Master’s Course in
Environmental and Land Engineering
Natural risks and civil protection
Academic year 2023/2024

Master’s degree thesis

Statistical analysis of Sirba River's rainfalls

Sahel, North-Western Africa, Burkina Faso & Niger

Professor:
Daniele Ganora PhD
Collaborator:
Francesco Saretto

Student:
Liviu-Toader Busteaga
Registration number:
s315296

Multumesc parintilor mei, care m-au sustinut atat prin grele,
cat si prin bune, in toti acesti ani de facultate.

I thank my parents, who have supported me through both the hard times
and the good times during all these years of university.

Summary

1: Introduction

With the ongoing changing climate, it is important to closely monitor weather patterns, especially rainfall. In Europe and developed countries, rainfall data is often easy to collect and analyze. However, in less developed countries, such data, particularly from rain gauges, can be limited, disorganized, or both. Without a clear understanding of past rainfall events, it is difficult to implement effective flood protection measures.

This paper aims to analyze rainfall series in the Sirba Basin, situated in the east of Burkina Faso, using hydrological statistical techniques.

The Polytechnic University of Turin has contributed to the ANADIA 2.0 project, between April 2017 and August 2021 [1], and to its successor, the SLAPIS Project (acronym for the French *Système Locale d'Alerte Précoce pour les Inondations au Sahel,* in English Local Early Warning System for Floods in the Sahel), aimed at building an early warning system for floods in the Sirba Basin [1]. It has been operative since 2019 and tested in 2020 for the said basin, the Niger River and the city capital of Niger, Niamey [1]. Many researchers from this institution have contributed, and continue to contribute, to work on this region, through various academic papers, making them a reliable source of information.

The Sahel region is particularly sensitive to weather conditions. Between 1970 and 1990, the region experienced severe droughts, followed by a recovery of precipitation [2]. This return to wetter conditions became evident with an increase in annual rainfall, which partly led to an increase in streamflows in Sahelian rivers [2]. However, the rise in rainfall alone cannot fully explain the frequency and intensity of flood events, as the current rainfall trend remains below pre-1970 levels [2] [3]. This phenomenon is referred to by some authors as the "Sahelian Paradox" [2] [4] [5].

The analysis of cause-and-effect factors is complex, but there are some contributing causes to this phenomenon. Many researchers argue that global climate change and uncontrolled land use are the main drivers [2] [3] [6].

The Sahel is home to many countries experiencing population growth, leading to increased demand for cropland and new settlements, which is often obtained through the clearing and deforestation of the natural habitat to make way to rain-fed cereal cultures, such as millet [2] [3] [6]. These land-use changes raise the runoff coefficient and reduce the soil's capacity to retain water, as Descroix et al. argue in [6], thus inducing more severe floods. The rivers' morphology, characterized by shallow and large river beds with light slopes [3], intensifies this problem, because in cases of even light rain the river beds are easily filled up.

Some authors analyze the problem through the climate change frame, finding a correlation between the rainfall and discharge series of the last three decades, as Aich et al. did in [7]. Nevertheless, the extent of causality is not fully understood, and further studies with improved data are necessary.

The area is also experiencing an aridification process. According to the study by Somé et al. [8], the Sirba Basin region has shown increasing temperature trends over the past few decades, leading to higher rates of evapotranspiration. While the number of rainy days appears to be growing, the region's water supply issues stem from high evaporation rather than a lack of rainfall. Solutions to address this issue include reforestation, underground waterworks, and methods to reduce albedo [8].

The objective of this paper is to conduct an Annual Maxima (AM) statistical analysis using station and grid data. The process begins with data sorting, followed by the extraction of annual maxima values. Next, moments are computed, and plotting positions are determined. Distribution fitting is then performed, selecting the best candidates for the entire region. Following this, Depth-Duration-Frequency (DDF) and Intensity-Duration-Frequency (IDF) curves are plotted. Finally, a data correction is applied to align the station and grid data, whose Moments and Rainfall Quantiles will be interpolated for the said basin.

2: Context & data

2.1: Study area

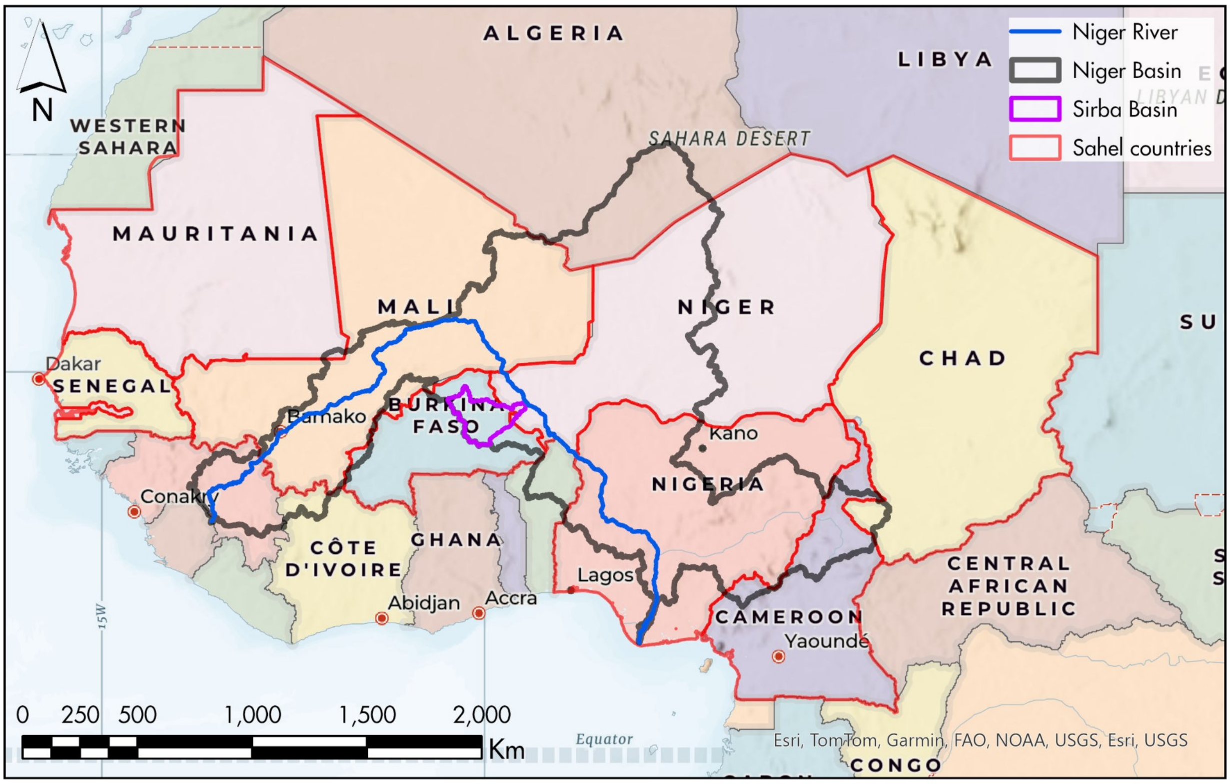


Figure 1: Regional administrative map of the Sahel, North-Western Africa.

The study area is the Sirba River basin, a major tributary of the Niger River, located in eastern Burkina Faso and partly in Niger (Figure 1 & Figure 2): 93% of its surface is into Burkina Faso while the rest is in Niger. These two countries are part of the region commonly known as the Sahel, thus the basin shows the typical Sahelian landscape in terms of land covering and vegetation, such as fallow savannah and shrub bushes [2].

The Sahel, as a region, cannot have fixed latitudinal limits because it is subject to changes driven by rainfall patterns. Today, the term "Sahel" also refers to a geopolitical entity called CLISS (French acronym for Permanent Interstate Committee for Drought Control), which includes nations like Burkina Faso and Niger, among many others. The member nations share not only climatic conditions and culture but also livelihood systems such as agriculture and livestock herding [9].

Given the vicinity with the Sahara Desert, the Sirba Basin has a semi-arid climate with considerable temperature fluctuations and distinct seasonal rainfall. Average temperatures typically range from about 20°C during the cooler months of December to February, rising to around 40°C in the hottest months of April to June. Rainfall is concentrated from June to September, with annual totals varying between 200 mm and 600 mm. Some argue that the notion of ‘normal rainfall’ may no longer be relevant in the context of the Sahel, given the frequency of extreme events. Droughts occur with varying severity in two out of every five years [9].

In the 20th century, the Sahel experienced three major droughts: 1910–1916, 1941–1945, and the prolonged "desiccation" from the 1970s to the 1990s. The 1983–1984 droughts were particularly severe, with some of the lowest rainfall ever recorded, affecting countries from Mauritania to Ethiopia and the southern Sahel [9].

The Sirba Basin is largely rural, with few urbanized areas. The communities here depend primarily on traditional activities such as farming, livestock raising, and fishing, rather than urban or industrial employment. The dry environment favors these rural livelihoods over city life, and population density remains low [10] [11].

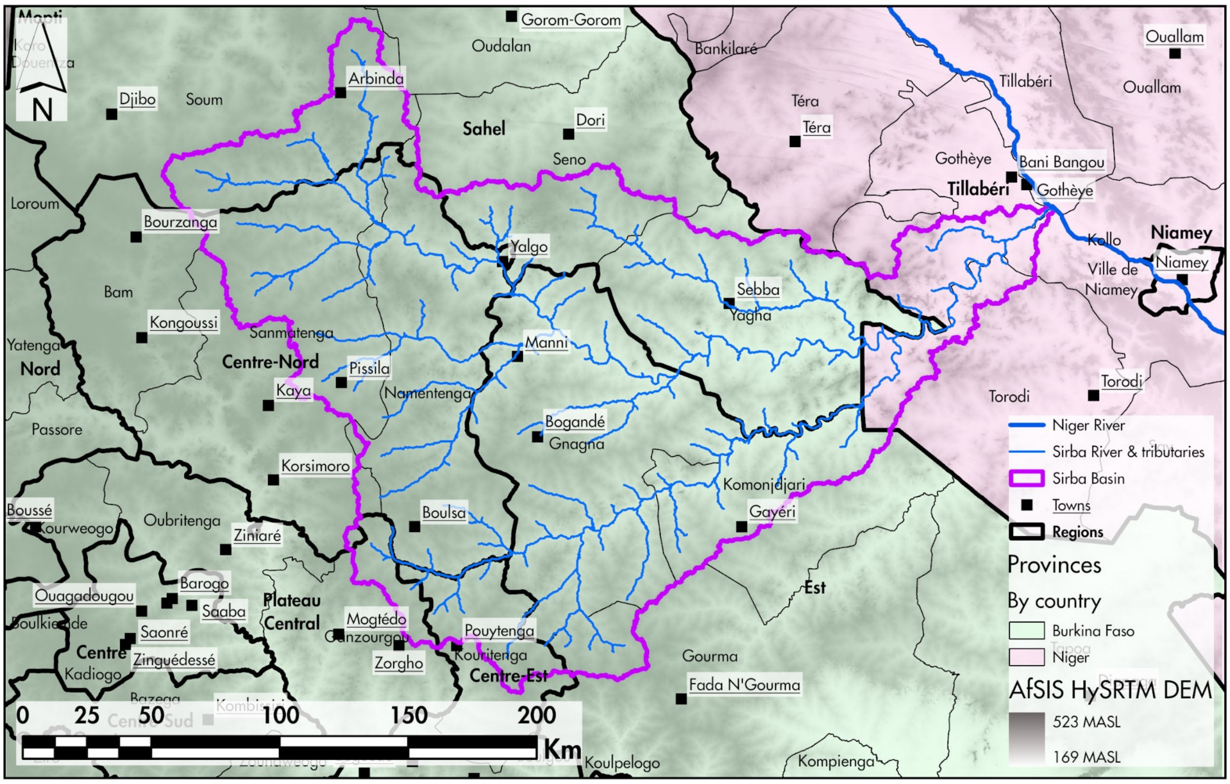


Figure 2: Local administrative map of the Sirba basin region.

Small towns and villages are scattered along the Sirba River, where people rely on seasonal floods or rainfalls to irrigate crops and support fishing. Some of the main settlements in and near the basin, relative to the latest population censuses [10] [11], have the flowing number of inhabitants:

- Aribinda: 45818;
- Bogande: 21443;
- Boulsa: 24200 (urban area only);
- Yalgo: 26340;
- Barsalogho: 98553 (with some rural surroundings);
- Pissila: 146671 (includes surrounding areas);
- Gayéri: 76218;
- Sebba: 34881;
- Gotheye: 7512.

Most of these settlements are small, with larger urban areas typically located outside the basin itself. While urban migration has increased across Burkina Faso, the Sirba Basin remains predominantly rural. This dynamic places additional demands on resources, especially water, which is needed for both local agriculture and the growing urban populations elsewhere, like for the capital cities of Ouagadougou and Niamey.

Since most farming relies on seasonal rainfall, societal systems in this region are heavily dependent on the ability to adapt to fluctuating water supply. Environmental pressures include expanding farmland, soil degradation, and climate variability, all of which affect water availability and the seasonal flooding essential for the basin's communities.

2.2: Station data

The data is gathered from pluviometers installed by the Burkina Faso and Niger State Meteorological Services, and also used by the SLAPIS Project. While many gauge stations were installed over the decades (some even recently as 2022), only a few provide continuous data due to frequent malfunctions and accessibility issues. These stations measure daily rainfall (a few at 15-minute intervals), temperatures, pressures, and relative humidity. For this paper, only daily precipitation data will be used.

Table 1: Descriptive table of the pluviometers' locations and time spans, that measure the daily rainfalls. Latitudes and longitudes are relative to the WGS1994 Reference System. Some stations were omitted due to insufficient data and for brevity also, given their high number.

Stations	Latitude [°]	Longitude [°]	First Year	Last Year	Discarded Years	N° of years
Kogho	12,65	-0,67	2022	2023	2024	2
Salogo	12,45	-0,62	2022	2023	2024	2
Zeguedeguen	12,97	-0,45	2022	2022	2022	None
Bogande	12,98	-0,16	1982	2022	None	41
Manni	13,25	-0,21	1980	2021	2009	41
Aribinda	14,23	-0,87	1960	2021	2003, 2004	60
Bani	13,72	-0,17	1960	2021	1980, 1981, 1994, 1999, 2000, 2001, 2002, 2003, 2005, 2014, 2015	51
Barsalogho	13,42	-1,06	1960	2022	None	63
Boulsa	12,66	-0,57	1960	2019	2001, 2002, 2004, 2005, 2006	55
Bouroum	13,62	-0,65	1968	2022	1979, 1980, 1981, 1982, 1983, 1984, 1985, 1990, 2002, 2005, 2006, 2008, 2020, 2021	41
Dakiri	13,29	-0,26	1962	2021	1976, 1978	58
Gayeri	12,65	0,49	1971	2022	1980, 1981, 1982, 1990, 1994, 1996, 1997, 1998, 2017	43
Kossougoudou	12,94	-0,23	1961	2021	2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011	53
Piela	12,70	-0,13	1961	2022	2001, 2020	60
Sebba	13,44	0,52	1960	2022	1984, 1986, 1987, 2001, 2002, 2003, 2004	56
Gotheye	13,86	1,57	1994	2023	2010	29
Karma	13,67	1,81	1994	2015	2006, 2007, 2008	19
Namaro	13,68	1,71	1994	2019	2002, 2003, 2004, 2005, 2006, 2007, 2015, 2016, 2018	17

For a better understanding, the time spans of the considered stations, were plotted also:

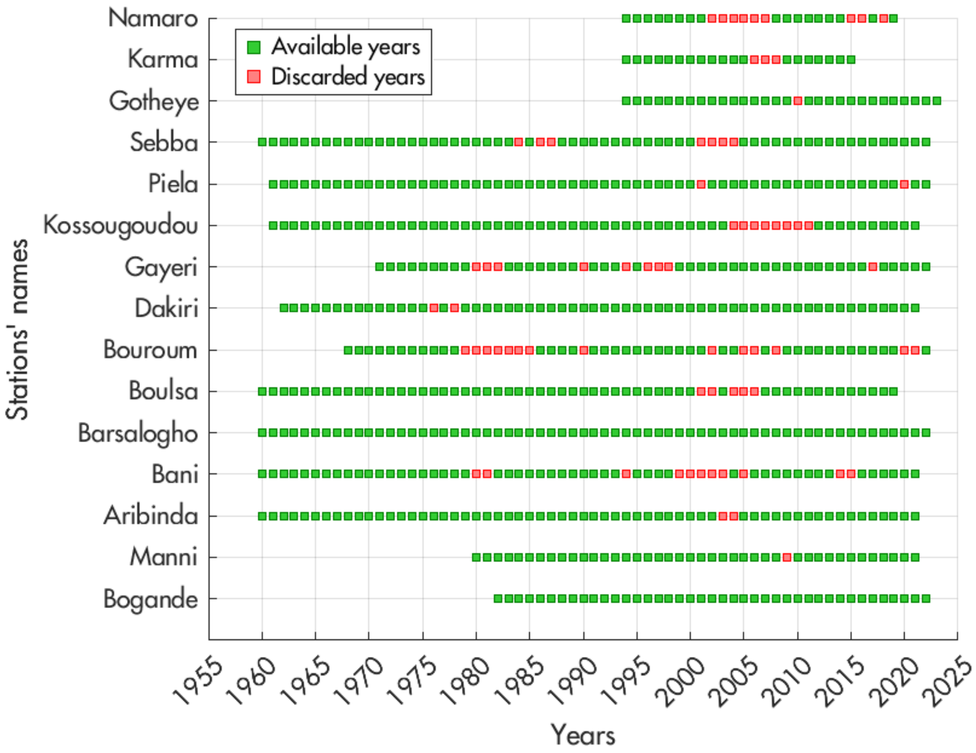


Figure 3: Time spans of the various considered stations, named after the localities' names.

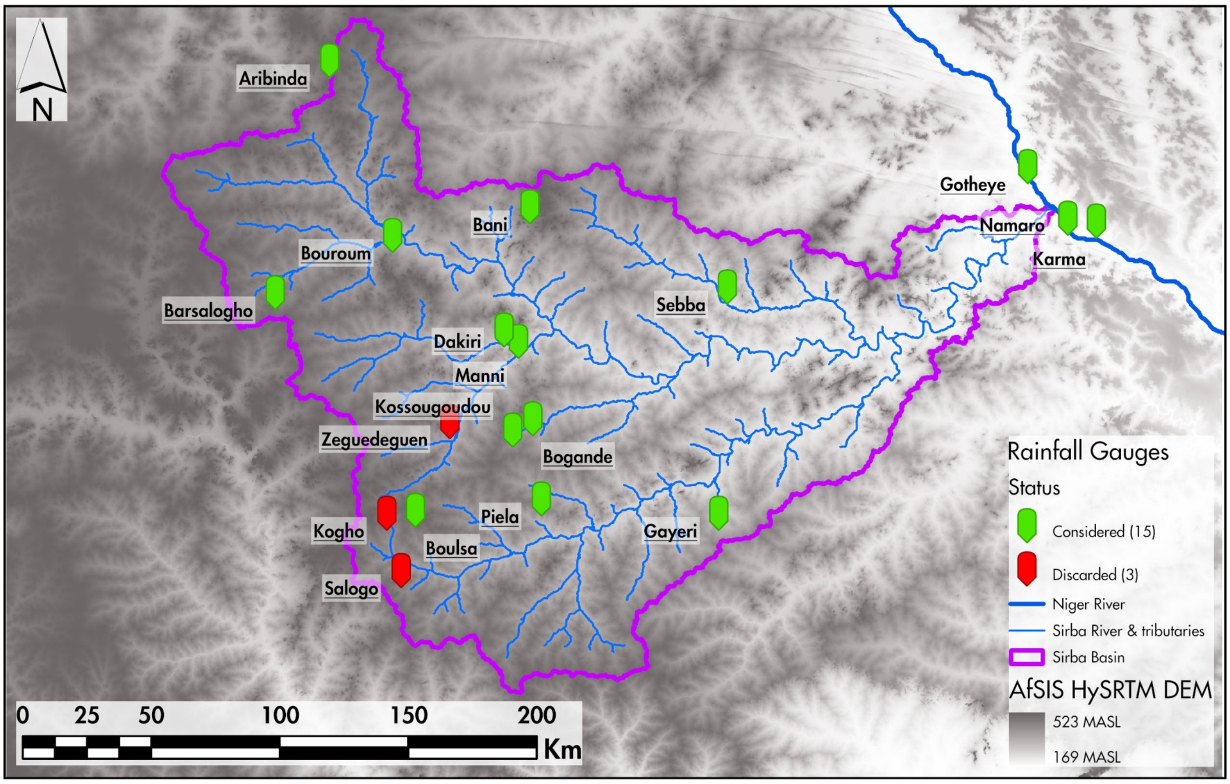


Figure 4: Map of the considered rainfall gauges (stations) relative to the Sirba basin.

Initially, the data was in a .xlsx file with a single sheet containing the headers: Station, Day, Month, and Rainfall (mm). The datasets were divided into sheets by station name (using the name of the locality for simplicity) for easier use. This division is useful because, for the SLAPIS Project, the towns of Bogande, Manni and Sebba are key locations, due to their population numbers and vicinity to the rivers.

As a rule of thumb, for each station, years with less than 10% of data (36 days with missing rainfall data) were discarded, as indicated in Table 1 & Figure 3.

This is particularly important to note, as the study area shows a pattern in rainfall distribution. In this part of the Sahel, the rainy season (or Monsoon season) usually spans from June to September (often from May to October). In most cases, interruptions do not occur during the Monsoon season, so useful data is not lost. The excluded years, in the majority of cases, show interruptions lasting 2 to 4 months, in some cases during the rainy season, making them unreliable.

The aim of this paper is to investigate the series of Annual Maximums (AM) of daily rainfall, thus the stations indicated in the previous table, such as Kogho, Salogo and Zeguedeguen among many others, unfortunately cannot be used because of insufficient data (see Table 1 & Figure 4). The small number of rainfall gauges will inevitably represent a serious challenge to map rainfall statistics, with various methods (see Spatial interpolation).

To get a general idea of how rainfall is distributed across all stations and years, some simplified plots can be generated. For each station, the mean cumulative monthly rainfall is calculated for each month over all available years, which helps avoid overly detailed and repetitive yearly plots for each station.

Here, only the two plots for the stations with the maximum and minimum cumulative rainfall peaks are shown to avoid an excess of graphs. The remaining plots are included in the Annex.

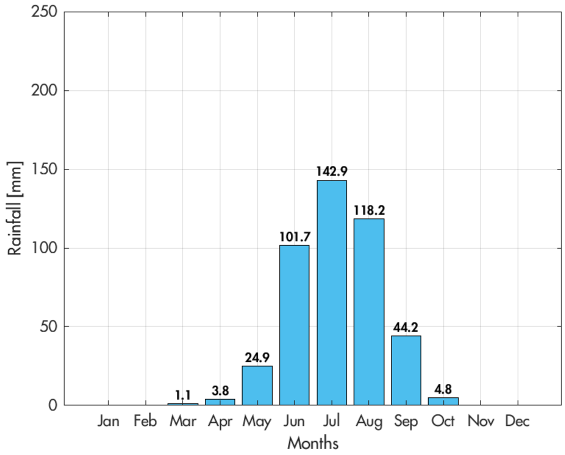


Figure 5: Averages of monthly cumulative rainfall: Salogo, with the min. cumulative rainfall peak.

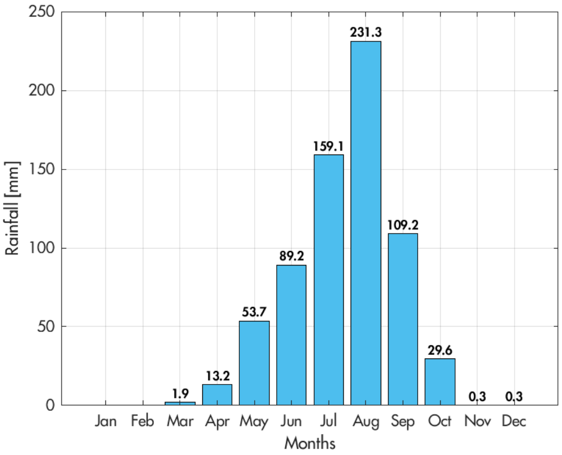


Figure 6: Averages of monthly cumulative rainfall: Gayeri, with the max. cumulative rainfall peak.

As shown in Figure 5 and Figure 6, and for all stations also, the most important months in terms of cumulative rainfall are July, August, and September. However, it is not guaranteed that these months will always yield the annual maximum daily rainfall (also known as Annual Maxima). It is evident that the months of May, June, and October typically show lesser amounts of rainfall compared to the previously mentioned months, highlighting the fact that the rainy season does not have a fixed period across all years.

The same approach can be applied to the maximum of daily rainfall values for each month across all years, providing insights into monthly extremes, but here the plots are omitted for brevity, though they are included in the Annex.

Table 2: Averages of yearly cumulative rainfalls, across the available years.

Stations	**Avg. Yearly Cumulative Rainfalls [mm]**
Kogho	693.85
Salogo	441.60
Bogande	619.38
Mane	666.50
Aribinda	463.43
Bani	511.20
Barsalogho	588.46
Boulsa	690.64
Bouroum	581.99
Dakiri	555.84
Gayeri	682.92
Kossougoudou	627.45
Piela	651.58
Sebba	577.83
Gotheye	457.80
Karma	511.46
Namaro	503.75

Also, the averages of cumulative rainfalls are computed across the available years. They display values between 400 and 700 mm, in line with the previously mentioned annual values in the region. The averages for Kogho and Salogo are relative to two years of data (see Table 1).

For a better understanding of rainfall evolution, individual yearly trends are included in the Annex section Yearly Cumulative Rainfalls.

2.3: Grid data: ERA5

A second type of data set is to be introduced.

ERA5 is the fifth generation ECMWF reanalysis that provides global climate and weather data from 1940 onward, replacing ERA-Interim. Reanalysis combines model data with global observations using data assimilation to create consistent datasets, similar to weather forecasting, but over longer periods and with reduced resolution. ERA5 offers hourly estimates of atmospheric, ocean-wave, and land-surface data, with uncertainty measured by a 10-member ensemble every three hours. The dataset is updated daily, with a five-day latency. It includes regridded data on a regular lat-lon grid and is divided into hourly and monthly products on pressure and single levels. This entry covers ERA5 hourly data on single levels from 1940 to the present [13].

