Rainfall intensification in tropical semi-arid regions: the Sahelian case.
Supplementary material.

Panthou G. et al.
1. Data

This section focuses on describing the rainfall data source and processing which support the results presented in this paper. The different stages of data processing – from the source to the final datasets used in this study – are described in the following sections. For explanation purposes, we reserve thereafter the use of the terms “database” and “dataset”: a database refers to a pool of quality-controlled and fully documented (missing values, changes over time, ...) rainfall series that share similar characteristics (time step, sensor used, source of data, ...); a dataset refers to a subset of a database that meet specifications (e.g. percentage of missing values, length of the series, period and area covered, ...) according to the objective of the analysis which made use of the dataset.

1.1. From raw data to database

Two databases have been used in this study: the “BADOPLU” database and the “AMMA-CATCH Niger” database.

The BADOPLU database contains daily rainfall series coming from West African meteorological services. This database initially originates from the CIEH-ORSTOM (old name of IRD – Institut de Recherche pour le Développement) which have collected rain gauges data from the beginning of the series to roughly 1985 over the West African region. Then, a complementary work by Le Barbé et al. (2002) led to extend in time a large number of series for the period 1985-2000. During the 2000s, data were provided by national services in the framework of the AMMA project (African Monsoon Multidisciplinary Analysis – Lebel et al. 2010). Since the beginning of the 2010s, a group of scientists from IGE (Institut des Géosciences de l’Environnement) committed themselves to maintaining the database and updating the series (via the national meteorological services). Numerous studies have been realized using this database (see e.g. Le Barbé & Lebel 1997, Le Barbé et al. 2002, Lebel & Ali 2009, Panthou et al. 2012, Panthou, Vischel & Lebel 2014, Blanchet et al. 2018, among many others). Each of these studies had specific scientific objectives. They have thus looked at the data with a particular view and have deployed specific algorithms to check and visualize the data. This leads to detect very different kinds of problems in the database (e.g. spurious and erroneous data – monthly data instead of daily, unit problem, gap-filled data, ... ; date problem – offset of one day, bad month number, ...). As a consequence, changes have been applied in the database in order to: better document (adding or removing some missing values) and clean (e.g. removing erroneous data) the database. All the changes applied are stored in a history of changes linked to the database (svn versioning). Throughout the long life cycle of the database, the maintenance team has therefore learned much of the different kinds of data problems/errors encountered in the past. It has thus developed a set of algorithms able to detect possible problems/errors in the data (e.g. day-by-day spatial coherency check, time-series analysis, distribution checking, ...). Nonetheless, these algorithms are just helping tools and it is the team that
looks at raw data and finally qualifies them. At present, these tools have been stabilized and run routinely each time the team receives new data. This permits a continuous, homogeneous, and fast processing of data integration. It also permits to seize all the opportunities for further controlling and improving the database: for example when the “same data” (series recorded by the same sensor) comes from different sources (same raw data but that have received a different post-processing). Now, any new series received by the team passes these filters and is therefore qualified. At the time of writing, the BADOPLU database over the Sahel box defined in the paper contains a total of 392 stations corresponding to 18753 station-years over the period 1900-2016. The locations of these stations as well the year by year number of stations in operation is plotted in figure S1.

![Figure S1. BADOPLU database over the Sahelian region: a) locations of stations (whatever the length of the series); b) evolution of the number of stations in the database.](image)

The AMMA-CATCH Niger database comes from the AMMA-CATCH observatory (Lebel et al. 2009). The first rainfall measurements were carried out in 1988, but the long-term rain gauge network started only in 1990. The observatory team is in charge of: the installation of the sensors (tipping bucket rain gauges), the collection of the raw data (tipping date of the buckets), and their treatment to obtain 5-minute rainfall series and derived products such as the daily series or spatial events (Vischel et al. 2011). These different series have been used in this study: 5-minutes and daily series series in the main paper; spatial events for producing the figure S5 of this supplementary. As for the BADOPLU database, numerous studies have used the AMMA-CATCH Niger database (see e.g. Lebel et al. 1992, Balme et al. 2005, Balme et al. 2006, Vischel & Lebel 2007, Vischel et al. 2009, Panthou, Vischel, Lebel, Quantin & Molinié 2014, among many others).

