



Sensitivity of SPI to Distribution Functions and Correlation Between its Values at Different Time Scales in Central Africa

G. M. Guenang¹ · M. A. J. Komkoua¹ · M. W. Pokam² · R. S. Tanessong³ · S. A. Tchakoutio⁴ · A. Vondou⁵ · A. T. Tamoffo⁵ · L. Djiotang⁵ · Z. Yepdo⁶ · K. F. Mkankam⁷

Received: 20 September 2018 / Accepted: 10 June 2019 / Published online: 22 June 2019
© King Abdulaziz University and Springer Nature Switzerland AG 2019

Abstract

This paper studies sensitivity of Standardized Precipitation Index (SPI) to statistical distribution functions used in SPI computation procedure, in order to find out which are more appropriate and to assess SPI shift if using inappropriate distribution functions. Results may explain one of the reasons why spatial SPI computed with unique distribution function as usually, sometimes does not give better drought description. Central Africa is chosen as the study area because of its importance in climate change perspective and the necessity to use improved tools for drought quantification in this region. Monthly precipitation data for the period 1951–2016, both from the Climatic Research Unit (CRU) and Global Precipitation Climatology Centre (GPCC) were used. They were first aggregated at various time scales and four statistical distribution functions (gamma, weibull, exponential and lognormal) were tested to select the best-fit. Next, SPI was calculated for various time scales using the best fit function at each grid point and results were compared to those computed assuming a same distribution function at all grid points. Results show that, from 1- to 9-month time scales, observed spatial patterns of distribution functions were more homogeneous and the weibull function had the highest extended spatial rate, followed successively by gamma, lognormal and exponential. From 12-month time scale, spatial patterns were inhomogeneous, and no gridded precipitation followed the exponential function. The study of cross-correlations showed significant resemblance between SPIs at different time scales, leading to reduce the studies from 1- to 15-month. SPI values were affected if inappropriate distribution functions were used and the shift increases in correlation with the increase of time scale. The two datasets CRU and GPCC showed similar results, but GPCC's SPIs were wetter and distribution functions somewhat more dispersed spatially from 12-month time scale. For SPI calculation in Central Africa, the weibull and gamma functions if used lead to good SPI results at short time scales (not more than 9-month) compared to SPI calculated using exponential or lognormal function. From 12-month time scale, it is recommended to choose the best fit distribution function at each grid point in order to expect good SPI results and then better drought description.

Keywords Statistical distribution function · Standardized Precipitation Index (SPI) · Drought · Central Africa

✉ G. M. Guenang
merlin.guenang@yahoo.fr

¹ Laboratory of Mechanics and Modeling of Physical Systems, Department of Physics, Faculty of Science, University of Dschang, Dschang, Cameroon

² Department of Physics, Higher Teacher Training College, University of Yaounde 1, Yaounde, Cameroon

³ School of Wood, Water and Natural Resources, Faculty of Agronomy and Agricultural Sciences, University of Dschang, Ebolowa, Cameroon

⁴ Department of Computer Science, Higher Technical Teachers Training College, University of Buea, Kumba, Cameroon

⁵ Laboratory for Environmental Modelling and Atmospheric Physics, Department of Physics, University of Yaounde 1, Yaounde, Cameroon

⁶ National Institute of Cartography, Yaounde, Cameroon

⁷ Université des Montagnes, Bangangté, Cameroon

1 Introduction

Prolonged drought leading to vegetation destruction and loss in biological and economic productivity of the dry land is cited by WMO (2006) as one of the causes of desertification. Given the ever-increasing threats of this phenomenon in many areas of the world as a result of climate change, scientists have invested in the quantification and categorization of drought for forecasting purposes. Despite these efforts on the part of the scientific community, one of the remaining challenges is to develop and improve efficient drought tools.

There are several indices developed for the quantification of drought, for instance the Palmer drought severity index (PDSI) (Palmer 1965), the standardized precipitation index (SPI) (McKee et al. 1993), the rainfall anomaly index (RAI), Bhalme and Mooley drought index (BMDI) and the Palmer drought index (PDI) (Oladipio 1985), the Z-score or standardized rainfall anomalies (Jones and Hulme 1996), the standardized precipitation evapotranspiration index (SPEI) (Sergio et al. 2008) and the precipitation anomaly index (PAI) (Zhang et al. 2013). Many other indices can be found in Amin et al. (2011). Several studies comparing results of these indices exist in the scientific literature (Oladipio 1985; Guttman 1998; Szalai and Szinell 2000; Lloyd-Hughes and Saunders 2002; Majid et al. 2017; Gustavo et al. 2018). It appears that these indices are efficient in quantifying drought, but with performances often changing slightly from one index to another (Alley 1984; Oladipio 1985), SPI being often more robust (Okpara and Tarhule 2015) and the PDSI sometimes problematic (Oladipio 1985; Alley 1984). Of all these indices, SPI, a probability index formulated by McKee et al. (1993) presents many advantages (Hayes et al. 1999): it is the most common and applicable tool, mainly in areas where access to data other than rainfall is difficult or even impossible (e.g., many African countries). Furthermore, it is not affected by topography and is suitable for quantifying most types of droughts (Lana et al. 2001; Szalai and Szinell 2000).

SPI which is the index studied in this paper for quantifying drought, needs first an investigation on appropriate statistical function fitting precipitation distribution. The gamma function has been found to fit the precipitation distribution quite well (Sharma and Singh 2010; Alvarez et al. 2011); it provides the best model for describing monthly precipitation over most of Europe (Lloyd-Hughes and Saunders 2002). In the region of West-Africa, particularly in the Niger River Basin, Okpara and Tarhule (2016) fitted to monthly rainfall series, five different distribution functions (lognormal, exponential, log-logistic,

Weibull and gamma) and the results of the case study showed the gamma-type two distributions to be the best fit over the upper Niger sub-basin. Other studies using SPI as drought indicator used the gamma distribution as the best fit only because it is commonly used (Dutra et al. 2013). Likewise, Okpara et al. (2017) also used this distribution (Gamma probability distribution type 2) to fit gauged-based monthly rainfall and derived SPI was used to investigate the potentials of SPI as standard measure for meteorological drought in West Africa. However, SPI derived from gamma distribution shows some good results. In Zambia, it successfully categorized extremely dry years (1992 and 2015), severely dry year (1995), moderately dry years (1972, 1980, 1987, 1999 and 2005) and 26 near normal years Libanda et al. (2019). Naumann et al. (2014) investigated the capability of different data sets (the ECMWF ERA-Interim reanalysis, the Tropical Rainfall Measuring Mission satellite monthly rainfall product 3B-43, the Global Precipitation Climatology Centre gridded precipitation data set, the Global Precipitation Climatology Project Global Monthly Merged Precipitation Analyses, and the Climate Prediction Center Merged Analysis of Precipitation) and drought indicators (the Standardized Precipitation Index, the Standardized Precipitation-Evaporation Index, and Soil Moisture Anomalies) to improve drought monitoring in Africa. The gamma function was chosen as distribution to fit precipitation and derive SPI. They found that for the areas affected by drought, all the drought indicators agree on the time of drought onset and recovery, and show higher uncertainties in regions with limited rain gauge data.

In many studies, the precipitation record is directly fitted to a gamma distribution before deriving SPI (Gidey et al. 2018; Mustafa and Rahman 2018; Spinoni et al. 2014; Amin et al. 2011; Lloyd-Hughes and Saunders 2002). Nevertheless, this function is not always the best-fit distribution in some areas (Guenang and Mkankam 2014) and very few studies consider this preliminary step that skipping may generate wrong SPI values.