Among the many variables available, the total precipitation one will be used. This parameter measures the total accumulated liquid and frozen water, including rain and snow, that reaches the Earth's surface. It combines large-scale precipitation, generated by the cloud scheme in the ECMWF forecasting system, and convective precipitation, generated by the convection scheme. Large-scale precipitation forms due to atmospheric changes on larger spatial scales, while convective precipitation represents smaller-scale convective processes [13].

This parameter excludes fog, dew, and precipitation that evaporates before reaching the ground. Accumulation is measured over specific time periods: 1 hour for reanalysis data and 3 hours for ensemble data. The values are expressed as water depth in meters, averaged over the model grid box, and should be compared cautiously with localized observations [13].

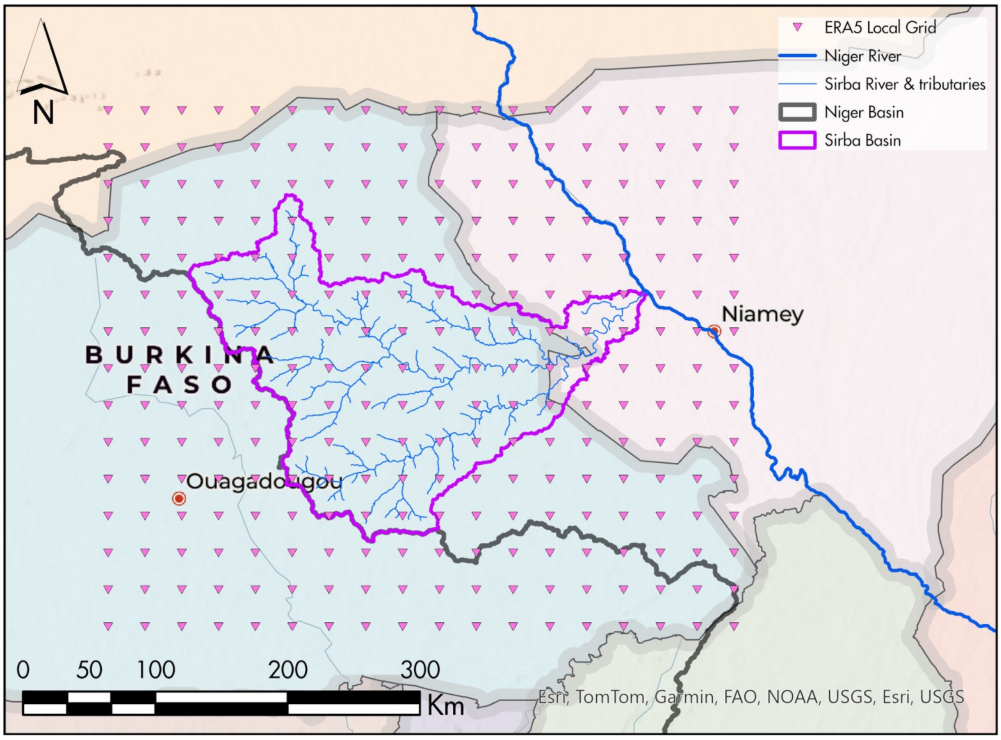


Figure 7: ERA5 Grid, overlapped with the Sirba basin.

As shown in Figure 7, the grid extent ranges from -2° to 2.25° longitude and 11.5° to 15° latitude, relative to the RS WGS1984 (EPSG 4326) coordinate system, providing ample coverage for the Sirba basin. The initial plan was to add two grid points beyond the basin's extents for additional coverage.

Since the dataset contains 270 points, no chronological rainfall plots will be made, only comparative box plots of statistical values between the dataset and the station data (see Initial data comparison).

The points are named simply by numbers on rows, with the number one assigned to the point situated in the north-western corner of the grid.

The chosen dataset can be easily downloaded from the Copernicus Catalog in batches of 12–13 years, in two different data formats: GRIB or NetCDF4 (.nc extension). The latter was chosen as it easy to import into Matlab. The data comes with three particularities: 1) the grid resolution is 0.25° lat-lon; 2) the time is measured in the standard UNIX system, the Epoch format, as seconds from January 1st, 1970; 3) the total precipitation values for each grid point are in meters, to be converted to millimeters. The batches of .nc files can be grouped together into a larger file that includes hourly gridded data from 1940 to 2023.

In simple terms, a .nc file can be conceptualized as a four-dimensional matrix: on the x-axis, the longitude; on the y-axis, the latitude; on the z-axis, the precipitation; and on the t-axis, the time in Epoch format, which can be converted into American format date.

To avoid multiple plots of the total precipitation, a single map for a specific date can be extracted through linear interpolation, for illustrative purposes. The most recent year has been chosen, with a month and day in the midst of the monsoon season, when there is a high chance of observing high values of hourly cumulative precipitation.

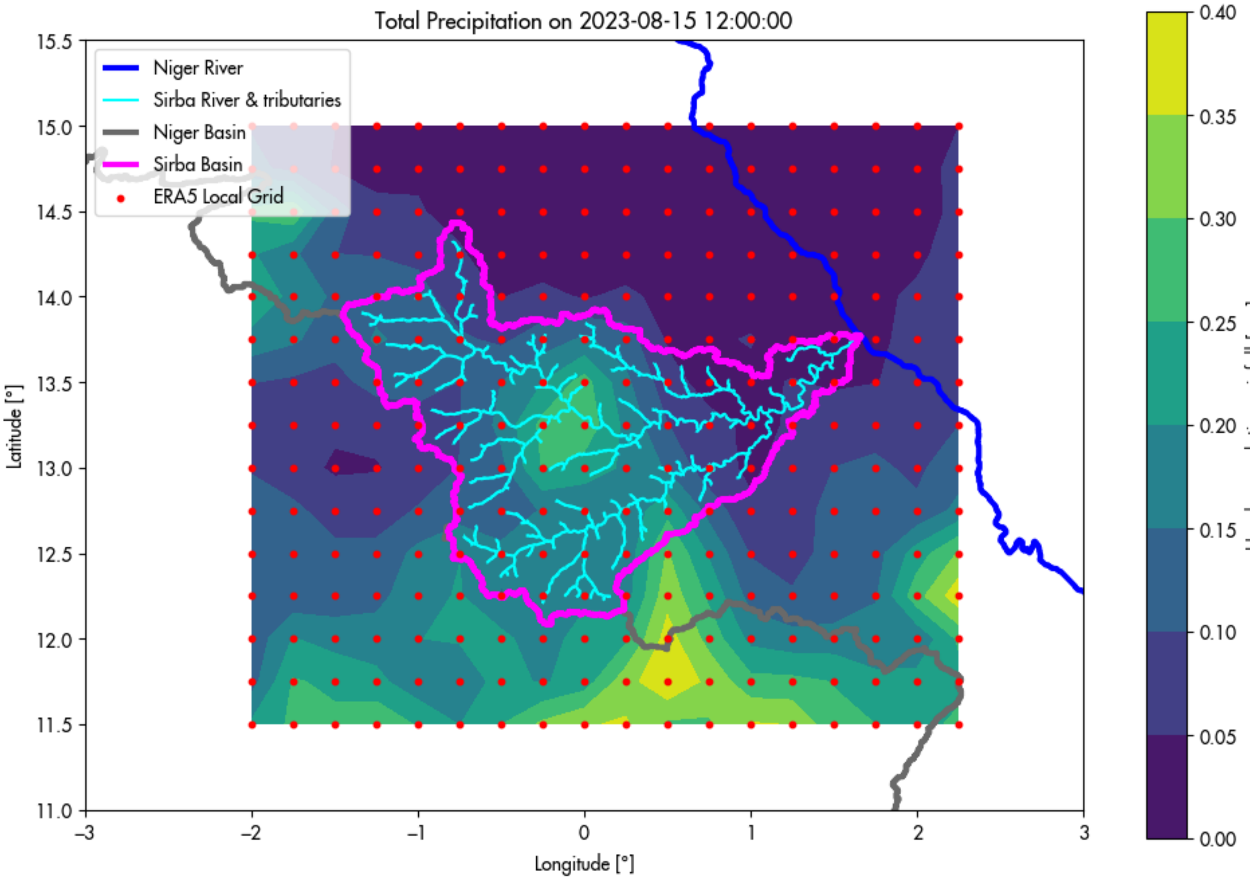


Figure 8: Example of ERA5 Grid values of total precipitation at a certain date.

As was done for the stations, for each grid point, the Annual Maxima of the daily total precipitation series can be extracted for the years 1940 and 2023, by using a 1 day time window. For each series, both Ordinary Moments and L-Moments can be calculated.

For the purposes of this paper, no distribution fitting, nor DDF and IDF curves, will be performed on the grid points. Instead, the grid points will be used solely for data correction (between the stations, as the reference, and the grid data itself) and spatial interpolation, given their regular configuration on a uniform grid.

The ERA5 grid shows some differences compared to the stations' data. Notably, the temporal span is larger and has no gaps in the data. However, the grid is a reanalysis dataset.

This type of dataset is computed using various variables, such as temperature, pressure, and humidity, among many others, and is spatially and temporally averaged to fit the 0.25° grid. While this approach effectively captures rainfall variations at a regional level, such as across the entire Sahel, aligning the data with regional trends, it may overlook local variations that can only be identified through ground data.

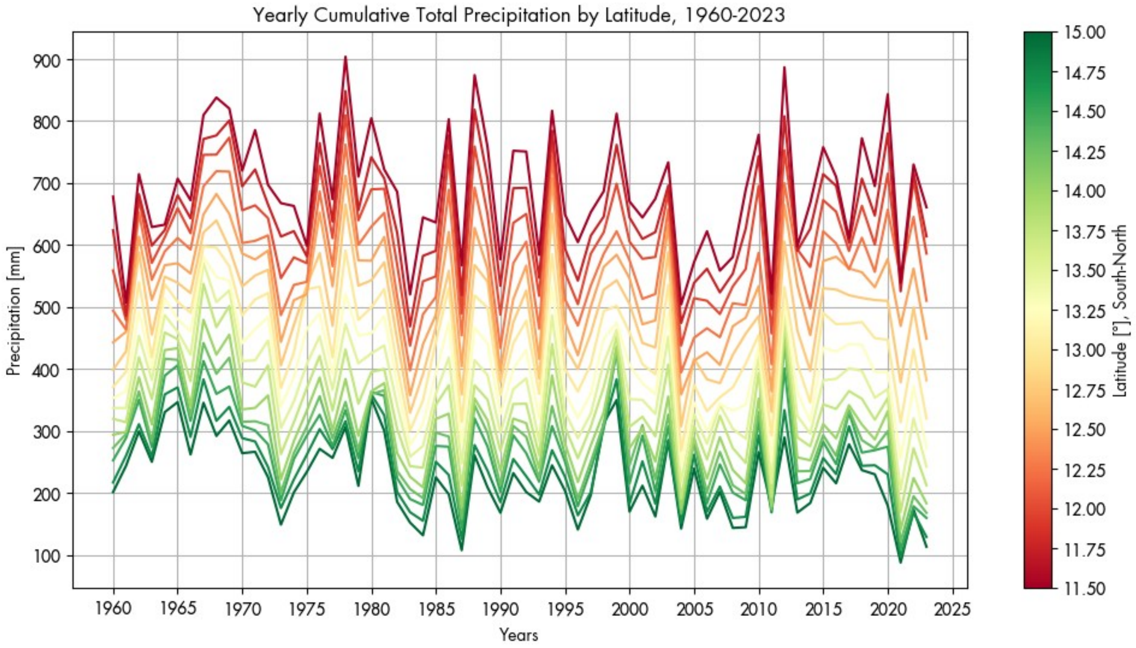


Figure 9: Yearly cumulative precipitation trends by latitude of the ERA5 local grid, between 1960 and 2023.

Some basic comparisons can be made using the average yearly rainfall. Evidently, the ERA5 grid underestimates the cumulative average yearly rainfall at smaller scales, such as the current basin, even when shorter time spans (from 1960 onward) are considered. Station values may also be higher because their time spans are shorter than those of the grid. This, at least in part, justifies the need to correct the grid data relative to the station measurements. Further explanations will be given in the sections Data correction and Discussion: Spatial interpolation.

Table 3: Average yearly cumulative rainfall of the closest ERA5 points to the stations, between 1960 and 2023. Note the difference from those in Table 2.

Stations	Closest ERA5 Point	Latitude [°]	Longitude [°]	Average Yearly Precipitation [mm]
Kogho	168	12.75	-0.75	513.87
Salogo	187	12.50	-0.50	515.08
Bogande	152	13.00	-0.25	451.81
Mane	134	13.25	-0.25	403.63
Aribinda	60	14.25	-0.75	321.27
Bani	98	13.75	-0.25	345.55
Barsalogho	113	13.50	-1.00	399.03
Boulsa	169	12.75	-0.50	487.31
Bouroum	114	13.50	-0.75	397.41
Dakiri	134	13.25	-0.25	403.63
Gayeri	173	12.75	0.50	462.03
Kossougoudou	152	13.00	-0.25	451.81
Piela	170	12.75	-0.25	476.97
Sebba	119	13.50	0.50	352.59
Gotheye	105	13.75	1.50	274.38
Karma	105	13.75	1.50	274.38
Namaro	106	13.75	1.75	280.14

3: Methods

3.1: River geometries

The basins and river geometries can be obtained using a Digital Terrain Model (DTM), also known as a Digital Elevation Model (DEM).

The one used in this case is the Africa Soil Information Service (AfSIS) Hydrologically Corrected/Adjusted SRTM DEM (AfrHySRTM), sourced from NASA's EarthData catalog in .tif format, a raster file [12].

The Africa Soil Information Service's Hydrologically Corrected/Adjusted SRTM DEM (AfrHySRTM) is a 90-meter resolution elevation raster in which depressions from the original DEM have been filled, allowing for internal drainage in certain landscapes [12]. Null cells were added to depressions over 20 meters deep with at least 1,000 pixels. This dataset, produced by the World Agroforestry Centre (ICRAF) in Nairobi, serves a wide user community for predicting soil properties and assessing erosion and landslide risks across Africa [12].

The original raster covered the northwestern part of Africa, including the Sahel region, so it was clipped for the specific study area.

GRASS GIS, a geographic information system application, provides a straightforward set of commands for basin and river network extraction. With the clipped raster, flow accumulation, drainage direction, and raster river network maps can be created using the `r.watershed` command. The level of detail is manually specified through the minimum size of the exterior watershed (for low detail, 100,000 pixels; for high detail, 10,000 pixels). This level of detail will apply to subsequent commands.

Once the drainage direction and raster river network maps are created, individual basins can be easily extracted using the `r.water.outlet` command by selecting the coordinates of the outlet point (a good approach is to use the river raster map as a visual reference, zoom in until the specific pixel is visible, and select it by clicking on it). The Sirba Basin and its sub-basins were obtained in this way.

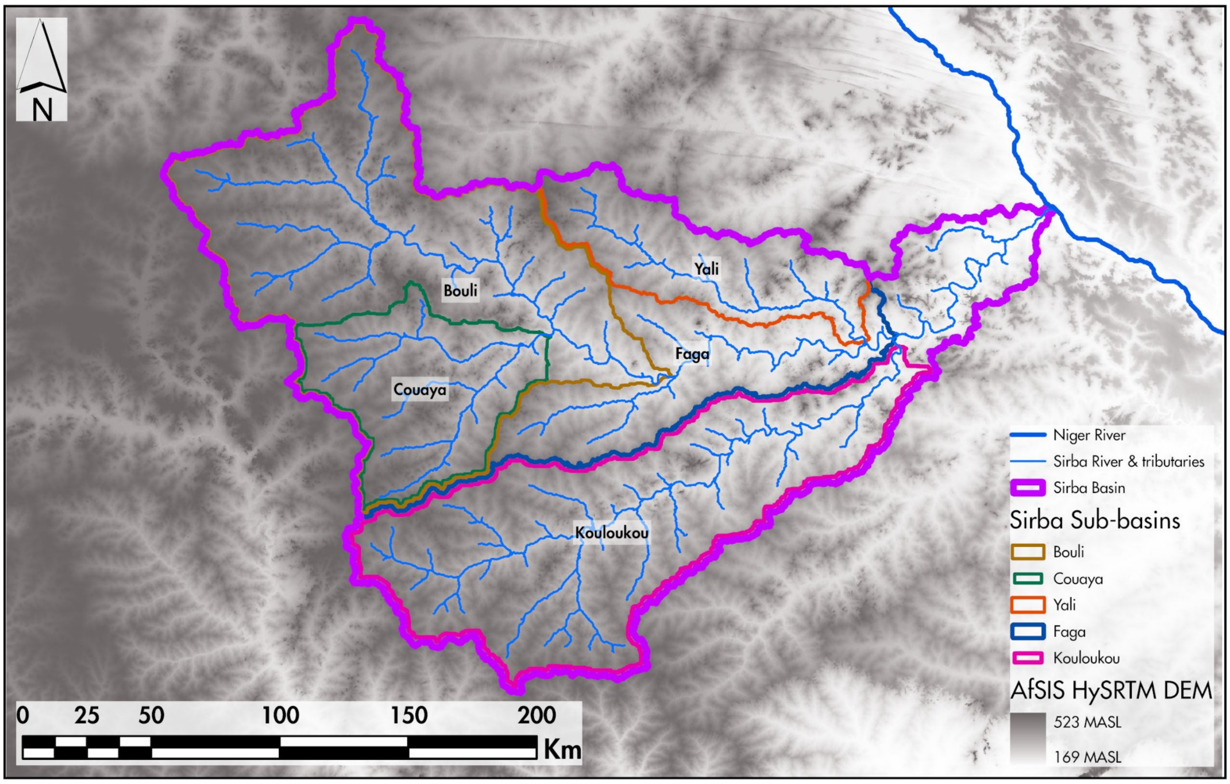


Figure 10: Sirba River and its major sub-basins, extracted using GRASS GIS.

The river network raster can be immediately converted to a vector format (such as a .shp file) or, for better results, by using the flow accumulation map through the `r.stream.extract` command to export a vector file. Any excess lines outside the larger basin can be trimmed off as needed. For simplicity, the Niger river and its basin were taken form the SLAPIS Project data catalog [14] [15]. The sub-basin's subdivisions displayed in Figure 10, take reference from Wikipedia [16] and the article written by Tiepolo et al. [17].

3.2: General statistics: Annual Maxima

Without further ado, the coding process must be explained. Most of the operations in this paper were performed using MATLAB, with data organized into .xlsx files.

In hydrology, for rainfall extraction, two methods are commonly used: the Annual Maxima and the Peak Over Threshold approaches.

For the Annual Maxima (AM), the main objective for each station (or grid point of the NetCDF dataset) is to obtain series of data relative to time windows. Since the maximum daily rainfall is being investigated, multiple moving windows were applied, ranging from 1 to 10 days, to extract the maximum cumulative values for each year, for each time window. This is crucial to understand, as the same format will be used for subsequent operations, from CDF fitting to the DDF and IDF curves. For each station, by having the yearly series of daily maximum rainfall, it is possible to compute some values and parameters that are going to be subsequently used for distribution fitting.

Since the ERA5 local grid contains 270 points, the tables for the ordinary moments and L-moments will be omitted. A general idea of the values will be given through a series of box plots (see Results: Initial data comparison).

The Peak Over Threshold (POT) approach looks at all rainfall events that exceed a set threshold, rather than focusing on just the largest event each year. This method can capture more extreme rainfall events within a year and is useful in regions with frequent or variable rainfall patterns. However, this approach was avoided because the data clearly distinguishes between the rainy and dry seasons, and rainfall on one day tends to influence nearby days, even when using different separation periods.

3.2.1: Ordinary Moments

French mathematician Pierre-Simon Laplace introduced the concept of conventional moments in the late 18th century, using them to explore characteristics of probability distributions, a foundational step for probability and statistics. Later, in the late 19th and early 20th centuries, Karl Pearson expanded on Laplace's work, formalizing moments as essential tools for describing distribution shapes, including key measures such as skewness and kurtosis.

Here are the formulas used to calculate the some basic stats and ordinary (or conventional) moments.

Table 4: Expressions of the ordinary moments.

Mean:	$\bar{\mu}=\frac{1}{n}\sum_{i=1}^{n} x_i$	Median:	$\tilde{\mu}=x_{(n+1)/2}$ *if n is odd* or $\tilde{\mu}=\frac{x_{n/2}+x_{(n/2)+1}}{2}$ *if n is even*
Variance:	$\sigma^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$	Unbiased variance:	$\sigma_{ub}^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$
Standard deviation:	$\sigma=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$	Coefficient of variation:	$c_v=\frac{\sigma}{\bar{x}}\times 100\,\%$
Skewness:	$S=\frac{1}{n}\frac{\sum_{i=1}^{n}(x_i-\bar{x})^3}{\sigma^3}$	Kurtosis:	$K=\frac{1}{n}\frac{\sum_{i=1}^{n}(x_i-\bar{x})^4}{\sigma^4}-3$

3.2.2: L-Moments

L-moments (or PWMs, Probability Weighted Moments) were introduced by James R. Hosking in 1990 as a robust alternative to conventional moments, designed especially for characterizing probability distributions and estimating parameters [18]. Hosking recognized that L-moments, which are based on linear combinations of order statistics, offer significant advantages over ordinary moments, including reduced sensitivity to outliers and lower sample variability [18].

Table 5: Expressions for the L-Moments.

L-Mean [mm]:	$L_1=b_0=\frac{1}{n}\sum_{i=1}^{n} x_i$
L-Scale [mm]:	$L_2=2b_1-b_0$ where $b_1=\frac{1}{n}\frac{\sum_{i=1}^{n}(i-1)}{(n-1)}x_i$
L_3, unscaled measure of skewness [mm]:	$L_3=6b_2-6b_1+b_0$ Where $b_2=\frac{1}{n}\frac{\sum_{i=1}^{n}(i-1)(i-2)}{(n-1)(n-2)}x_i$
L_4, unscaled measure of kurtosis [mm]:	$L_4=20b_3-30b_2+12b_1-b_0$ where $b_3=\frac{1}{n}\frac{\sum_{i=1}^{n}(i-1)(i-2)(i-3)}{(n-1)(n-2)(n-3)}x_i$
Tau, L_{CV} [ ]:	$\tau=\frac{L_2}{L_1}$
Tau-3, L-Skewness [ ]:	$\tau_3=\frac{L_3}{L_2}$
Tau-4, L-Kurtosis [ ]:	$\tau_4=\frac{L_4}{L_2}$

3.2.3: Empirical series

Plotting positions

Before proceeding with the CDF fitting, the empirical data series from the stations must be represented in order to be compared or fitted to different distributions. By having the series of annual maxima for each station, each series can be sorted, and its size (N) counted.

To avoid the maximum values of the series having a cumulative probability of 1 (which would imply they cannot be surpassed), various expressions, often called plotting position, can be used. Each expression makes different assumptions about the weight applied to each probability position in the series (i) [19] [20]. Among many there are:

Table 6: Expressions of some commonly known plotting positions.

Weibull:	$\phi_i=\frac{i}{N+1}$	Beard:	$\phi_i=\frac{i-0.14}{N+0.12}$
Hazen:	$\phi_i=\frac{i-0.5}{N}$	Gringorten:	$\phi_i=\frac{i-0.44}{N+0.12}$
Chebyshev:	$\phi_i=\frac{0.5+(i-1)}{N}$	Cunnane:	$\phi_i=\frac{i-0.4}{N+0.2}$
California:	$\phi_i=\frac{i-0.31}{N+0.28}$	Chegodayev:	$\phi_i=\frac{i-0.35}{N+0.3}$

The Gringorten and Cunnane methods are popular because they provide accurate estimates for extreme values, such as rainfall or flood data [20] [21] [22].

The Gringorten plotting position was designed to give unbiased probability estimates for normally distributed data (in some cases showing optimization for the Gumbel distribution), balancing cumulative probabilities near the median, while the Cunnane method, similar to Gringorten, applies a slightly different adjustment that favors mid-range probabilities.

For the purposes of this paper, and after numerous trials (that for brevity will be omitted), the Gringorten plotting position was chosen. In the many plots, Gringorten and Cunnane show very similar values in terms empirical probability. Of course among the available methods, these two approaches offer reasonable assumptions and more conservative probability estimates compared to the others, like the Beard plotting position (the overestimation of the return period T should be avoided, with a maximum assumed to be reasonably around 100 years, since the data series show series of data between 30 and 80 years).

Return period

To represent these empirical values, we need to understand a key concept called the period of return, often labeled as T.

The period of return is a statistical measure that estimates how often a particular event might happen over time. It's commonly used in fields like hydrology, meteorology, and risk analysis to estimate how frequently extreme events, such as floods, storms, or earthquakes, can occur.

The period of return tells us the average time between events of a certain size or greater [23]. For example, a 100-year event has a return period of 100 years, meaning that, on average, such event is expected once every 100 years.

To calculate the return period for a specific event, typically is needed a time series of past occurrences [19]. Events are ranked by size, and plotting positions are used to assign cumulative probabilities [19]. The return period T is then estimated as:

$$T=\frac{1}{1-F} \quad \text{and} \quad F=1-\frac{1}{T}$$

where F is the cumulative probability of the event (it can be either an empirical, fitted or fixed one).

By the definition itself, the return period is the average interval of time within an event will be equaled or exceeded [23]. It's important to remember that the return period does not guarantee when events will happen. For example, two 100-year events could occur in back-to-back years or be far apart. Additionally, the return period can be affected by changes in climate or environment, so historical return periods might not always predict future events accurately.

The Annual Maxima rainfall series can be normalized relative to the simple mean of each station's series, the ones form Table 11, and can be plotted together with the cumulative probability and the T (see Figure 12 to Figure 15).

The T does not completely characterize the risk in project terms. Given T, the residual risk is used, defined by the following expression:

$$\text{Residual risk:} \quad R_r=1-\left(1-\frac{1}{T}\right)^L=1-F^L$$

where L is the considered interval and R_r values essentially measure the probability of a certain event, having a return period T, of being surpassed in the L years [23].

In some cases if L << T, the residual risk can be computed as:

$$R_r=\frac{L}{T}$$

thus considering L as risk multiplier of the natural risk [23].

3.3: Diagnostic analysis

The last step before proceeding to the distribution fitting is to make a preliminary diagnostic graph. To achieve this, it is possible to represent 3 Par. Distributions in polynomial forms relative to certain L-moments, such as tau-3 and tau-4. This will provide an initial idea of which distribution best represents the data sample. Hosking and Wallis introduced this concept in their 1997 study [24] and proposed the following polynomial approximations:

Table 7: Approximations proposed by Hosking and Wallis for various distributions and lower bound.

