructions: positive shift in the mean (blue),**

347 **negative shift in the mean (red). Stations are ordered by country (BF: Burkina Faso, BJ: Benin, CF: Central African**

348  **Republic, CG: Democratic Republic of Congo, CI: Cote d´Ivoire, CM: Cameroon, GA: Gabon, GH: Ghana, GN: Guinea**

349  **Conakry, ML: Mali, NG: Nigeria, SN: Senegal, TD: Chad, TO: Togo). The black curve on top presents the temporal**

350  **distribution of change-points over the study area.**

351  Gradual changes (trends) are investigated in mean annual reconstructed streamflow time series (MICE and

352  RF) over the periods defined by the change-points analysis, which highlights two major discontinuities at

353  the regional scale (1970 and 1993): 1950-1970 (wet conditions), 1970-1993 (drought conditions), 1993-

354  2005 (partial recovery). Figure 7 presents the correlation between the results from both reconstruction

355  methods for the different time intervals. Both reconstruction methods show similar streamflow trends at the

356  regional scale (Figure 7). However, although the results from both methods remain significantly correlated

357  ($p \leq 0.1$), trends differ slightly in the post-1990 period, mainly due to the limited ability of MICE to

358  extrapolate beyond the range of observed values, highlighting that hydrological regimes may have changed

359  in the 1993–2005 period.

360  **Figure 7: Spatial correlation between normalized trends calculated using both reconstructed datasets, for the three**

361  **periods of investigation: 1950-1970 (red), 1970-1993 (green) and 1993-2005 (blue).**

362  Trend analysis over the three different time intervals revealed that, during the 1950–1970 period, even

363  though mean annual streamflow values are at the highest, streamflow trends are significantly negative (up

364  to -4% per year) over the Sahelian and Sudanian regions of West Africa (Figure 8a-b): this suggests that the

365  step change observed around 1970 in this region was mainly induced by a gradual aridification pattern.

366  During the same period, significant positive trends are identified over Central Africa (up to +2.5% per year)

367  (Figure 7a-b). At the regional scale, 35% and 30% of trends are significant for MICE and RF respectively.

368  Among those significant trends, 52% and 40% are positive mainly in Sudanian and coastal regions (Figure

369  8a-b) for MICE and RF respectively. Most of the significant negative trends are in the Sahelian region,

370  driven by dryer conditions in the end of the 1960s compared to the 1950s (Figure 8a-b).

371  These negative streamflow trends along the Sahelian band spread toward the Gulf of Guinea and over

372  Central Africa during the well-known drought period of the 1970s and 1980s (Dai *et al.,*2004; Lebel*,* 2003;

373  Nicholson, 2013; Figure 8c-d), marking a stronger spatial coherence. During this dry period, mean annual

374 streamflow values decrease by up to 69% compared to the 1950s and 1960s. Also, more than 90% of all

375 significant trends (40% and 38% using MICE and RF, respectively) are negative (up to -5% per year), as a

376 result of intensified dry conditions from the end of the 1960s (Figure 8c-d).

377 The last period (1993-2005) is characterized by a reduction in significant trends [MICE (26%) and RF

378 (8%)] and contrasting patterns mainly due to the limited ability of MICE to fully capture complex

379 streamflow interactions (Figure 8e-f). Compared to the previous period (1970-1993) mean annual

380 streamflow values mark an increase of at least 15% over more than half of the study area and a decrease of

381 around 7% in some regions (Figure 8c-f). Significant positive trends on the Niger River, as shown using

382 RF, would however be consistent with the "Sahelian paradox" (Descroix *et al.,* 2013; Mahé *et al.,* 2005),

383 with a higher flow contribution from the Sahelian tributaries. Despite positive rainfall trends in some

384 Sudanian areas (Northern Ghana and Ivory Coast), which are detected using both MICE and RF,

385 streamflow trends remain negative (Figure 8e-f). This might have resulted from severe groundwater

386 depletion during the dry periods 1970s and 1980s (Mahé *et al.,* 2005), but this needs further research.

387 **Figure 8: Streamflow trends estimated for both reconstructed datasets, upward triangles for positive trends and**

388 **downward triangles for negative trends, filling highlights the significance of trend at 10% (negative trends in red and**

389 **positive trends in blue). River basins are greyed and the river network in blue.**

390 **4.2 Observed climatic trends between 1950 and 2005**

391 **4.2.1 Trends in annual precipitations**

392 Annual rainfall trends for the 1950–1970 period decline by ~10 mm yr-1 (significant for around 34% of the

393 study area) along the entire Sahelian band, but locally increase in parts of the Central African Republic and

394 Democratic Republic of Congo (Figure 9a). This suggests that the drying trends might have started earlier

395 than hitherto recognized. The negative trends observed along the Sahelian band then spread towards the

396 Gulf of Guinea during the 1970–1993 period (Figure 9b), similarly to the pattern observed in streamflow

397 (Figure 8 c-d).

398    However, although this period is widely recognized to be extremely dry from comparisons of mean values,

399    we find here that only 11.5% of the study area show significantly negative precipitation trends.

400    Interestingly, however, significant positive trends are identified in the Congo River basin (Figure 9b). This

401    highlights a potential hiatus in the regional drying trend during the 1970s and 1980s, supporting earlier

402    studies (Le Barbé and Lebel, 1997; D'Amato and Lebel, 1998). These could result from increasing quasi-

403    decadal rainfall variability as suggested in Dieppois *et al.* (2013, 2015). In the post-1993 period, we note an

404    increase of annual precipitation compared to the previous period (trends significant for 11% of the study

405    area), corroborating previous findings (Biasutti, 2013; Lebel and Ali, 2009; Nicholson *et al.*, 2000). This

406    potential annual rainfall recovery (~ +11.5 mm yr-1) is particularly pronounced in western and eastern

407    Sahel and Liberia (Figure 9c), which agrees with the findings of Ogungbenro and Morakinyo (2014) in

408    northern Nigeria. At the same time, regions in northern Cameroon and in the Democratic Republic of

409    Congo, are characterized by significant negative trends (~ -7 mm yr-1, to ~ -30 mm yr-1), in agreement

410    with the recent study of central African rainfall by Maidment *et al.* (2015).

411    The same analysis, conducted using the GPCC V7 datasets, show similar patterns. The relationships are,

412    however, slightly more significant over the study area for the three periods (35%, 11.43%, and 14.65% for

413    the 1950-1970, 1970-1993 and post-1993 periods, respectively; not shown). In addition, during the post-

414    1993 period, the GPCC V7 dataset underlines a significant decreasing trend in Guinea (which,

415    interestingly, does not appear in the CRU dataset) and a wider spatial extent of negative trends in

416    Cameroon and Central African Republic. Despite these slight differences probably resulting from the

417    greater number of observation stations used to generate the GPCC V7 dataset, agreement between

418    precipitation and streamflow trends remains strong.

419    Overall, there is a good agreement between annual streamflow and precipitation trends over the entire study

420    area highlighting the importance of precipitation in driving hydrological systems. However, quantifying

421    runoff response to increasing precipitation is likely to be a complex task since rising temperatures and

422    potential evapotranspiration could offset increasing precipitation. This issue is addressed in the following

423    section by investigating trends in temperatures and potential evapotranspiration and their impact on runoff

424    responses.

### 4.2.2    Trends in mean annual minimum and maximum temperatures, and potential evapotranspiration

427    As widely accepted, temperatures over the African continent have been increasing during the 20th century

428    (since 1950), and this is primarily associated with anthropogenic causes (*e.g.* IPCC, 2014; Stott *et al.,*

429    2010). However, here we aim to discuss temperature trends in term of impact on water resources, through

430    its impact on evapotranspiration and effective rainfall.

431    Trends in annual minimum and maximum temperatures over the study area show different amplitude and

432    spatial extents. For instance, in West and Central Africa, the 1950–1970 period is characterized by positive

433    trends (+0.5 to +1.5°C) in minimum annual temperatures (significant for 32.5% of the study area).

434    However, weaker and spatially less coherent trends are detected for annual maximum temperatures (~

435    +0.5°C; significant for 9.5% of the study area). Maximum values are reported only in the western Sahel

436    (Figure 9d, g). The rest of the study area shows few significant trends, apart from some significant negative

437    trends in both minimum and maximum annual temperatures (Figure 9d, g). According to the CRU potential

438    evapotranspiration estimates, the patterns in both minimum and maximum temperatures could have resulted

439    in significant positive evapotranspiration trends (~ +2.5 mm yr-1) in western and central Sahel, and

440    significant decreasing trends (~ -3.75 mm yr-1) over the Gulf of Guinea and Central Africa regions (Figure

441    9j).

442    The 1970–1993 period is marked by a homogeneous increase in annual minimum temperatures, which is

443    significant over 63% of the study area (including regions in the Congo River basin, where significant

444    cooling is identified; Figure 9e). These trends contrast with annual maximum trends, which are negative in

445    the Sahelian region (~ -1°C), but positive in the Gulf of Guinea coastal regions and Central Africa (Figure

446    9h). This configuration is, however, consistent with a weaker meridional thermal gradient, which

447    characterizes a southward shift of the ITCZ and dry conditions in the Sahel (Chiang and Friedman, 2012;

18

448  Webster *et al.,* 1998). The fluctuation of temperature range during this period has driven a uniform

449  decrease in potential evapotranspiration over the Sahelian band but increased significant positive trends in

450  the Gulf of Guinea and Central Africa (Figure 9k).

451  Since 1993, greater spatial coherence emerges, with increasing trends of both annual minimum

452  temperatures (significant for 65% of the study area) and maximum temperatures (significant for 85% of the

453  study area; Figure 9f, i). Trends in annual maximum temperatures, however, are more pronounced (~0.1°C

454  higher in average) than in annual minimal temperature (Figure 9i). This could be an artefact of the baseline

455  period used in our study, as this result contrasts with those revealed in some other studies (*e.g.* Funk *et al.,*

456  2012; Ringard *et al*., 2016), which suggested that minimum temperatures have risen faster compared to

457  maximum temperatures in the post-1990 period. Nonetheless, temperature trends are consistent with trends

458  in potential evapotranspiration (Figure 9l), which highlight a uniformly significant (for around 46.8% of the

459  study area) and positive trend (~ < +3.8 mm yr-1) over almost the entire eastern part of the study region.