This paper aims for providing appropriate distribution functions that better fit multi-scale aggregated precipitation at each grid point in Central Africa and investigate on the derived SPIs if using non appropriate distribution functions; it also investigates the possibility to reduce the number of time scales in a multiscalar drought study to guide users on the choice of a minimum and efficient time scales for its activity sector. Study domain, data and methodology are described in the next section. Sections 3 and 4 present the results and discussion, respectively, and the work ends with conclusions.

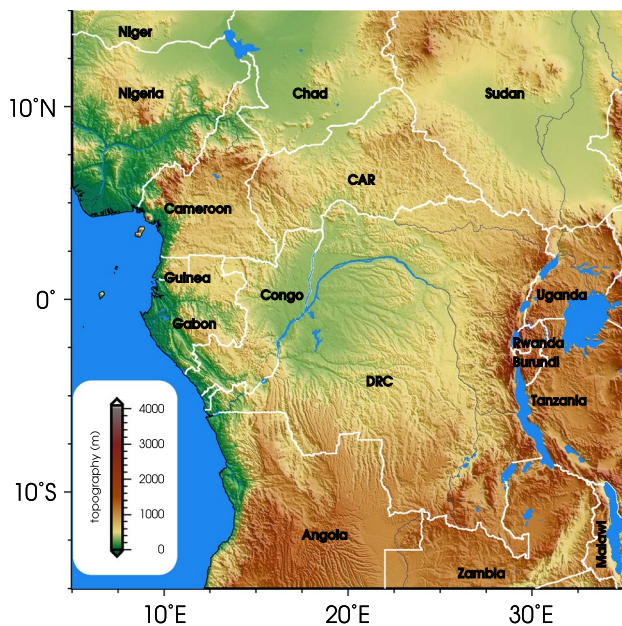


Fig. 1 Study area and topography

2 Study Domain, Data and Methodology

2.1 Study Domain

The domain of this study is delimited by the coordinates 15°S–15°N and 5°E–35°E (Fig. 1) comprising southern part of Niger, Nigeria, Chad, Sudan, Cameroon, Central African Republic (CAR), Guinea, Gabon, Congo, The Democratic Republic of Congo (DRC), Angola, Zambia, Malawi, Tanzania, Burundi, Rwanda, and Uganda. Figure 1 show that the topography is diversified with highlands, basins, plateaus and plains. Its hydrological network is also important. The northern part of the domain (> 7°N of latitude and including northern Nigeria, Chad, north Cameroon and Sudan) is subject to desert conditions and the southern part to tropical rain-forest. The intermediate zone is dominated by an equatorial climate. More information about the description of the region can be found in Fotso-Nguemo et al. (2016). This area is a center of interest for the study of drought as its plays an important role for the global climate change perspective. Hence the need to develop and improve drought quantification tools for control purposes and alerts in case of risk.

This domain is located between the tropics where solar radiations are constantly high during the year (Qiang 2003). The southern part of the domain is the Congo basin known as the second largest rainforest in the world, a significant carbon sink in global climate perspective. The northern part (Sahel region) is defined as an arid region. The western part is bordered by Atlantic Ocean. The relief of the area is diversified and characterized by many high lands.

2.2 Data Used

Two datasets ranging from 1951 to 2016 were used in this study:

- Gridded monthly precipitation data version 4.01, from the Climatic Research Unit (CRU) released the 20 September 2017 (Harris et al. 2014). They have $0.5^\circ \times 0.5^\circ$ spatial resolution and cover the time period ranging from 1951 to 2010. These data (CRU TS 4.01) are downloadable free of charge from the CRU website. Version 3 of these data was used in a part of this region especially in Cameroon to compute SPI and results were satisfactory similar to those from observation stations (Guenang and Mkankam 2014). CRU data have also been used successfully for computing total water storage change and establishing the link with onset and retreat dates of the rainy season in the same area (Guenang et al. 2016).
- Full monthly precipitation dataset version 8, from the Global Precipitation Climatology Centre (GPCP) which is a German contribution to the World Climate Research Program (WCRP) and to the Global Climate Observing System (GCOS) (Schneider et al. 2016, 2017, 2018). These data based on quality-controlled data from all stations in GPCP's data base are on a regular grid with a spatial resolution of $0.5^\circ \times 0.5^\circ$ latitude by longitude. They are optimized for best spatial coverage and used for water budget studies and available through the GPCP Product Access Page.

2.3 Methodology

First, gridded monthly precipitation time series (CRU TS 4.01) were aggregated on 1-, 3-, 6-, 9-, 12-, ..., 48-month time scales. Second, four probability distribution functions (gamma, weibull, exponential and lognormal) estimated by the method of the maximum-likelihood (ML) (Thom 1958) were tested on time series of each aggregated precipitation and the Kolmogorov–Smirnov goodness-of-fit test statistic (Frank and Massey 1951) was applied.

The Kolmogorov–Smirnov test (KS-test) is a non-parametric test used to decide if a sample comes from a population following a specific distribution. The KS-test statistic, D , for a given cumulative distribution function F and empirical cumulative distribution function F_n is defined as the maximum absolute difference between $F(x)$ and $F_n(x)$ for the same x .

$$D = \sup_{-\infty < x < \infty} |F_n(x) - F(x)|, \quad (1)$$

where n is the number of independent and identically distributed observations x of a continuous variable, and

Table 1 Drought classification according to SPI ranges

| SPI range | Drought category |
|---------------------------------|------------------|
| $-0.99 \leq \text{SPI} < 0$ | Mild drought |
| $-1.49 \leq \text{SPI} < -1$ | Moderate drought |
| $-1.99 \leq \text{SPI} < -1.50$ | Severe drought |
| $\text{SPI} \leq -2$ | Extreme drought |

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{(-\infty, x]}(X_i) \tag{2}$$

According to Glivenko–Cantelli theorem (Howard 1959), D almost converges to zero for $n \rightarrow \infty$, if empirical sample data follow the distribution F . At each grid point, the

best-fit distribution function among the four used was that having the lowest KS-test statistic. Next, the best-fit function was used to calculate the cumulative distribution of data, which was finally transformed into standardized normal variate (Abramowitz and Stegun 1965) of zero mean and unity variance. This Standardized Precipitation Index quantifies precipitation deficits and value less 0 (zero) indicates a drought condition. The drought severity is categorized into four classes shown in Table 1. More details on the method, theory and the main mathematical tools used can be found in Guttman (1999), Lloyd-Hughes and Saunders (2002) and Guenang and Mkankam (2014).

Spatial SPIs using each of the four distribution functions at all grid points were computed and compared to those derived from appropriate distribution functions at each grid point in order to evaluate the impact of the choice of