GPA:	$\tau_4=0.20196\tau_3+0.95924\tau_3^2-0.20096\tau_3^3+0.04061\tau_3^4$
GEV:	$\tau_4=0.10701+0.11090\tau_3+0.84838\tau_3^2-0.06669\tau_3^3+0.00567\tau_3^4-0.04208\tau_3^5+0.03763\tau_3^6$
GLO:	$\tau_4=0.16667+0.83333\tau_3^2$
LN3P:	$\tau_4=0.12282+0.77518\tau_3^2+0.12279\tau_3^4-0.13638\tau_3^6+0.11368\tau_3^8$
PT3:	$\tau_4=0.12240+0.30115\tau_3^2+0.95812\tau_3^4-0.57488\tau_3^6+0.19383\tau_3^8$
Overall lower bound:	$\tau_4=0.20196\tau_3+0.95924\tau_3^2-0.20096\tau_3^3+0.04061\tau_3^4$

2 Par. Distributions are represented through points [24].

In this manner, it is also possible to represent the Log-Pearson Type 3 distribution. In their 2007 study, Griffis and Stedinger specifically describe the Log-Pearson Type 3 representation for various parameter configurations [25]. The matrix A_{LP3} contains the coefficients used in the cubic polynomial approximation of tau-4 as a function of tau-3 for the Log-Pearson Type 3 (LPT3) distribution. This approximation is particularly relevant for modeling the behavior of skewed distributions, where skewness plays a critical role in defining the shape of the distribution.

These coefficients are derived from Table A3 of the study [25], which explores various skewness values in log-space. The specific entries in the matrix correspond to different values of the log-space skewness parameter, denoted as γ.

Each row of the matrix A_{LP3} represents the coefficients for a cubic polynomial of the form:

$$\tau_4=A_0+A_1\tau_3+A_2\tau_3^2+A_3\tau_3^3 \quad \text{, where:}$$

- A_0 is the constant term (1st column of the A_{LP3} matrix);
- A_1 is the coefficient of τ_3 (the linear term and 2nd column);
- A_2 is the coefficient of τ_3^2 (the quadratic term and 3rd column);
- A_3 is the coefficient of τ_3^3 (the cubic term and 4th column).

Coefficients by skewness value:

- For γ = -1.4: coefficients: [0.0602, -0.1673, 0.8010, 0.2897];
- For γ = -1.0: coefficients: [0.0908, -0.1267, 0.7636, 0.2562];
- For γ = -0.5: coefficients: [0.1166, -0.0439, 0.6247, 0.2939];
- For γ = 0.0 (no skewness): coefficients: [0.1220, 0.0238, 0.6677, 0.1677];
- For γ = 0.5: coefficients: [0.1152, 0.0639, 0.7486, 0.0645];
- For γ = 1.0: coefficients: [0.1037, 0.0438, 0.9327, -0.0951];
- For γ = 1.4: Coefficients: [0.0776, 0.0762, 0.9771, -0.1394].

By acknowledging the precedent, the distributions can be plotted using MATLAB, together with the samples, as points. The MATLAB Online Community provides code to create the tau-4/tau-3 graph, developed by Guillaume Talbot [26], who specifically references the [18], [24] and [25] studies. The code was modified and adapted to meet the data requirements.

3.4: CDF Fitting

3.4.1: Distributions

The Annual Maxima series of each station can be fitted using various distributions. Many authors have contributed to their development. Given the large number of available distributions, only a select few can be chosen, including:

- Normal (Gaussian);
- Log-Normal (2 Parameters);
- Log-Normal (3 Parameters);
- Exponential;
- Logistic;
- Gamma;
- Pearson Type III (PT3);
- Log-Pearson Type III (LPT3);
- Gumbel (Extreme Value Type I, EV1);
- Generalized Extreme Value (GEV);
- Generalized Logistic (GLO);
- Generalized Pareto (GPA);
- Burr Type XII (BTXII).

For each of these distributions, the necessary parameters will be calculated using the previously calculated L-Moments (of all the considered stations, see Table 10), as they more effectively describe short sample datasets. All distributions will undergo statistical tests, such as Anderson-Darling and Kolmogorov-Smirnov, followed by Information Criteria calculations, to better determine which distributions most accurately represent the various rainfall series.

The preceding distributions have either 2 to 3 parameters, a probability density function (or PDF) a cumulative probability function (CDF or F), and an inverse CDF. The latter will be used to calculate the rainfall as a function of the empirical cumulative probability (the ϕ from the plotting position expression), to be used for the QQ plots as direct comparison with the sample rainfalls. Distributions with "log" in the prefix require the L-moments to be calculated from log(x) instead of the raw value x (i.e., the AM rainfall series). This said, all the necessary expressions, for each distribution are indicated in the Annex: Distributions.

Some distributions, such as the Pearson types, GLO, GPA, and Burr, are more complex to fit than others. Therefore, existing code was used as a reference for all distributions and subsequently modified to enhance compatibility with the dataset and improve graph readability. The same author previously cited, provided an extensive library of distributions implemented in MATLAB [26], and the advisor of this study supplied the necessary R code snippet (subsequently translated into Matlab syntax) for fitting the Burr distribution [27].

3.4.2: Error indicators

For the CDF fitting, various commonly known error indicators were used, including root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2).

With a small digression, the evaluation metrics of RMSE, MAE, MAPE, and R^2 have rich historical roots in the development of statistics. RMSE originated from least squares regression, a method formalized in the early 19th century by Carl Friedrich Gauss and Adrien-Marie Legendre, who focused on minimizing residuals in data fitting. MAE, stemming from Pierre-Simon Laplace's work with absolute deviations, gained popularity in time series analysis for its straightforward approach to representing model accuracy. MAPE, widely used in economic forecasting, was popularized in the 1970s by M. J. De Muth for its intuitive percentage-based interpretation. Lastly, R^2 was introduced by Karl Pearson and later refined by Ronald Fisher in early 20th-century regression analysis to measure explained variance in statistical models. These metrics have since become essential tools for evaluating models across numerous disciplines.

Given that the original sample is expressed with x and the computed one with x̂, here are the expressions:

$$errors = x - \hat{x}$$

Table 8: Expressions of the error indicators.

RMSE:	$MSE = \frac{1}{n}\sum_{i=1}^{n}(errors_i)^2 \quad ; \quad RMSE = \sqrt{MSE}$
MAE:	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert(errors_i)\rvert$
MAPE:	$MAPE_{gauss} = \frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{errors_i}{x_i}\right\rvert \cdot 100$
R^2:	$SS_{res} = \sum_{i=1}^{n}(errors_i)^2 \quad ; \quad SS_{tot} = \sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n}x_i\right)^2 \quad ; \quad R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$

3.4.3: Statistical tests

After fitting the CDF to the empirical values, the goodness-of-fit must be quantified. This is achieved using statistical tests. Although many types exist, and some newer ones have been developed in the last decade, only the two most used tests will be applied for the purposes of this paper.

Kolmogorov-Smirnov

This test directly compares the computed values with the sample values across the entire CDF [28] [29]. It calculates a statistic D, which, if large enough, will lead to the rejection of the null hypothesis, thus rejecting the theoretical assumption of the test [28] [29]. The KS test is commonly used to assess normality and is more sensitive to differences around the median of the distribution; however, one limitation is that it is less sensitive to deviations in the tails, making it less effective for small samples [28].

$$D_{plus} = max(\phi - F) \qquad D_{minus} = max(F - \phi)$$

where ϕ is the empirical cumulative probability while F is the computed one

$$D = max(D_{plus}, D_{minus})$$

$$KS_{stat} = D \cdot \sqrt{(N)} + \frac{1}{g \cdot \sqrt{(N)}}$$

$p_{V,KS} = 1 - \exp(-2 \cdot KS_{stat}^2)$; if $P_{V,KS} > 1$, then cap the value at $P_{V,KS} = 1$

if $P_{V,KS} < \alpha = 0.05$, then reject null hypothesis: data does not fit the distribution

Anderson-Darling

In contrast, this test gives more weight to differences in the tails, making it particularly useful when deviations in the tails are expected and the sample size is relatively small [30]. A modified version of this test has been used:

$$i=(1,\ldots,N) \quad \text{; AD statistic:} \quad A_{sq}=-N-\frac{1}{N}\sum_{i=1}^{N}(2i-1)\cdot(\log F_i+\log(1-F_{N-i+1}))$$

$$\text{Adjust the statistic for N size:} \quad A_{stat}=A_{sq}\cdot\left(1+\frac{0.75}{N}+\frac{2.24}{N^2}\right)$$

$$\text{if } AD_{stat} < 0.2\text{, then} \quad p_{V,AD}=1-\exp(-13.436+101.14\cdot AD_{stat}-223.73\cdot AD_{stat}^2)$$

$$\text{if } AD_{stat} < 0.34\text{, then} \quad p_{V,AD}=1-\exp(-8.318+42.796\cdot AD_{stat}-59.938\cdot AD_{stat}^2)$$

$$\text{if } AD_{stat} < 0.6\text{, then} \quad p_{V,AD}=1-\exp(0.9177+4.279\cdot AD_{stat}-1.38\cdot AD_{stat}^2)$$

$$\text{else, then} \quad p_{V,AD}=1-\exp(-1.2937+5.709\cdot AD_{stat}-0.0186\cdot AD_{stat}^2)$$

if $P_{V,AD} < \alpha = 0.05$, then reject null hypothesis: data does not fit the distribution

3.4.4: Information Criteria

Without delving too deeply into information theory and Bayesian statistics, there are certain quantities, known as information criteria (IC), that can be calculated to compare different statistical models fitted to a given dataset. To compute these criteria, the maximized log-likelihood of each distribution for each station dataset must first be calculated using f_i, which represents the individual values of the probability density function (PDF) for each distribution. A higher likelihood value indicates a better fit.

Commonly used ICs are the Akaike and the Bayesian one.

$$\hat{L}=\sum_{i=1}^{N}\log(f_i) \quad \text{where N is the sample size}$$

Akaike information criterion

Introduced in the 1974 by H. Akaike, it takes into account a penalty factor, $2\cdot n_{par}$, that increases with the number of parameters, thus discouraging over-fitting [31].

$$AIC=2\cdot n_{par}-2\cdot\hat{L}$$

For small datasets, the previous expression can become overly biased, favoring complex, over-fitted models. Therefore, an additional penalty factor is introduced, which increases when the sample size N is close to the number of parameters n_{par}. As N grows, this factor becomes negligible, causing the standard and corrected information criteria (IC) to converge. A simple guideline is that if $N < 40$ or $N/n_{par} < 40N$, the corrected version should also be computed.

$$AIC_c=2\cdot n_{par}-2\cdot\hat{L}+\frac{2\cdot n_{par}\cdot(n_{par}+1)}{N-n_{par}-1}$$

The AIC is a comparative measure, to be used on a group of fitted models. With all the previous considerations, a lower values of AIC indicates a balance between goodness of with and model complexity in terms of parameters.

Bayesian information criterion

First introduced by G. Schwarz in 1978, this IC is not based in information theory really but on Bayesian statistics. It applies a more stringent penalty factor, that grows more rapidly relative to sample size, thus favoring even simpler models [32]. Here also the preferred model has a lowest BIC value [32].

$$BIC=-2\cdot\hat{L}+n_{par}\cdot\log(N)$$

The differences between the two information criteria (ICs) are as follows: AIC focuses on finding the best data-generating process and is therefore best suited for predictive purposes, while BIC, based on Bayesian statistics, assumes that the correct model is among the models considered, making it more appropriate for model selection [33] [34]. Then for larger samples, BIC often favors models that may have less predictive power but greater interpretability, while AIC tends to favor models with higher predictive accuracy [33] [34].

From Table 25 to Table 30 are presented the values of log-L and ICs, for the previously mentioned distributions and stations. The Log-Likelihood that has a particularly low value are represented through the -∞ symbol. Extremely low values in log-L (thus PDF values close to 0) will bring forth very high values for the AIC, corrected AIC and BIC, represented through the +∞ symbol, thus rendering the distribution, for a given station, out of bounds.

3.4.5: Confidence Intervals

The confidence intervals for each station can be computed using the Monte Carlo Method.

It performs 5,000 simulations (N_{SIM}), generating sorted random probabilities matrix (N by N_{SIM}), having as number of rows the length of the AM Series of each station (N). In each simulation, a random sample of sorted rainfall heights is generated using initial parameters of each station and the random probabilities, validated against theoretical constraints derived from L-Moments (tau and tau-3), and refined to estimate distribution parameters (λ, k, c). Valid samples are stored, and confidence intervals (2.5% and 97.5% quantiles) are calculated from the estimated parameters using the inverse CDF applied to the sorted random probabilities, resulting as rainfall heights.

The process ensures consistency with Burr distribution properties by incorporating predefined domain bounds for validation. A helper function manages sample generation, L-Moment computation, and validation within each simulation.

The confidence intervals can be plotted using the same return period T of the empirical series for better comparison, as shown from Figure 17 to Figure 21.

3.5: DDF and IDF Curves

Since the dataset has been organized using moving windows, the maximum cumulative rainfall for 1 to 10 days can be extracted for all available years and stations. From this, the DDF and IDF curves can be computed.

3.5.1: Average curves

The Depth-Duration-Frequency and Intensity-Duration-Frequency curves are common expressions to compute the cumulative precipitation height and intensity over different time windows, first introduced by Robert E. Norton in 1932 [35]. Their expressions are:

$$\text{DDF:} \quad h=a\cdot d^{n} \qquad \text{IDF:} \quad i=a\cdot d^{(n-1)}$$

where "a" and "n" are parameters, and d is the considered time window

Commonly, these curves are used for hourly windows, from 1 to 24 hours, as the US agencies did in 1955 [36]. However, since the study data is at a daily resolution, the curves will be computed at this resolution as well; an uncommon approach, but not incorrect.

The series extracted with the moving windows span multiple years, equal to the available years for each station; these are referred to as durations. For each duration, as with the CDF fitting for daily AM, various parameters, such as ordinary moments and L-moments, can be calculated.

Initially, the average curves must be calculated, specifically the parameters "a" (compatible with d, so mm/day) and "n". For all curve forms (for different stations and distributions), the initial reference is the average DDF, whose parameters can be easily found by fitting a simple linear model to the logarithm of average height (relative to each duration) and the logarithm of each duration. Thus, the same parameters used for the average DDF curve will be used for the average IDF curve without additional fitting. To assess the goodness of fit, the previously mentioned error indicators will be used.

3.5.2: Growth factors

Having introduced the concepts of CDF fitting and average curves, we now turn to the next topic. The average DDF and IDF curves alone indicate only the average cumulative rainfall and intensity for a given dataset. This approach is somewhat reductive, as rainfall computed through distribution fitting can also be expressed in relation to the return period T. This is achievable here as well, using the Dalrymple method [37], also known as the index method, introduced in 1960, initially used for flood frequency analysis.

This method uses the average values of rainfall height and intensity as a reference index (over a certain period, in this case for the available years of each station), to be multiplied by a growth factor (or function), a dimensionless quantity, expressed in terms of T.

$$h(T)=h\cdot K(T)=a\cdot d^{n}\cdot K(T) \quad ; \quad i(T)=i\cdot K(T)=a\cdot d^{(n-1)}\cdot K(T)$$

The method assumes regional homogeneity, meaning it works best when applied to sites within a region that share similar hydrological and climatological characteristics; it is straightforward and leverages regional data, making it suitable for areas with short or sparse records [37].

K(T) is nothing more than another way to express the inverse CDF, but in this case is relative to all time windows, from 1 to 10 days, not only just for the daily cumulative rainfall.

As did before, the distributions necessitate different parameters, calculated through the L-moments. The approach is as follows:

- By using the rainfall series, for each station, the L-moments and the parameters, can be computed and plotted, relative to their duration (see Annex: L-Moments trends, By rainfall durations);
- The rainfall series of each duration, for each station, must be normalized to their mean;
- With the normalized series, the dimensionless L-Moments across all duration can be calculated;
- After this, for each duration, from 1 to 10 days, there will be 7 L-moments values (from L_1 to tau-4). The values can be organized into a 7x10 matrix (rows for L-Moment types, columns for the days). For each value and type (rows), the mean across all the time windows (columns) must be computed. The new average L-moments will be used to compute the parameters for each distribution, as did before, but making them dimensionless now.

The previously mentioned distributions used for CDF fitting can also be applied to express K(T). Since the same assumptions, conditions, and parameter expressions apply (except for the current approach regarding L-moments), they will be omitted here, with only the expressions for K(T) presented.

Obviously, for the Log-distributions, the L-moments for all durations are calculated on the logarithm of the normalized series, and from there the passages remain the same as for the other distributions.

Different return periods can be chosen, but since the datasets are short, as did for the plotting position, and overestimation of T must be avoided, the maximum T is 100 years:

$$T=[5,10,20,50,100]$$

Since the K(T) is an inverse CDF, it needs fixed values of cumulative probability, then:

$$\overline{F}=1-\frac{1}{T}$$

All the expressions are included in the Annex: Growth factors.

3.6: Data correction

In the Results section Initial data comparison, are shown the comparative box-plots between the Stations and ERA Local Grid. The two datasets in most cases appears shifted to one another.

Given this difference, the values must be corrected accordingly. Since the mean values of Mean (L-Mean), Median, and L-Scale values for the station data appear to be larger than those of the local grid, the correction will be made with reference to the values from the first dataset as the goal.

First, for each station point is associated the closest point form the ERA5 grid data, as shown in Table 9.

The procedure is done considering the following values, that are also mentioned in Table 4 and Table 5:

- Mean, L-Mean;
- Variance, standard and unbiased;
- Standard deviation;
- Skewness;
- Kurtosis;
- L-Scale;
- L_3;
- L_4.

For each of these values, the ratios between the all station data points and the grid one are computed (see Table 42 and Table 43).

Because of their proximity, it is possible that different station points can have the same associated grid data point, such as for Bogande and Kossougoudou, Gotheye and Karma, or Manni and Dakiri.

Table 9: Station data points with the associated grid data points.

Stations	Latitude [°]	Longitude [°]	ERA5 Points	Latitude [°]	Longitude [°]	Couples' n°
Bogande	12.98	-0.16	152	13.00	-0.25	1
Manni	13.25	-0.21	134	13.25	-0.25	2
Aribinda	14.23	-0.87	60	14.25	-0.75	3
Bani	13.72	-0.17	98	13.75	-0.25	4
Barsalogho	13.42	-1.06	113	13.50	-1.00	5
Boulsa	12.66	-0.57	169	12.75	-0.50	6
Bouroum	13.62	-0.65	114	13.50	-0.75	7
Dakiri	13.29	-0.26	134	13.25	-0.25	8
Gayeri	12.65	0.49	173	12.75	0.50	9
Kossougoudou	12.94	-0.23	152	13.00	-0.25	10
Piela	12.70	-0.13	170	12.75	-0.25	11
Sebba	13.44	0.52	119	13.50	0.50	12
Gotheye	13.86	1.57	105	13.75	1.50	13
Karma	13.67	1.81	105	13.75	1.50	14
Namaro	13.68	1.71	106	13.75	1.75	15

The Max ratios are avoided because they are simply the maximum values in the AM series of rainfall for each data point. They could be considered independent, at least for the purpose of this analysis, but not in theory. At a large scale, the maximum values for each station and grid point are linked due to the regional climate phenomenon.

Additionally, the ratios for the coefficient of variation and the L-Moments' ratios, such as tau, tau-3, and tau-4, are avoided because they are simply calculated from the previously mentioned values as ratios.

After computing the ratios for each coupled point and each value, their averages are calculated and then multiplied by the original grid data to complete the correction (see Ratios at Table 42 and Table 43).

The error indicators between the station data and both the original and corrected grid data can be computed and presented in tables, from Table 44 to Table 47. The error reduction will then be shown by comparing the error indicators for the original grid data versus the corrected grid data (see Table 48 and Table 49).

3.7: Spatial interpolation

The interpolation of the station's Ordinary Moments and L-Moments can, in theory, be performed using multiple methods, such as Spline, Natural Neighbor, Thiessen polygons, Inverse Distance Weighting (IDW), and Kriging, among others, to compute values at ungauged points. Obviously, deterministic methods like Spline must be avoided because their uncertainties cannot be quantified and because they generate unrealistic results, at least for the stations' dataset (negative values for rainfall).

Therefore, to represent the Burr Type XII quantiles, growth factors, and the L-moments used to compute them, maps of the stations' Thiessen polygons were used (see Annex's section, Basic spatial interpolation: Thiessen Polygons).

Thiessen polygons (or Voronoi cells) are constructed by first creating a Delaunay triangulation between the considered stations, which forms triangles that are as close to equilateral as possible [38] [39].

The polygon boundaries are then defined by the perpendicular bisectors of the triangle edges, passing through the midpoints of those edges. These polygons are trimmed to a 30 km boundary outside the Sirba Basin, and the corresponding values of quantiles and K(T) are applied to them [38] [39].

ArcGIS offers a comprehensive toolbox with various Kriging methods, including Ordinary Kriging, Simple Kriging, and Co-Kriging, among others, called Geostatistical Wizard, which was used in this paper [40] [41] [42].

Unfortunately, the stations are too few and unevenly spaced to effectively apply some methods. The data lacks clear spatial correlation, which is critical for interpolation techniques like IDW and Kriging that rely on such correlations.

Several attempts were made to interpolate the station data using the Ordinary Kriging method, including:

- Interpolation of the Ordinary Moments and L-Moments of the Annual Maxima of daily rainfall series;
- Interpolation of cumulative annual rainfall for the available stations (noting that some stations have discarded years);
- Interpolation of cumulative monthly rainfall for each year (again, accounting for some stations with discarded years or interruptions lasting up to 36 days).

Most of these attempts yielded unsatisfactory results in terms of QQ plots, density plots, and error indicators. As a result, these attempts are not included in this paper. An exception to this was the Kriging of cumulative rainfall for June and July of 1974, which achieved an R^2 value of 0.50, but alone by themselves are not enough.

For these reasons, the ERA5 local grid was used as an alternative. The grid is regular and provides better spatial correlation, making it a more reliable source for this analysis. In the same instances of Ordinary Kriging previously mentioned, the ERA5 local grid produced very good results. However, since the ERA5 local grid is a reanalysis dataset, it does not accurately represent ground data from the stations, which exhibit higher values for some Ordinary Moments and L-Moments of the Annual Maxima series. Therefore, the grid data was adjusted relative to the station data, as outlined in the previous section, Data correction.

Before explaining the steps involved in working with the corrected ERA5 data, it is necessary to provide a brief explanation of the Ordinary Kriging method.

3.7.1: Ordinary Kriging

Geostatistics is a vast subject that holds particular importance in environmental analysis. The core objective is to robustly predict values at unobserved points using a given spatial dataset. This dataset is typically represented in three dimensions: x (or longitude), y (latitude), and z (the dataset values to be interpolated).

Kriging involves some important steps:

- Expressing the sample data through a semi-variance estimator, often referred to as the experimental variogram;
- Fitting a semi-variogram model to this estimator;
- Using the semi-variogram model function to calculate weights and perform interpolation.

Kriging takes its roots from the 1951 paper by D. Kriege, who introduced the method for the gold deposits in Witwatersrand, South-Africa [43] [44] [45]. G. Matheron improved the method and introduced the term "Kriging" in honor of the original author [43].

It is clear that there are similarities between Kriging and CDF fitting, such as the semi-variance estimator and plotting position, probability distribution, and semi-variance model function. Similarly, the parameters of the variogram model are critical, as they significantly affect the results. These parameters vary depending on the configuration of the original dataset and its sampling [43].

For a given phenomenon, a dataset under certain assumptions can be considered stationary through a random process. This process is defined as the sum of a constant mean and a random quantity with a mean of zero. Matheron introduced this concept in 1963, although it is an approximation of real-world phenomena where the mean is not constant and covariance does not exist [43] [45].

Matheron initially purposed this initial estimator of the variogram:

$$\text{semi-variance:} \quad \gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [(Z(x_i) - Z(x_i + h))^2] \quad \text{where}$$

N(h) is the number of pairs separated by the lag distance h, $Z(x_i)$ and $Z(x_i+h)$ are values at x_i and x_i+h

The semi-variance can be computed in different directions, along x, y, or all directions, by dividing the space into bins described by radial coordinates. The lag h requires careful consideration to provide reliable estimators, taking into account factors such as sample size, spacing, and robust measurements at each site [43].

In situations where the data lacks such reliability, exhibits clear skewness, or contains outliers, alternative estimators to Matheron's have been proposed. These include methods developed by Cressie & Hawkins (1980), Dowd (1984), and Genton (1998). These estimators account for the presence of outliers and apply corrections based on either the median or the normally distributed differences over lag distance [43]. These estimators yielded the best results for the few significant Ordinary Kriging attempts.

There are multiple variogram model functions:

- Bounded: for example, Spherical, Circular, Gaussian, and bounded asymptotically, Exponential;
- Unbounded: Power function and Linear;

All of these functions have some parameters in common that describe the spatial structure of a dataset [43]:

- The nugget variance "c_0" represents the variability at very short distances, typically near zero, and accounts for small-scale fluctuations or measurement errors that are not captured by the spatial process being modeled. It reflects micro-scale variations or noise.
- The range "a" is the distance at which the variogram reaches its maximum value, or the sill. Beyond this distance, spatial correlation between points diminishes, and the data points become essentially uncorrelated.

- The sill variance "c_0+c_1" represents the total variability in the dataset, indicating the upper limit of the semi-variance after which further increases in distance do not add significant new variability.

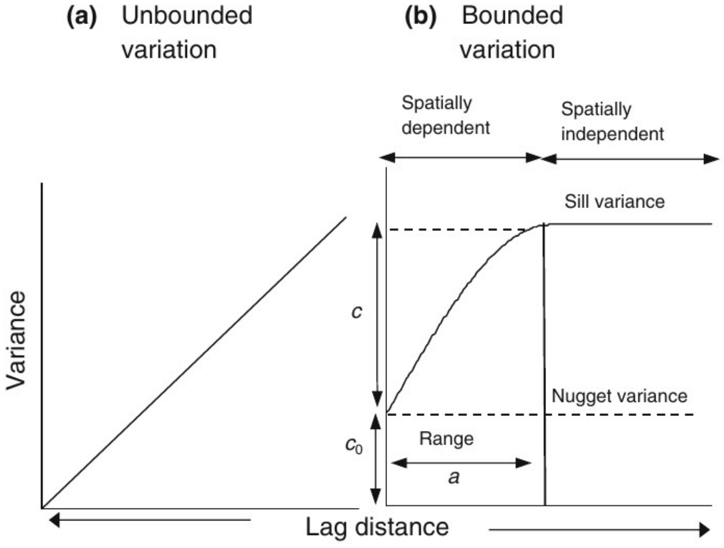


Figure 11: Types of semi-variograms and their limits, source [43].