460  Regions in western, eastern Sahel and part of the Gulf of Guinea, however, show non-significant negative

461  trends (Figure 9l), which may result from the spurious trends (above) in minimum temperatures and errors

462  resulting from the use of the same monthly wind speed values (1961-1990) for each year.

463  Trends in effective rainfall, approximated here as the difference between rainfall totals and potential

464  evapotranspiration are presented in Figure 9m-o. Over the two first periods (1950-1970 and 1970-1993),

465  these trends are similar to precipitation trends: this suggests the limited effect of potential

466  evapotranspiration on the relationship between rainfall and streamflow (Figure 9m-n). However, from

467  1993, the situation is reversed, mainly in the eastern part of the Sahel (eastern Niger, Chad and northern

468  Nigeria), where high potential evapotranspiration rates significantly subdue the potential impact of the

469  rainfall recovery (Figure 9o) on streamflow. This might help explain, at least partially, why the rainfall

470  recovery over these regions is not associated with significant positive streamflow trends (Figure 8c-d).

471  Over Central Africa (areas in the Congo basin), it also appears from trends in effective rainfall that during

19

472    recent decades, the decrease in rainfall is exacerbated by increased evapotranspiration (Figure 9c, i, o). This

473    suggests enhanced climatic stress on Central African streamflow in relation to warming temperatures.

474    **Figure 9: Hydroclimatic trends over the study area for three different time intervals (1950-1970, 1970-1993 and 1993-**

475    **2005) according to the CRU.TS. V4.00 dataset: a-c) Annual precipitation trends d-f) Minimum annual temperature**

476    **trends g-i) Maximum annual temperature trends j-l) Annual potential evapotranspiration trends m-o) Annual effective**

477    **rainfall trends. Sen's slope values are displayed through a red-white-blue color scale. Solid red lines highlight trend**

478    **significance at $p \leq 0.1$ according to a modified MK trend test accounting for serial correlation in the time series.**

479    ### 4.3  Precipitation and streamflow variability

480    Standard trend analysis methods assess the slope of the considered variables over the period of

481    investigation. The value of the slope is, however, highly dependent on the selected time window and

482    changes significantly for different start and end dates, mainly because of internal variability. Such

483    limitations are tackled in the multitemporal trend analysis method (Liebmann *et al.*,2010; McCabe and

484    Wolock, 2002). We used the Liebmann *et al*. (2010) approach, to calculate precipitation and streamflow

485    trends for all possible segments of 5 to 56 years between 1950 and 2005. The results are stored in two-

486    dimensional diagrams (*e.g.* Figure 11), which have been analyzed using multiscale bootstrapped

487    agglomerative hierarchical clustering.

488    Clustering streamflow variability diagrams resulted in three main clusters, which are highly significant ($p \leq$

489    *0.1*) based on the multiscale bootstrapping test, with associated spatial distributions presented in Figure 10,

490    identified using hierarchical clustering.

491    **Figure 10: Spatial distribution of streamflow variability (1950–2005) clusters based on multi-temporal trend analysis**

492    **superimposed on the river network (blue) and major river basins (grey shaded). All the clusters are highly significant at p**

493    **$\leq 0.1$ according to the multiscale bootstrapping test. Different colours displayed the location of the different clusters.**

494    Overall, these three clusters show decreasing flow trends over the entire period (1950–2005), but we also

495    identify decadal periods of alternating positive and negative trends with different amplitudes, modulating

496    the general trend, according to the three clusters (Figure 11). For instance, a pronounced positive trend in

497    the mid-1970s during the drought period emerges in cluster 1 (Congo Basin at Brazzaville), which

498  progressively disappears in cluster 2 (lower Niger River, Benue and stations in the upper Congo basin) and

499  cluster 3 (all the other stations; Figure 11). This emphasizes the importance of decadal variability in

500  modulating streamflow trends (which has hitherto been little studied) and provides a new picture of the

501  behaviour of hydrological systems in West and Central Africa.

502  These differences in the contribution of interannual to decadal variability could, however, arise from

503  differences in the large-scale climate drivers. According to Mahé *et al.* (2013), Cluster 1, which is located

504  at the outlet of the Congo Basin at Brazzaville, could be more sensitive to changes in the thermal gradient

505  between the Atlantic and Indian oceans resulting in a unique runoff variability. Such decadal fluctuations

506  have also been reported for eastern Sahel rainfall in Dieppois *et al.* (2013, 2015), suggesting that

507  differences between clusters should at least partly be related to different interactions with catchment

508  properties (*e.g.* reduction in soil water holding capacity) and water management. In addition, while trend

509  amplitude is a distinctive element between clusters, the sign and temporal scale during the humid period

510  (1950-1960) and the recovery period (post-1990) are also very important. For instance, stations in clusters 1

511  and 2 are characterized by wet conditions in the 1950s-1960s, whereas most of the stations in cluster 3

512  show decreasing trends during the same period (Figure 11). Furthermore, cluster 3 highlights less intense

513  dry conditions in the 1980s and a more pronounced recovery in the recent years compared to the first two

514  clusters (Figure 11). A further classification of the stations in cluster 3 is provided as supplementary

515  materials. The significant negative trend (observed in the 1980s) in stations of cluster 2, for instance might

516  have been partly accentuated by large dams in Nigeria (*e.g.* the Dadin Kowa Dam and the Kiri dam, on a

517  main tributary of the Benue river).

518  **Figure 11: Multi-temporal diagrams of the different cluster centroids: trends in m³/s are presented in red (negative) –**

519  **white (null) – blue (positive) colour scale, contours lines represent trend significance at $p \leq 0.1$.**

520  Applying the same clustering method to gridded annual rainfall, variability diagrams resulted in 12 major

521  clusters ($p \leq 0.1$) and few grid points with lower probabilities ($p \leq 0.2$) and therefore unclassified (Figure

522  12).

523 **Figure 12: Clusters of rainfall variability generated using CRU TS V4.00 (2.5°x2.5°) on the period 1950-2005: colours**

524 **and numbers from 1 to 12 refer to the grid points within the 12 initial clusters at $p \leq 0.1$. Red boxes represent grid points**

525 **which did not fall within the clusters. All the clusters are highly significant at $p \leq 0.1$ according to the multiscale**

526 **bootstrapping test.**

527 West African regions predominantly fall within clusters 11, 8, 2 and 1 (Figures 12, 13) which are mainly

528 characterized by persistent dry conditions from the end of the 1960s, and positive trends starting earlier in

529 clusters 2 and 8 (1970s) compared to clusters 1 and 11 (end of 1980s). Comparing, for instance, patterns

530 observed in streamflow cluster 3 and rainfall cluster 11, it appears that the significant negative rainfall trend

531 in the 1980s is attenuated in the streamflow signal and, furthermore, the observed streamflow recovery is

532 more widespread compared to the recovery observed in rainfall. This suggests a combination of drivers

533 which might have enhanced the runoff response, described by some authors as the "Sahelian paradox"

534 (Descroix *et al.,* 2013; Mahé *et al.,* 2005) which refers to a counterintuitive increase in runoff coefficient

535 despite decreasing rainfall. In fact, parts of this region are known to have experienced drastic changes in

536 land cover resulting from several coupled interactions between increasing cultivated areas (Cappelaere *et*

537 *al.,*2009), and natural vegetation changes after the 1970s and 1980s major drought periods (Gal *et al.,*

538 2017).

539 The clustering underlines a high variability in rainfall over the western part of West Africa, where some

540 grid points are left outside the clusters. Some parts of this region are characterized by the pattern observed

541 in cluster 9 (Figures 12-13). After the humid period of the 1960s, rainfall is characterized by decreasing

542 trends until the 1990s (Figure 13). From the end of the 1990s rainfalls largely returned to their level of the

543 1960s as a result of a recovery which started in the 1980s (Figure 13). From these different clusters, it

544 appears that most regions over western Africa have experienced improved streamflow conditions because

545 of the recent rainfall recovery even though long-term trends remain negative.

546 Over Central Africa, rainfall shows high decadal variability (succession of wet and dry periods) with no

547 clear long-term trends (clusters 4, 5, 6, 10 and 12; Figure 13). This region is characterized by a humid

548     period from the mid-1950s to the 1970s, with dry episodes around 1980 (Figure 13). In cluster 4, for

549     instance, recent conditions (1990s-2000s) are almost as wet as the humid period, which is not the case for

550     cluster 6 where recent conditions remain relatively drier (Figure 13). The streamflow variability displayed

551     in cluster 1 (Congo basin at Brazzaville) appears to have resulted from a combination of rainfall clusters 6,

552     10 and 12, highlighting the diverse climatic influences in this basin (Figure 11-13). Rainfall-runoff

553     relations over this region suggest that rainfall is the main driving factor, with no, or limited, effect from

554     other moderating factors (*e.g*. land use change, intensification of agriculture, deforestation, and warming

555     temperatures). The change in seasonal rainfall distribution may likely be the major factor related to climatic

556     change in this area to have an impact on discharges' seasonal regimes. This can be observed at the scale of

557     small basins like the Kienke at Kribi in the South coastal Cameroon, where the small dry season

558     disappeared during the last decades (Lienou *et al.,* 2008), and at the larger scale for the Ogooue river in

559     Gabon, where the Spring flood lost 30% of discharge value after the rainfall regimes slightly changed over

560     past decades (Mahé *et al.*, 2013), the same being observed to a lesser extent for the whole Congo basin

561     (Alsdorf *et al.,* 2016; Tshimanga *et al.,* 2016).

562     **Figure 13: Multi-temporal diagrams of the 12 rainfall variability clusters derived from the multi-scale bootstrap**

563     **clustering: trends (mm) are presented in red (negative) – white (null) - blue (positive) color scale, contours lines represent**

564     **trend significance at *p ≤ 0.1*.**

565     **5. CONCLUSION**

566     Using parametric (MICE) and non-parametric (RF) gap filling methods, a new and complete streamflow

567     dataset, spatially distributed over West and Central Africa and encompassing different climatic zones and

568     hydrological regimes has been generated. Gap filling results highlighted that both methods performed well,

569     though, in general, MICE was slightly outperforming RF. However, due to its parametric nature, MICE

570     analyses, in some cases, failed to capture changes in streamflow conditions (case of Niamey on the Niger

571     River). The complete streamflow dataset (RF method) was then used to investigate streamflow changes and

572     variability and their interactions with key climatic variables (P, T, PET) over West and Central Africa

573  between 1950 and 2005.