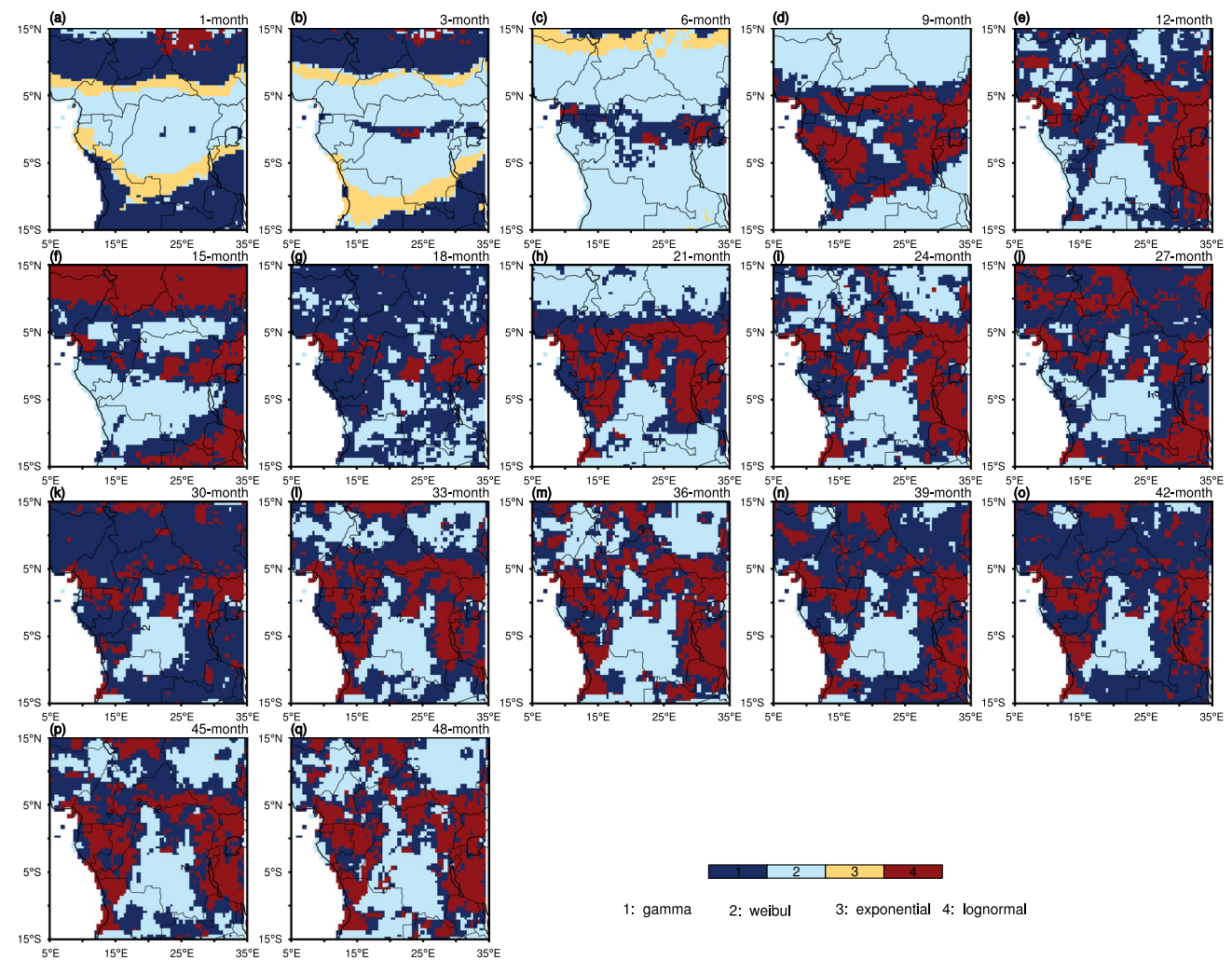


Fig. 2 Distribution functions fitting several time scales aggregated precipitation from CRU and for the time period 1951–2016. The time scales concerned are 1-, 3-, 6-, ..., 48-month. Four distribution func-

tions were tested and the best fit at each grid point was represented. So four different colors were used for the four distribution functions

distribution on SPI values. A classification of drought events according to SPI ranges (Lloyd-Hughes and Saunders 2002) was adopted (Table 1). The interpretation of SPI is done on the basis of the following illustration: the 3-month SPI provides a comparison of the precipitation over a specific 3-month period with the precipitation totals from the same 3-month period for all the years included in the historical record. Similar interpretation was done for n -month SPI value, where n is the number of months used as time scale.

After, calculation of cross-correlations between multiscale SPIs were undertaken to objectively reduce the number of time scales without losing important information. So, the total surface average of SPIs for each time scale was calculated and used for computing cross-correlation.

3 Results

3.1 Determination of Statistical Distribution Functions Fitting Various Time Scales Aggregated Precipitation

Figures 2 and 3 show the spatial distribution of statistical functions that better fit the time series of precipitation over the period 1951–2016, for several time scales ranging from 1- to 48-month, and for CRU and GPCP precipitation data, respectively.

For short time scales ranging from 1- to 6-month (Fig. 2a, b, c), the four distribution functions tested are spatially well represented, but with a strong preponderance of gamma and weibull functions at 1- and 3-month, and only weibull largely predominates at 6-month, the lognormal function

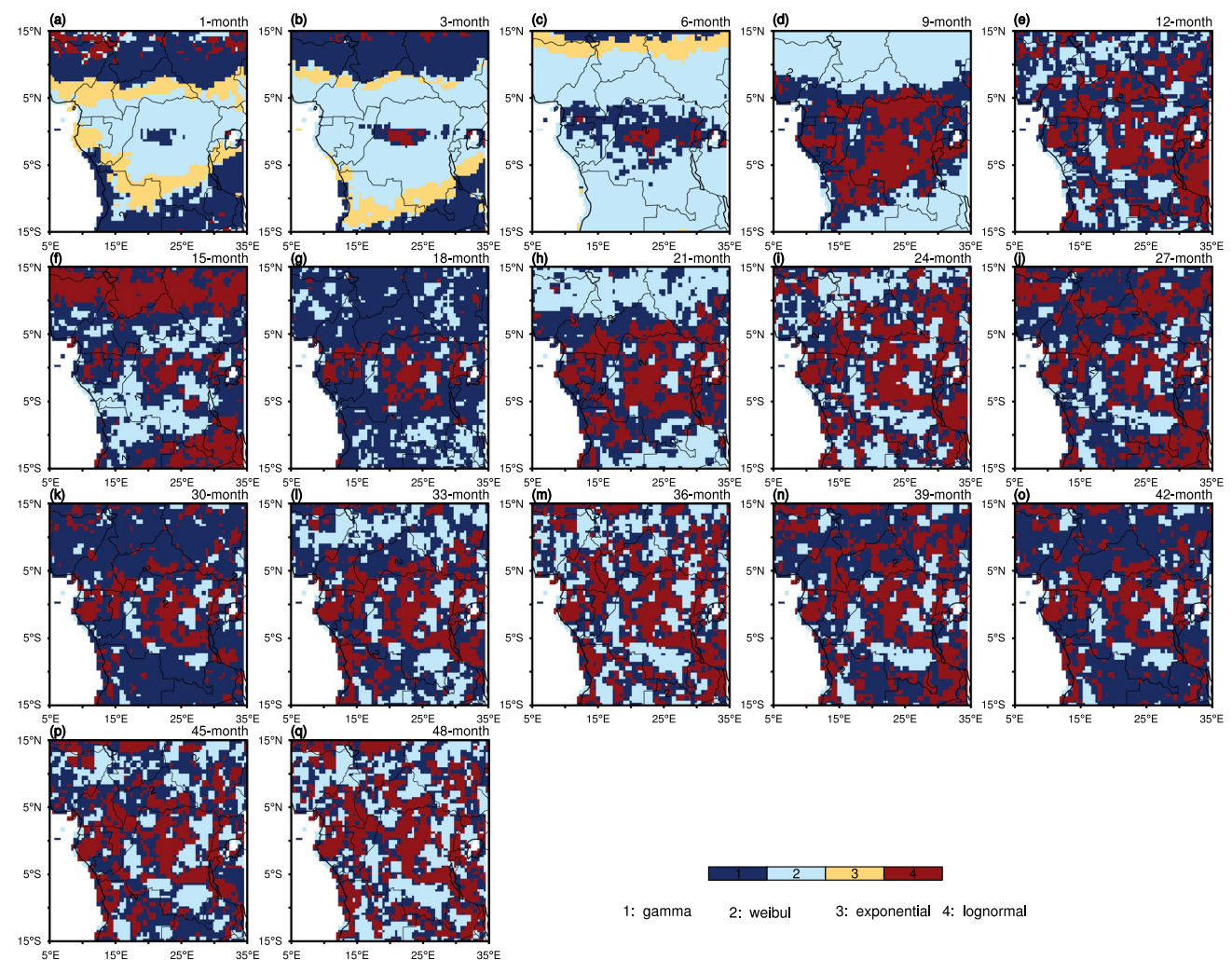


Fig. 3 Distribution functions fitting several time scales aggregated precipitation from GPCP and for the time period 1951–2016. The time scales concerned are 1-, 3-, 6-, ..., 48-month. Four distribution

functions were tested and the best fit at each grid point was represented. So four different colors were used for the four distribution functions