Though the mathematical details are omitted here, the key idea is that the process incorporates the inherent spatial correlation of the data to generate the most reliable estimate.

The predicted value at an unsampled location is obtained by taking a weighted average of the sampled values. The weights assigned to each sampled point are derived from the spatial autocorrelation structure of the data, which is modeled using the semi-variogram [43].

The core principle behind this method is that points that are closer to the unsampled location are more strongly correlated with it, and therefore, they have a greater influence on the predicted value than points that are farther away [43].

These weights are calculated by solving a system of equations, which ensures that the prediction is unbiased and accurately reflects the spatial distribution of the data. This system of equations guarantees that the predicted value at an unsampled location is a linear combination of the sampled values, with the weights dependent on the spatial structure as encoded in the semi-variogram [43].

The adopted procedure for Ordinary Kriging involves specific steps, following the order proposed by ArcGIS's Geostatistical Wizard Interface [41], which has been used for all instances of such interpolation:

- A Stable semi-variogram function, optimized for the dataset used, represented by Matheron's estimator;
- A circular searching neighborhood with four sectors at 45° angles, utilizing the semi-variance values derived from the stable function;
- Cross-validation graphs, including the normal QQ plot, density of the predicted and measured dataset, predicted vs. measured QQ plot, and error indicators.

For the Moments values, since these errors are based on the corrected grid itself as the initial dataset, they cannot be considered reliable in relation to the station values. The true error indicators to consider are those outlined in the Grid data correction section Error indicators.

For the purposes of this paper, the last step will be omitted. However, as previously mentioned, the ERA5 corrected grid performs very well in terms of predicted values derived from the initial dataset, for both the Moments and Quantile values.

3.7.2: ERA5 quantiles

The final step is to calculate the quantiles (by return period T) of the ERA5 corrected grid using the grid's corrected L-Moments and then interpolate them with Ordinary Kriging.

The considerations made for the station data, including error indicators, statistical tests, and information criteria that led to the selection of the Burr Type XII distribution, cannot be applied to the corrected grid data. This is because only the corrected moments' values are analyzed, rather than the entire AM series of total precipitation for each grid point. Nevertheless, three distributions are used: Gumbel, GEV, and Burr Type XII, as they were the best-fitting distributions for the stations. Thus, for the purpose of this paper, the various error indicators, statistical tests, and information criteria are not applied to the corrected grid data.

The Burr Type XII distribution is particularly complex to apply, as its parameters are valid only within certain intervals of tau and tau-3. If this condition is not respected, the distribution becomes either a Generalized Pareto or Weibull distribution (GEV with the shape parameter set to 0). This condition was, however, respected for the corrected grid points. Given the poor correlation of tau and tau-3 (ratios from L_1, L_2 and L_3) for the station and the corrected grid (see Table 45 and Table 47), this may increase uncertainties in the quantiles.

As a precaution, two commonly used distributions in hydrology, the Gumbel (EV1) and Generalized Extreme Value (GEV) distributions, are also used to compute the quantiles. Although these distributions were not ultimately chosen to fit the rainfall samples from the station, they generally show good fitting performance (see Table 16 and Table 28). However, they are less robust for the Anderson-Darling test (see Table 22). Additionally, the GEV uses only one L-Moments ratio, tau-3, which may reduce uncertainties when fitting the corrected grid quantiles (see Annex section: Distributions, 3 Par. for the expressions).

4: Results

4.1: General statistics

4.1.1: Ordinary Moments

The maximum values of the AM's series of daily rainfall, for each station, was also extracted.

Table 10: Values of max. and ordinary moments, for the AM series of daily maximum rainfall of the gauge stations.

Stations	Max [mm]	$\bar{\mu}$ [mm]	$\tilde{\mu}$ [mm]	σ^2 [mm^2]	σ_{ub}^2 [mm^2]	σ [mm]	c_v [ ]	S [ ]	K [ ]
Bogande	125.80	61.99	55.30	322.66	314.79	17.96	0.29	1.42	5.38
Manni	151.10	70.04	63.00	593.65	579.17	24.36	0.35	1.41	4.85
Aribinda	135.70	58.33	54.50	449.16	441.67	21.19	0.36	1.26	5.20
Bani	118.50	59.34	56.00	408.27	400.27	20.21	0.34	1.01	3.71
Barsalogho	122.90	62.40	57.00	352.21	346.62	18.77	0.30	1.31	4.64
Boulsa	239.40	66.25	59.90	838.24	823.00	28.95	0.44	4.00	24.40
Bouroum	123.40	59.40	55.80	331.44	323.36	18.21	0.31	1.89	7.45
Dakiri	104.90	59.71	57.45	323.64	318.06	17.99	0.30	0.51	2.75
Gayeri	109.00	65.08	61.50	232.62	227.21	15.25	0.23	1.11	4.30
Kossougoudou	175.00	61.92	56.00	637.27	625.25	25.24	0.41	2.76	11.47
Piela	170.40	63.30	56.25	591.95	582.08	24.33	0.38	2.24	8.96
Sebba	143.50	60.90	53.05	529.18	519.73	23.00	0.38	1.17	4.62
Gotheye	130.50	59.08	60.00	405.07	391.10	20.13	0.34	1.53	6.58
Karma	135.00	59.37	52.00	627.49	594.46	25.05	0.42	1.48	5.58
Namaro	190.00	66.81	57.00	1429.43	1345.34	37.81	0.57	2.31	7.83

4.1.2: L-Moments

Table 11: Values of L-moments and their ratios, for the AM series of daily maximum rainfall of the gauge stations.

Stations	L_1 [mm]	L_2 [mm]	L_3 [mm]	L_4 [mm]	τ [ ]	τ_3 [ ]	τ_4 [ ]
Bogande	61.99	9.52	2.53	1.68	0.15	0.27	0.18
Manni	70.04	12.79	3.77	2.77	0.18	0.29	0.22
Aribinda	58.33	11.41	2.26	1.90	0.20	0.20	0.17
Bani	59.34	11.13	2.31	1.43	0.19	0.21	0.13
Barsalogho	62.40	10.02	2.62	1.59	0.16	0.26	0.16
Boulsa	66.25	12.30	3.90	3.55	0.19	0.32	0.29
Bouroum	59.40	9.14	2.41	2.10	0.15	0.26	0.23
Dakiri	59.71	10.11	1.24	1.39	0.17	0.12	0.14
Gayeri	65.08	8.26	1.71	1.54	0.13	0.21	0.19
Kossougoudou	61.92	11.05	4.39	4.36	0.18	0.40	0.39
Piela	63.30	11.60	4.14	3.47	0.18	0.36	0.30
Sebba	60.90	12.49	2.75	1.63	0.21	0.22	0.13
Gotheye	59.08	10.53	2.03	1.89	0.18	0.19	0.18
Karma	59.37	13.08	3.73	4.07	0.22	0.29	0.31
Namaro	66.81	17.34	9.29	6.52	0.26	0.54	0.38

The L-moments can be plotted as a function of the years for each station. These plots provide information on how yearly rainfall extremes vary over the decades and are included in the Annex: L-Moments trends.

4.1.3: Empirical series

Here are some possible representation of the AM empirical series, relative to the Gringorten method.

Plotting position

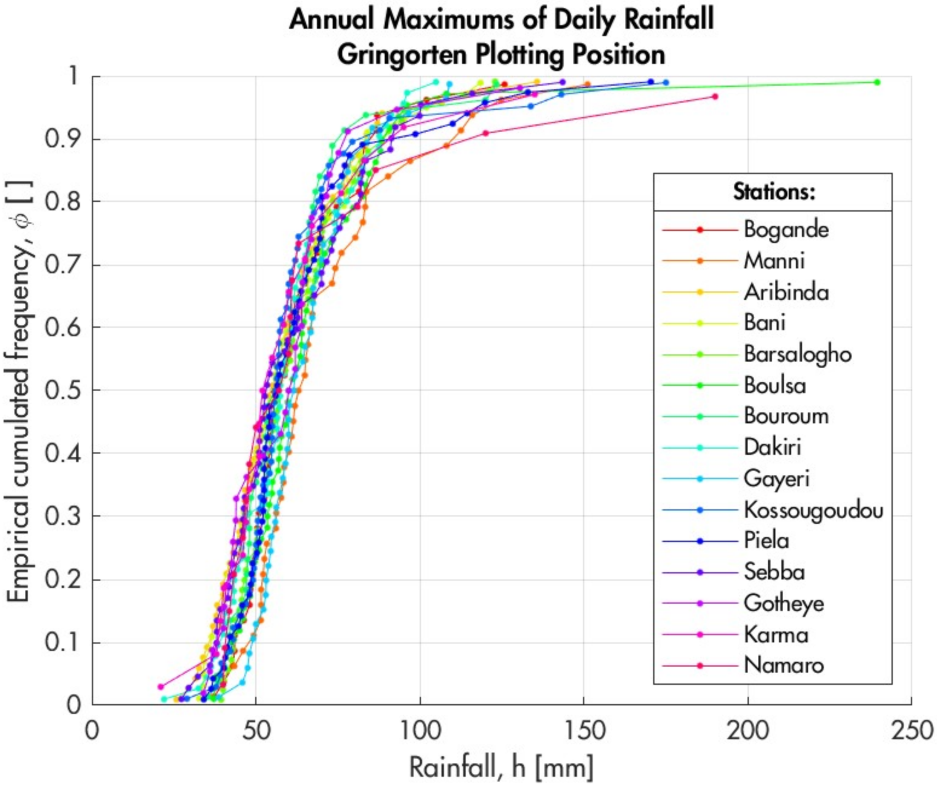


Figure 12: ϕ vs sample AM series, for the considered stations.

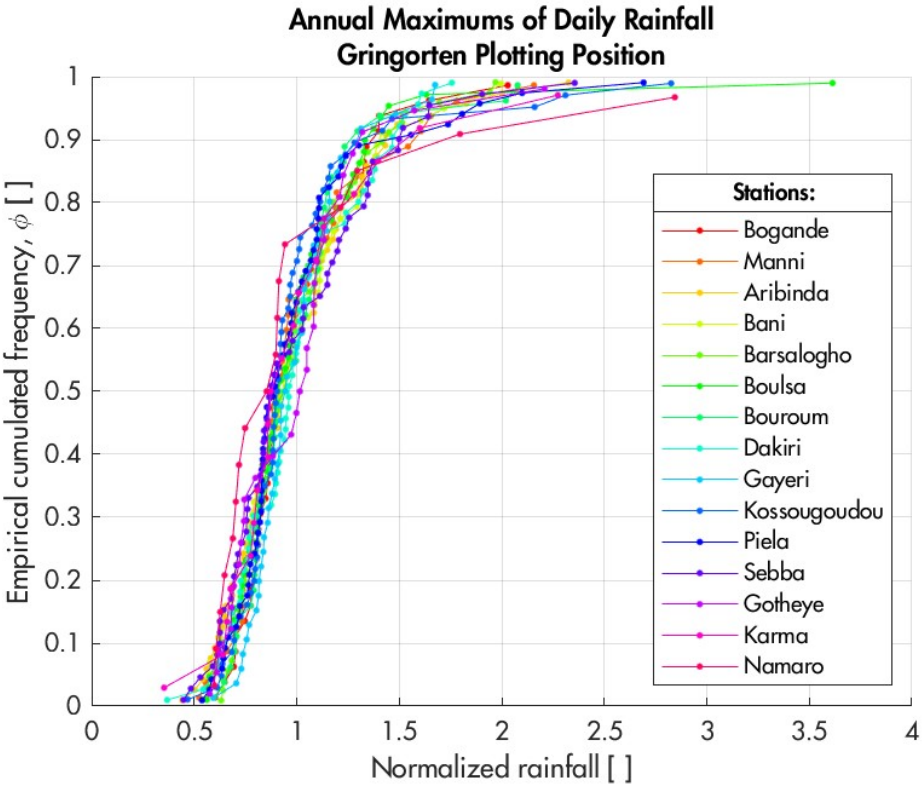


Figure 13: ϕ vs normalized sample AM series, for the considered stations.

Return periods

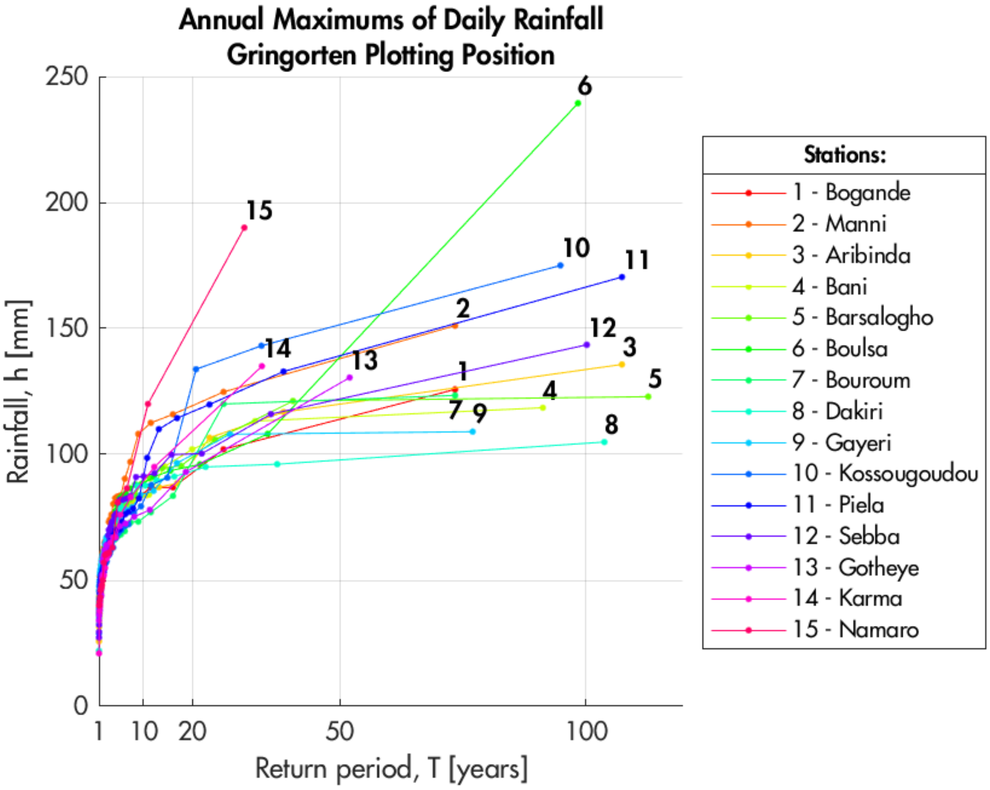


Figure 14: Sample AM series vs return period T, for the considered stations.

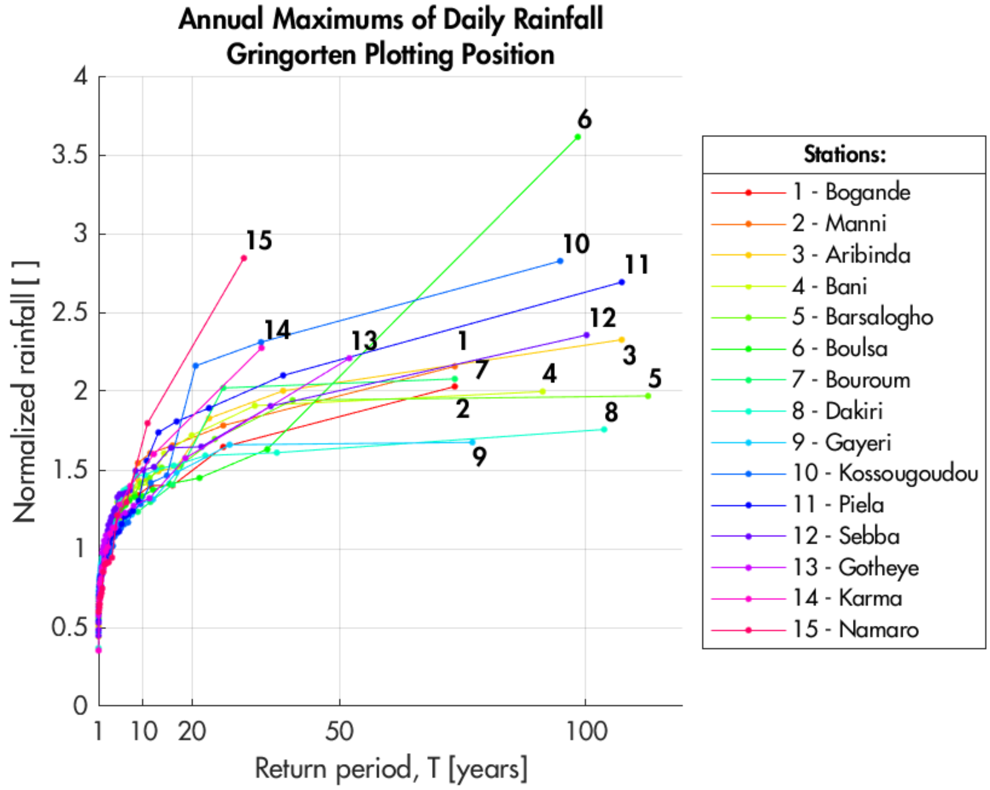


Figure 15: Normalized sample AM series vs return period T, for the considered stations.

4.2: Diagnostic analysis

Table 12: Values of tau-3 and tau-4 for the considered stations.

Sample nr	Stations	τ_3 [ ]	τ_4 [ ]
1	Bogande	0.27	0.18
2	Manni	0.29	0.22
3	Aribinda	0.2	0.17
4	Bani	0.21	0.13
5	Barsalogho	0.26	0.16
6	Boulsa	0.32	0.29
7	Bouroum	0.26	0.23
8	Dakiri	0.12	0.14
9	Gayeri	0.21	0.19
10	Kossougoudou	0.4	0.39
11	Piela	0.36	0.3
12	Sebba	0.22	0.13
13	Gotheye	0.19	0.18
14	Karma	0.29	0.31
15	Namaro	0.54	0.38

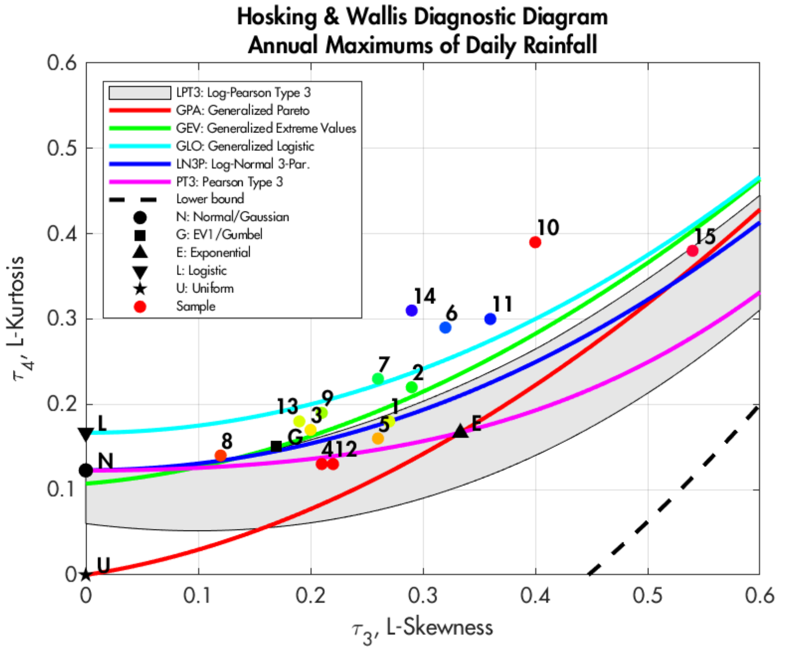


Figure 16: Diagnostic graph of tau-4 as function of tau-3.

The Log-Pearson Type III distribution, covering a wide range of skewness values, is shown as an area with boundaries at skewness values γ from -1.4 to 1.4, while two-parameter distributions are displayed simply as points.

In Figure 16, some samples align with specific distribution curves, such as Sample 1 (Bogande) and Sample 15 (Namaro), both fitting with the Log-Pearson Type III distribution. Other samples, however, appear closer to multiple distribution curves, though with less evident overlap, as seen with Sample 3 (Aribinda). Samples 6, 10, 11, and 14 show notably high tau-4 values, exceeding 0.25, and do not align with any curve. Given the assumptions of this graph, these samples could pose challenges for distribution fitting.

It is important to consider that, due to the small sample size (15 stations), the uncertainties associated with tau-3 and tau-4 may be significant. Nonetheless, the data appears reasonably realistic and discretely clustered for the values of tau-3 and tau-4, which supports the accuracy of the calculations.

4.3: CDF Fitting

4.3.1: Error indicators

The expressions of the error indicators are in Table 8.

Table 13: Error indicators for the Normal, Log-Normal 2 Par. and Log-Normal 3 Par. distributions.

Distributions	**Normal**				**Log-Normal 2 Par.**				**Log-Normal 3 Par.**			
Stations & Err. Ind.	**RMSE**	**MAE**	**MAPE [%]**	**R^2**	**RMSE**	**MAE**	**MAPE [%]**	**R^2**	**RMSE**	**MAE**	**MAPE [%]**	**R^2**
Bogande	6.18	4.40	7.31	0.88	4.00	2.61	3.96	0.95	2.01	1.30	1.99	0.99
Manni	8.68	6.91	10.16	0.87	5.23	3.71	4.82	0.95	2.67	1.88	2.76	0.99
Aribinda	6.17	3.90	7.25	0.91	2.45	1.18	1.66	0.99	1.90	0.98	1.31	0.99
Bani	5.52	3.92	7.34	0.92	2.26	1.57	2.74	0.99	2.02	1.30	2.08	0.99
Barsalogho	6.25	4.47	7.41	0.89	3.61	2.40	3.71	0.96	2.00	1.10	1.54	0.99
Boulsa	17.59	7.14	9.52	0.62	15.52	3.76	3.50	0.71	11.78	3.94	4.08	0.83
Bouroum	7.77	4.81	7.57	0.81	6.11	2.88	3.95	0.88	4.32	2.36	3.26	0.94
Dakiri	3.40	2.86	5.16	0.96	2.93	2.12	3.53	0.97	2.63	2.09	3.52	0.98
Gayeri	4.33	3.12	4.73	0.92	2.85	1.70	2.38	0.96	2.00	1.22	1.75	0.98
Kossougoudou	14.25	8.78	12.82	0.68	12.39	5.76	7.00	0.75	6.24	3.97	5.87	0.94
Piela	11.69	8.17	12.27	0.77	9.05	5.09	6.50	0.86	3.61	2.46	3.65	0.98
Sebba	6.76	4.58	8.87	0.91	2.67	1.98	3.30	0.99	2.22	1.67	2.62	0.99
Gotheye	7.60	4.71	8.02	0.85	5.47	3.52	5.58	0.92	4.93	3.48	5.45	0.94
Karma	9.28	6.53	11.17	0.86	5.68	3.79	6.87	0.95	4.16	2.86	6.26	0.97
Namaro	21.24	15.50	23.22	0.66	17.53	9.76	12.02	0.77	7.18	3.82	4.05	0.96

Table 14: Error indicators for the Exponential, Gamma and Logistic distributions.

Distributions	**Exponential**				**Gamma**				**Logistic**			
Stations & Err. Ind.	**RMSE**	**MAE**	**MAPE [%]**	**R^2**	**RMSE**	**MAE**	**MAPE [%]**	**R^2**	**RMSE**	**MAE**	**MAPE [%]**	**R^2**
Bogande	2.51	1.84	3.09	0.98	4.50	3.18	5.14	0.94	6.00	4.36	7.41	0.89
Manni	2.93	2.29	3.77	0.99	5.95	4.70	6.67	0.94	8.40	6.75	10.21	0.88
Aribinda	3.58	2.85	5.57	0.97	3.54	2.01	3.27	0.97	5.93	3.80	7.62	0.92
Bani	3.81	2.69	4.69	0.96	3.00	2.09	3.80	0.98	5.69	3.98	7.77	0.92
Barsalogho	2.45	1.61	2.58	0.98	4.21	2.90	4.74	0.95	6.26	4.52	7.72	0.89
Boulsa	12.71	4.88	5.62	0.80	15.69	5.73	6.79	0.70	16.95	6.91	9.49	0.65
Bouroum	4.49	3.10	4.85	0.94	6.43	3.70	5.50	0.87	7.46	4.57	7.42	0.83
Dakiri	6.09	3.95	7.28	0.88	2.52	2.13	3.63	0.98	3.67	3.01	5.72	0.96
Gayeri	3.08	2.25	3.58	0.96	3.23	2.19	3.21	0.95	4.17	2.99	4.70	0.92
Kossougoudou	8.42	5.43	7.73	0.89	12.26	7.32	10.11	0.76	13.59	8.37	12.47	0.70
Piela	4.98	3.38	4.98	0.96	9.27	6.40	9.22	0.85	11.15	7.79	11.93	0.79
Sebba	3.93	2.91	4.91	0.97	3.71	2.51	4.61	0.97	6.70	4.62	9.37	0.91
Gotheye	5.09	4.12	6.77	0.93	6.03	3.86	6.22	0.91	7.34	4.60	8.21	0.86
Karma	4.64	3.44	7.92	0.96	6.74	4.61	7.72	0.92	8.72	6.07	10.72	0.87
Namaro	11.37	7.01	9.29	0.90	16.59	11.53	16.60	0.80	20.74	15.12	22.82	0.68

Table 15: Error indicators for the Pearson Type III and Log-Pearson Type III distributions.