574  Majority of streamflow stations over the study area present a step change in 1970 mainly induced by a

575  gradual aridification pattern. In the 1990s a positive shift in mean discharge is observed, but it is difficult to

576  conclude whether this change is led by positive rainfall trends or single wet episodes amplified by land use

577  change, warming temperature and evapotranspiration reduction. In general, there is a good agreement

578  between streamflow and precipitation trends, with an offsetting effect of potential evapotranspiration

579  observed in some regions. Over the study area, the period 1950–1970 was characterized by negative

580  streamflow trends in Sahelian and Sudanian regions of West Africa, which seems counterintuitive

581  considering that this period was the wettest on record. The opposite is observed over Central Africa where

582  significant positive streamflow trends emerge. The following period (1970–1993), is marked mostly by

583  negative trends due to dryer conditions. This pattern is reversed during the 13-year period 1993–2005, with

584  mainly positive trends resulting from increased rainfall and changes in land use in Sahelian regions. Annual

585  streamflow trends reflect annual precipitation trends which decrease from the 1950s to 1980s and increase

586  from the 1990s. More importantly, the study showed that, even though the 1950s and 1960s were the

587  wettest decades in terms of total rainfall amounts, decreasing annual rainfall trends were more prominent,

588  suggesting an earlier start of the drought. The drought peaked during the 1970s/80s, over most of western

589  Africa, but the reduced negative trends in precipitation suggest a hiatus, which have resulted from quasi-

590  decadal variability.

591  Furthermore, over most of the study area, hydrological regimes during the recent period have been

592  impacted by the rainfall recovery which is not limited to the west African Sahel. Even though other

593  climatic variables such as wind speed and vapor pressure deficit might have played an important role,

594  temperature trends appeared to be highly related to trends in potential evapotranspiration, which seem to

595  have hampered the effect of the rainfall recovery on hydrological regimes in some areas over the eastern

596  Sahel (eastern Niger, Chad and northern Nigeria).

597  Building significantly on previous studies, which generally provide trend estimates over a certain period,

24

598 we have provided novel information and analyses of the impact of internal variability using the

599 multitemporal trend analysis method. The results highlight strong interannual to decadal signals which

600 clearly modulate streamflow and precipitation trends. In West Africa, for instance, the 1970-1989 period is

601 characterized by two main dry events (1972-1973 and 1983-1984) separated by a wet period (Nicholson *et*

602 *al.,* 2000; Dai *et al.,* 2004). This probably resulted in increased runoff coefficients in Sahelian catchments,

603 as observed by Albergel (1987) in Burkina Faso over the period (1969-1984) and later in larger Sahelian

604 catchments (Descroix *et al.,* 2013; Mahé *et al.,* 2005). Such a rainfall-runoff response (referred to as the

605 Sahelian paradox) indeed seems paradoxical when considering long-term trends but becomes less

606 counterintuitive when investigating variability in precipitation and streamflow time series. Therefore, rather

607 than describing the "Sahelian paradox" as an increase in runoff despite reduced rainfall since 1970, it

608 should be considered as enhancing runoff response to positive rainfall anomalies, as a result of changes in

609 land-surface properties.

610 If flow trends can be largely explained by decadal variability in rainfall (Dieppois *et al.,* 2013), influence of

611 other driving factors should also be considered at the catchment level (such as geology, soils, agricultural

612 land use change, water consumption and urbanization). For instance, large dams constructed in the 1980s in

613 Nigeria (*e.g.* the Dadin Kowa Dam and the Kiri dam, on a main tributary of the Benue river), might have

614 affected to some extent the variability of the lower Niger river, but this is beyond the scope of the present

615 paper.

616 This study has shed light on hydroclimatic variability and its associated impact on streamflow regimes over

617 large, key parts of West and Central Africa over recent decades, and also provides water practitioners with

618 reconstructed streamflow time series which can be used as input for water balance models to develop sound

619 water and agricultural management policies. These useful time series here can form the basis of future

620 developments, to include updating of the streamflow datasets through national water offices. This should

621 further improve the quality of the reconstructions and open up investigations of more recent conditions. In

622 addition, future in-depth studies are required of climate processes (*e.g.* sea-surface temperature,

623 atmospheric circulation), catchment land use properties, and water management policies, all of which can

624 drive streamflow variability at interannual to decadal timescales. As these potentially modulate the climate

625 signal, such work is required to further improve our understanding of hydrological variability in West and

626 Central Africa, and our ability to model hydrological changes in this region.

627 **Acknowledgements**

633 **Appendix A: List of reconstructed streamflow time series**

| ID | Basin | Station name | latitude | Longitude |
|---|---|---|---|---|
| BFQ0010 | LERABA | YENDERE au pont | 10.1667 | -5.0683 |
| BFQ0060 | VOLTA | WAYEN | 12.3789 | -1.08 |
| BFQ0064 | VOLTA | BOROMO | 11.7833 | -2.9167 |
| BFQ0065 | VOLTA | DAPOLA | 10.5667 | -2.9167 |
| BFQ0072 | VOLTA | NWOKUY | 12.5278 | -3.55 |
| BFQ0074 | VOLTA | SAMANDENI | 11.4667 | -4.4667 |
| BJQ0009 | SOTA | COUBERI | 11.74 | 3.3333 |
| BJQ0022 | COUFFO | LANHOUNTA - LANTA | 7.1 | 1.8833 |
| BJQ0028 | MONO | ATHIEME | 6.9167 | 1.6667 |
| BJQ0033 | OUEME | BONOU | 6.9 | 2.45 |
| BJQ0036 | OUEME | HETIN SOTA | 6.6 | 2.5 |
| BJQ0047 | OKPARA | KABOUA | 8.25 | 2.7167 |
| BJQ0050 | SOTA | RTE KANDI-SEGBANA AMONT | 10.9833 | 3.25 |
| BJQ0075 | WE-WE | WE-WE | 9.1667 | 2.1083 |
| BJQ1000 | PENDJARI | PORGA | 10.99401 | 0.9773 |
| BJQ2000 | NIGER | MALANVILLE | 11.888 | 3.383 |
| BJQ2004 | OUEME | PONT DE BETEROU | 9.199179 | 2.267582 |
| BJQ2005 | OUEME | PONT DE SAVE | 8 | 2.4167 |
| BJQ2006 | ZOU | ATCHERIGBE | 7.5333 | 2.0333 |
| CFQ0005 | OUHAM | BOSSANGOA | 6.4667 | 17.45 |
| CFQ0025 | OUBANGUI | ZINGA TRANSIT | 3.713833 | 18.58716 |
| CFQ0027 | MBOMOU | ZEMIO | 5.028726 | 25.1471 |
| CFQ0028 | BANGUI-KETTE | ALINDAO | 5.04457 | 21.20172 |

| CFQ0034 | LOBAYE | M'BATA | 3.666296 | 21.98114 |
|---|---|---|---|---|
| CFQ0040 | M'POKO | BOSSELE-BALI | 4.530737 | 18.46878 |
| CFQ0057 | SANGHA | SALO | 3.181621 | 16.11362 |
| CFQ2000 | OUBANGUI | BANGUI | 4.364275 | 18.59487 |
| CGQ0003 | ALIMA | TCHIKAPIKA | -1.26385 | 16.16937 |
| CGQ0013 | LEFINI | BWEMBE | -2.9167 | 15.6308 |
| CGQ0014 | LIKOUALA | ETOUMBI | 0.0167 | 14.95 |
| CGQ0015 | LIKOUALA | MAKOUA | 0.00167 | 15.633 |
| CGQ0017 | N'KENI | GAMBOMA | -1.9 | 15.85 |
| CGQ0020 | KOUYOU | LINNEGUE | -0.5 | 15.9333 |
| CGQ0026 | LIKOUALA | BOTOUALI | -0.55 | 17.45 |
| CGQ2000 | CONGO | BEACH - V.N. Brazzaville | -4.27285 | 15.29392 |
| CGQ2001 | SANGHA | OUESSO | 1.6167 | 16.05 |
| CIQ0007 | BANDAMA | MBRIMBO | 6.0125 | -4.425 |
| CIQ0013 | BANDAMA | KIMOUKRO  BALISE  10201 | 6.5056 | -5.3053 |
| CIQ0032 | MARAOUE | RTE BEOUMI-SEGUELA - KONGASSO 10145 | 7.8319 | -6.2542 |
| CIQ0033 | MARAOUE | BOUAFLE     10147 | 6.979988 | -5.75437 |
| CIQ0058 | NZI | BOCANDA | 7.0442 | -4.52 |
| CIQ0061 | NZI | DIMBOKRO  10141 | 6.6358 | -4.71 |
| CIQ0154 | KOUROUKELE | IRADOUGOU | 9.7069 | -7.8028 |
| CIQ0292 | KAVI | MBESSE | 5.8386 | -4.2961 |
| CIQ0312 | CAVALLY | FLAMPLEU | 7.2833 | -8.0583 |
| CIQ0314 | CAVALLY | TAI | 5.86 | -7.45 |
| CIQ0319 | NSE | TAI 1 (TAI PONT) | 5.875 | -7.4583 |
| CIQ4020 | BANDAMA | BADA | 8.1069 | -5.4972 |
| CIQ4022 | BANDAMA | TIASSALE     10144 | 5.8947 | -4.8178 |
| CIQ4025 | NZI | FETEKRO | 7.8106 | -4.6875 |
| CIQ4026 | NZI | MBAHIAKRO     10133 | 7.4458 | -4.3556 |
| CIQ4027 | NZI | NZIENOA     10136 | 5.9964 | -4.8125 |
| CIQ4028 | COMOE | ANIASSUE PONT 10138 | 6.6375 | -3.7126 |
| CIQ4029 | COMOE | MBASSO | 6.125 | -3.48 |
| CIQ4030 | COMOE | SEREBOU | 7.9383 | -3.9419 |
| CIQ4031 | SASSANDRA | SEMIEN     10130 | 7.7083 | -7.0669 |
| CIQ4032 | SASSANDRA | SOUBRE | 5.7833 | -6.6131 |
| CIQ4033 | BAFING | BAFINGDALA (BADALA)  BIANKOUMA 10162 | 7.841611 | -7.66658 |
| CIQ4034 | LOBO | NIBEHIBE (NIBEIGBEU) | 6.8003 | -6.7 |
| CIQ4035 | COMOE | AKAKOMOEKRO  10149 | 7.447418 | -3.5094 |
| CMQ0008 | DOUME | DOUME | 4.2333 | 13.45 |
| CMQ0029 | SANAGA | NACHTIGAL | 4.35 | 11.6333 |
| CMQ0030 | SANAGA | NANGA EBOKO | 4.7 | 12.3833 |
| CMQ0038 | MBAM | BAC DE GOURA | 4.5667 | 11.3667 |
| CMQ0071 | NYONG | DEHANE | 3.5667 | 10.1167 |
| CMQ5001 | VINA NORD | PONT DE BEREM | 7.55 | 13.95 |
| CMQ5005 | DJA | SOMALOMO | 3.3667 | 12.7333 |