Obviously, the authors acknowledge that – despite all the good care and rigor provided to collect, process, qualify, and format the data – both databases are probably not totally error free. Authors are aware that new minor errors are likely to be detected in the future, as knowledge and tools improve. In fact, these statements are valid for most (if not all) hydro-meteorological databases. Nevertheless, authors are confident that these two databases are sufficiently clean (gross error and problem removed) and
whitened (missing values well documented) to meet the requirements that permit to obtain reliable climatic analyses when using suitable dataset extractions from these databases. Note that at this stage, the databases contain only two types of data: positive float values and missing values (nan).

1.2. Extraction of datasets from database

Extracting a dataset depends on the purpose of the study. For this study, the following criteria were considered:

(i) covering the longest period (ending after 2010 to document the recent years);
(ii) having the largest number of stations within each box;
(iii) having the longest series (with the fewest possible missing values).

These different criteria must be satisfied for the four boxes (West Sahel, Central Sahel, East Sahel, and AMMA-CATCH Niger). Criterion (i) is well constrained by the data availability of the two databases, and leads to retain the period 1950-2014 for the BADOPLU extractions and 1990-2016 for the AMMA-CATCH extractions. Criteria (ii) and (iii) involve making a compromise between the completeness of the series retained and the number of stations selected. The adjustment variable is the authorized number of years classified as missing for each series: the more strict is the criterion for considering a series as complete, the smaller is the number of stations that will be used. This requires to classify missing/valid years at each station. For a given year, the annual aggregation of missing is done by attributing a flag value at each of the following steps:

- if its annual totals is 2 (resp. 5) times below or higher the mean annual total of the whole series, then flag equal 1(resp. 2) else flag=0;
- same algorithm for the mean number of wet days;
- same algorithm for the mean intensity of wet days;
- a interannual average of monthly totals is computed, and rainy/dry month are classified according to a threshold of 90 mm, then the flag value is equal to the number of months that belong to the interannual rainy months but that have not recorded rains.

The sum of these 4 flags gives the aggregated flag of the considered year. If the aggregated flag is $\leq 1$, then the year is tagged as valid else is tagged as missing. This procedure is somewhat strict but leads to be very confident with the quality of the information available for a year qualified as valid.

Different percentages of missing years for a series have been tested (results remain very similar to those presented in this study), and lead to retain a threshold of 10%. Imposing that a station must have less than 10% of missing years, 71 (resp. 30) stations have been selected for computing indexes over the period 1950-2014 (resp. 1990-2016) for the Sahel box (resp. ACN box) extracted from the BADOPLU (resp. AMMA-CATCH Niger) database. In the resulting theoretical dataset over the Sahel box (resp.
Rainfall intensification in tropical semi-arid regions: the Sahelian case. Supplementary material.

ACN box) of 71 stations × 65 years = 4615 station-years (resp. 30 × 27 = 810), there are 120 (resp. 12) missing station-years, which corresponds to less than 3 % (resp. 2%) of missing station-years. The complete charts about missing/valid years of each dataset extracted for the study are plotted in figure S2.

Note that, altogether, for the retained station-years, there are only 1% of missing daily data during the core of the rainy season (JJAS, details for each month and each dataset are given in table S1). These daily missing values are excluded when computing annual statistics.

Figure S2. Network operation of the datasets used in this study. The left y-axis gives the station names, the total number of valid years for a station is given in the right y-axis.