Distributions	**Pearson Type III**				**Log-Pearson Type III**			
Stations & Err. Ind.	**RMSE**	**MAE**	**MAPE [%]**	**R^2**	**RMSE**	**MAE**	**MAPE [%]**	**R^2**
Bogande	2.23	1.52	2.39	0.98	1.96	1.28	1.96	0.99
Manni	2.75	2.13	3.43	0.99	2.75	1.94	2.75	0.99
Aribinda	2.25	1.21	1.69	0.99	1.85	0.95	1.29	0.99
Bani	1.76	1.14	1.80	0.99	2.34	1.32	2.09	0.99
Barsalogho	1.68	0.93	1.28	0.99	2.31	1.11	1.53	0.98
Boulsa	12.88	4.79	5.39	0.80	12.88	3.49	3.23	0.80
Bouroum	4.66	2.69	3.79	0.93	4.44	2.30	3.10	0.94
Dakiri	2.58	2.07	3.53	0.98	2.60	2.08	3.53	0.98
Gayeri	2.10	1.34	1.98	0.98	2.00	1.22	1.74	0.98
Kossougoudou	7.49	5.18	7.81	0.91	7.59	4.25	5.81	0.91
Piela	4.68	3.28	5.03	0.96	3.93	2.60	3.68	0.97
Sebba	2.27	1.73	2.79	0.99	2.40	1.75	2.68	0.99
Gotheye	5.13	3.52	5.37	0.93	4.95	3.45	5.45	0.94
Karma	4.79	3.57	7.64	0.96	4.65	3.38	6.61	0.96
Namaro	5.86	4.12	5.77	0.97	3.98	2.77	3.50	0.99

Table 16: Error indicators for the Gumbel and GEV distributions.

Distributions	**Gumbel, EV1**				**Generalized Extreme Values**			
Stations & Err. Ind.	**RMSE**	**MAE**	**MAPE [%]**	**R^2**	**RMSE**	**MAE**	**MAPE [%]**	**R^2**
Bogande	3.00	2.10	3.30	0.97	1.99	1.23	1.85	0.99
Manni	4.27	3.36	4.69	0.97	2.91	1.88	2.56	0.99
Aribinda	2.25	1.16	1.74	0.99	1.78	0.95	1.34	0.99
Bani	2.14	1.47	2.55	0.99	2.19	1.42	2.31	0.99
Barsalogho	2.67	1.70	2.81	0.98	2.33	1.27	1.79	0.98
Boulsa	14.37	4.85	5.29	0.75	11.16	3.52	3.49	0.85
Bouroum	5.22	2.75	3.86	0.92	4.18	2.21	3.03	0.95
Dakiri	3.18	2.17	3.56	0.97	2.58	2.06	3.47	0.98
Gayeri	2.09	1.22	1.76	0.98	1.98	1.17	1.65	0.98
Kossougoudou	10.71	6.28	8.40	0.82	6.03	3.41	4.90	0.94
Piela	7.58	5.18	7.30	0.90	3.47	2.22	3.18	0.98
Sebba	2.65	2.00	3.52	0.99	2.29	1.73	2.78	0.99
Gotheye	5.11	3.53	5.62	0.93	4.84	3.46	5.51	0.94
Karma	5.75	3.94	6.92	0.94	3.94	2.53	5.60	0.97
Namaro	15.93	10.97	15.75	0.81	10.11	5.05	5.06	0.92

Table 17: Error indicators for the GLO and GPA distributions.

Distributions	**Generalized Logistic**				**Generalized Pareto**			
Stations & Err. Ind.	**RMSE**	**MAE**	**MAPE [%]**	**R^2**	**RMSE**	**MAE**	**MAPE [%]**	**R^2**
Bogande	2.17	1.41	2.14	0.99	2.68	1.78	2.83	0.98
Manni	3.44	2.08	2.59	0.98	2.95	2.37	3.84	0.99
Aribinda	1.70	1.13	2.00	0.99	3.93	2.50	3.99	0.96
Bani	3.03	1.99	3.29	0.98	2.14	1.55	2.51	0.99
Barsalogho	3.08	1.69	2.43	0.97	1.84	1.09	1.67	0.99
Boulsa	10.56	3.16	3.01	0.86	13.12	4.98	5.67	0.79
Bouroum	4.07	2.02	2.82	0.95	5.15	3.12	4.49	0.92
Dakiri	3.26	2.34	3.67	0.97	3.14	2.30	4.54	0.97
Gayeri	2.19	1.17	1.58	0.98	2.76	1.94	2.94	0.97
Kossougoudou	6.10	3.14	4.41	0.94	6.82	4.57	6.84	0.93
Piela	3.64	2.11	2.93	0.98	4.40	3.09	4.72	0.97
Sebba	2.81	2.28	3.82	0.98	3.21	2.10	3.60	0.98
Gotheye	4.48	3.26	5.52	0.95	6.22	3.96	5.61	0.90
Karma	3.92	2.54	5.24	0.97	5.26	4.02	8.48	0.95
Namaro	10.84	5.35	5.31	0.91	8.23	4.12	4.20	0.95

Table 18: Error indicators for the Burr Type XII distribution.

Distribution	**Burr Type XII**			
Stations & Err. Ind.	**RMSE**	**MAE**	**MAPE [%]**	**R^2**
Bogande	2.26	1.43	2.16	0.98
Manni	3.51	2.04	2.47	0.98
Aribinda	1.71	1.15	2.03	0.99
Bani	3.09	2.15	3.60	0.98
Barsalogho	3.23	1.85	2.74	0.97
Boulsa	10.32	3.29	3.25	0.87
Bouroum	3.98	2.06	2.93	0.95
Dakiri	3.04	2.36	3.79	0.97
Gayeri	2.21	1.23	1.66	0.98
Kossougoudou	5.31	3.27	4.74	0.95
Piela	3.09	1.96	2.76	0.98
Sebba	2.87	2.36	3.96	0.98
Gotheye	4.53	3.25	5.52	0.95
Karma	3.90	2.46	4.95	0.97
Namaro	8.15	4.23	4.51	0.95

4.3.2: Statistical tests

In the following tables A-D stands for Anderson-Darling and K-S for Kolmogorov-Smirnov. P-values lower that 0.05 are highlighted in red. Since some p-values are very low, a precision of four digits has been adopted.

Table 19: P-Values for the Normal, Log-Normal 2 Par. and Log-Normal 3 Par. distributions.

Distributions	**Normal**		**Log-Normal 2 Par.**		**Log-Normal 3 Par.**	
Stations & p-values	**A-D**	**K-S**	**A-D**	**K-S**	**A-D**	**K-S**
Bogande	0.0014	0.8714	0.1513	0.7072	0.8584	0.3075
Manni	0.0001	0.9598	0.1139	0.7819	0.4524	0.4610
Aribinda	0.0128	0.5730	0.9976	0.1335	0.9990	0.2150
Bani	0.0102	0.7060	0.8020	0.2188	0.9550	0.1733
Barsalogho	0.0001	0.8881	0.1010	0.5847	0.9751	0.1776
Boulsa	0.0000	0.9133	0.2665	0.4378	0.1744	0.4592
Bouroum	0.0005	0.7166	0.1561	0.4445	0.5024	0.3069
Dakiri	0.0181	0.7957	0.3110	0.5333	0.2983	0.4456
Gayeri	0.0148	0.6672	0.5178	0.3979	0.9142	0.1729
Kossougoudou	0.0000	0.9955	0.0002	0.8953	0.0006	0.8348
Piela	0.0000	0.9787	0.0018	0.7808	0.0273	0.6736
Sebba	0.0021	0.9265	0.4873	0.5970	0.7493	0.4030
Gotheye	0.0245	0.5925	0.1385	0.5676	0.1425	0.5650
Karma	0.0289	0.5977	0.4602	0.2951	0.0260	0.4257
Namaro	0.0000	0.9457	0.0148	0.7508	0.8928	0.3181

Table 20: P-Values for the Exponential, Gamma and Logistic distributions.

Distributions	**Exponential**		**Gamma**		**Logistic**	
Stations & p-values	**A-D**	**K-S**	**A-D**	**K-S**	**A-D**	**K-S**
Bogande	0.0353	0.4573	0.0453	0.7377	0.0014	0.9137
Manni	0.0009	0.6270	0.0200	0.8578	0.0002	0.9662
Aribinda	0.0000	0.6389	0.7189	0.2111	0.0222	0.6421
Bani	0.0000	0.4443	0.4054	0.3504	0.0103	0.7512
Barsalogho	0.0196	0.4951	0.0168	0.6705	0.0001	0.9177
Boulsa	0.0000	0.6892	0.0056	0.7004	0.0001	0.9136
Bouroum	0.0019	0.5282	0.0271	0.5149	0.0032	0.7903
Dakiri	0.0000	0.9552	0.2454	0.5301	0.0151	0.7542
Gayeri	0.0001	0.4875	0.2057	0.4614	0.0307	0.7353
Kossougoudou	0.0000	0.8991	0.0000	0.9724	0.0000	0.9950
Piela	0.0003	0.8028	0.0000	0.9184	0.0000	0.9638
Sebba	0.0197	0.4951	0.1774	0.7145	0.0011	0.9500
Gotheye	0.0006	0.8405	0.1154	0.5223	0.0474	0.6853
Karma	0.0182	0.6077	0.2694	0.3675	0.0563	0.5920
Namaro	0.0374	0.6678	0.0028	0.8318	0.0001	0.9513

Table 21: P-Values for the Pearson Type III and Log-Pearson Type III distributions.

Distributions	**Pearson Type III**		**Log-Pearson Type III**	
Stations & p-values	**A-D**	**K-S**	**A-D**	**K-S**
Bogande	0.3758	0.3867	0.8675	0.3221
Manni	0.0220	0.6099	0.5542	0.4505
Aribinda	0.9944	0.1799	0.9992	0.2012
Bani	0.9789	0.1505	0.9523	0.1909
Barsalogho	0.9878	0.3036	0.9770	0.1849
Boulsa	0.0002	0.6406	0.6544	0.2909
Bouroum	0.1619	0.4360	0.5526	0.2634
Dakiri	0.2861	0.4645	0.2950	0.4779
Gayeri	0.4795	0.2366	0.9369	0.1644
Kossougoudou	0.0000	0.9527	0.0059	0.7749
Piela	0.0000	0.8485	0.0167	0.6043
Sebba	0.7356	0.3684	0.7630	0.3830
Gotheye	0.1596	0.5653	0.1356	0.5459
Karma	0.0312	0.5153	0.4199	0.3364
Namaro	0.0000	0.4903	0.9229	0.2645

Table 22: P-Values for the Gumbel and GEV distributions.

Distributions	**Gumbel, EV1**		**Generalized Extreme Values**	
Stations & p-values	**A-D**	**K-S**	**A-D**	**K-S**
Bogande	0.3613	0.5814	0.8775	0.3523
Manni	0.1571	0.7421	0.7528	0.3881
Aribinda	0.9968	0.1483	0.9990	0.2355
Bani	0.8674	0.1836	0.9277	0.1934
Barsalogho	0.3063	0.4251	0.9291	0.2163
Boulsa	0.0780	0.4941	0.5211	0.3498
Bouroum	0.2635	0.3639	0.5794	0.2567
Dakiri	0.2698	0.5856	0.3130	0.4493
Gayeri	0.9176	0.2038	0.9661	0.1442
Kossougoudou	0.0001	0.9247	0.0017	0.7394
Piela	0.0016	0.8273	0.1843	0.5479
Sebba	0.4444	0.5860	0.6597	0.4220
Gotheye	0.1406	0.5301	0.1302	0.5668
Karma	0.4455	0.3276	0.3808	0.3743
Namaro	0.0051	0.8166	0.8127	0.2459

Table 23: P-Values for the GLO and GPA distributions.

Distributions	**Generalized Logistic**		**Generalized Pareto**	
Stations & p-values	**A-D**	**K-S**	**A-D**	**K-S**
Bogande	0.7335	0.4532	0.0350	0.5118
Manni	0.8132	0.4141	0.0109	0.6732
Aribinda	0.9884	0.3047	0.0100	0.3200
Bani	0.6544	0.2739	0.4968	0.3845
Barsalogho	0.6501	0.2870	0.0744	0.4766
Boulsa	0.8552	0.2306	0.0001	0.6714
Bouroum	0.6462	0.2711	0.0123	0.5677
Dakiri	0.2291	0.3659	0.0124	0.6710
Gayeri	0.9818	0.2035	0.0211	0.4436
Kossougoudou	0.0384	0.6763	0.0000	0.9082
Piela	0.3891	0.4340	0.0000	0.8175
Sebba	0.2997	0.5156	0.0076	0.3097
Gotheye	0.0991	0.6182	0.0327	0.5784
Karma	0.6474	0.2969	0.0167	0.5712
Namaro	0.7666	0.2256	0.8940	0.2904

Table 24: P-Values for the Burr Type XII distribution.

Distribution	**Burr Type XII**	
Stations & p-values	**A-D**	**K-S**
Bogande	0.7397	0.4457
Manni	0.9005	0.3541
Aribinda	0.9860	0.3080
Bani	0.5260	0.3043
Barsalogho	0.4868	0.3534
Boulsa	0.7855	0.2414
Bouroum	0.5450	0.3361
Dakiri	0.1628	0.4134
Gayeri	0.9725	0.2223
Kossougoudou	0.0148	0.8209
Piela	0.3864	0.5505
Sebba	0.2319	0.5442
Gotheye	0.0831	0.6504
Karma	0.8690	0.2358
Namaro	0.7843	0.2885

4.3.3: Information Criteria

Table 25: Log-Likelihood and Information Criteria for the Normal, Log-Normal 2 Par. and Log-Normal 3 Par. distributions.

Distributions	**Normal**				**Log-Normal 2 Par.**				**Log-Normal 3 Par.**			
Stations & ICs	**log-L**	**AIC**	**AICc**	**BIC**	**log-L**	**AIC**	**AICc**	**BIC**	**log-L**	**AIC**	**AICc**	**BIC**
Bogande	-176.20	356.39	356.71	359.82	-170.57	345.15	345.46	348.58	-168.70	343.40	344.05	348.54
Manni	-188.74	381.48	381.80	384.91	-182.09	368.18	368.50	371.61	-181.71	369.42	370.07	374.57
Aribinda	-267.95	539.89	540.10	544.08	-260.66	525.32	525.54	529.51	-260.39	526.79	527.22	533.07
Bani	-225.18	454.35	454.60	458.22	-219.70	443.41	443.66	447.27	-219.03	444.06	444.57	449.85
Barsalogho	-273.76	551.51	551.71	555.80	-265.11	534.23	534.43	538.51	-261.47	528.94	529.35	535.37
Boulsa	-267.66	539.31	539.54	543.33	-242.04	488.07	488.30	492.09	-238.74	483.48	483.95	489.50
Bouroum	-177.12	358.23	358.55	361.66	-169.34	342.68	343.00	346.11	-167.26	340.51	341.16	345.65
Dakiri	-249.40	502.81	503.03	506.93	-248.19	500.39	500.61	504.51	-247.83	501.66	502.10	507.84
Gayeri	-177.71	359.42	359.72	362.94	-174.01	352.01	352.31	355.53	-173.83	353.67	354.28	358.95
Kossougoudou	-249.56	503.13	503.37	507.07	-229.28	462.56	462.80	466.50	-219.73	445.46	445.95	451.37
Piela	-277.85	559.70	559.91	563.89	-260.84	525.69	525.90	529.87	-252.29	510.58	511.01	516.86
Sebba	-254.60	513.21	513.43	517.26	-247.55	499.10	499.33	503.15	-247.02	500.04	500.50	506.12
Gotheye	-127.80	259.60	260.06	262.33	-123.63	251.25	251.71	253.98	-123.24	252.47	253.43	256.58
Karma	-87.69	179.39	180.14	181.28	-85.02	174.04	174.79	175.93	-91.71	189.42	191.02	192.25
Namaro	-85.96	175.92	176.78	177.59	-78.56	161.12	161.98	162.79	-72.95	151.90	153.75	154.40

Table 26: Log-Likelihood and Information Criteria for the Exponential, Gamma and Logistic distributions.

Distributions	**Exponential**				**Gamma**				**Logistic**			
Stations & ICs	**log-L**	**AIC**	**AICc**	**BIC**	**log-L**	**AIC**	**AICc**	**BIC**	**log-L**	**AIC**	**AICc**	**BIC**
Bogande	-∞	+∞	+∞	+∞	-172.02	348.04	348.36	351.47	-174.75	353.51	353.82	356.93
Manni	-∞	+∞	+∞	+∞	-183.72	371.45	371.76	374.87	-187.19	378.38	378.70	381.81
Aribinda	-∞	+∞	+∞	+∞	-261.99	527.97	528.18	532.16	-265.94	535.87	536.08	540.06
Bani	-∞	+∞	+∞	+∞	-220.92	445.84	446.09	449.71	-224.89	453.78	454.03	457.64
Barsalogho	-∞	+∞	+∞	+∞	-267.36	538.72	538.92	543.00	-272.02	548.04	548.24	552.33
Boulsa	-∞	+∞	+∞	+∞	-246.91	497.82	498.05	501.83	-248.89	501.77	502.00	505.79
Bouroum	-∞	+∞	+∞	+∞	-171.21	346.41	346.73	349.84	-172.95	349.90	350.22	353.33
Dakiri	-∞	+∞	+∞	+∞	-247.75	499.50	499.72	503.63	-250.12	504.25	504.47	508.37
Gayeri	-∞	+∞	+∞	+∞	-174.91	353.82	354.12	357.34	-176.60	357.20	357.50	360.72
Kossougoudou	-∞	+∞	+∞	+∞	-233.36	470.71	470.95	474.65	-236.21	476.43	476.67	480.37
Piela	-∞	+∞	+∞	+∞	-264.74	533.47	533.69	537.66	-269.40	542.80	543.01	546.99
Sebba	-∞	+∞	+∞	+∞	-248.97	501.94	502.17	505.99	-253.79	511.57	511.80	515.62
Gotheye	-∞	+∞	+∞	+∞	-124.52	253.04	253.50	255.77	-126.02	256.05	256.51	258.78
Karma	-∞	+∞	+∞	+∞	-85.35	174.71	175.46	176.59	-86.54	177.09	177.84	178.98
Namaro	-77.28	158.56	159.42	160.23	-80.51	165.02	165.88	166.69	-83.36	170.71	171.57	172.38

Table 27: Log-Likelihood and Information Criteria for the Pearson Type III and Log-Pearson Type III distributions.

Distributions	**Pearson Type III**				**Log-Pearson Type III**			
Stations & ICs	**log-L**	**AIC**	**AICc**	**BIC**	**log-L**	**AIC**	**AICc**	**BIC**
Bogande	-∞	+∞	+∞	+∞	-168.72	343.43	344.08	348.57
Manni	-∞	+∞	+∞	+∞	-181.18	368.37	369.02	373.51
Aribinda	-260.36	526.72	527.15	533.00	-260.44	526.87	527.30	533.16
Bani	-218.29	442.57	443.08	448.37	-219.04	444.09	444.60	449.88
Barsalogho	-259.88	525.77	526.18	532.20	-261.50	528.99	529.40	535.42
Boulsa	-∞	+∞	+∞	+∞	-238.00	482.00	482.47	488.02
Bouroum	-∞	+∞	+∞	+∞	-167.18	340.36	341.00	345.50
Dakiri	-247.96	501.93	502.37	508.11	-247.84	501.67	502.11	507.85
Gayeri	-∞	+∞	+∞	+∞	-173.57	353.14	353.75	358.42
Kossougoudou	-∞	+∞	+∞	+∞	+∞	+∞	+∞	+∞
Piela	-∞	+∞	+∞	+∞	-266.60	539.19	539.62	545.47
Sebba	-246.91	499.81	500.28	505.89	-247.08	500.16	500.62	506.23
Gotheye	-122.98	251.97	252.93	256.07	-123.29	252.57	253.53	256.68
Karma	-∞	+∞	+∞	+∞	-85.27	176.53	178.13	179.37
Namaro	-∞	+∞	+∞	+∞	-72.55	151.11	152.95	153.61

Table 28: Log-Likelihood and Information Criteria for the Gumbel and GEV distributions.

Distributions	**Gumbel, EV1**				**Generalized Extreme Values**			
Stations & ICs	**log-L**	**AIC**	**AIC_C**	**BIC**	**log-L**	**AIC**	**AIC_C**	**BIC**
Bogande	-169.85	343.70	344.01	347.12	-168.77	343.53	344.18	348.67
Manni	-181.92	367.83	368.15	371.26	-180.79	367.58	368.23	372.72
Aribinda	-260.80	525.60	525.81	529.79	-260.53	527.07	527.50	533.35
Bani	-219.80	443.60	443.85	447.46	-219.37	444.75	445.26	450.54
Barsalogho	-264.21	532.42	532.62	536.70	-262.18	530.35	530.76	536.78
Boulsa	-241.74	487.48	487.71	491.49	-237.89	481.77	482.24	487.79
Bouroum	-168.37	340.74	341.06	344.17	-167.16	340.31	340.96	345.45
Dakiri	-248.51	501.01	501.23	505.13	-247.72	501.43	501.88	507.61
Gayeri	-173.31	350.63	350.93	354.15	-173.38	352.75	353.37	358.04
Kossougoudou	-227.69	459.37	459.61	463.31	-243.47	492.94	493.43	498.85
Piela	-260.10	524.21	524.42	528.39	-256.51	519.01	519.44	525.30
Sebba	-247.92	499.84	500.06	503.89	-247.30	500.60	501.07	506.68
Gotheye	-123.57	251.13	251.60	253.87	-123.38	252.75	253.71	256.85
Karma	-84.94	173.88	174.63	175.77	-85.78	177.56	179.16	180.39
Namaro	-79.86	163.72	164.58	165.39	-74.03	154.07	155.91	156.57

Table 29: Log-Likelihood and Information Criteria for the GLO and GPA distributions.

Distributions	**Generalized Logistic**				**Generalized Pareto**			
Stations & ICs	**log-L**	**AIC**	**AIC_C**	**BIC**	**log-L**	**AIC**	**AIC_C**	**BIC**
Bogande	-169.23	344.46	345.11	349.60	-∞	+∞	+∞	+∞
Manni	-180.71	367.41	368.06	372.55	-∞	+∞	+∞	+∞
Aribinda	-261.27	528.54	528.96	534.82	-∞	+∞	+∞	+∞
Bani	-220.61	447.21	447.72	453.01	-∞	+∞	+∞	+∞
Barsalogho	-263.47	532.94	533.35	539.37	-∞	+∞	+∞	+∞
Boulsa	-237.40	480.80	481.27	486.82	-∞	+∞	+∞	+∞
Bouroum	-167.23	340.47	341.12	345.61	-∞	+∞	+∞	+∞
Dakiri	-248.45	502.91	503.35	509.09	-∞	+∞	+∞	+∞
Gayeri	-173.30	352.60	353.22	357.88	-∞	+∞	+∞	+∞
Kossougoudou	-213.78	433.57	434.06	439.48	-∞	+∞	+∞	+∞
Piela	-255.41	516.81	517.24	523.09	-∞	+∞	+∞	+∞
Sebba	-248.46	502.93	503.39	509.00	-∞	+∞	+∞	+∞
Gotheye	-123.87	253.74	254.70	257.84	-∞	+∞	+∞	+∞
Karma	-84.80	175.59	177.19	178.43	-∞	+∞	+∞	+∞
Namaro	-74.18	154.36	156.21	156.86	-72.73	151.47	153.32	153.97

Table 30: Log-Likelihood and Information Criteria for the Burr Type XII distribution.

Distribution	**Burr Type XII**			
Stations & ICs	**log-L**	**AIC**	**AIC_C**	**BIC**
Bogande	-169.35	344.69	345.34	349.83
Manni	-180.64	367.28	367.93	372.42
Aribinda	-261.29	528.57	529.00	534.85
Bani	-220.77	447.54	448.05	453.33
Barsalogho	-263.98	533.97	534.37	540.40
Boulsa	-238.14	482.29	482.76	488.31
Bouroum	-167.59	341.18	341.83	346.32
Dakiri	-248.36	502.72	503.16	508.90
Gayeri	-173.33	352.66	353.28	357.95
Kossougoudou	-223.12	452.24	452.72	458.15
Piela	-255.44	516.88	517.31	523.17
Sebba	-248.60	503.20	503.66	509.27
Gotheye	-123.95	253.91	254.87	258.01
Karma	-84.43	174.86	176.46	177.69
Namaro	-73.90	153.79	155.64	156.29

4.3.4: Distribution plots

On the sole basis of the Error Indicators, most distributions for most stations represent good candidates for fitting, minus some isolated cases for the Normal, Log-Normal 2 Par., Gamma, and Logistic distributions, with R^2 values lower than 0.80. Obviously, this alone is not sufficient.

Some distributions, like the Exponential, Pearson Type III, and Generalized Pareto, in most cases, show clear signs of an out-of-bounds condition because the values of the density function are very small, thus making log-L = -∞. The Log-Pearson Type III shows an out-of-bounds condition for the Kossougoudou station. Excluding the mentioned cases, the ICs show comparable values across the remaining distributions, for most stations.

The Anderson-Darling test is particularly stringent compared to the Kolmogorov-Smirnov test, and unfortunately, there are no cases where a single distribution passes both tests for all stations.

Considering all that was indicated in the previous tables of the Error Indicators, P-values, and ICs, the Burr Type XII distribution is a valid candidate to represent the data for all stations. It shows low errors (see Table 18), reasonable values for the Log-L and ICs, passing both statistical tests for all stations except one: Kossougoudou, which fails the Anderson-Darling test. Kossogoudou station shows low values for this statistical tests, for most distributions, due to the particular shape of its empirical series: very steep around the flex point and almost flat at the tails.

Given the high number of plots, only those for some stations will be shown. Two alternatives are displayed: some of the best fittings and the worst one. Aribinda shows good fitting, at least from a visual inspection, for most distributions. Bogande, Manni and Sebba show good fitting for some distributions, while Kossougoudou has acceptable visual fitting for only a small number of distributions. As said before, Bogande, Manni and Sebba must be included since they are considered key localities for the SLAPIS project.