27

| CMQ5006 | BENOUE | BUFFLE NOIR | 8.1167 | 13.8333 |
|---|---|---|---|---|
| CMQ5007 | BENOUE | GAROUA | 9.294019 | 13.4041 |
| CMQ5015 | MAPE | AU PONT DE MAGBA  AMONT | 5.9833 | 11.2667 |
| CMQ5016 | VINA DU SUD | LAHORE | 7.25 | 13.5667 |
| CMQ5018 | LOBE | BAC KRIBI-CAMPO | 2.8667 | 9.8833 |
| CMQ5019 | LOKOUNDJE | LOLODORF | 3.2333 | 10.7333 |
| CMQ5038 | MUNGO | MUNDAME | 4.5667 | 9.5333 |
| CMQ5040 | NTEM | BAC DE NGOAZIK | 2.1333 | 11.3 |
| CMQ5044 | LOM | BETARE OYA | 5.9167 | 14.1333 |
| CMQ5047 | KIENKE | KRIBI SCIERIE | 2.9333 | 9.9 |
| CMQ5050 | KADEI | BATOURI | 4.4167 | 14.3167 |
| GAQ0006 | OGOOUE | BOOUE (LMNG) | -0.1025 | 11.9367 |
| GAQ0015 | OGOOUE | NDJOLE OPERATIONNEL | -0.455 | 10.4025 |
| GAQ0028 | IVINDO | MAKOKOU (LMNG) | 0.5689 | 12.8611 |
| GAQ0041 | NGOUNIE | FOUGAMOU  S H O  (LMNG) | -1.2156 | 10.5908 |
| GAQ0046 | NGOUNIE | MOUILA VAL MARIE | -1.8869 | 11.0558 |
| GHQ0045 | NASIA | NASIA | 10.15 | -0.8 |
| GHQ1005 | VOLTA NOIRE | BUI AMONT | 8.2833 | -2.2333 |
| GNQ0015 | NIGER | FARANAH | 10.03744 | -10.7495 |
| GNQ0016 | NIGER | KOUROUSSA | 10.65169 | -9.87096 |
| GNQ0018 | NIGER | TIGUIBERY  (Siguiri) | 11.3545 | -9.16459 |
| GNQ0026 | MILO | KANKAN | 10.3833 | -9.3 |
| GNQ0030 | NIANDAN | BARO | 10.6167 | -9.7 |
| GNQ0034 | NIANDAN | KISSIDOUGOU (NIANDAN SCIERIE) | 9.25 | -10.0167 |
| GNQ0200 | BADI | BAC DE BADI | 10.2833 | -13.4 |
| GNQ0204 | KONKOURE | PONT DE LINSAN | 10.3 | -12.4167 |
| MLQ0009 | NIGER | DIRE | 16.27595 | -3.395 |
| MLQ0012 | NIGER | KE MACINA | 13.95831 | -5.35896 |
| MLQ0019 | NIGER | KOULIKORO | 12.85727 | -7.55811 |
| MLQ0022 | NIGER | MOPTI | 14.49605 | -4.20127 |
| MLQ0036 | NIGER | TOSSAYE | 16.9333 | -0.5833 |
| MLQ0052 | DIAKA | KARA | 14.1667 | -5.0167 |
| MLQ0091 | BANI | SOFARA | 14.01393 | -4.2429 |
| MLQ0123 | SENEGAL | GALOUGO | 13.8333 | -11.1333 |
| MLQ0130 | SENEGAL | BAFING MAKANA | 12.55 | -10.2667 |
| MLQ0131 | SENEGAL | SOUKOUTALI | 13.2 | -10.4167 |
| MLQ0134 | BAKOYE | OUALIA | 13.6 | -10.3833 |
| MLQ0135 | BAKOYE | TOUKOTO | 13.45 | -9.8833 |
| MLQ0137 | FALEME | FADOUGOU | 12.5167 | -11.3833 |
| MLQ0145 | BAOULE | SIRAMAKANA  (Balenda) | 13.5833 | -9.8833 |
| MLQ2003 | NIGER | KENIEROBA | 12.1 | -8.3167 |
| MLQ2007 | SANKARANI | SELINGUE | 11.5833 | -8.1667 |
| MLQ2008 | BANI | DOUNA | 13.21385 | -5.90311 |
| MLQ2064 | SENEGAL | DAKA SAIDOU | 11.95 | -10.6167 |

| MLQ2066 | SENEGAL | DIBIA | 13.2333 | -10.8 |
|---------|---------|-------|---------|-------|
| MLQ2069 | FALEME | GOURBASSY | 13.4 | -11.6333 |
| MLQ2070 | SENEGAL | KAYES | 14.45 | -11.45 |
| NEQ2000 | NIGER | NIAMEY | 13.5016 | 2.105 |
| NGQ0001 | BENUE | MAKURDI | 7.75 | 8.5333 |
| NGQ0002 | NIGER | ONITSHA | 6.166667 | 6.75 |
| NGQ2000 | NIGER | LOKOJA | 7.8 | 6.7667 |
| NGQ2004 | NIGER | IDAH | 7.1 | 6.716667 |
| SNQ2039 | GAMBIE | KEDOUGOU | 12.55 | -12.1833 |
| SNQ2045 | GAMBIE | MAKO | 12.8667 | -12.35 |
| SNQ2055 | GAMBIE | SIMENTI | 13.0333 | -13.3 |
| SNQ2060 | GAMBIE | WASSADOU-AMONT | 13.35 | -13.3667 |
| SNQ2062 | GAMBIE | WASSADOU-AVAL | 13.35 | -13.3833 |
| SNQ2063 | SENEGAL | BAKEL | 14.9 | -12.45 |
| SNQ2064 | SENEGAL | DAGANA | 16.5167 | -15.5 |
| SNQ2065 | FALEME | KIDIRA  UHEA | 14.45466 | -12.205 |
| SNQ2066 | SENEGAL | MATAM | 15.65 | -13.25 |
| SNQ2067 | DOUE | NGOUI | 16.15 | -13.9167 |
| SNQ2068 | SENEGAL | SALDE | 16.16325 | -13.8795 |
| TDQ0004 | CHARI | SARH (EX.FORT-ARCHAMBAULT) | 9.15 | 18.4167 |
| TDQ0009 | CHARI | MAILAO | 11.6 | 15.2833 |
| TDQ0013 | BAHR-SARA | MANDA | 9.1833 | 18.2 |
| TDQ0014 | BAHR-SARA | MOISSALA | 8.3333 | 17.7667 |
| TDQ0036 | LIM | OULI BANGALA | 7.8333 | 15.8333 |
| TDQ0041 | PENDE | GORE | 7.95 | 16.6167 |
| TDQ0043 | TANDJILE | TCHOA | 9.3333 | 16.0833 |
| TDQ2011 | CHARI | BOUSSO | 10.5 | 16.7167 |
| TDQ5004 | LOGONE | KATOA | 10.8333 | 15.0833 |
| TDQ5005 | LOGONE | LAI (MISSION) | 9.4 | 16.3 |
| TDQ5006 | LOGONE | LOGONE-GANA | 11.55 | 15.15 |
| TOQ0006 | KARA | LAMA KARA 1 | 9.5333 | 1.1833 |
| TOQ0037 | SIO | KPEDJI | 6.5317 | 1.0083 |
| TOQ0042 | MONO | CORREKOPE | 7.8 | 1.3 |
| TOQ0043 | MONO | DOTAIKOPE | 7.8167 | 1.2667 |
| TOQ0046 | MONO | TETETOU | 7.0167 | 1.5333 |
| TOQ0048 | AMOU | AMOU OBLO | 7.4 | 0.8667 |
| TOQ0053 | ANIE | PONT C F T | 7.7333 | 1.2 |
| TOQ0056 | KOLOWARE | KOLOWARE | 8.9667 | 1.2833 |
| TOQ0057 | NA | PARATAO | 8.95 | 1.1833 |
| TOQ0059 | OGOU | SIRKA | 7.9167 | 1.3667 |

634

635 **REFERENCES**

636 Adeaga, O., Mahé, G., Dieulin, C., Elbaz-Poulichet, F., Rouche, N., Seide, J. L., & Servat, E., 2012.   Rainfall-Runoff

Simulation in Part of Lower Niger Basin. *Journal of Environmental Science and Engineering. B*, *1*(6B), 812.

Albergel, J., 1987. Sécheresse, désertification et ressources en eau de surface : application aux petits bassins du Burkina Faso. *The influence of climate change and climatic variability on the hydrologic regime and water resources*, *168*, 355-365.