Table 2 Cross-correlation between SPIs computed with CRU precipitation data for several time scales and for total area

| SPI _n | 1 | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 27 | 30 | 33 | 36 | 39 | 42 | 45 | 48 |
|------------------|---|-------|--------------|--------|-------|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 1 | 1 | 0.294 | 0.757 | 0.133 | 0.013 | 0.070 | 0.684 | 0.111 | -0.009 | 0.042 | 0.571 | 0.083 | -0.009 | 0.033 | 0.490 | 0.081 | -0.002 |
| 3 | | 1 | 0.209 | 0.458 | 0.042 | -0.209 | 0.189 | 0.368 | 0.008 | -0.165 | 0.133 | 0.256 | -0.007 | -0.124 | 0.108 | 0.211 | -0.013 |
| 6 | | | 1 | -0.017 | 0.031 | 0.246 | 0.854 | -0.021 | -0.010 | 0.188 | 0.727 | 0.002 | -0.020 | 0.141 | 0.618 | 0.002 | -0.027 |
| 9 | | | | 1 | 0.154 | -0.473 | 0.118 | 0.813 | 0.049 | -0.339 | 0.126 | 0.637 | 0.058 | -0.254 | 0.125 | 0.508 | 0.029 |
| 12 | | | | | 1 | 0.284 | 0.150 | 0.290 | 0.682 | 0.378 | 0.245 | 0.354 | 0.520 | 0.325 | 0.223 | 0.302 | 0.371 |
| 15 | | | | | | 1 | 0.281 | -0.314 | 0.231 | 0.845 | 0.290 | -0.196 | 0.149 | 0.671 | 0.253 | -0.161 | 0.105 |
| 18 | | | | | | | 1 | 0.234 | 0.149 | 0.311 | 0.910 | 0.245 | 0.116 | 0.264 | 0.799 | 0.207 | 0.082 |
| 21 | | | | | | | | 1 | 0.374 | -0.100 | 0.310 | 0.855 | 0.304 | -0.031 | 0.314 | 0.718 | 0.237 |
| 24 | | | | | | | | | 1 | 0.516 | 0.325 | 0.489 | 0.815 | 0.528 | 0.357 | 0.473 | 0.647 |
| 27 | | | | | | | | | | 1 | 0.453 | 0.061 | 0.440 | 0.875 | 0.441 | 0.085 | 0.355 |
| 30 | | | | | | | | | | | 1 | 0.433 | 0.320 | 0.450 | 0.919 | 0.393 | 0.265 |
| 33 | | | | | | | | | | | | 1 | 0.565 | 0.197 | 0.484 | 0.888 | 0.478 |
| 36 | | | | | | | | | | | | | 1 | 0.628 | 0.437 | 0.597 | 0.855 |
| 39 | | | | | | | | | | | | | | 1 | 0.566 | 0.276 | 0.573 |
| 42 | | | | | | | | | | | | | | | 1 | 0.552 | 0.436 |
| 45 | | | | | | | | | | | | | | | | 1 | 0.659 |
| 48 | | | | | | | | | | | | | | | | | 1 |

Values showing significant correlation between SPIs are indicated in bold

Table 3 Cross-correlation between SPIs computed with GPCP precipitation data for several time scales and for total area

| SPI _n | 1 | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 27 | 30 | 33 | 36 | 39 | 42 | 45 | 48 |
|------------------|---|-------|--------------|--------|-------|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 1 | 1 | 0.311 | 0.775 | 0.178 | 0.026 | 0.045 | 0.736 | 0.165 | 0.022 | 0.032 | 0.634 | 0.136 | 0.029 | 0.037 | 0.561 | 0.126 | 0.029 |
| 3 | | 1 | 0.196 | 0.491 | 0.082 | -0.244 | 0.192 | 0.399 | 0.073 | -0.176 | 0.157 | 0.312 | 0.063 | -0.115 | 0.155 | 0.269 | 0.053 |
| 6 | | | 1 | -0.029 | 0.067 | 0.267 | 0.866 | -0.019 | 0.026 | 0.213 | 0.759 | 0.008 | 0.022 | 0.173 | 0.667 | 0.020 | 0.022 |
| 9 | | | | 1 | 0.193 | -0.490 | 0.130 | 0.836 | 0.105 | -0.361 | 0.144 | 0.674 | 0.106 | -0.257 | 0.150 | 0.555 | 0.097 |
| 12 | | | | | 1 | 0.323 | 0.184 | 0.365 | 0.774 | 0.416 | 0.273 | 0.439 | 0.676 | 0.425 | 0.300 | 0.444 | 0.596 |
| 15 | | | | | | 1 | 0.274 | -0.311 | 0.283 | 0.868 | 0.292 | -0.181 | 0.226 | 0.723 | 0.283 | -0.114 | 0.192 |
| 18 | | | | | | | 1 | 0.237 | 0.174 | 0.304 | 0.925 | 0.242 | 0.151 | 0.276 | 0.833 | 0.221 | 0.129 |
| 21 | | | | | | | | 1 | 0.419 | -0.109 | 0.319 | 0.875 | 0.367 | -0.014 | 0.343 | 0.760 | 0.324 |
| 24 | | | | | | | | | 1 | 0.527 | 0.331 | 0.541 | 0.871 | 0.567 | 0.395 | 0.575 | 0.784 |
| 27 | | | | | | | | | | 1 | 0.438 | 0.060 | 0.480 | 0.899 | 0.452 | 0.123 | 0.427 |
| 30 | | | | | | | | | | | 1 | 0.420 | 0.337 | 0.448 | 0.935 | 0.398 | 0.302 |
| 33 | | | | | | | | | | | | 1 | 0.605 | 0.212 | 0.493 | 0.910 | 0.543 |
| 36 | | | | | | | | | | | | | 1 | 0.653 | 0.464 | 0.670 | 0.913 |
| 39 | | | | | | | | | | | | | | 1 | 0.563 | 0.314 | 0.621 |
| 42 | | | | | | | | | | | | | | | 1 | 0.553 | 0.464 |
| 45 | | | | | | | | | | | | | | | | 1 | 0.707 |
| 48 | | | | | | | | | | | | | | | | | 1 |

Values showing significant correlation between SPIs are indicated in bold

being poorly represented. At 9-month time scale and above, areas where precipitation time series are described by these functions decrease in the benefit of the lognormal function which below 9-month time scale was very weakly represented. For 12-month time scale and more, the lognormal function predominates except at 18- and 21-month where

both gamma and weibull outperform, respectively (Fig. 2g, h). We note that between 12- and 48-month, the three functions, namely gamma, weibull and exponential are all represented at competing portions, but spatially scattered. The exponential function has the lowest spatial representation at 1-, 3- and 6-month, but none of the gridded time series