Aribinda

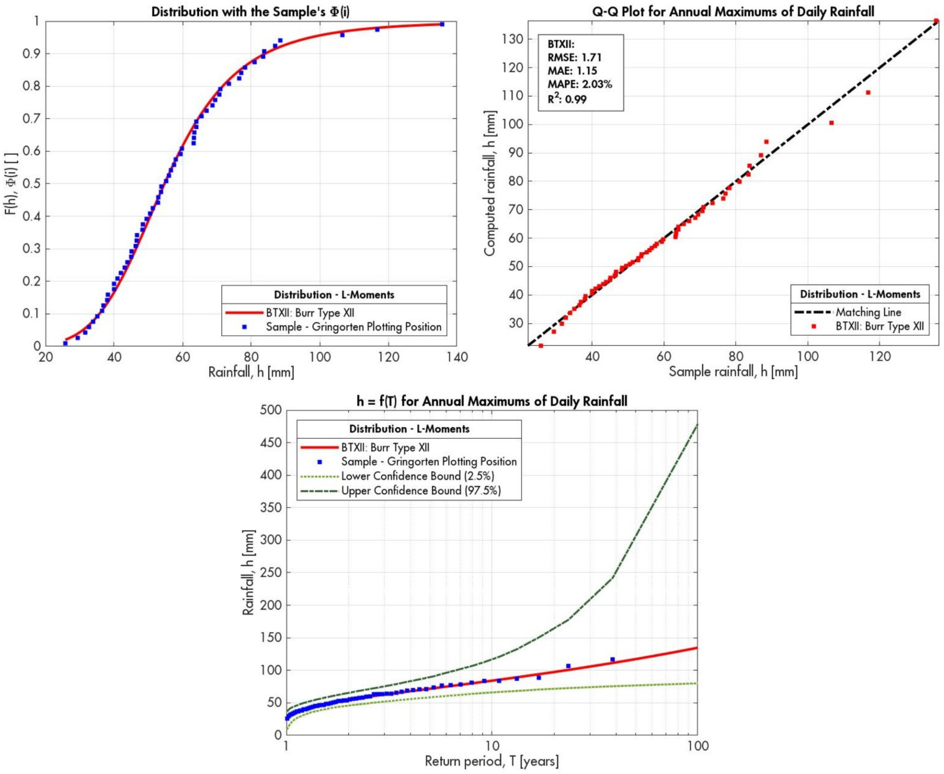


Figure 17: Distribution fitting between the Burr Type XII and Aribinda's rainfall sample.

Bogande

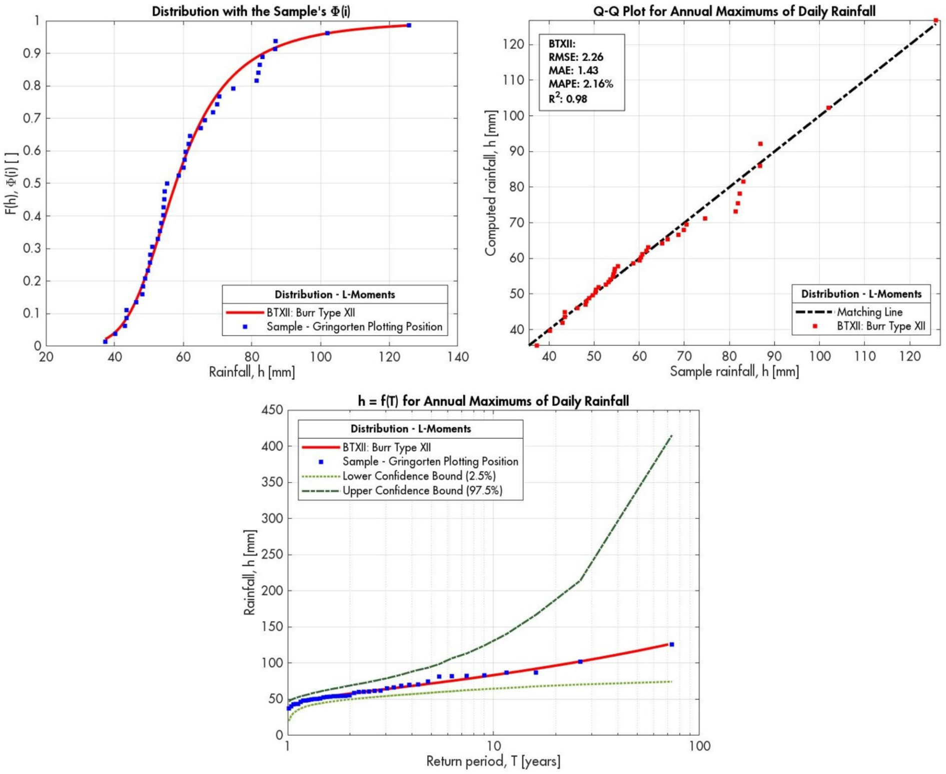


Figure 18: Distribution fitting between the Burr Type XII and Bogande's rainfall sample.

Manni

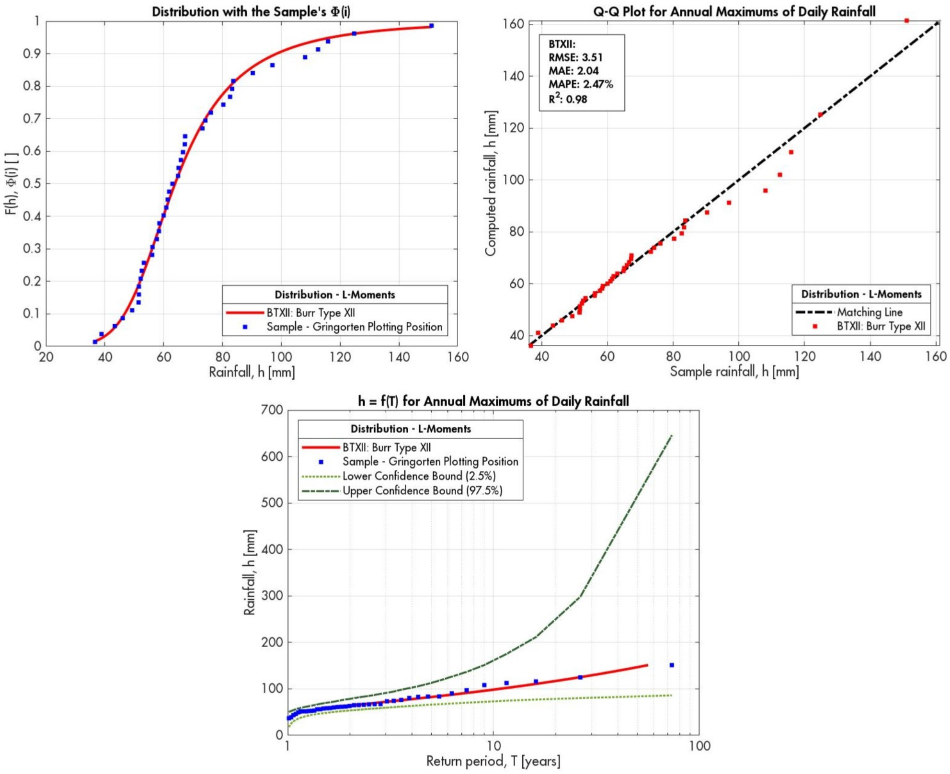


Figure 19: Distribution fitting between the Burr Type XII and Manni's rainfall sample.

Sebba

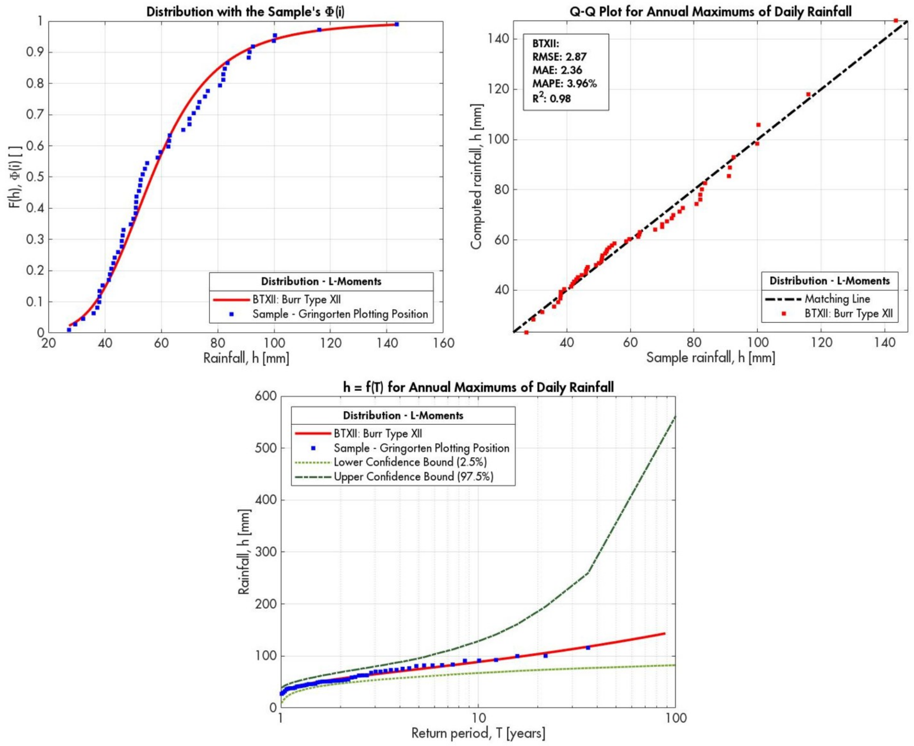


Figure 20: Distribution fitting between the Burr Type XII and Sebba's rainfall sample.

Kossogoudou

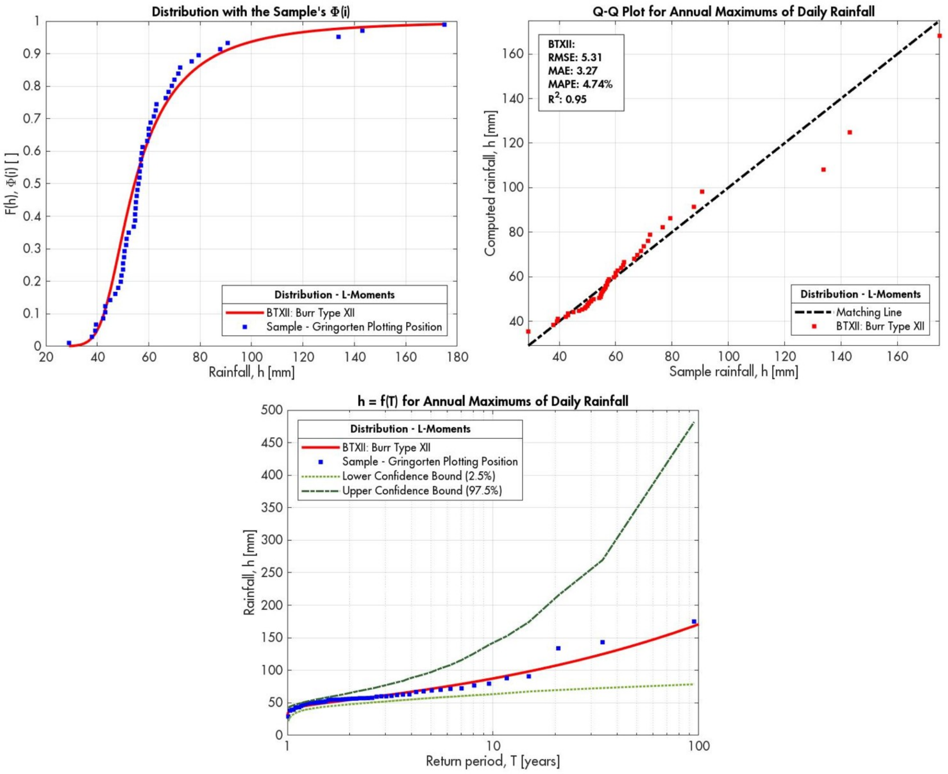


Figure 21: Distribution fitting between the Burr Type XII and Kossogoudou's rainfall sample.

4.3.5: Quantiles and residual risk

For completeness, the tables of the computed rainfalls (quantiles) and residual risk for different return periods T are also included.

Table 31: Computed rainfalls for the Burr Type XII distribution, for different Return periods.

Rainfalls heights, h [mm], for different Return periods T					
Distribution	**Burr Type XII**				
Stations & T [years]	**5**	**10**	**20**	**50**	**100**
Bogande	71.85	83.37	96.51	117.02	135.36
Manni	82.55	98.37	116.91	146.75	174.26
Aribinda	71.68	84.02	97.30	117.23	134.64
Bani	71.96	84.00	97.10	116.93	134.34
Barsalogho	72.78	84.72	98.32	119.57	138.63
Boulsa	77.90	93.66	112.40	142.96	171.48
Bouroum	68.88	79.90	92.45	112.03	129.54
Dakiri	72.52	81.88	91.10	103.78	114.04
Gayeri	74.31	83.46	93.43	108.31	121.09
Kossougoudou	71.43	87.40	106.93	139.59	170.77
Piela	73.73	89.63	108.88	140.80	171.03
Sebba	74.89	88.61	103.68	126.77	147.29
Gotheye	71.32	82.55	94.63	112.67	128.33
Karma	72.21	87.58	105.53	134.59	161.68
Namaro	76.59	101.90	135.58	197.74	263.09

Since the fixed cumulative probabilities F for different return periods can be computed, it will obviously have the same value for all stations, as did for the DDF and IDF curves for different values of T (see the Empirical series and Growth factors sections). This fixed F must not be mistaken for the different CDF values that each station has for different rainfall values. Each station has different values of cumulative probability because the parameters are computed based on each station's rainfall series. Fixing the same return periods for all stations allows the calculation of different rainfall heights, ensuring that all stations are assessed under the same theoretical conditions. This approach is standard practice in hydrology and risk assessment.

As mentioned previously, the return period alone is not sufficient to quantify rainfall occurrences. For simplicity, the residual risk can be computed using the fixed probability F. However, it is important to note that this method assumes the parameters of the probability distribution for each station are valid and well-calibrated. Additionally, while residual risk can be simplified using the fixed F, this approach may not fully account for uncertainties in the statistical model or varying environmental factors.

Table 32: Values of residual risk for different return periods and reference periods.

Residual Risk for different Return periods T and Ref. periods L					
L & T [years]	**5**	**10**	**20**	**50**	**100**
5	67.23%	40.95%	22.62%	9.61%	4.90%
10	89.26%	65.13%	40.13%	18.29%	9.56%
20	98.85%	87.84%	64.15%	33.24%	18.21%
50	100.00%	99.48%	92.31%	63.58%	39.50%
100	100.00%	100.00%	99.41%	86.74%	63.40%

From the precedent table, the following can be deduced: for example, over a time span of 100 years, a rainfall event with T = 5 years has a 100% probability of being exceeded. Similarly, an event with T = 100 years has a probability of almost 5% of being surpassed within a 5-year time frame. This aligns with the definition of cumulative probability, which inherently accounts for the probability of non-exceedance.

4.4: DDF and IDF curves

4.4.1: Average curves

Table 33: "a" and "n" parameters used for all the DDF and IDF curves.

Stations	a [mm/day]	n [ ]
Bogande	55.82	0.34
Manni	64.94	0.31
Aribinda	54.25	0.32
Bani	54.22	0.33
Barsalogho	57.28	0.34
Boulsa	61.98	0.32
Bouroum	54.03	0.38
Dakiri	55.67	0.32
Gayeri	59.71	0.34
Kossougoudou	57.49	0.33
Piela	58.93	0.34
Sebba	56.33	0.33
Gotheye	54.09	0.28
Karma	54.66	0.35
Namaro	62.41	0.26

Again, to assess the goodness of fit, the previously mentioned error indicators will be used. For all stations, the log-linear fitting, although simple, has been shown to be adequate.

Here are the various plots, in the order shown in the previous table, with the necessary box-plots for each time window:

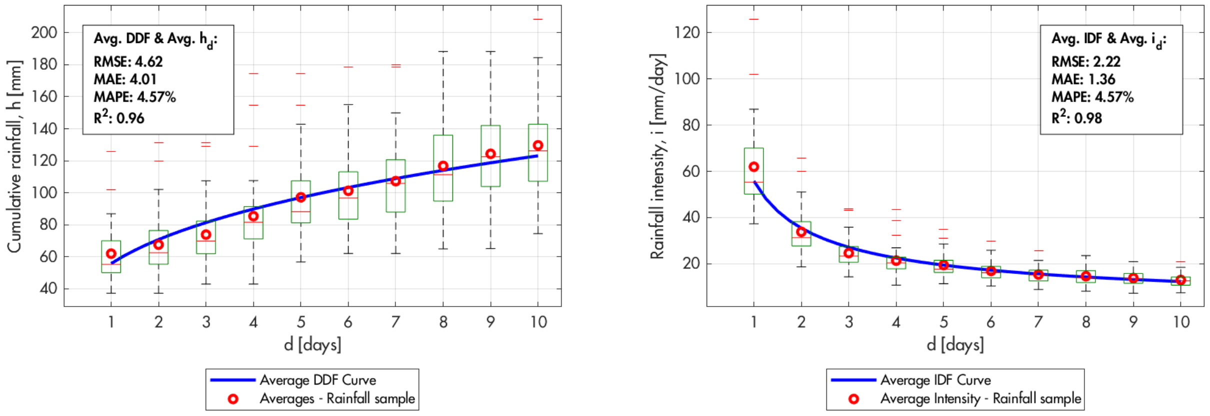


Figure 22: Average DDF and IDF curves: Bogande.

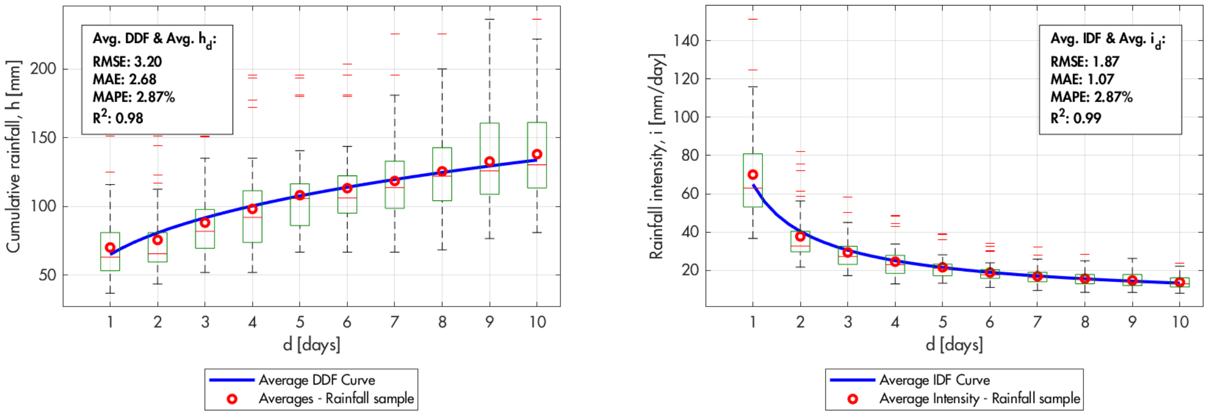


Figure 23: Average DDF and IDF curves: Manni.

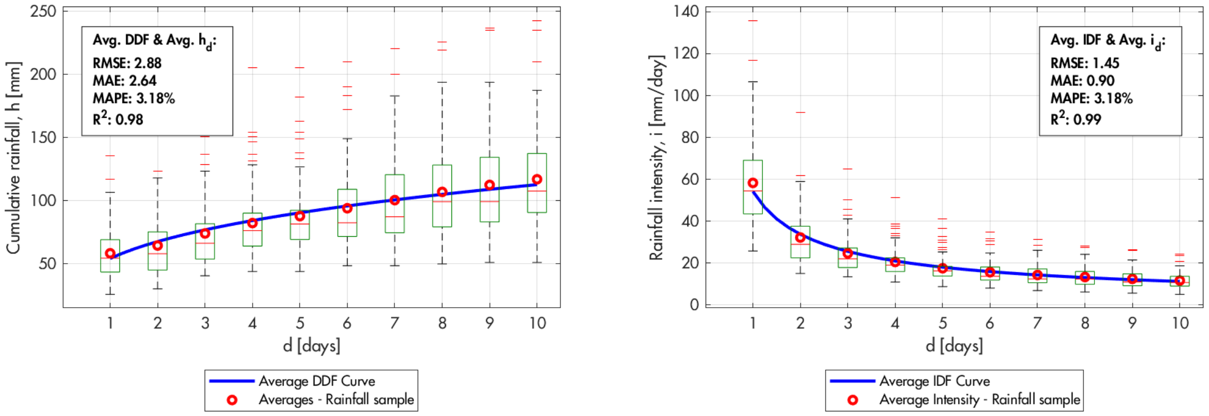


Figure 24: Average DDF and IDF curves: Aribinda.

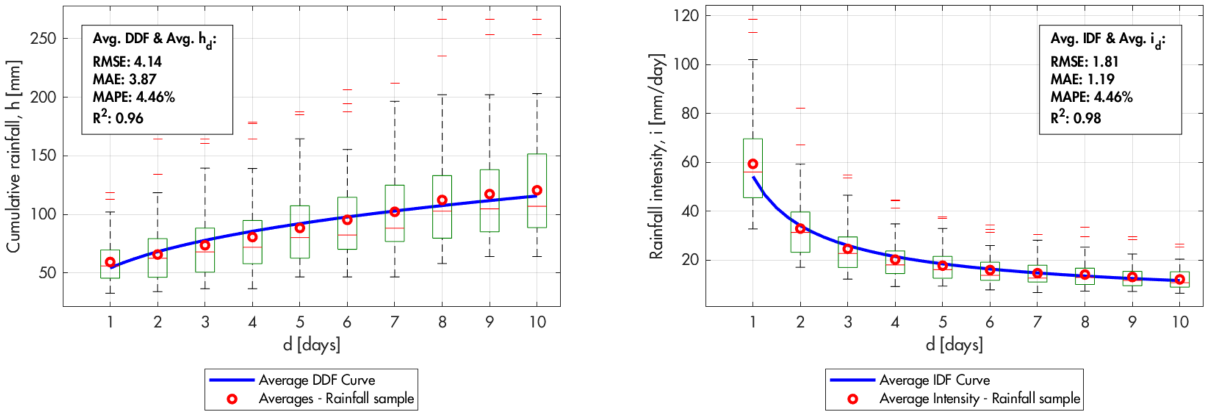


Figure 25: Average DDF and IDF curves: Bani.

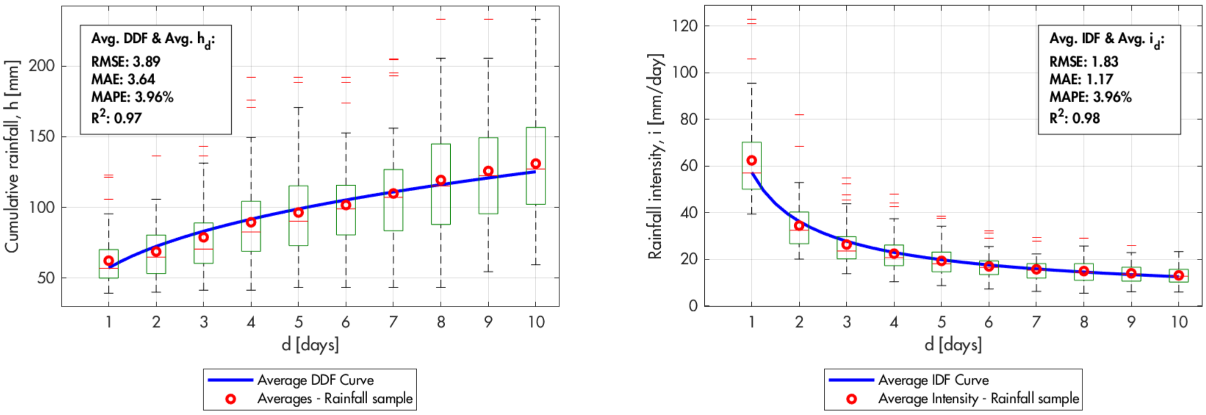


Figure 26: Average DDF and IDF curves: Barsalogho.

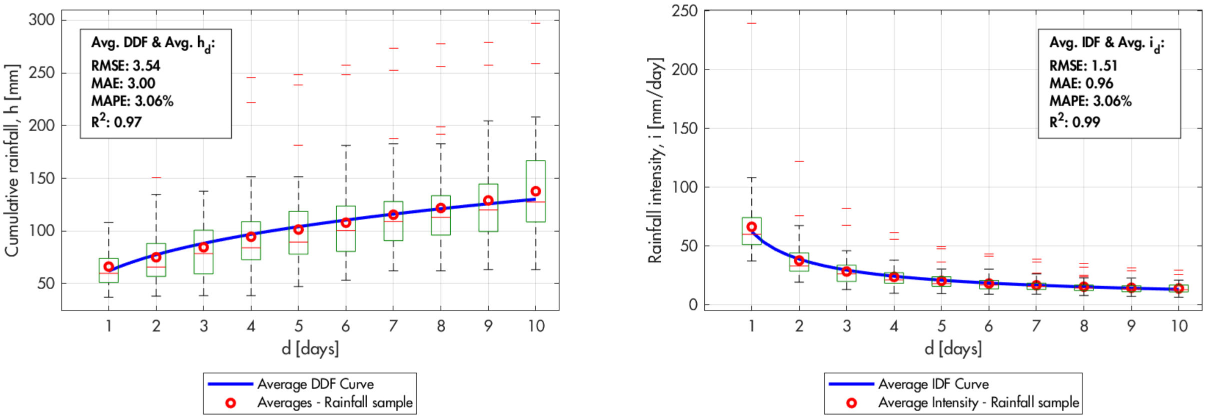


Figure 27: Average DDF and IDF curves: Boulsa.

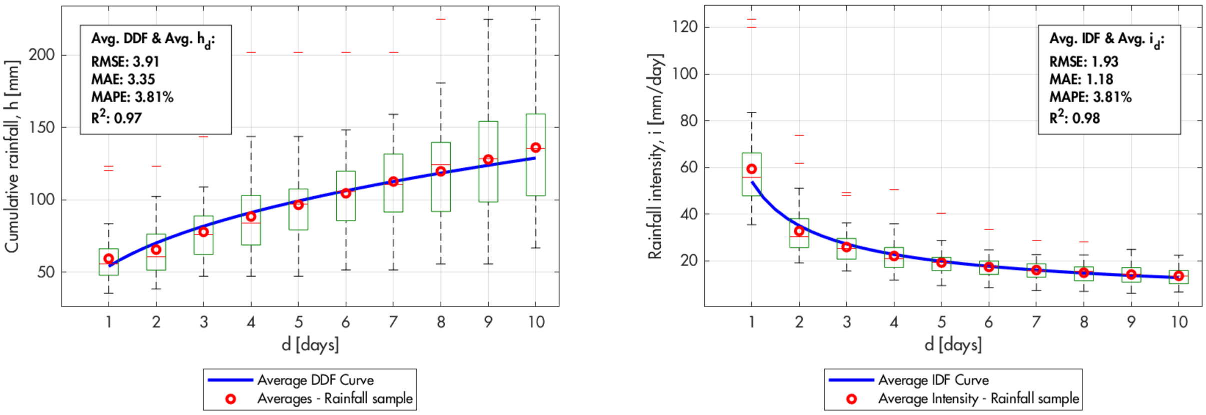


Figure 28: Average DDF and IDF curves: Bouroum.