Ali, A., Lebel, T., 2009. The Sahelian standardized rainfall index revisited. Int. J. Climatol. 29, 1705–1714. https://doi.org/10.1002/joc.1832

Alsdorf, D., Beighley, E., Laraque, A., Lee, H., Tshimanga, R., O'Loughlin, F., Mahé, G., Dinga, B., Moukandi, G., Spencer, R.G.M., 2016. Opportunities for hydrologic research in the Congo Basin. Rev. Geophys. https://doi.org/10.1002/2016RG000517

Amogu, O., Descroix, L., Yéro, K.S., Le Breton, E., Mamadou, I., Ali, A., Vischel, T., Bader, J.-C., Moussa, I.B., Gautier, E., Boubkraoui, S., Belleudy, P., 2010. Increasing River Flows in the Sahel? Water 2, 170–199. https://doi.org/10.3390/w2020170

Auger, I.E., Lawrence, C.E., 1989. Algorithms for the optimal identification of segment neighborhoods. Bull. Math. Biol. 51, 39–54. https://doi.org/10.1007/BF02458835

Badr, H.S., Dezfuli, A.K., Zaitchik, B.F., PETers-Lidard, C.D., 2016. Regionalizing Africa: Patterns of precipitation variability in observations and global climate models. J. Clim. 29, 9027–9043. https://doi.org/10.1175/JCLI-D-16-0182.1

Biasutti, M., 2013. Forced Sahel rainfall trends in the CMIP5 archive. J. Geophys. Res. Atmos. 118, 1613–1623. https://doi.org/10.1002/jgrd.50206

Boyer, J. F., Dieulin, C., Rouche, N., Cres, A., Servat, E., Paturel, J. E., Mahé, G., 2006. SIEREM: An environmental information system for water resources. *Climate Variability and Change—Hydrological Impacts,* S. Demuth et al., Eds., IAHS Publication 308, IAHS/AISH, 19–25

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Cappelaere, B., Descroix, L., Lebel, T., Boulain, N., Ramier, D., Laurent, J.P., Favreau, G., Boubkraoui, S., Boucher, M., Bouzou Moussa, I., Chaffard, V., Hiernaux, P., Issoufou, H.B.A., Le Breton, E., Mamadou, I., Nazoumou, Y., Oi, M., Ottlé, C., Quantin, G., 2009. The AMMA-CATCH experiment in the cultivated Sahelian area of south-west Niger - Investigating water cycle response to a fluctuating climate and changing environment. J. Hydrol. 375, 34–51. https://doi.org/10.1016/j.jhydrol.2009.06.021

Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., 2014. **NbClust** : An *R* Package for Determining the Relevant Number of Clusters in a Dataset. J. Stat. Softw. 61. https://doi.org/10.18637/jss.v061.i06

Chiang, J.C.H., Friedman, A.R., 2012. Extratropical Cooling, Interhemispheric Thermal Gradients, and Tropical Climate Change. Annu. Rev. Earth Planet. Sci. 40, 383–412. https://doi.org/10.1146/annurev-earth-042711-105545

D'amato, N., Lebel, T., 1998. On the characteristics of the rainfall events in the Sahel with a view to the analysis of climatic variability. Int. J. Climatol. 18, 955–974. https://doi.org/10.1002/(SICI)1097-0088(199807)18:9<955::AID-JOC236>3.0.CO;2-6

Dai, A., Lamb, P.J., Trenberth, K.E., Hulme, M., Jones, P.D., Xie, P., 2004. The recent Sahel drought is real. Int. J. Climatol. 24, 1323–1331. https://doi.org/10.1002/joc.1083

Descroix, L., Mahé, G., Lebel, T., Favreau, G., Galle, S., Gautier, E., Olivry, J.C., Albergel, J., Amogu, O., Cappelaere, B., Dessouassi, R., Diedhiou, A., Le Breton, E., Mamadou, I., Sighomnou, D., 2009. Spatio-temporal variability of hydrological regimes around the boundaries between Sahelian and Sudanian areas of West Africa: A synthesis. J. Hydrol. 375, 90–102. https://doi.org/10.1016/j.jhydrol.2008.12.012

Descroix, L., Moussa, I. B., Genthon, P., Sighomnou, D., Mahé, G., Mamadou, I., Vandervaere, J.-P., Gautier, E., Maiga, O. F., Rajot, J.-L., Abdou, M. M., Dessay, N., Ingatan, A., Noma, I., Yéro, K.S., Karambiri, H., Fensholt, R., Albergel, J., and Olivry, J.-C.: Impact of Drought and Land – Use Changes on Surface – Wa-ter Quality and Quantity:

The Sahelian Paradox, Curr. Perspect.Contam. Hydrol. Water Res. Sustain., in: Current Perspectives in Contaminant Hydrology and Water Resources Sustainability, chap. 10, edited by: Bradley, P. M., 243–271, doi:10.5772/54536, 2013

Dieppois, B., Diedhiou, A., Durand, A., Fournier, M., Massei, N., Sebag, D., Xue, Y., Fontaine, B., 2013. Quasi-decadal signals of Sahel rainfall and West African monsoon since the mid-twentieth century. J. Geophys. Res. Atmos. 118, 12587–12599. https://doi.org/10.1002/2013JD019681

Dieppois, B., Durand, A., Fournier, M., Diedhiou, A., Fontaine, B., Massei, N., Nouaceur, Z., Sebag, D., 2015. Low-frequency variability and zonal contrast in Sahel rainfall and Atlantic sea surface temperature teleconnections during the last century. Theor. Appl. Climatol. 121, 139–155. https://doi.org/10.1007/s00704-014-1229-5

Ekström, M., Jones, P.D., Fowler, H.J., Lenderink, G., Buishand, T. a., Conway, D., 2007. Regional climate model data used within the SWURVE project &ndash; 1: projected changes in seasonal patterns and estimation of PET. Hydrol. Earth Syst. Sci. 11, 1069–1083. https://doi.org/10.5194/hess-11-1069-2007

FAO: Sahel: situation météorologique et état des cultures, Rapport du 11 juin 2004, http://www.fao.org/docrep/006/j2517f/j2517f00.HTM (last access: October 2016), 2004.

Funk, C., Michaelsen, J., Marshall, M., 2012. Mapping recent decadal climate variations in precipitation and temperature across Eastern Africa and the Sahel. Drought and Water Crises, pp. 331–358. CRC Press.

Gal, L., Grippa, M., Hiernaux, P., Pons, L., Kergoat, L., 2017. The paradoxical evolution of runoff in the pastoral Sahel: Analysis of the hydrological changes over the Agoufou watershed (Mali) using the KINEROS-2 model. Hydrol. Earth Syst. Sci. 21, 4591–4613. https://doi.org/10.5194/hess-21-4591-2017

Gyau-Boakye, P., Schultz, G.A., 1994. Filling gaps in runoff time series in west africa. Hydrol. Sci. J. 39, 621–636. https://doi.org/10.1080/02626669409492784

Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. Int. J. Climatol. 34, 623–642. https://doi.org/10.1002/joc.3711

Herrmann, S.M., Mohr, K.I., 2011. A continental-scale classification of rainfall seasonality regimes in Africa based on gridded precipitation and land surface temperature products. J. Appl. Meteorol. Climatol. 50, 2504–2513. https://doi.org/10.1175/JAMC-D-11-024.1

Hubert, P., Carbonnel, J.P., Chaouche, A., 1989. Segmentation des séries hydrométéorologiques - application à des séries de précipitations et de débits de l'afrique de l'ouest. J. Hydrol. 110, 349–367. https://doi.org/10.1016/0022-1694(89)90197-2

Hulme, M., Doherty, R., Ngara, T., New, M., Lister, D., 2001. African climate change: 1900-2100. Clim. Res. 17, 145–168. https://doi.org/10.3354/cr017145

Hulme, M., Osborn, T.J., Johns, T.C., 1998. Precipitation sensitivity to global warming: Comparison of observations with HadCM2 simulations. Geophys. Res. Lett. 25, 3379. https://doi.org/10.1029/98GL02562

Ibrahim, B., Karambiri, H., Polcher, J., 2015. Hydrological Impacts of the Changes in Simulated Rainfall Fields on Nakanbe Basin in Burkina Faso. Climate 3, 442–458. https://doi.org/10.3390/cli3030442

IPCC, 2014. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge Univ. Press 688. https://doi.org/10.1017/CBO9781107415324.004

Janicot, S., 1992. Spatiotemporal variability of West African rainfall. Part I: Regionalizations and Typings. J. Clim. 5, 489–497. https://doi.org/10.1175/1520-0442(1992)005<0489:SVOWAR>2.0.CO;2

Jones, P.D., 1994. Hemispheric surface air temperature variations: a reanalysis and an update to 1993. J. Clim. 7, 1794–1802. https://doi.org/10.1175/1520-0442(1994)007<1794:HSATVA>2.0.CO;2

Jones, P.D., Moberg, A., 2003. Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. J. Clim. 16, 206–223. https://doi.org/10.1175/1520-0442(2003)016<0206:HALSSA>2.0.CO;2

Jury, M.R., 2013. A return to wet conditions over Africa: 1995-2010. Theor. Appl. Climatol. 111, 471–481. https://doi.org/10.1007/s00704-012-0677-z

Kalteh, A.M., Hjorth, P., 2009. Imputation of missing values in a precipitation–runoff process database. Hydrol. Res. 40, 420. https://doi.org/10.2166/nh.2009.001

Kendall, M., 1975. *Multivariate analysis*. Charles Griffin.

Killick, R., Eckley, I., 2014. changepoint: An R package for changepoint analysis. *Journal of statistical software*, *58*(3), 1-19.

Kim, J.W., Pachepsky, Y.A., 2010. Reconstructing missing daily precipitation data using regression trees and artificial neural networks for SWAT streamflow simulation. J. Hydrol. 394, 305–314. https://doi.org/10.1016/j.jhydrol.2010.09.005

Kling, H., Fuchs, M., Paulin, M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. J. Hydrol. 424–425, 264–277. https://doi.org/10.1016/j.jhydrol.2012.01.011

Kundzewicz Z.W., Radziejewski M., 2006. Methodologies for trend detection. In: Climate Variability and Change-Hydrological Impacts (FRIEND). S. Demuth, A. Gustard, E. Planos, F. Scatena & E. Servat (Eds), IAHS Publ. 308, 538-550.

L'Hôte, Y., Dubreuil, P., Lerique, J. ,1996. Carte des types de climats "en Afrique Noire à l´ouest du Congo". Rappels et extension aux régimes hydrologiques. In: *L´hydrologie Tropicale: Geoscience et Outil Pour le Développement* (Mélanges à la mémoire de Jean Rodier. Paris, mai 1995), 55-65. *IAHS P*ubl. no. 238.