Table 4 Correspondence between SPIs at different time scales and for CRU precipitation data

| SPI_n | SPI_n Correlated to | |
|--------------|--|-----------------------|
| | SPI_n in 1st column | SPI_n in 2nd column |
| SPI_1^* | SPI_6 (76%); SPI_{18} (68%); SPI_{30} (57%) | |
| SPI_3^* | | |
| SPI_6^* | SPI_{18} (85%); SPI_{30} (73%); SPI_{42} (62%) | |
| SPI_9^* | SPI_{21} (81%); SPI_{33} (64%); SPI_{45} (51%) | |
| SPI_{12}^* | SPI_{24} (68%); SPI_{36} (52%) | |
| SPI_{15}^* | SPI_{27} (85%); SPI_{39} (67%) | |
| SPI_{18} | SPI_6^* (85%); SPI_1^* (68%) | |
| SPI_{21} | SPI_9^* (81%) | |
| SPI_{24} | SPI_{12}^* (68%) | |
| SPI_{27} | SPI_{15}^* (85%) | |
| SPI_{30} | SPI_{18} (91%); SPI_6^* (73%); SPI_1^* (57%) | |
| SPI_{33} | SPI_{21} (86%); SPI_9^* (64%) | |
| SPI_{36} | SPI_{24} (82%) | SPI_{12}^* (68%) |
| SPI_{39} | SPI_{27} (88%) | SPI_{15}^* (85%) |
| SPI_{42} | SPI_{30} (92%) | SPI_6^* (73%) |
| SPI_{45} | SPI_{33} (89%) | SPI_9^* (64%) |
| SPI_{48} | SPI_{36} (86%) | SPI_{12}^* (52%) |

Correlation values are shown in parenthesis. Symbol (*) indicates the selected SPIs significantly correlated to other SPIs. SPI_n represents SPI at n-month time scale

precipitation follows this distribution at longer time scales. In many cases, the Republic Democratic of Congo (DRC) area globally shows preference to weibull function. GPCC precipitation data show similar results (Fig. 3), but with distribution functions somewhat more dispersed spatially from 12-month time scale.

Overall, from 1- to 6-month, the four distribution functions are well grouped by zone and the weibull function is the most extended spatially. The gamma function for 1-month time scale is preferentially located in the belt 10°N–15°N of latitude, and in the southern part of the domain. In the band 7°N–10°N, and in the loop crossing the southern parts of Gabon and the DRC, a thin layer of the exponential function is revealed, surrounding a space largely dominated by the weibull function.

3.2 Study of the links between SPIs at different time scales

Tables 2 and 3 show cross-correlations between SPI values at different time scales (1-, 3-, 6-, 9-, 12-, ..., 48-month) and computed using appropriate distribution functions at each grid point. Results give a symmetric matrix and

only the part upper the diagonal is shown. Correlations between two SPIs at the same time scale are obviously perfect ($r = 1$). Correlations between SPI from 1- to 6-month time scales and SPI at other time scales are generally too weak ($|r| < 0.1$) except few cases. This shows less similarity between SPI at short time scales. However, from 9 to 36 months, cross-correlations between SPIs are higher and ranged from 0.1 to 0.9. Using CRU data (Table 2), SPI_1 and SPI_6 are significantly correlated and one can account for about 76% of the variance of the other. SPI_6 is also significantly correlated with SPI_{18} at 85%. By reasoning in a similar way for other time scales, we establish a correspondence table (Table 4) to better understand the links between SPIs at different scales. In this table, SPI ran at 9-month time scale (SPI_9 in column 1) can significantly explain 81% of the variance of SPI_{21} (column 2), 64% for SPI_{33} (column 2) and 51% for SPI_{45} (column 2); SPI_{18} and SPI_{30} can be skipped in the benefit of SPI_6^* which is already selected, because SPI_{18} (column 1) can be significantly explained by 85% of SPI_6^* (column 2). SPI_{42} (column 1) can be explained at 92% by SPI_{30} (column 2) which can in turn be explained by 73% of SPI_6^* (column 3). In the same way as previously, we can skip SPI_{39} , SPI_{42} , SPI_{45} and SPI_{48} in the benefit of SPI_{15}^* , SPI_6^* , SPI_9^* and SPI_{12}^* , respectively. In addition, spatial patterns of SPIs from 18- to 48-month show resemblances to at least one from 1- to 15-month time scale (Fig. 5: (i), (m) and (q) are similar to (e), (j), (k), (l), (n), (o) and (p) are similar to (c) but with a slight overestimation of SPI absolute value). Precipitation data from GPCC gives similar results to those from CRU, but with correlation values all slightly larger, increasing the number of significant correlations (> 0.5) between time scales (from 32 to 37 significant values). The links that are added are those between 24- and 33-month, 1- and 42-month, 12- and 48-month, 24- and 45-month, and 33- and 48-month. It remains true that SPIs at these scales are represented between 1 and 15 months. The correspondence table (Table 4) established for CRU data remains valid for GPCC. Finally, SPIs from 1- to 15-month time scales can significantly explain the variances of SPIs at other time scales. Therefore, the rest of the study will be limited to time scales ranging from 1 to 15 month.

3.3 Sensitivity of SPI to statistical distribution functions

Figures 4 and 5 present multiscalar SPI means (sub-figures (a1), ..., (a6)) using appropriate distribution functions at each grid point and SPI biases ((b1), ..., (b6); (c1), ..., (c6); (d1), ..., (d6); (e1), ..., (e6)) if using one of the four distribution functions (gamma, exponential, lognormal and weibull, respectively) at all grid points. 1- to 15-month time scales aggregated monthly precipitation data from CRU and GPCC