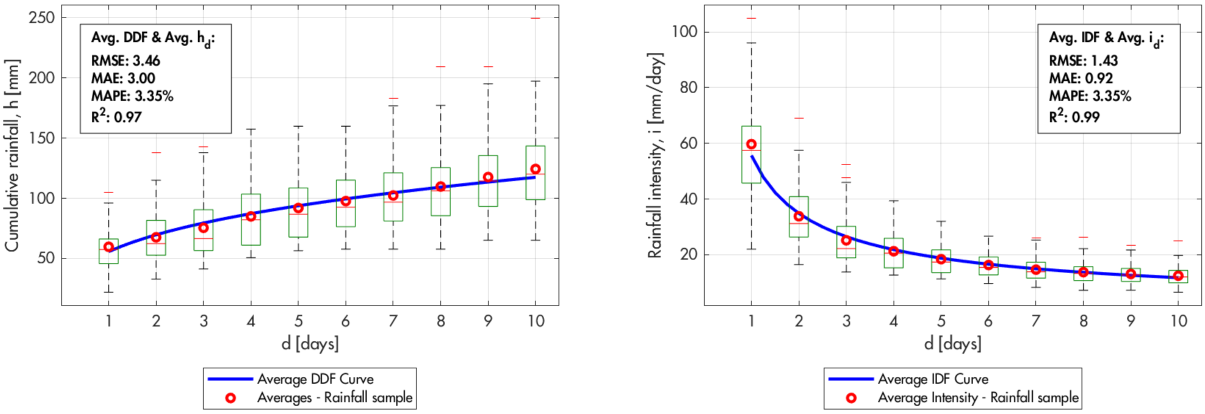


Figure 29: Average DDF and IDF curves: Dakiri.

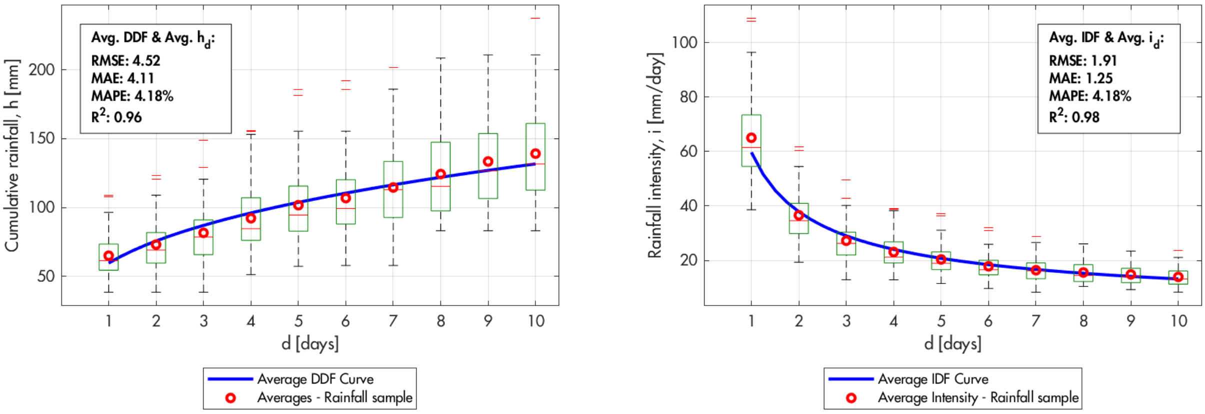


Figure 30: Average DDF and IDF curves: Gayeri.

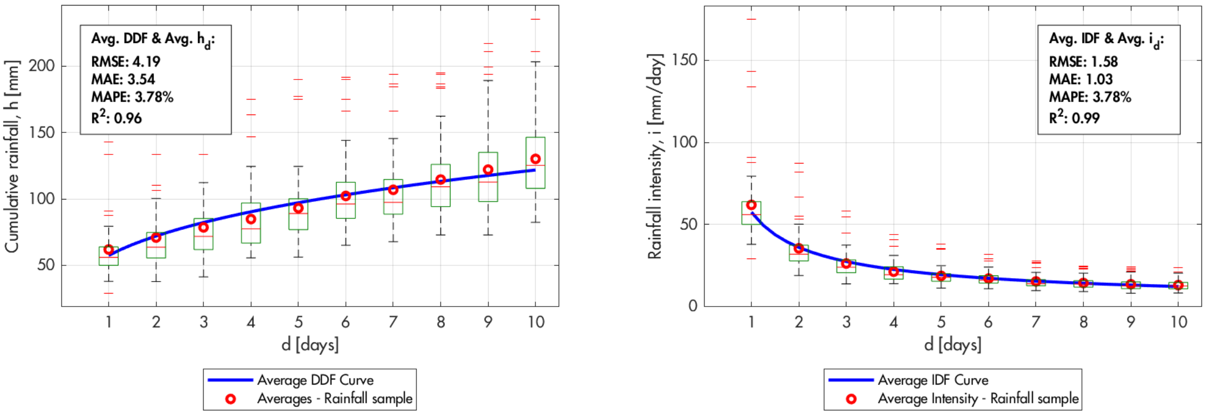


Figure 31: Average DDF and IDF curves: Kossougoudou.

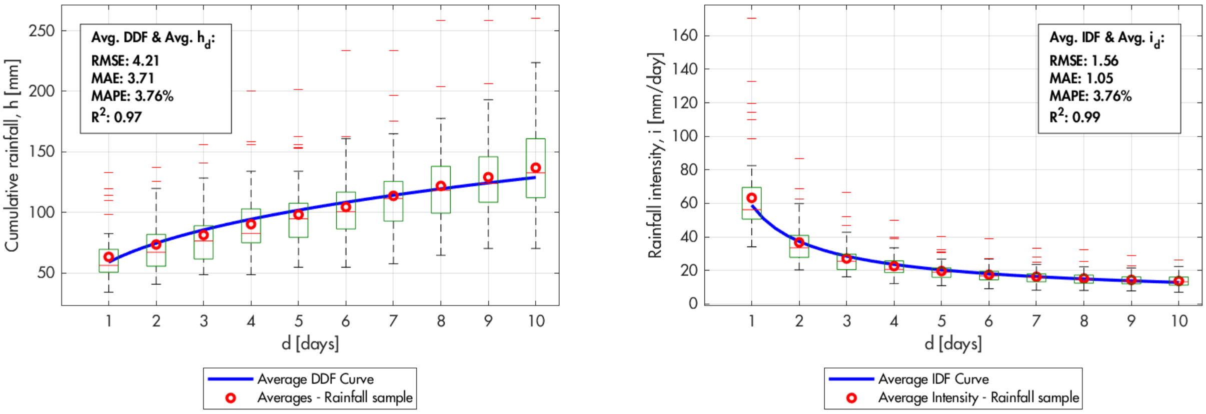


Figure 32: Average DDF and IDF curves: Piela.

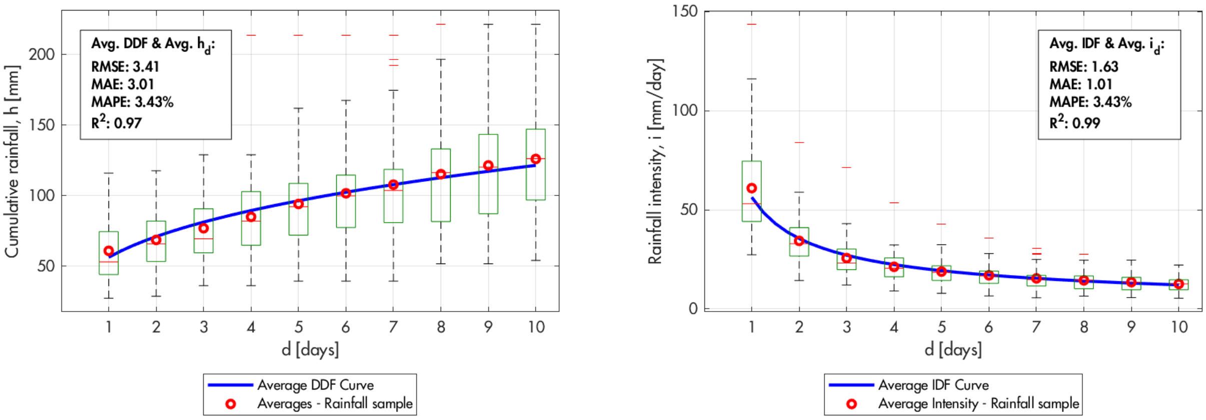


Figure 33: Average DDF and IDF curves: Sebba.

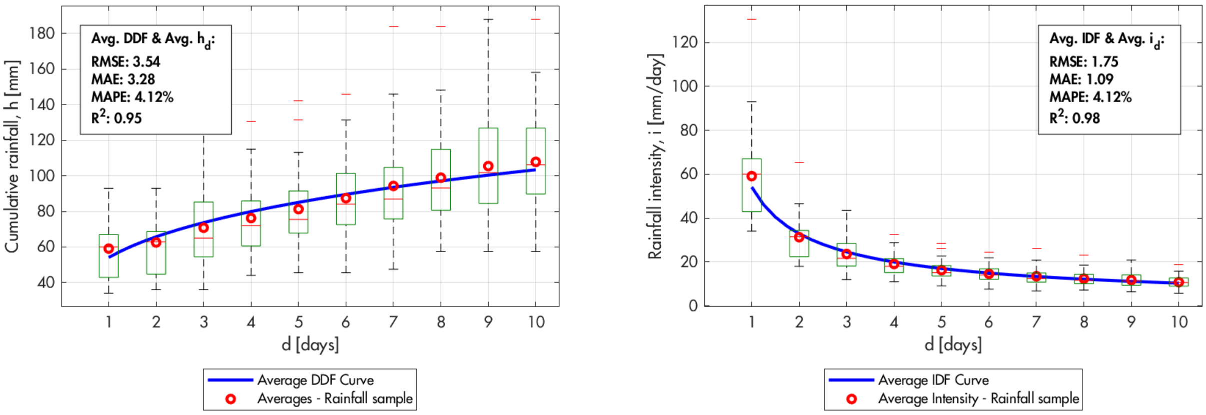


Figure 34: Average DDF and IDF curves: Gotheye.

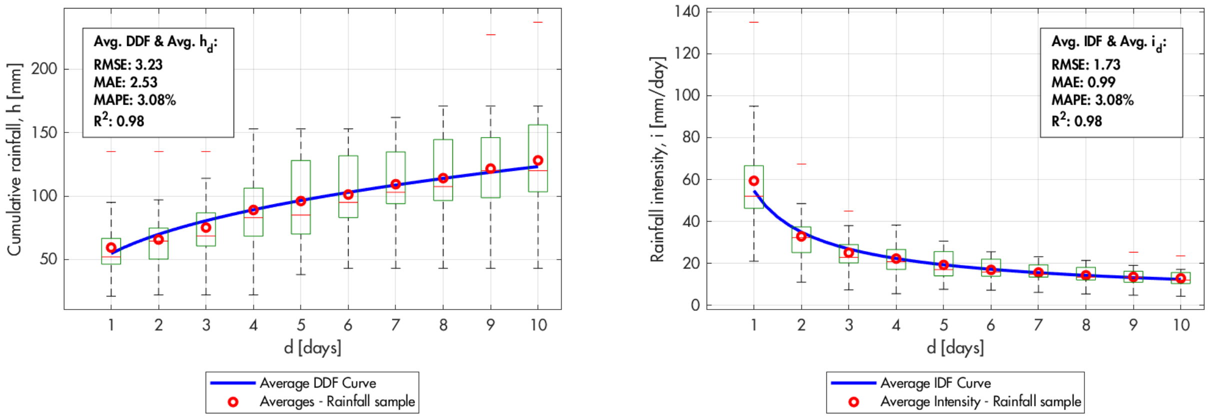


Figure 35: Average DDF and IDF curves: Karma.

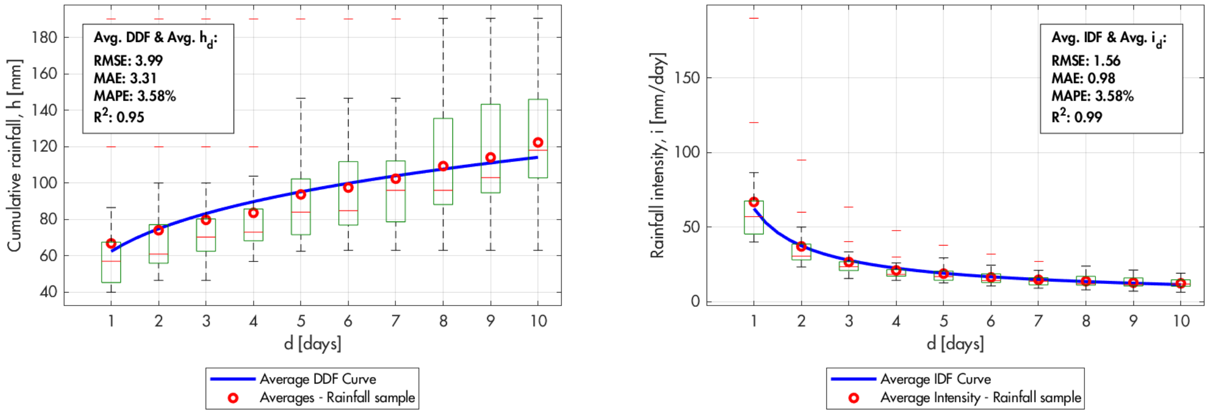


Figure 36: Average DDF and IDF curves: Namaro.

4.4.2: Growth factors

For each distribution these plots can be made:

- Dimensionless parameters for each duration (1-10 days);
- DDF curves, standard and/or in natural logarithmic scale;
- IDF curves, standard and/or in natural logarithmic scale.

Given the excessive number of possible plots, only the tables of all growth factors, relative to each station, distribution and considered period of return T, will be shown in this paper. If the reader will want to graph the growth factor DDF and IDF curves, can easily use the K(T) values form the following tables, the "a" and "n" parameters from Table 33, following the indications from the Methods' section Growth factors.

The average L-Moments across all durations (from the normalized series) were used to calculate the different K(T)s, but the L-Moments across all duration from the initial rainfall series are shown with their trends in the Annex's section L-Moments trends By rainfall durations.

Table 34: Growth factors K(T) for the Normal, Log-Normal 2 Par. and Log-Normal 3 Par. Distributions.

Distributions	**Normal**					**Log-Normal 2 Par.**					**Log-Normal 3 Par.**				
Stations & T [years]	**5**	**10**	**20**	**50**	**100**	**5**	**10**	**20**	**50**	**100**	**5**	**10**	**20**	**50**	**100**
Bogande	1.21	1.32	1.42	1.52	1.59	1.20	1.33	1.46	1.62	1.73	1.19	1.34	1.49	1.68	1.83
Manni	1.24	1.37	1.47	1.59	1.67	1.21	1.37	1.52	1.70	1.84	1.21	1.40	1.59	1.85	2.06
Aribinda	1.29	1.44	1.57	1.71	1.80	1.25	1.44	1.63	1.87	2.04	1.24	1.47	1.72	2.06	2.33
Bani	1.31	1.47	1.60	1.75	1.85	1.27	1.48	1.69	1.95	2.15	1.25	1.51	1.76	2.12	2.41
Barsalogho	1.26	1.39	1.50	1.63	1.71	1.23	1.41	1.58	1.79	1.94	1.23	1.42	1.60	1.84	2.02
Boulsa	1.27	1.42	1.53	1.67	1.76	1.23	1.41	1.58	1.79	1.94	1.22	1.45	1.69	2.03	2.30
Bouroum	1.24	1.36	1.46	1.58	1.65	1.22	1.38	1.53	1.71	1.85	1.22	1.38	1.53	1.71	1.85
Dakiri	1.25	1.38	1.49	1.61	1.69	1.23	1.40	1.56	1.76	1.91	1.23	1.41	1.58	1.80	1.97
Gayeri	1.22	1.33	1.43	1.53	1.61	1.20	1.35	1.48	1.64	1.76	1.20	1.36	1.51	1.71	1.86
Kossougoudou	1.23	1.35	1.45	1.57	1.64	1.20	1.34	1.48	1.64	1.76	1.18	1.38	1.59	1.90	2.15
Piela	1.24	1.36	1.47	1.58	1.66	1.21	1.37	1.51	1.69	1.82	1.21	1.39	1.57	1.80	1.99
Sebba	1.28	1.43	1.55	1.68	1.77	1.25	1.45	1.64	1.88	2.06	1.25	1.45	1.64	1.88	2.06
Gotheye	1.24	1.37	1.47	1.59	1.67	1.23	1.39	1.55	1.74	1.88	1.22	1.39	1.55	1.76	1.91
Karma	1.28	1.42	1.54	1.67	1.76	1.26	1.47	1.67	1.92	2.11	1.26	1.45	1.62	1.84	2.01
Namaro	1.28	1.42	1.54	1.68	1.77	1.23	1.40	1.56	1.77	1.93	1.19	1.45	1.73	2.17	2.54

Table 35: Growth factors K(T) for the Exponential, Logistic and Gamma distributions.

Distributions	**Exponential**					**Logistic**					**Gamma**				
Stations & T [years]	**5**	**10**	**20**	**50**	**100**	**5**	**10**	**20**	**50**	**100**	**5**	**10**	**20**	**50**	**100**
Bogande	1.17	1.37	1.57	1.83	2.03	1.20	1.31	1.42	1.55	1.65	1.21	1.34	1.45	1.59	1.68
Manni	1.20	1.42	1.65	1.95	2.17	1.23	1.36	1.48	1.63	1.75	1.23	1.39	1.52	1.69	1.80
Aribinda	1.24	1.51	1.77	2.13	2.40	1.27	1.43	1.57	1.75	1.89	1.28	1.47	1.63	1.84	1.98
Bani	1.25	1.54	1.83	2.21	2.49	1.29	1.46	1.61	1.81	1.95	1.29	1.50	1.68	1.91	2.07
Barsalogho	1.21	1.45	1.69	2.01	2.25	1.24	1.38	1.51	1.67	1.79	1.25	1.41	1.56	1.73	1.86
Boulsa	1.22	1.48	1.73	2.07	2.32	1.25	1.40	1.54	1.71	1.84	1.26	1.44	1.60	1.79	1.92
Bouroum	1.19	1.41	1.63	1.92	2.14	1.22	1.35	1.47	1.62	1.73	1.23	1.38	1.51	1.66	1.77
Dakiri	1.20	1.44	1.67	1.98	2.21	1.23	1.37	1.49	1.65	1.77	1.24	1.40	1.54	1.71	1.83
Gayeri	1.18	1.38	1.59	1.86	2.06	1.20	1.32	1.43	1.57	1.67	1.21	1.35	1.47	1.61	1.71
Kossougoudou	1.19	1.41	1.62	1.91	2.12	1.22	1.34	1.46	1.61	1.72	1.22	1.37	1.50	1.65	1.76
Piela	1.19	1.42	1.64	1.93	2.15	1.22	1.35	1.47	1.62	1.73	1.23	1.38	1.51	1.67	1.78
Sebba	1.23	1.49	1.75	2.09	2.35	1.26	1.41	1.55	1.73	1.86	1.27	1.45	1.61	1.81	1.94
Gotheye	1.20	1.42	1.65	1.95	2.17	1.23	1.36	1.48	1.63	1.75	1.23	1.39	1.52	1.69	1.80
Karma	1.23	1.48	1.74	2.08	2.34	1.26	1.41	1.55	1.72	1.85	1.26	1.44	1.60	1.80	1.93
Namaro	1.23	1.48	1.74	2.08	2.34	1.26	1.41	1.55	1.72	1.85	1.26	1.44	1.60	1.80	1.94

Table 36: Growth factors K(T) for the Pearson Type III and Log-Pearson Type III distributions.

Distributions	**Pearson Type III**					**Log-Pearson Type III**				
Stations & T [years]	**5**	**10**	**20**	**50**	**100**	**5**	**10**	**20**	**50**	**100**
Bogande	1.20	1.35	1.49	1.67	1.80	1.19	1.35	1.49	1.69	1.83
Manni	1.22	1.41	1.59	1.83	2.00	1.20	1.40	1.59	1.86	2.07
Aribinda	1.25	1.49	1.73	2.03	2.25	1.23	1.48	1.73	2.09	2.39
Bani	1.27	1.53	1.77	2.09	2.32	1.25	1.51	1.79	2.19	2.52
Barsalogho	1.24	1.43	1.60	1.83	1.99	1.23	1.42	1.61	1.86	2.05
Boulsa	1.24	1.47	1.70	1.99	2.21	1.22	1.44	1.67	1.99	2.26
Bouroum	1.22	1.38	1.53	1.71	1.83	1.22	1.38	1.53	1.71	1.85
Dakiri	1.23	1.41	1.58	1.79	1.94	1.23	1.41	1.59	1.83	2.02
Gayeri	1.20	1.36	1.51	1.70	1.83	1.20	1.36	1.52	1.73	1.89
Kossougoudou	1.20	1.40	1.60	1.87	2.06	1.18	1.38	1.59	1.89	2.14
Piela	1.22	1.40	1.57	1.79	1.94	1.21	1.38	1.55	1.78	1.96
Sebba	1.26	1.46	1.64	1.87	2.03	1.25	1.45	1.64	1.88	2.06
Gotheye	1.23	1.40	1.55	1.75	1.89	1.23	1.39	1.55	1.74	1.88
Karma	1.26	1.45	1.62	1.83	1.99	1.27	1.45	1.60	1.78	1.91
Namaro	1.22	1.49	1.76	2.12	2.39	1.20	1.46	1.75	2.21	2.61

Table 37: Growth factors K(T) for the EV1 and GEV distributions.

Distributions	**Gumbel, EV1**					**Generalized Extreme Values**				
Stations & T [years]	**5**	**10**	**20**	**50**	**100**	**5**	**10**	**20**	**50**	**100**
Bogande	1.19	1.34	1.49	1.68	1.83	1.19	1.34	1.49	1.69	1.84
Manni	1.22	1.39	1.56	1.78	1.94	1.20	1.39	1.59	1.86	2.09
Aribinda	1.26	1.47	1.67	1.93	2.13	1.23	1.46	1.71	2.07	2.38
Bani	1.28	1.50	1.72	1.99	2.20	1.24	1.49	1.76	2.14	2.45
Barsalogho	1.23	1.42	1.60	1.83	2.00	1.23	1.42	1.60	1.85	2.04
Boulsa	1.24	1.44	1.63	1.88	2.06	1.21	1.43	1.68	2.04	2.35
Bouroum	1.21	1.38	1.55	1.76	1.92	1.22	1.38	1.53	1.72	1.85
Dakiri	1.22	1.41	1.58	1.81	1.98	1.22	1.41	1.58	1.81	1.98
Gayeri	1.20	1.35	1.51	1.70	1.85	1.19	1.35	1.51	1.72	1.87
Kossougoudou	1.21	1.38	1.54	1.75	1.91	1.17	1.37	1.58	1.91	2.20
Piela	1.21	1.39	1.55	1.77	1.93	1.20	1.38	1.57	1.81	2.01
Sebba	1.25	1.45	1.65	1.90	2.09	1.25	1.45	1.64	1.89	2.07
Gotheye	1.22	1.39	1.56	1.78	1.95	1.22	1.39	1.56	1.76	1.91
Karma	1.25	1.45	1.64	1.89	2.07	1.25	1.45	1.62	1.85	2.01
Namaro	1.25	1.45	1.64	1.89	2.08	1.18	1.42	1.70	2.17	2.60

Table 38: Growth factors K(T) for the GLO and GPA distributions.

Distributions	**Generalized Logistic**					**Generalized Pareto**				
Stations & T [years]	**5**	**10**	**20**	**50**	**100**	**5**	**10**	**20**	**50**	**100**
Bogande	1.17	1.32	1.48	1.71	1.92	1.23	1.38	1.49	1.60	1.66
Manni	1.18	1.37	1.57	1.88	2.16	1.24	1.44	1.60	1.78	1.89
Aribinda	1.21	1.43	1.69	2.09	2.46	1.27	1.52	1.74	1.99	2.15
Bani	1.22	1.46	1.73	2.15	2.54	1.29	1.55	1.78	2.04	2.21
Barsalogho	1.21	1.39	1.59	1.88	2.13	1.27	1.46	1.61	1.75	1.83
Boulsa	1.19	1.41	1.65	2.05	2.42	1.25	1.49	1.70	1.96	2.14
Bouroum	1.20	1.36	1.52	1.75	1.95	1.26	1.42	1.52	1.61	1.66
Dakiri	1.20	1.38	1.57	1.84	2.07	1.27	1.45	1.58	1.70	1.77
Gayeri	1.18	1.33	1.50	1.74	1.95	1.23	1.39	1.51	1.63	1.69
Kossougoudou	1.16	1.34	1.56	1.91	2.25	1.21	1.41	1.61	1.85	2.02
Piela	1.19	1.36	1.55	1.84	2.09	1.24	1.43	1.57	1.72	1.81
Sebba	1.23	1.42	1.63	1.92	2.18	1.30	1.50	1.64	1.77	1.84
Gotheye	1.20	1.37	1.54	1.79	2.01	1.26	1.43	1.55	1.66	1.72
Karma	1.23	1.42	1.61	1.89	2.12	1.31	1.49	1.61	1.73	1.78
Namaro	1.16	1.39	1.68	2.16	2.64	1.21	1.48	1.75	2.13	2.43

Table 39: Growth factors K(T) for the Burr Type XII distribution.