L'Hôte, Y., Mahé, G., Somé, B., Triboulet, J.P., 2002. Analysis of a Sahelian annual rainfall index from 1896 to 2000; the drought continues. Hydrol. Sci. J. 47, 563–572. https://doi.org/10.1080/02626660209492960

Le Barbé, L., Lebel, T., 1997. Rainfall climatology of the HAPEX-Sahel region during the years 1950-1990. J. Hydrol. 188–189, 43–73. https://doi.org/10.1016/S0022-1694(96)03154-X

Le Barbé, L., Lebel, T., Tapsoba, D., 2002. Rainfall variability in West Africa during the years 1950-90. J. Clim. 15, 187–202. https://doi.org/10.1175/1520-0442(2002)015<0187:RVIWAD>2.0.CO;2

Lebel, T., 2003. Seasonal cycle and interannual variability of the Sahelian rainfall at hydrological scales. J. Geophys. Res. 108, 8389. https://doi.org/10.1029/2001JD001580

Lebel, T., Ali, A., 2009. Recent trends in the Central and WesternSahel rainfall regime (1990-2007). J. Hydrol. 375, 52–64. https://doi.org/10.1016/j.jhydrol.2008.11.030

Lehner, B., Liermann, C.R., Revenga, C., Vörömsmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J.C., Rödel, R., Sindorf, N., Wisser, D., 2011. High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. Front. Ecol. Environ. https://doi.org/10.1890/100125

Liebmann, B., Dole, R.M., Jones, C., Bladé, I., Allured, D., 2010. Influence of choice of time period on global surface temperature trends estimates. Bull. Am. Meteorol. Soc. 1485–1491. https://doi.org/10.1175/2010BAMS3030.1

Lienou, G., Mahe, G., Paturel, J.E., Servat, E., Sighomnou, D., Ekodeck, G.E., Dezetter, A., Dieulin, C., 2008. Evolution des régimes hydrologiques en région équatoriale camerounaise: Un impact de la variabilité climatique en Afrique équatoriale? Hydrol. Sci. J. 53, 789–801. https://doi.org/10.1623/hysj.53.4.789

Machiwal, D., Jha, M.K., 2006. Time series analysis of hydrologic data for water resources planning and management: a review. J. Hydrol. Hydromechanics 54, 237–257.

Mahé, G., L'Hôte, Y., Olivry, J.C., Wotling, G., 2001. Trends and discontinuities in regional rainfall of West and Central Africa: 1951–1989. Hydrol. Sci. J. 46, 211–226. https://doi.org/10.1080/02626660109492817

Mahé, G., Lienou, G., Descroix, L., Bamba, F., Paturel, J.E., Laraque, A., Meddi, M., Habaieb, H., Adeaga, O., Dieulin, C., Chahnez Kotti, F., Khomsi, K., 2013. The rivers of Africa: Witness of climate change and human impact on the environment. Hydrol. Process. 27, 2105–2114. https://doi.org/10.1002/hyp.9813

Mahé, G., Paturel, J.E., 2009. 1896-2006 Sahelian annual rainfall variability and runoff increase of Sahelian Rivers. Comptes Rendus - Geosci. 341, 538–546. https://doi.org/10.1016/j.crte.2009.05.002

Mahé, G., Paturel, J.E., Servat, E., Conway, D., Dezetter, A., 2005. The impact of land use change on soil water holding capacity and river flow modelling in the Nakambe River, Burkina-Faso. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2004.04.028

Maidment, R.I., Allan, R.P., Black, E., 2015. Recent observed and simulated changes in precipitation over Africa. Geophys. Res. Lett. 42, 8155–8164. https://doi.org/10.1002/2015GL065765

Mann, H.B., 1945. Nonparametric Tests Against Trend. Econometrica 13, 245. https://doi.org/10.2307/1907187

McCabe, G.J., Wolock, D.M., 2002. A step increase in streamflow in the conterminous United States. Geophys. Res. Lett. 29, 38-1-38–4. https://doi.org/10.1029/2002GL015999

McCartney, M., Forkuor, G., Sood, A., Amisigo, B., Hattermann, F., Muthuwatta, L., 2012. The water resource implications of changing climate in the Volta River Basin, IWMI Research Report. https://doi.org/10.5337/2012.219

Mitchell, T.D., Jones, P.D., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int. J. Climatol. 25, 693–712. https://doi.org/10.1002/joc.1181

Moalafhi, D.B., Sharma, A., Evans, J.P., 2017. Reconstructing hydro-climatological data using dynamical downscaling of reanalysis products in data-sparse regions – Application to the Limpopo catchment in southern Africa. J. Hydrol. Reg. Stud. 12, 378–395. https://doi.org/10.1016/j.ejrh.2017.07.001

Nicholson, S.E., 2013. The West African Sahel: A Review of Recent Studies on the Rainfall Regime and Its Interannual Variability. ISRN Meteorol. 2013, 1–32. https://doi.org/10.1155/2013/453521

Nicholson, S.E., Some, B. and Kone, B. (2000). An Analysis of Recent Rainfall Conditions in West Africa, Including the Rainy Seasons of the 1997 El Niño and the 1998 La Niña Years. Journal of Climate, 13, 2628-2640. https://dx.doi.org/10.1175/1520-0442(2000)013<2628:AAORRC>2.0.CO;2

Nka, B.N., Oudin, L., Karambiri, H., Paturel, J.E., Ribstein, P., 2015. Trends in floods in West Africa: Analysis based on 11 catchments in the region. Hydrol. Earth Syst. Sci. 19, 4707–4719. https://doi.org/10.5194/hess-19-4707-2015

Ogungbenro, S.B., Morakinyo, T.E., 2014. Rainfall distribution and change detection across climatic zones in Nigeria. Weather Clim. Extrem. 5, 1–6. https://doi.org/10.1016/j.wace.2014.10.002

Ozer, P., Erpicum, M., Demarée, G., Vandiepenbeeck, M., 2003. The Sahelian drought may have ended during the 1990s. Hydrol. Sci. J. 48, 489–492. https://doi.org/10.1623/hysj.48.3.489.45285

Page, E.S., 1954. Continuous Inspection Schemes. Biometrika 41, 100. https://doi.org/10.2307/2333009

Peterson, T.C., Vose, R.S., 1997. An Overview of the Global Historical Climatology Network Temperature Database. Bull. Am. Meteorol. Soc. 78, 2837–2849. https://doi.org/10.1175/1520-0477(1997)078<2837:AOOTGH>2.0.CO;2

Ringard, J., Dieppois, B., Rome, S., Diedhiou, A., Pellarin, T., Konaré, A., Diawara, A., Konaté, D., Dje, B.K., Katiellou, G.L., Seidou Sanda, I., Hassane, B., Vischel, T., Garuma, G.F., Mengistu, G., Camara, M., Diongue, A., Gaye, A.T., Descroix, L., 2016. The intensification of thermal extremes in west Africa. Glob. Planet. Change 139, 66–77. https://doi.org/10.1016/j.gloplacha.2015.12.009

Roudier, P., Ducharne, A., Feyen, L., 2014. Climate change impacts on runoff in West Africa: A review. Hydrol. Earth Syst. Sci. https://doi.org/10.5194/hess-18-2789-2014

877
878 Rubin, D.B., 1987. Introduction, in: Multiple Imputation for Nonresponse in Surveys. pp. 1–26.
879 https://doi.org/10.1002/9780470316696.ch1
880
881 Schönbrodt, F.D., Perugini, M., 2013. At what sample size do correlations stabilize? J. Res. Pers. 47, 609–612.
882 https://doi.org/10.1016/j.jrp.2013.05.009
883
884 Stekhoven, D.J., Bühlmann, P., 2012. Missforest-Non-parametric missing value imputation for mixed-type data.
885 Bioinformatics 28, 112–118. https://doi.org/10.1093/bioinformatics/btr597
886
887 Stott, P.A., Gillett, N.P., Hegerl, G.C., Karoly, D.J., Stone, D.A., Zhang, X., Zwiers, F., 2010. Detection and attribution
888 of climate change: A regional perspective. Wiley Interdiscip. Rev. Clim. Chang. 1, 192–211.
889 https://doi.org/10.1002/wcc.34
890
891 Suzuki, R., Shimodaira, H., 2006. Pvclust: An R package for assessing the uncertainty in hierarchical clustering.
892 Bioinformatics 22, 1540–1542. https://doi.org/10.1093/bioinformatics/btl117
893
894 Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., Altman, R.B., 2001.
895 Missing value estimation methods for DNA microarrays. Bioinformatics 17, 520–525.
896 https://doi.org/10.1093/bioinformatics/17.6.520
897
898 Tshimanga, R.M., Tshitenge, J.M., Kabuya, P., Alsdorf, D., Mahe, G., Kibukusa, G., Lukanda, V., 2016. A Regional
899 Perceptive of Flood Forecasting and Disaster Management Systems for the Congo River Basin, in: Flood
900 Forecasting: A Global Perspective. pp. 87–124. https://doi.org/10.1016/B978-0-12-801884-2.00002-5
901
902 University of East Anglia Climatic Research Unit; Harris, I.C.; Jones, P.D. (2017): CRU TS4.00: Climatic Research
903 Unit (CRU) Time-Series (TS) version 4.00 of high resolution gridded data of month-by-month variation in climate
904 (Jan. 1901- Dec. 2015). Centre for Environmental Data Analysis, *25 May 2017*
905
906 Van Buuren, S., Oudshoorn, K. (1999). Flexible multivariate imputation by MICE. *Leiden, The Netherlands: TNO*
907 *Prevention Center*.
908
909 Von Storch, H., 1995. Misuses of Statistical Analysis in Climate. Anal. Clim. Var. Appl. Stat. Tech. 11–26.
910 https://doi.org/10.1007/978-3-662-03744-7
911
912 Webster, P.J., Magaña, V.O., Palmer, T.N., Shukla, J., Tomas, R.A., Yanai, M., Yasunari, T., 1998. Monsoons:
913 Processes, predictability, and the prospects for prediction. J. Geophys. Res. Ocean. 103, 14451–14510.
914 https://doi.org/10.1029/97JC02719
915
916 Yue, S., Pilon, P., Phinney, B., Cavadias, G., 2002. The influence of autocorrelation on the ability to detect trend in
917 hydrological series. Hydrol. Process. 16, 1807–1829. https://doi.org/10.1002/hyp.1095
918

919 HIGHLIGHTS

920 • The first imputed streamflow dataset for West and central Africa

921 • Good agreement between historical trends in streamflow and rainfall

922 • Partial modulations of post-1990s rainfall recovery by enhancing evapotranspiration