Table S1. Percentage of missing daily data contained in the different datasets during the rainy months.

| Month | AMMA-CATCH niger | Central Sahel | East Sahel | West Sahel |
|-------|------------------|---------------|------------|------------|
| 6     | 3.7%             | 0.1%          | 0.2%       | 1.0%       |
| 7     | 3.4%             | 0.0%          | 0.0%       | 0.6%       |
| 8     | 2.9%             | 0.2%          | 0.1%       | 0.6%       |
| 9     | 3.0%             | 0.3%          | 0.3%       | 0.6%       |
2. Methods

2.1. Regional averaging

The three following steps are used to compute the regional annual means – \( N_S(y), R_S(y), \) and \( AR_S(y) \):

(i) Computing the mean interannual values at each station \( k \) (\( \bar{N}_k, \bar{R}_k, \) and \( \bar{AR}_k \) – c.f. equations S1, S5, and S9). These values are obtained using all the available years \( Y_k \) at a given station \( k \). In the case of no missing values, \( Y \) is a vector of \( |Y| = 65 \) values ranging from 1950 to 2014 for the BADOPLU datasets (resp. \( |Y| = 27 \) values from 1990 to 2016 for the AMMA-CATCH datasets).

(ii) Computing the mean regional interannual values (\( \bar{N}_S, \bar{R}_S, \) and \( \bar{AR}_S \) – c.f. equations S2, S6, and S10).

(iii) Computing the mean regional anomaly for year \( y \) and adding this value to the mean regional interannual values in order to obtain \( \bar{N}_S(y), \bar{R}_S(y), \) and \( \bar{AR}_S(y) \) (equations S3, S7, and S11).

Working on anomalies is a way for filtering the latitudinal gradient in case of one or more missing stations for a given year \( y \) (if working on raw values, a missing station in the south of the study domain would have more weight than if it is located in the north).

The corresponding equations for \( N \) are below:

\[
\bar{N}_k = \frac{1}{|Y_k|} \sum_{y} N_k(y) \quad \forall k \in K \tag{S1}
\]
\[
\bar{N}_S = \frac{1}{|K|} \sum_{k} N_k \tag{S2}
\]
\[
\bar{N}_S(y) = \bar{N}_S + \frac{1}{|K|} \left[ N_k(y) - N_k \right] \quad \forall y \in Y \tag{S3}
\]

where \( N_k(y) \) is the number of rainy days at station \( k \) for year \( y \), \( K \) represents a vector containing all the stations, and \( \frac{1}{|K|} \) is an averaging operator that is either an arithmetic mean or a kriging average. Note that a day is considered rainy when the daily accumulation exceeds 1 mm.

The corresponding equations for \( R \) are:

\[
R_k(y) = \frac{1}{N_k(y)} \sum_{d=1}^{N_k(y)} r_k(d, y) \quad \forall y \in Y_k; \forall k \in K \tag{S4}
\]
\[
\bar{R}_k = \frac{1}{|Y_k|} \sum_{y} R_k(y) \quad \forall k \in K \tag{S5}
\]
\[
\bar{R}_S = \frac{1}{|K|} \sum_{k} R_k \tag{S6}
\]
\[
\bar{R}_S(y) = \bar{R}_S + \frac{1}{|K|} \left[ R_k(y) - R_k \right] \quad \forall y \in Y \tag{S7}
\]

where \( r_k(d, y) \) is the daily rainfall of rainy day \( d \) of year \( y \) at station \( k \).
A similar set of equations is used for $AR$:

$$AR_k(y) = \sum_{d=1}^{N_k(y)} r_k(d, y) \quad \forall y \in Y_k; \forall k \in K$$  \hfill (S8)

$$\overline{AR}_k = \frac{1}{|Y_k|} \sum_y^{Y_k} AR_k(y) \quad \forall k \in K$$  \hfill (S9)

$$\overline{AR}_S = \frac{K}{E} [\overline{AR}_k]$$  \hfill (S10)

$$\overline{AR}_S(y) = \overline{AR}_S + \frac{K}{E} [AR_k(y) - \overline{AR}_k] \quad \forall y \in Y$$  \hfill (S11)

The hydro-climatic intensity $HCI$ is computed using the previous results obtained for $R$ and $N$, and applying the formula $HCI = R/N$.