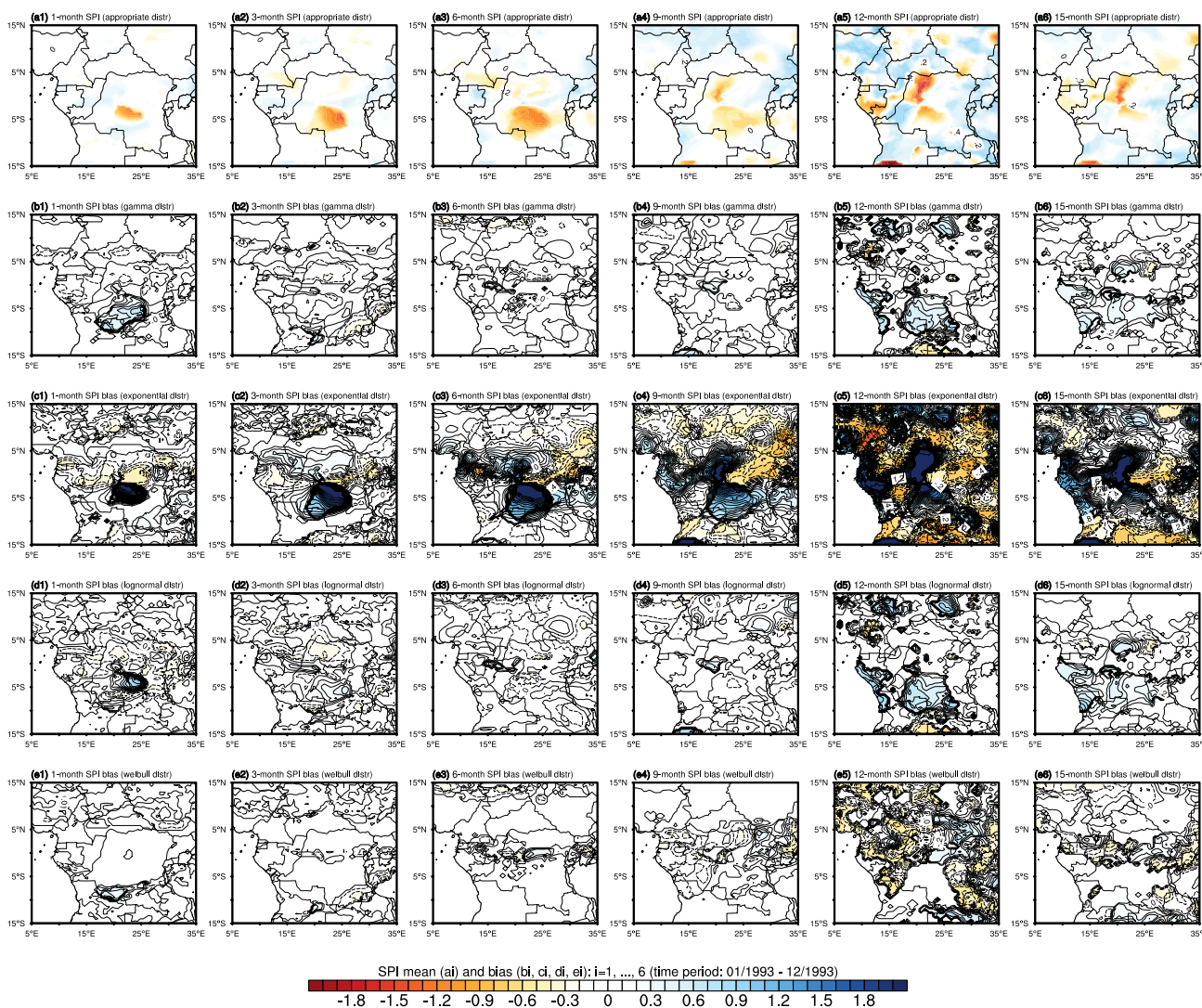


Fig. 4 Multiscalar SPI mean from January to December 1993 ((a1), ..., (a6)) using appropriate distribution functions at each grid point and SPI biases ((b1), ..., (b6); (c1), ..., (c6); (d1), ..., (d6); (e1), ..., (e6);) if using one of the four distribution functions (gamma, expo-

nential, lognormal and weibull, respectively) at all grid points. Aggregated CRU monthly precipitation data from 1- to 15-month time scales were used

were used. The time period 01/1993–12/1993 was chosen within the decade 1990s known as one of the driest in Central Africa.

Figure 4 ((a1), ..., (a6)) indicates in Congo Basin an area where SPIs are negative and of the order of -1.2 , showing the Moderate drought type, that widens from 1- to 6-month time scale. From 9-month, it decreases in mild drought ($SPI > -1$) while another similar drought pole appears north of the basin and intensifies at higher timescales. The study of SPIs trends indicates that the Congo Basin in general has recorded a significant drying that deserves attention in further investigation. The other parts of the domain generally have much lower and positive SPIs, indicating an increase in precipitation. Spatial patterns of SPIs with GPCC data are

similar to that of CRU (Fig. 5 (a1), ..., (a6)) but with lower values and highest dispersion from 12-month time scale. So, GPCC underestimates droughts and spatially more disperses them as compared to CRU.

If the same distribution function is applied at all grid points, instead of the appropriate distribution functions at each grid point, the results of SPIs obtained become biased. These are the cases shown in Figs. 4 and 5 ((bi), (ci), (di), (ei), $i=1, \dots, 6$) where the gamma, exponential and weibull distribution functions were used, respectively. The bias is weaker at 1-month time scale. It is generally observed that the intensity and the spatial extent of biases increase as the time scale increases and is much more

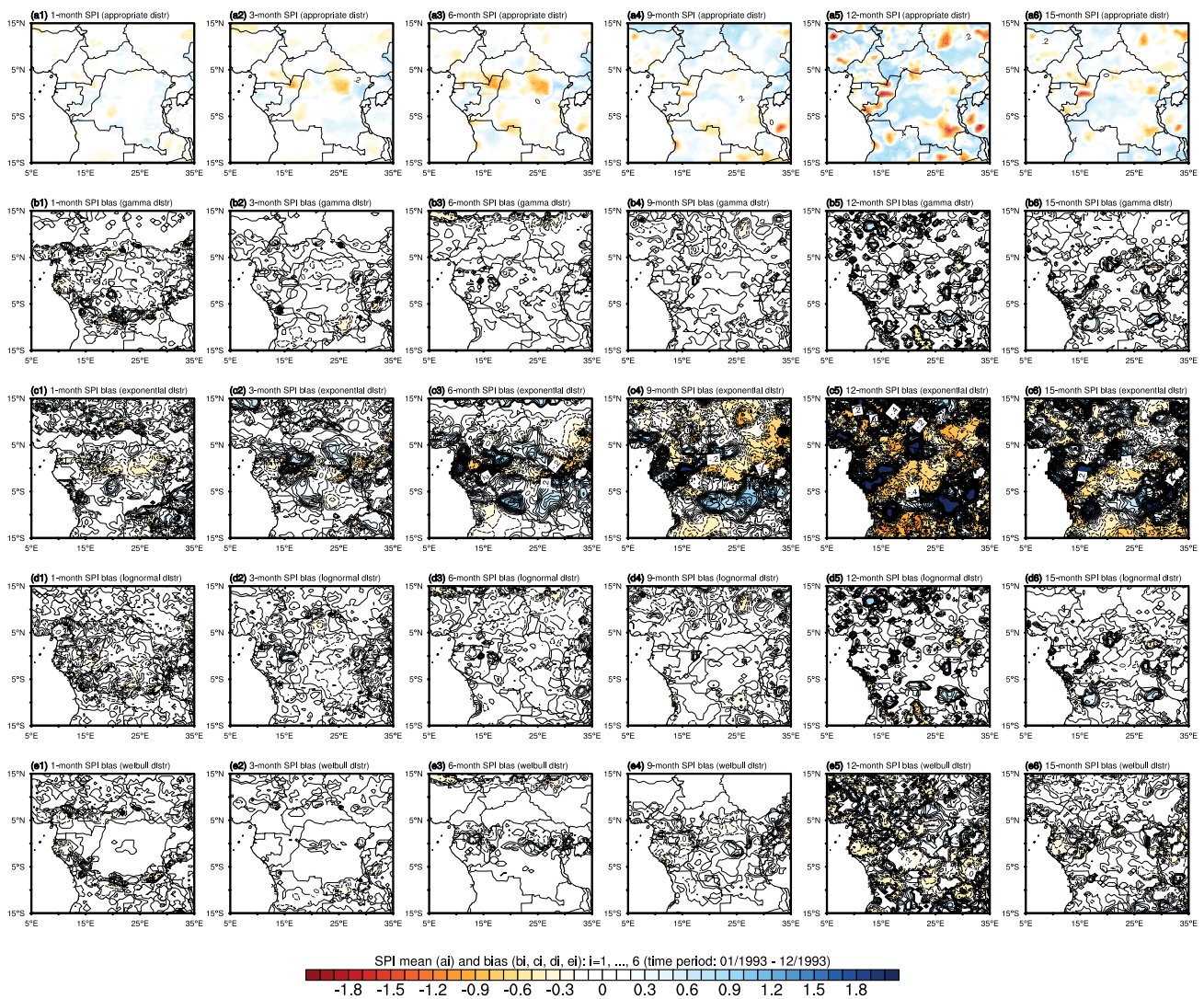


Fig. 5 Multiscalar SPI mean from January to December 1993 ((a1), ..., (a6)) using appropriate distribution functions at each grid point and SPI biases ((b1), ..., (b6); (c1), ..., (c6); (d1), ..., (d6); (e1), ..., (e6);) if using one of the four distribution functions (gamma, expo-

ponential, lognormal and weibull, respectively) at all grid points. Aggregated GPCP monthly precipitation data from 1- to 15-month time scales were used

localized in the center of the domain, particularly where grid points have inappropriate distribution functions. The exponential function, the least suitable distribution in the domain, presents the highest biases (c1, ..., c6), whereas the most representative weibull function (d1, ..., d6) spatially has the weakest bias. These biases sometimes going up to 2 can unfortunately give rise to a type of drought rather than another.

In all, SPI is sensitive to distribution function used to fit precipitation. However, weibull followed by gamma shows best performances for computing SPI, while the exponential most of the time shows significant difference. From

12-month time scale and above, the best fit distribution function at each grid point will be benefited to expect good SPI computation and then better drought description.