923 • Decadal modulations of Trends in hydroclimatic trends

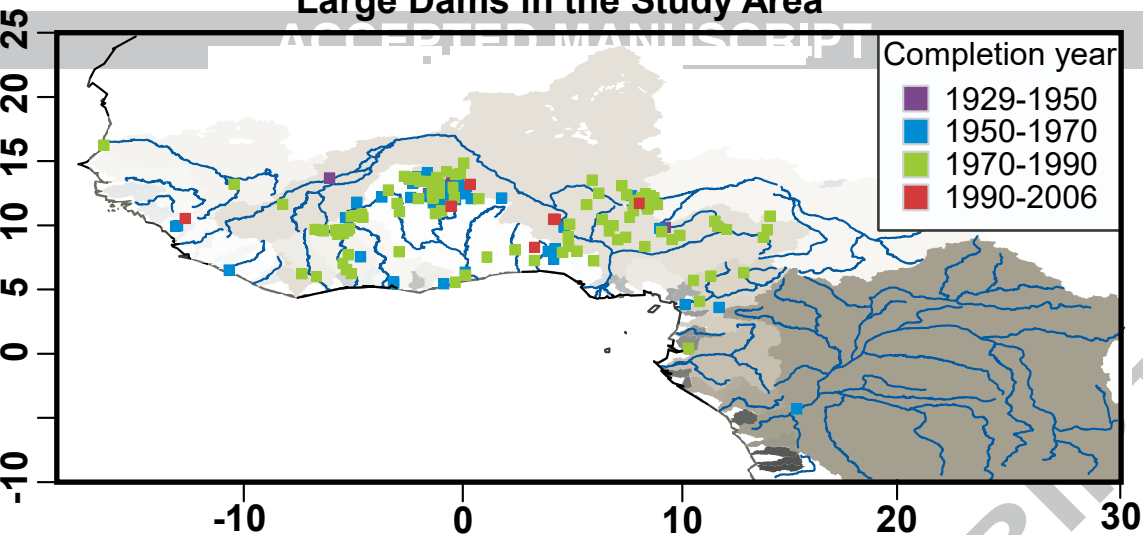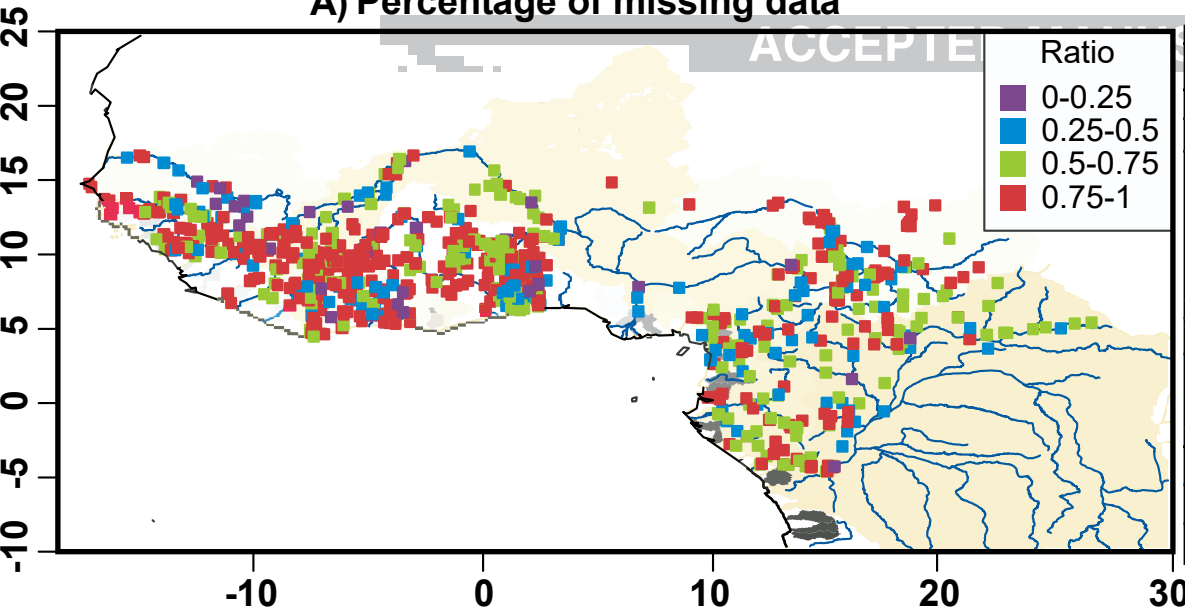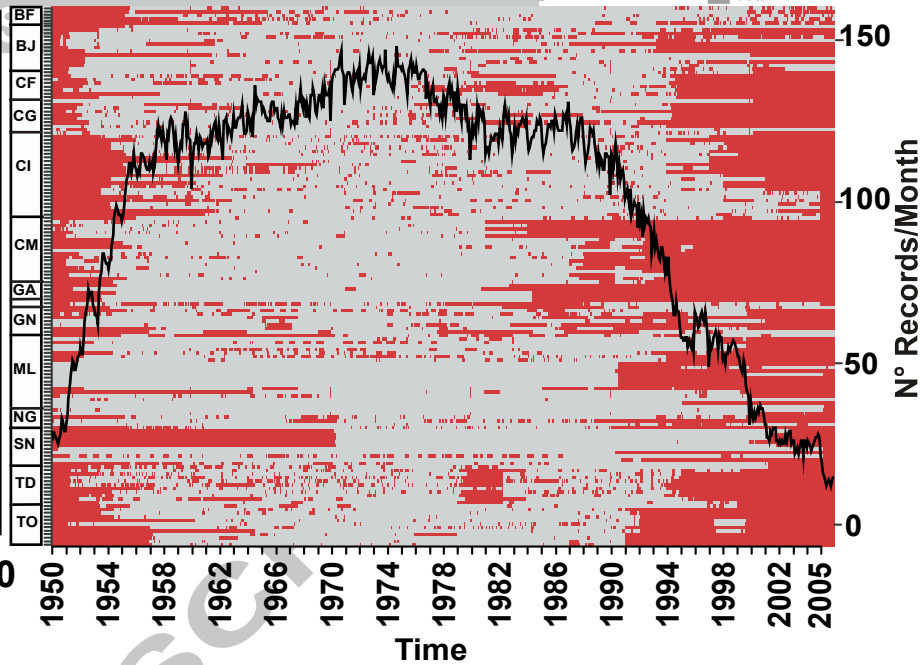924 • Homogenous zones of streamflow and precipitation variability

925
926

# STUDY AREA

- Gauging stations
- Rivers
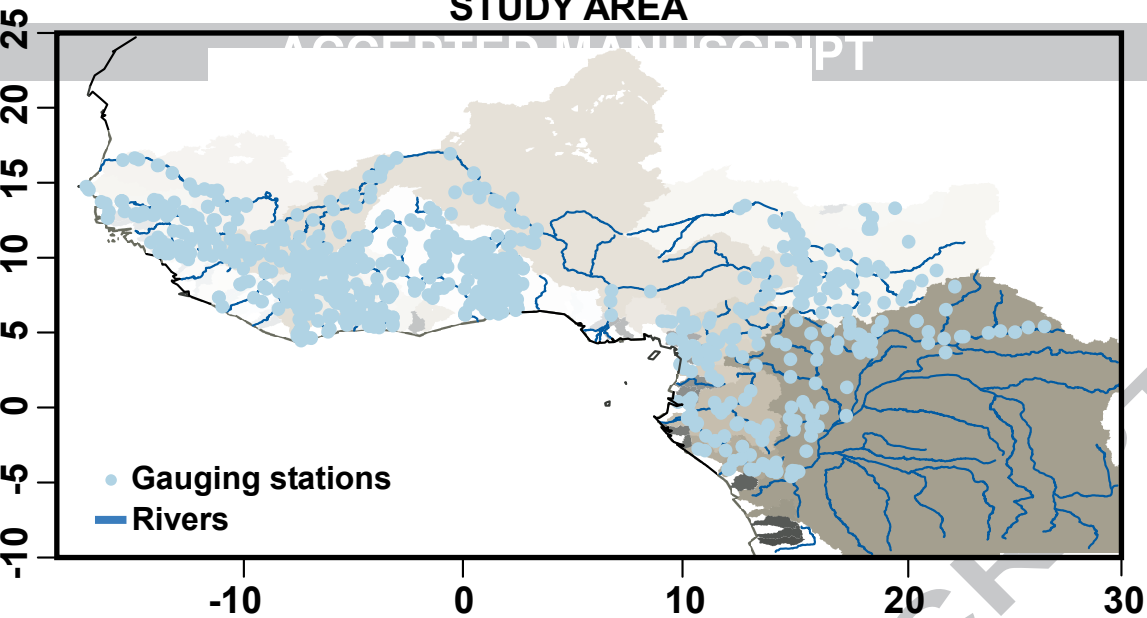
**Large Dams in the Study Area**

| Completion year | |
|---|---|
| 1929-1950 | (purple) |
| 1950-1970 | (blue) |
| 1970-1990 | (green) |
| 1990-2006 | (red) |

**A) Percentage of missing data**

Ratio
- 0-0.25
- 0.25-0.5
- 0.5-0.75
- 0.75-1

**B) Location of missing values**

No data
data

N° Records/Month

Time

Periods
1950-1970
R²=0.996
1970-1993
R²=0.99
1993-2005
R²=0.76

**MICE: Trends 1950-1970**     **RF: Trends 1950-1970**

a)

- ▲ Positive trend significant at 10%
- △ Positive trend
- ▽ Negative trend
- ▼ Negative trend significant at 10%

b)

- ▲ Positive trend significant at 10%
- △ Positive trend
- ▽ Negative trend
- ▼ Negative trend significant at 10%

**MICE: Trends 1970-1993**     **RF: Trends 1970-1993**

c)

- ▲ Positive trend significant at 10%
- △ Positive trend
- ▽ Negative trend
- ▼ Negative trend significant at 10%

d)

- ▲ Positive trend significant at 10%
- △ Positive trend
- ▽ Negative trend
- ▼ Negative trend significant at 10%

**MICE: Trends 1993-2005**     **RF: Trends 1993-2005**

e)

- ▲ Positive trend significant at 10%
- △ Positive trend
- ▽ Negative trend
- ▼ Negative trend significant at 10%

f)

- ▲ Positive trend significant at 10%
- △ Positive trend
- ▽ Negative trend
- ▼ Negative trend significant at 10%