Note that the spatial averaging of the values computed for each individual station – denoted $E$ – was carried out using both a simple arithmetic mean operator and a kriging operator (ordinary kriging for anomalies – equations S3, S7, and S11; and universal kriging for interannual values in order to take into account the north-south gradient – equations S2, S6, and S10). The results of both methods are equivalent when it comes to study the year to year evolution and the long-term trends of the different rainfall regime statistics used here (which is the main objective of this study). However, they differ when it comes to estimate the mean regional interannual values ($\overline{N}_S$, $\overline{R}_S$, $\overline{AR}_S$, and $\overline{HCI}_S$) – especially when the gauge network is not evenly distributed over the box (as for the Central Sahel box where gauges are mainly located in the south). In such a case, the universal kriging permits to reduce the network distribution effect by estimating an external drift along latitude and longitude, thus allowing for taking into account the north-south gradient of rainfall statistics. The results presented in the main text are consequently based on a kriging of the point values to estimate the different spatial averages ($K$ in equations S2, S3, S6, S7, S10, and S11).

### 2.2. Uncertainty assessment

The uncertainty of the indexes due the spatial averaging ($K$ in equations S2, S3, S6, S7, S10, and S11) was estimated using a non-parametric bootstrap (Efron & Tibshirani 1994). It consists of the following steps:

(i) The vector of station $K$ is resampled with replacement (Monte Carlo resampling) until its length equals the length of the original vector. The obtained bootstrap vector of stations is named $K_{\text{boot}}$.

(ii) Mean regional interannual values are computed using equations S2, S6, and S10 applied on the $K_{\text{boot}}$ vector of stations.

(iii) Mean regional annual values are then obtained using equations S3, S7, and S11 applied on the $K_{\text{boot}}$ vector of stations.

(iv) The obtained annual indexes for this bootstrap – $\overline{N}_S(y)_{\text{boot}}$, $\overline{R}_S(y)_{\text{boot}}$, $\overline{AR}_S(y)_{\text{boot}}$, $\overline{HCI}_S(y)_{\text{boot}}$ – are stored.
These four steps are repeated 200 times leading to generate 200 $K_{\text{boot}}$ vectors of stations. For a given index, a vector of 200 values is thus obtained for each year. The confidence intervals were computed on these 200 values. Note that these confidence intervals represent the uncertainty only due to the network configuration.

2.3. Statistical testing

Statistical tests were used in this study. Two kinds of statistical tests have been applied on indexes to test: i) whether the central tendency of two samples are significantly different (table 1), and ii) if an index presents a significant trend (figures 2, 3, and 4 of the main paper and figures S3, S4, and S5 of this supplementary) or is correlated with another one (table 2). For both kinds, parametric and non-parametric tests can be found. The parametric ones assume the normality of the data in the case of a central tendency comparison (Student’s t-test, Student 1908) or the normality of the residuals in the case of correlation testing (Pearson’s $r$ correlation, Benesty et al. 2009). The non-parametric counterparts – Wilcoxon rank-sum test for central tendency comparison (Wilcoxon 1945, Mann & Whitney 1947), Spearman $\rho$ for correlation testing (Mann 1945, Lehmann & D’Abrera 1998) – are rank-based and do not require to fulfill the normality assumption. Since some samples and residuals do not pass the normality tests, the results of the non-parametric tests were presented in this study. In fact, the parametric ones give very similar results in term of statistics and significance.

Note that the trend lines plotted on the different figures are estimated using the Theil-Sen algorithm (Sen 1968, Theil 1992).
3. Additional Figures

**Figure S3.** Evolution of the average of the annual maximum rainfall intensity (mm h$^{-1}$) extracted independently for different durations (SAM samples): 5 minutes, 10 minutes, 30 minutes, and 1 hour. Trends (Theil-Sen estimator) and significance (Spearman test, 1%***, 5%**, 10%*) are given in the bottom legend. Error bars (resp. shaded area) delineate 80% confidence intervals for annual values (resp. 11-year rolling mean).