4 Discussion

Several studies show that the choice of methods to calculate drought characteristics can introduce uncertainties in drought projections in the future periods (Okpara and Tarhule 2011; Mo 2008; Sheffield and Wood 2007). One of the sources of error in SPI results could come from the fact that in many computations (Lloyd-Hughes and Saunders

2002; Amin et al. 2011; Spinoni et al. 2014; Gidey et al. 2018; Mustafa and Rahman 2018), the precipitation record is directly fitted to a gamma distribution before deriving SPI, yet there are other distribution functions. In a part of the current study, we tested four distribution functions (gamma, exponential, lognormal and weibull) and chose the suitable fit at each grid point for better computing SPI. We found that the functions weibull and gamma are concurrently the best fits in most area over Central Africa. These results corroborate those previously obtained in Cameroon by Guenang and Mkankam (2014), where using both station and CRU datasets, they found that distribution function fitting precipitation changes with space and depends on the aggregated precipitation time scales. Thus the gamma distribution is not only the best-fit function in Africa contrary to Europe (Sharma and Singh 2010; Alvarez et al. 2011). Danielle et al. (2015) also fitted and tested three distributions functions, including lognormal, gamma and generalized extreme value to identify suitable parameters for the 3-month, 6-month and 12-month SPI. They found that the gamma distribution was suitable for precipitation in most area over the globe. Unfortunately afterward, the SPI calculation was done only with the best distribution statistically at all grid points, which could sometimes skew the results in areas where a different distribution function was suitable. Other investigations (Sienz et al. 2012) have noted limitations with the gamma distribution, particularly under extreme climates, and the three-parameter distributions such as the Pearson Type three have been proposed as an alternative (Guttman 1999).

There is a challenge to reduce the number of time scale in multi-scalar drought study or to find optimal time scale to be used if studying many drought indices. In the first case, considering different time scales, we found that some SPIs are highly correlated and show similar patterns. Thus the choice of time scales is important to avoid redundant drought interpretation. In the second case, Javier et al. (2018) found that 6-month time scale best reproduces observed events across all the drought indices used.

The current study has also shown that CRU TS 4.01 and GPCC datasets corroborate, but GPCC's SPIs are wetter and slightly more dispersed at longer time scales. The version of CRU precipitation data used have been improved comparatively to version 3.10.01 or older whose a wet bias (due to a sharp decline in its number of rain-gauges for the recent decades) was underlined by Trenberth et al. (2014). This bias is due to a sharp decline in its number of rain-gauges for the recent decades. Furthermore, CRU data show high correlation with the GRACE data in Central Africa, especially south of the equator (Aiguo and Tianbao 2016). Otherwise, some investigations showed reliability of GPCC data and recommended to use them where there is poor data coverage since the 1990s (Aiguo and Tianbao 2016).

This study also noticed a trend to drying in Congo basin in particular, reinforcing the idea that the central African rain-forests have experienced a long-term drying (Wenjian et al. 2016; Spinoni et al. 2014). Such climatic condition is very alarming in Jordan according to Mustafa and Rahman (2018) who detected an increase in drought magnitude by applying both the Mann–Kendall trend test to precipitation recorded at several stations of the country and the regression test to the SPI values. Northern Ethiopia is not spared by this hazards in the sense that Gidey et al. (2018) detected a spatial and temporal increase in drought with incidences high during the past three decades. Over Africa in general, many different measures of drought all found widespread drying from 1950 to 2012 (Aiguo and Tianbao 2016). Dai (2013) tries to explain that the increase of drying over many land areas in the globe in general is due to the warming since the 1980s. It would be necessary to go deeper into this study in further investigations.

5 Conclusions

In a context where the climate is changing, it is of utmost importance to follow the dynamic of indices that can inform on the state of the climate in various regions of the world. Therefore, this study was undertaken to examine SPI calculation method and to show the sensitivity of results to the choice of distribution function used to fit precipitation data in one of the stages of SPI computation. Study domain was Central Africa and two monthly gridded precipitation datasets, both from the CRU and GPCC, were used on the time period 1951–2016. Four statistical distribution functions (gamma, weibull, exponential and lognormal) were tested to determine the best fit for multi-scale precipitation accumulation ranging from 1-, 3-, 6-, ... to 48-month. Next the multi-scale SPIs were computed and the sensitivity of results to distribution functions was examined. The study also attempted to establish relationship between SPIs computed at various timescales.

The results let appear that for short time scales (not more than 9-month), most gridded precipitation data preferably follow the weibull distribution function. Many other grid points show interest for gamma, few for lognormal and very few for exponential. Observed spatial pattern at 1-month time scale progressively shifts northward along longer time scales. Spatial distribution of functions scattered at 12-month, and no grid point follows the exponential function. Results of cross-correlations showed links between SPIs at different time scales and led to the necessity to reduce the study from 1- to 15-month. Overall, SPI values are sensitive to distribution function used to fit precipitation. The functions weibull and gamma showed lower bias in computing SPI if used at all grid points while the

exponential most of the time shows significant difference. For long time scale (12-month and more), the best fit distribution function chosen at each grid point will be benefit to expect good SPI computation and then better drought description.

The SPI was recommended in 2009 as the best meteorological index (Hayes et al. 2011). One of the advantages of SPI is its ability to compute drought levels for different time scales ranging from 3, 6, 12, 24, and 48-month periods (Amin et al. 2011), but the pattern of increase in the duration and magnitude of droughts resulting from higher temperatures can not be identified by SPI, so SPEI was developed to overcome this shortcoming (Vicente-Serrano et al. 2010).

Acknowledgements Many thanks to the Climatic Research Unit of the University of East Anglia and to the Global Precipitation Climatology Centre that produced CRU and GPCC gridded monthly precipitation data, respectively.