**Clusters of streamflow variability**
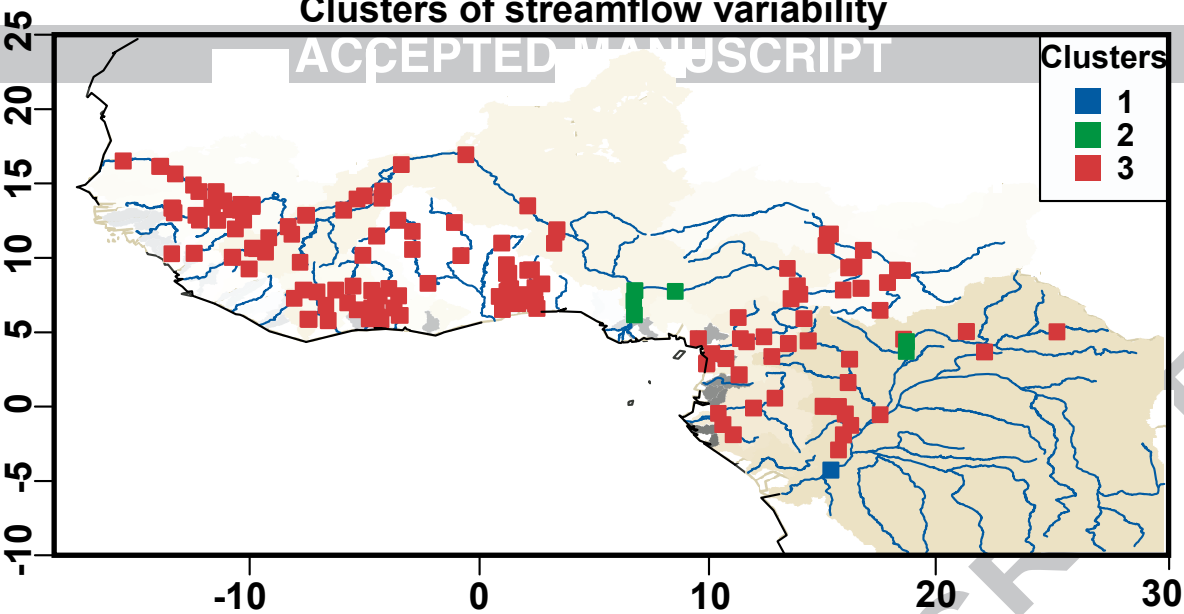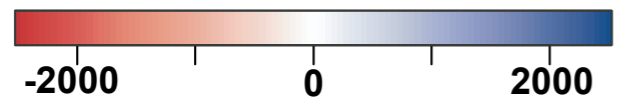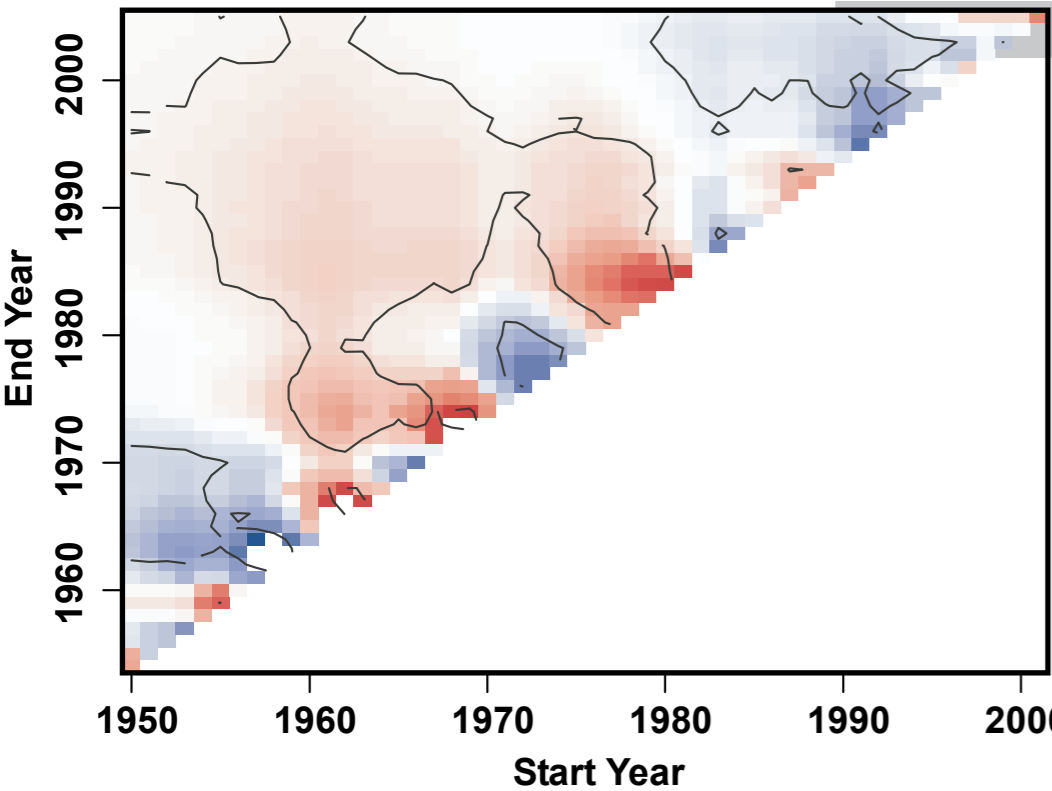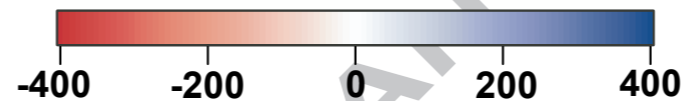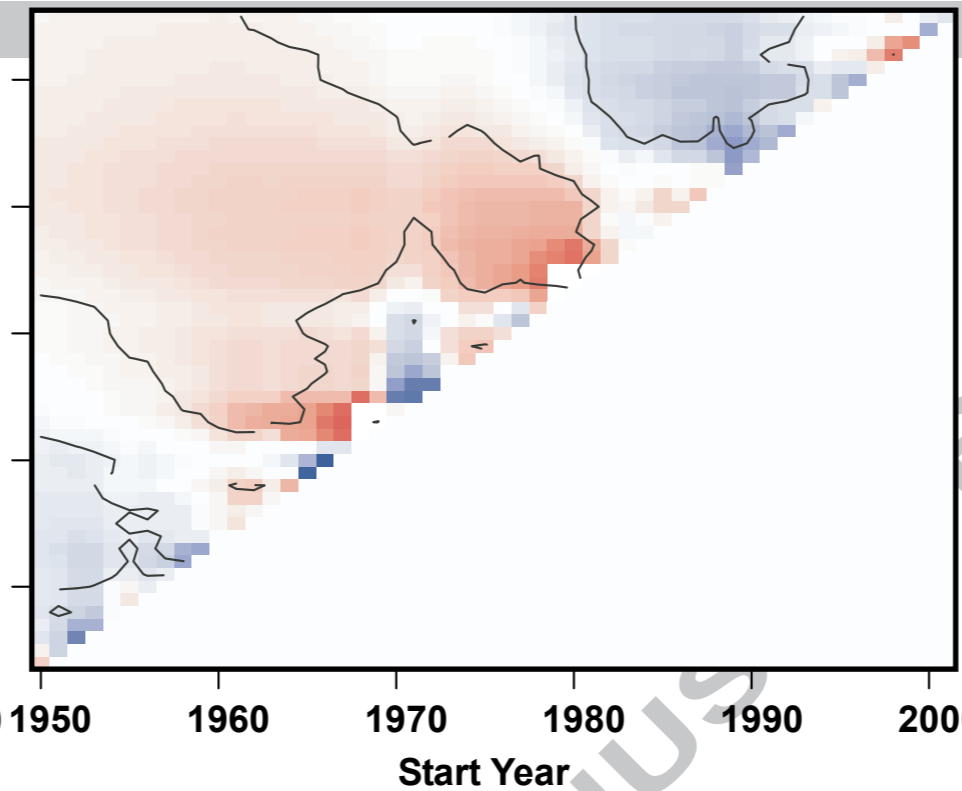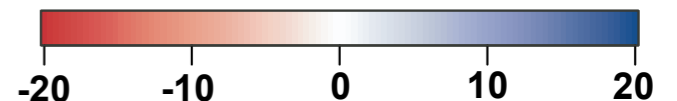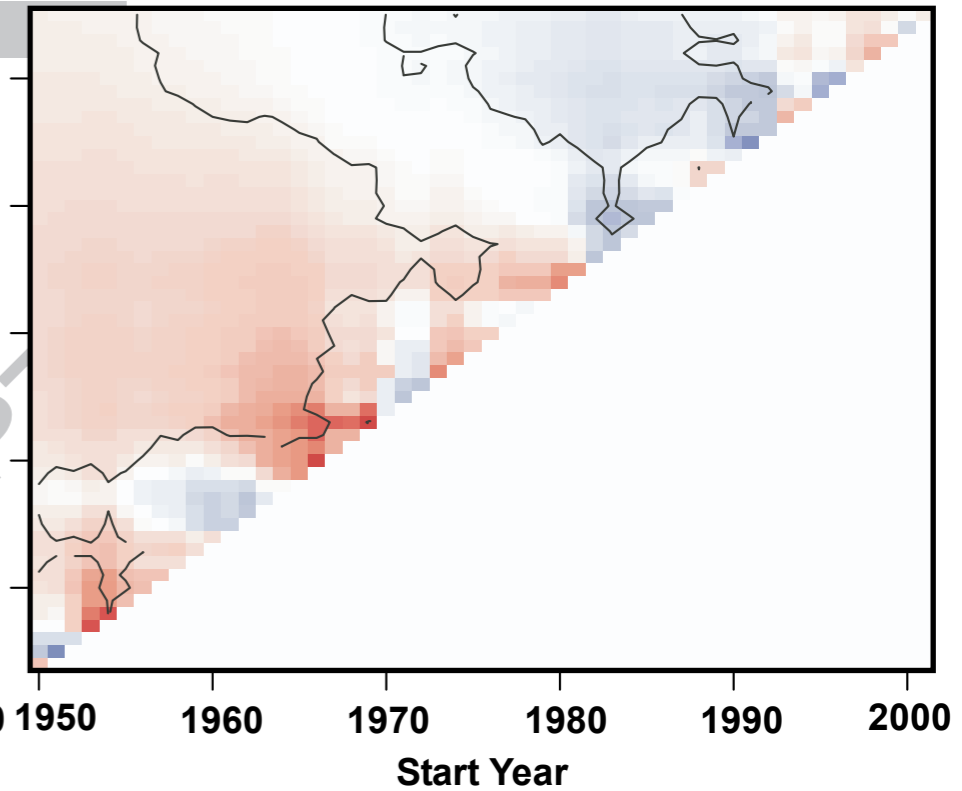
Clusters
- 1
- 2
- 3

**Spatial distribution of rainfall clusters**

Clusters

| | | |
|---|---|---|
| ■ 1 | ■ 5 | ■ 9 | ■ Unclassified |
| ■ 2 | ■ 6 | ■ 10 |
| ■ 3 | ■ 7 | ■ 11 |
| ■ 4 | ■ 8 | ■ 12 |