**Figure S4.** Evolution of the annual maximum maximorum rainfall intensity (mm h$^{-1}$) recorded over all the stations of the AMMA-CATCH network (spatial maxima of SAM) for different durations: 5 minutes, 10 minutes, 30 minutes, and 1 hour (it is not necessarily the same station that records the maximum maximorum at all 4 timescales). Trends (Theil-Sen estimator) and significance (Spearman test, 1%***, 5%**, 10%*) are given in the side legend.
Rainfall intensification in tropical semi-arid regions: the Sahelian case. Supplementary material.

Figure S5. Evolution of the mean annual maximum rainfall amount (mm) obtained from spatial rainfall events (Vischel et al. 2011). Trends (Theil-Sen estimator) and significance (Spearman test, 1%***, 5%**, 10%*) are given in the inset legend.

References

Balme M, Galle S & Lebel T 2005 Secheresse 16, 15–22.
Balme M, Vischel T, Lebel T, Peugeot C & Galle S 2006 Journal of Hydrology 331(1-2), 336–348.
Benesty J, Chen J, Huang Y & Cohen I 2009 in ‘Noise reduction in speech processing’ Springer pp. 1–4.
Blanchet J, Aly C, Vischel T, Panthou G, Sané Y & Diop Kane M 2018 Journal of Geophysical Research: Atmospheres.
Efron B & Tibshirani R 1994 An introduction to the bootstrap Chapman & Hall New York.
Le Barbé L & Lebel T 1997 Journal of Hydrology 188-189, 43–73.
Le Barbé L, Lebel T & Tapsoba D 2002 Journal Of Climate 15, 187–202.
Lebel T & Ali A 2009 Journal of Hydrology 375(1-2), 52–64.
Lebel T, Cappelaere B, Galle S, Hanan N, Kergoat L, Levis S, Vieux B, Descroix L, Gosset M, Mougin E, Peugeot C & Séguis L 2009 Journal of Hydrology 375, 3–13.
Lebel T, Parker D J, Flamant C, Bouriès B, Marticorena B, Mougin E, Peugeot C, Diedhiou A, Haywood J M, Ngamini J B, Polcher J, Redelsperger J L & Thorncroft C D 2010 Quarterly Journal of the Royal Meteorological Society 136(S1), 8–33.
Lebel T, Sauvageot H, Hoepffner M, Desbois M, Guillot B & Hubert P 1992 Hydrological Sciences 37, 201–215.
Lehmann E L & D’Abrera H J M 1998 Nonparametrics: statistical methods based on ranks Prentice Hall Upper Saddle River, N.J.
Mann H 1945 Econometrica: Journal of the Econometric Society 13(3), 245–259.
Mann H B & Whitney D R 1947 The Annals of Mathematical Statistics 18(1), 50–60.
Panthou G, Vischel T & Lebel T 2014 International Journal of Climatology 34(15), 3998–4006.
Panthou G, Vischel T, Lebel T, Blanchet J, Quantin G & Ali A 2012 Water Resources Research 48(8), 1–19.
Panthou G, Vischel T, Lebel T, Quantin G & Molinié G 2014 Hydrology and Earth System Sciences 18(12), 5093–5107.
Sen P 1968 Journal of the American Statistical Association 63(324), 1379–1389.
Student 1908 Biometrika pp. 1–25.
Theil H 1992 in ‘Henri Theil’s contributions to economics and econometrics’ Springer pp. 345–381.
Vischel T & Lebel T 2007 Journal of Hydrology 333(2-4), 340–355.
Vischel T, Lebel T, Massuel S & Cappelaere B 2009 Journal of Hydrology 375(1-2), 273–286.
Vischel T, Quantin G, Lebel T, Viarre J, Gosset M, Cazenave F & Panthou G 2011 Journal of Hydrometeorology 12(6), 1465–1482.
Rainfall intensification in tropical semi-arid regions: the Sahelian case. Supplementary material.

Wilcoxon F 1945 *Biometrics Bulletin* 1(6), 80.