References

- Abramowitz M, Stegun Ae (1965) Handbook of mathematical formulas, graphs, and mathematical tables. Dover Publications Inc, New York
- Aiguo D, Tianbao Z (2016) Uncertainties in historical changes and future projections of drought. Part I: estimates of historical drought changes. *Clim Change*. <https://doi.org/10.1007/s10584-016-330>
- Alley W (1984) The Palmer drought severity index: limitations and assumptions. *J Clim Appl Meteorol* 23:1100–1109
- Alvarez B, Ferreira G, Hube M (2011) A proposed reparametrization of gamma distribution for the analysis of data of rainfall-runoff driven pollution. *Proyecciones Journal of Mathematics* 30(3):415–439
- Amin Z, Rehan S, Bahman N, Faisal IK (2011) A review of drought indices. *Environ Rev* 19:333–349
- Dai A (2013) Increasing drought under global warming in observations and models. *Nat Clim Chang* 3:52–58
- Danielle T, Moetasim A, Munir AN, Shih-Chieh K, Noah SD (2015) A multi-model and multi-index evaluation of drought characteristics in the 21st century. *Journal of Hydrology* 526:196–207
- Dutra E, Giuseppe F, Wetterhall F, Pappenberger F (2013) Seasonal forecasts of droughts in African basins using the standardized precipitation index. *Hydrol Earth Syst Sci* 17:2359–2373. <https://doi.org/10.5194/hess-17-2359-2013>
- Fotso-Nguemo T, Andreas H, Derbetini AV, Clement T (2016) Assessment of simulated rainfall and temperature from the regional climate model REMO and future changes over Central Africa. *Clim Dyn*. <https://doi.org/10.1007/s00382-016-3294-1>
- Frank J, Massey J (1951) The Kolmogorov-Smirnov test for goodness of fit. *J Am Stat Assoc* 46(253):68–78. <https://doi.org/10.1080/01621459.1951.10500769>
- Gidey E, Dikinya O, Sebego R, Segesebe E, Zenebe A (2018) Modeling the spatio-temporal meteorological drought characteristics using the standardized precipitation index (SPI) in raya and its environs, Northern Ethiopia. *Earth Syst Environ* 2(2):281–292. <https://doi.org/10.1007/s41748-018-0057-7>
- Guenang G, Mkankam KF (2014) Computation of the standardized precipitation index (SPI) and its use to assess drought occurrences in Cameroon over recent decades. *J Appl Meteorol Climatol* 53:2310–2324
- Guenang G, Vondou DA, Mkankam KF (2016) Total water storage change in Cameroon: calculation, variability and link with onset and retreat dates of the rainy season. *Hydrology* 3(36):1–20
- Gustavo M, Nicola F, Ashok KM (2018) Investigating drought in Apulia region. *Theoretical Appl Climatol Italy SPI RDI*. <https://doi.org/10.1007/s00704-018-2604-4>
- Guttman N (1999) Accepting the Standardized Precipitation Index: a calculation algorithm. *J Amer Water Resour Assoc* 35(2):311–322
- Guttman NB (1998) Comparing the Palmer drought index and the standardized precipitation index. *J Am Water Resour Assoc* 34(1):113–121
- Harris I, Jones P, Osborn T, Lister D (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>
- Hayes M, Svoboda M, Wilhite D, Vanyarkho O (1999) Monitoring the 1996 drought using the standardized precipitation index. *Bull Am Meteorol Soc* 80:429–438
- Hayes M, Svoboda M, Wall N, Widhalm M (2011) The Lincoln declaration on drought indices: universal meteorological drought index recommended. *Bull Amer Meteorol* 92(4):485–488. <https://doi.org/10.1175/2010BAMS3103.1>
- Howard GT (1959) A generalization of the Glivenko-Cantelli theorem. *Ann Math Stat* 30:828–830. <https://doi.org/10.1214/aoms/1177706212>
- Javier FS, Deng P, Luzia F, Boris O, Javier GH, Frederic J, Christoph H, Fangwei Z, Jijun X (2018) Searching for the optimal drought index and timescale combination to detect drought: a case study from the lower Jinsha River basin, China. *Hydrology and Earth Systems* 22:889–910
- Jones P, Hulme M (1996) Calculating regional climatic time series for temperature and precipitation: methods and illustrations. *Int J Clim* 16:361–377
- Lana X, Serra C, Burgueno A (2001) Patterns of monthly rainfall shortage and excess in terms of the standardized precipitation index for Catalonia (NE Spain). *Int J Climatol* 21:1669–1691. <https://doi.org/10.1002/joc.697>
- Libanda B, Zheng M, Ngonga C (2019) Spatial and temporal patterns of drought in Zambia. *J Arid Land*. <https://doi.org/10.1007/s40333-019-0053-2>
- Lloyd-Hughes B, Saunders MA (2002) A drought climatology for Europe. *Int J Clim* 22:1571–1592
- Majid M, Babak A, Rizwan N (2017) A Monte Carlo Simulation-Based Approach to Evaluate the Performance of three Meteorological Drought Indices in Northwest of Iran. *Water Resour Manage* 31, 4:1323–1342. <https://doi.org/10.1007/s11269-017-1580-2>
- McKee T, Doesken N, Kliest J (1993) The relationship of drought frequency and duration to time scales. In Proceedings of the 8th conference of applied climatology, 17–22 January, Anaheim, CA. American Meteorological Society, Boston, Massachusetts pp 179–184
- Mo K (2008) Model-based drought indices over the United States. *J Hydrometeorol* 9:1212–1230
- Mustafa A, Rahman G (2018) Assessing the spatio-temporal variability of meteorological drought in Jordan. *Earth Syst Environ* 2:247–264. <https://doi.org/10.1007/s41748-018-0071-9>
- Naumann G, Dutra E, Barbosa P, Pappenberger F, Wetterhall F, Vogt J (2014) Comparison of drought indicators derived from multiple data sets over Africa. *Hydrol Earth Syst Sci* 18:1625–1640
- Okpara J, Tarhule A (2011) Drought under global warming: a review. *Wiley Interdiscip Rev Clim Chang* 2:45–65
- Okpara J, Tarhule A (2015) Evaluation of Drought indices in the Niger Basin, West Africa. *J Geogr Earth Sci* 3(2):1–32

- Okpara J, Tarhule A (2016) Evaluation of drought indices in the Niger River Basin. *West Africa J Geogr Earth Sci* 3(2):1–32
- Okpara J, Afiesimama E, Anuforom A, Owino A, Ogunjobi K (2017) The applicability of standardized precipitation index: drought characterization for early warning system and weather index insurance in West Africa. *Nat Hazards* 89(2):555–583. <https://doi.org/10.1007/s11069-017-2980-6>
- Oladipio E (1985) A comparative performance analysis of three meteorological drought indices. *Int J Climatol* 5:655–664
- Palmer W (1965) Meteorological drought. Research Paper No. 45. US Weather Bureau, Washington
- Qiang F (2003) Radiation (solar). Elsevier Science Ltd, Amsterdam, pp 1859–1863
- Schneider U, Ziese M, Meyer-Christoffer A, Finger P, Rustemeier E, Becker A (2016) The new portfolio of global precipitation data products of the Global Precipitation Climatology Centre suitable to assess and quantify the global water cycle and resources. *Proc IAHS* 374:29–34
- Schneider U, Finger P, Meyer-Christoffer A, Rustemeier E, Ziese M, Becker A (2017) Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the global precipitation climatology centre (GPCC). *Atmosphere* 8(3):52. <https://doi.org/10.3390/atmos8030052>
- Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M (2018) GPCP full data monthly product version 2018 at 0.5°: monthly land-surface precipitation from rain-gauges built on GTS-based and historical data. https://doi.org/10.5676/DWD_GPCP/FD_M_V2018_050
- Sergio MVS, Begueria S, Juan ILM (2008) A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J Clim* 23:1696–1718
- Sharma M, Singh J (2010) Use of probability distribution in rainfall analysis. *New York Science Journal* 3(9):40–49
- Sheffield J, Wood E (2007) Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Clim Dyn* 31:79–105
- Siens F, Bothe O, Fraedrich K (2012) Monitoring and quantifying future climate projections of dryness and wetness extremes: SPI bias. *Hydrol Earth Syst Sci* 16(7):2147–2157
- Spinoni J, Naumann G, Carrao H, Barbosa P, Vogt J (2014) World drought frequency, duration, and severity for 1951–2010. *Int J Climatol* 34:2792–2804
- Szalai S, Szinell C (2000) Comparison of two drought indices for drought monitoring in Hungary—a case study. In: Vogt JV, Somma F (eds) *Drought and drought mitigation in Europe*. Kluwer, Dordrecht, pp 161–166
- Thom H (1958) A note on the gamma distribution. *Monthly Weather Review* 86:117–122
- Trenberth K, Dai A, Van Der Schrier G, Jones P, Barichivich J, Briffa K, Sheffield J (2014) Global warming and changes in drought. *Nat Clim Change* 4:17–22
- Vicente-Serrano S, Begueria S, Lopez-Moreno J (2010) A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J Clim* 23(7):1696–1718. <https://doi.org/10.1175/2009JCLI2909.1>
- Wenjian H, Liming Z, Haishan C, Sharon EN, Ajay R, Yan J (2016) Possible causes of the Central Equatorial African long-term drought. *Environ Res Lett*. <https://doi.org/10.1088/1748-9326/11/12/124002>
- WMO (2006) Impacts of desertification and drought and other extreme meteorological events. Report of the Working Group on the Impacts of Desertification and Drought and of Other Extreme Meteorological Events WMO/TD No. 1343, Geneva, Switzerland pp 1–32
- Zhang Q, Li J, Singh V, Xu C, Deng J (2013) Influence of ENSO on precipitation in the East River basin, South China. *J Geophys Res D: Atmos* 118:2207–2219