Distribution	**Burr Type XII**				
Stations & T [years]	**5**	**10**	**20**	**50**	**100**
Bogande	1.17	1.31	1.47	1.70	1.90
Manni	1.18	1.36	1.56	1.88	2.16
Aribinda	1.20	1.42	1.68	2.09	2.46
Bani	1.22	1.45	1.72	2.14	2.53
Barsalogho	1.20	1.39	1.58	1.87	2.12
Boulsa	1.18	1.40	1.66	2.06	2.44
Bouroum	1.20	1.35	1.51	1.73	1.91
Dakiri	1.20	1.38	1.56	1.84	2.07
Gayeri	1.17	1.32	1.49	1.73	1.94
Kossougoudou	1.16	1.36	1.59	1.96	2.29
Piela	1.18	1.35	1.54	1.83	2.07
Sebba	1.23	1.42	1.61	1.89	2.12
Gotheye	1.20	1.37	1.54	1.78	1.99
Karma	1.23	1.41	1.59	1.83	2.04
Namaro	1.17	1.42	1.73	2.24	2.73

For illustrative purposes, the possible plots will be shown for some stations and one distribution, as proof that the procedure has been followed. Since the Burr Type XII distribution can represent the AM rainfall series of most stations, it will also be used for the DDF and IDF curves of the Bogande, Manni and Sebba stations, given their importance.

4.4.3: Plots by Return period

Bogande

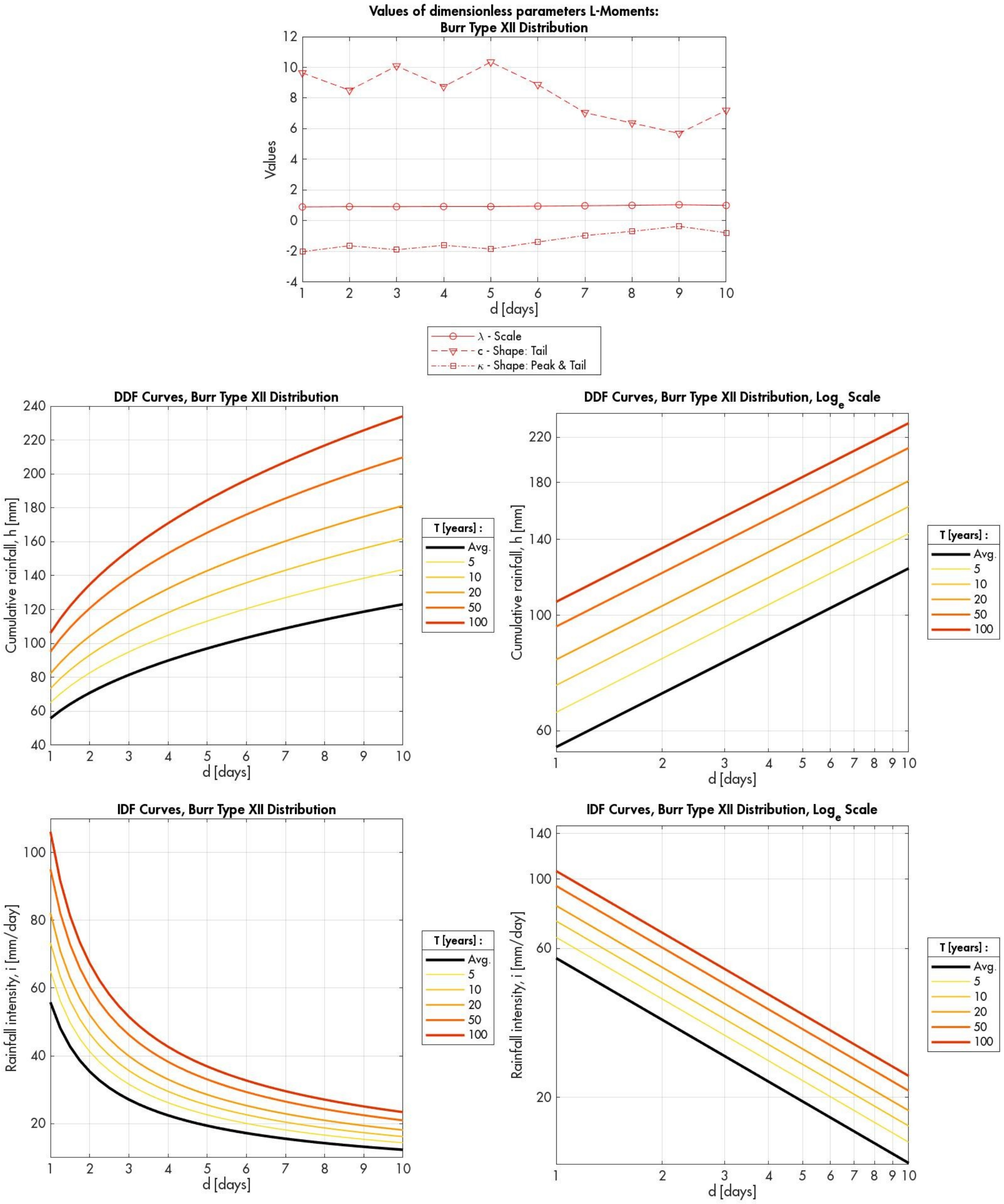


Figure 37: Dimensionless parameters, DDF and IDF curves of the Burr Type XII distribution, for the Bogande station.

Manni

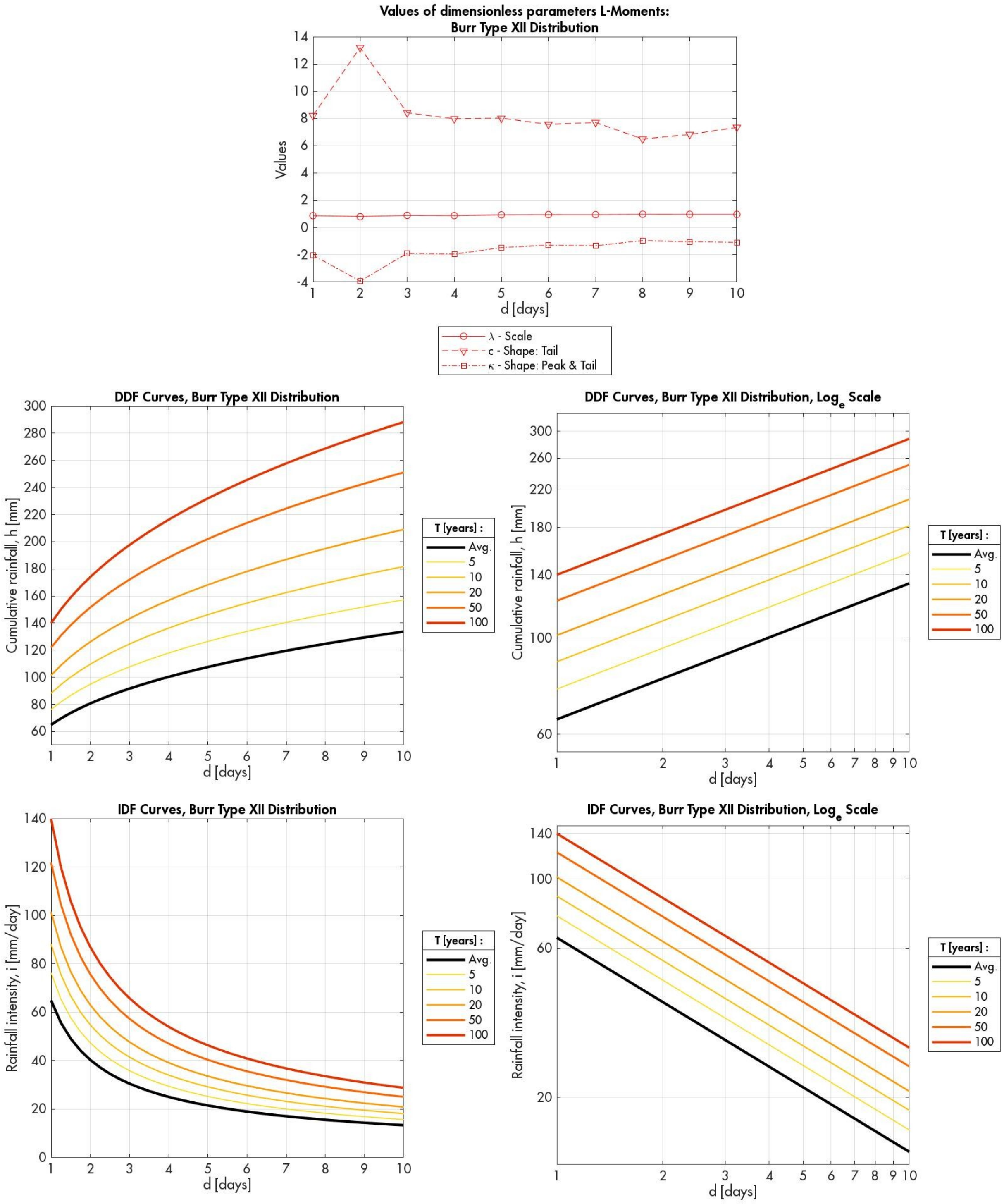


Figure 38: Dimensionless parameters, DDF and IDF curves of the Burr Type XII distribution, for the Manni station.

Sebba

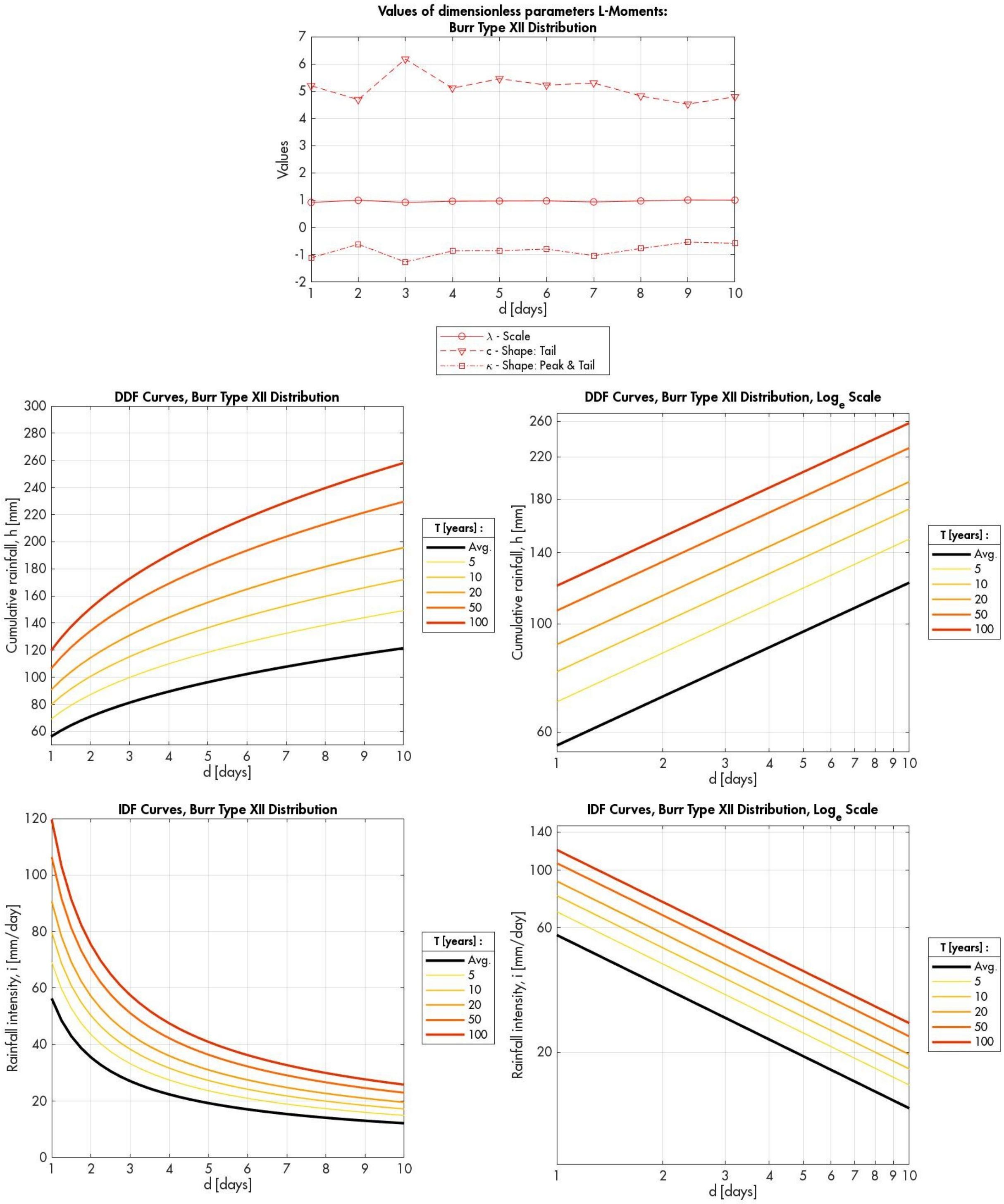


Figure 39: Dimensionless parameters, DDF and IDF curves of the Burr Type XII distribution, for the Sebba station.

4.5: Initial data comparison

Here are the comparative box-plots between the station (or pluviometer) data and the ERA5 local grid (Figure 7), for the values displayed in Table 10 and Table 11.

Figure 40: Comparative box-plots, Max.

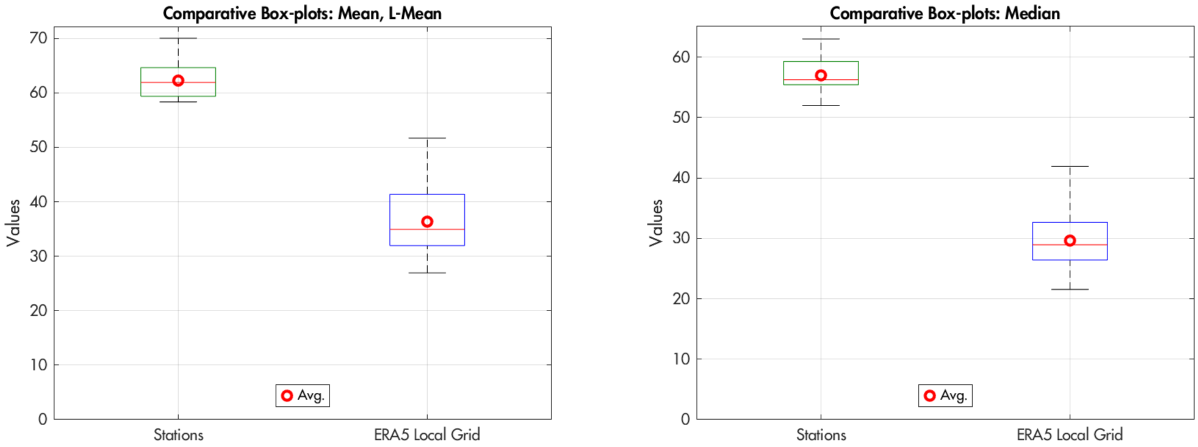


Figure 41: Comparative box-plots, Mean and Median.

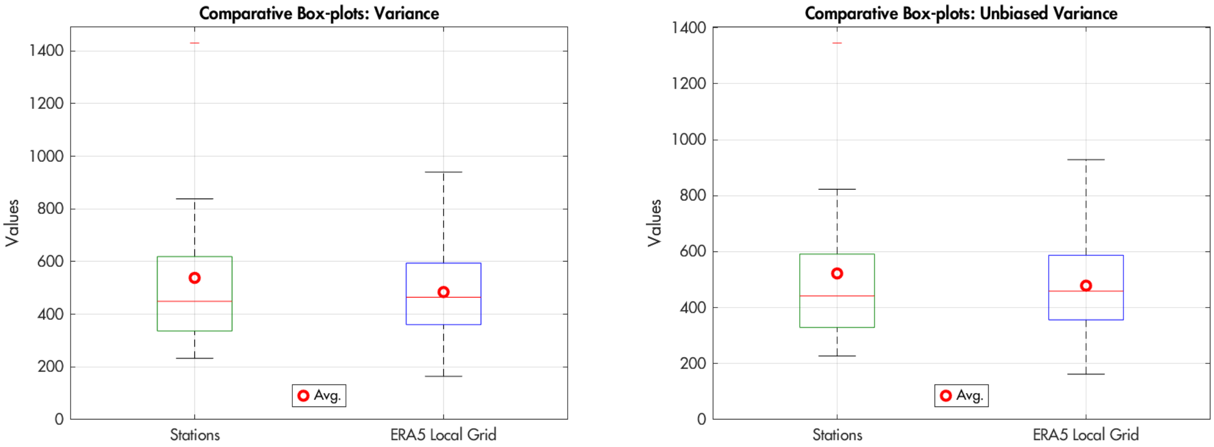


Figure 42: Comparative box-plots, Variance and Unbiased Variance.

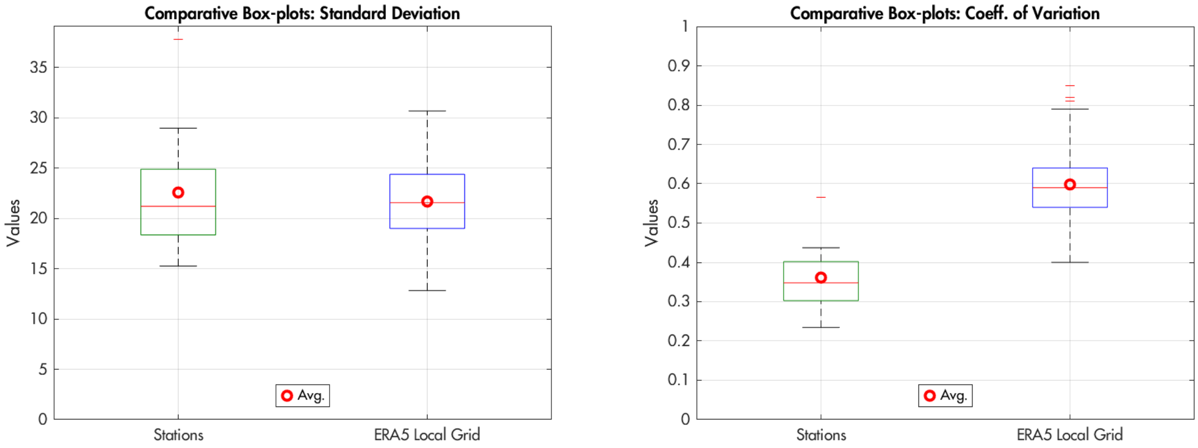


Figure 43: Comparative box-plots, Standard Deviation and Coefficient of Variation.

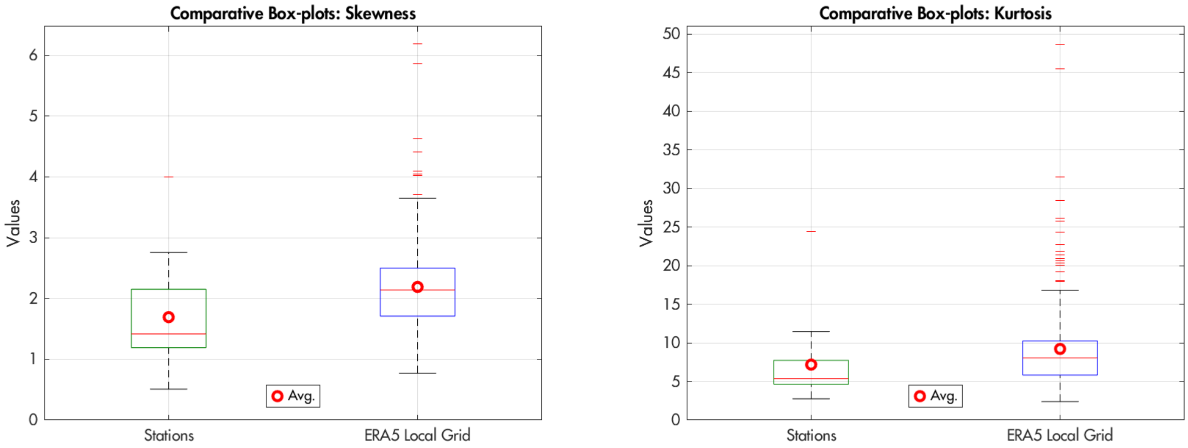


Figure 44: Comparative box-plots, Skewness and Kurtosis.

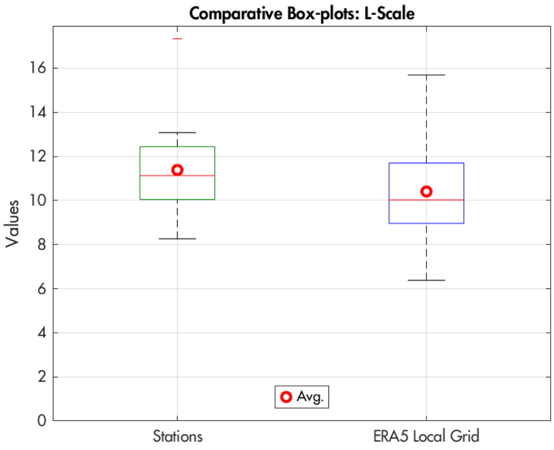


Figure 45: Comparative box-plots, L-Scale.

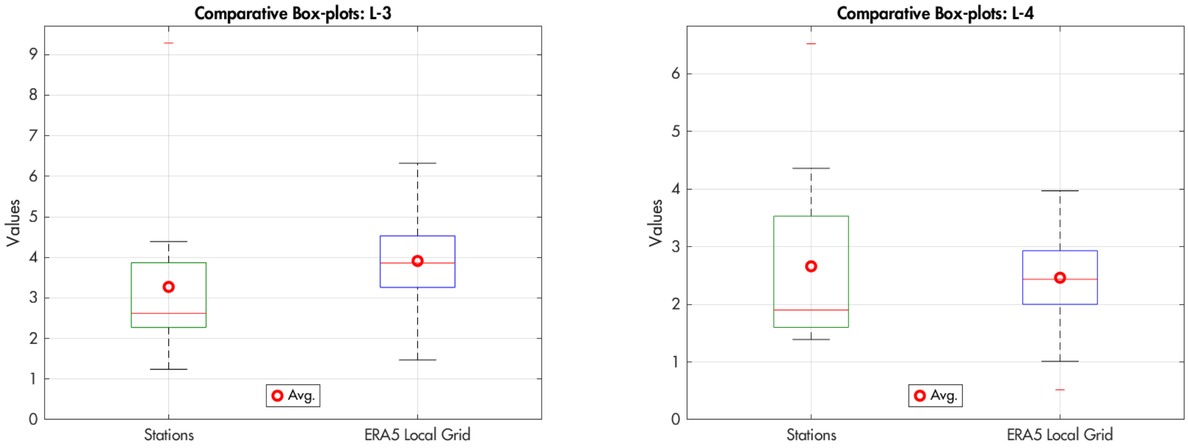


Figure 46: Comparative box-plots, L-3 and L-4.

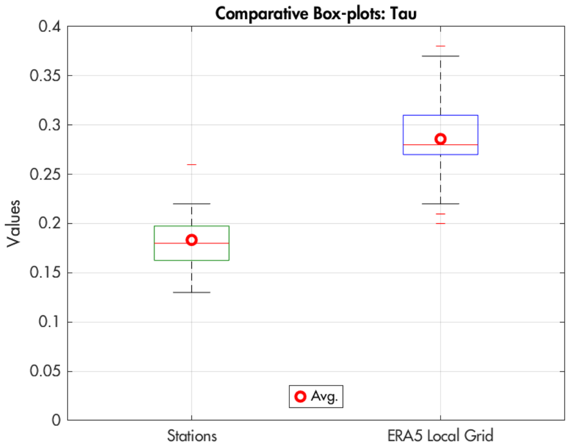


Figure 47: Comparative box-plots, Tau.

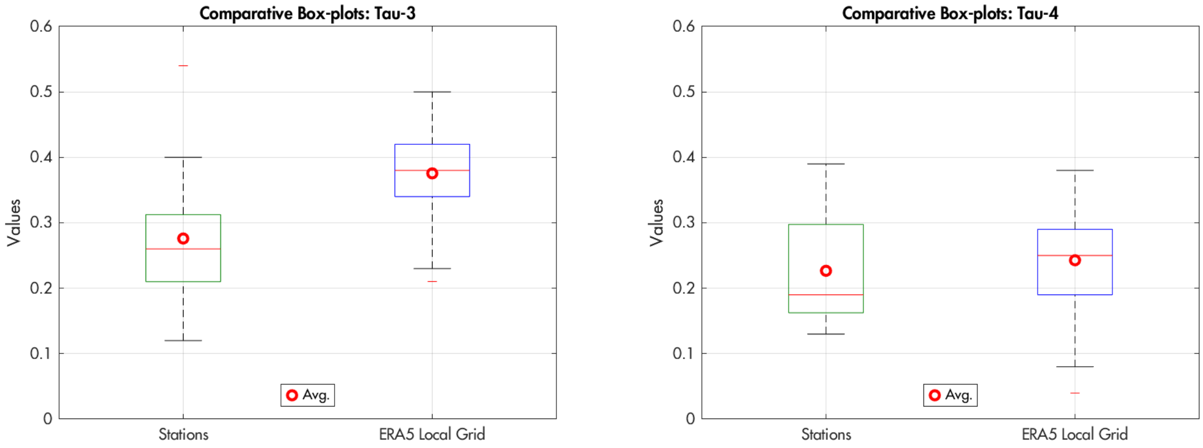


Figure 48: Comparative box-plots, Tau-3 and Tau-4.

From Figure 40 and Figure 48 it can be seen that the two datasets are shifted to each other. If in some instances like for the Variance, Unbiased Variance, Standard Deviation, L-Mean and L-Skewness, the means of the datasets seem comparable, ERA5 has much more data points thus covering more values. From Figure 44, it seems that the ERA5 data appears to be more asymmetric around the mean values and flat in proximity to the tails.

4.6: Grid data correction

4.6.1: Associated grid points

For the associated station and grid data points see Table 9.

Table 40: Ordinary moments for the closest grid data points relative to the stations.

ERA5 Points	$\bar{\mu}$ [mm]	$\tilde{\mu}$ [mm]	σ^2 [mm^2]	σ_{ub}^2 [mm^2]	σ [mm]	S [ ]	K [ ]
152	36.18	29.63	461.47	455.97	21.48	2.70	11.23
134	35.04	28.95	378.41	373.90	19.45	2.09	7.79
60	32.03	26.32	482.17	476.43	21.96	2.61	9.88
98	33.98	28.61	411.17	406.27	20.28	1.84	5.86
113	33.63	29.87	275.89	272.60	16.61	2.33	9.93
169	40.33	32.18	587.62	580.63	24.24	2.34	9.08
114	34.80	29.69	335.71	331.71	18.32	1.69	5.71
134	35.04	28.95	378.41	373.90	19.45	2.09	7.79
173	36.27	31.88	420.61	415.60	20.51	2.76	14.40
152	36.18	29.63	461.47	455.97	21.48	2.70	11.23
170	40.65	32.51	504.30	498.30	22.46	1.85	6.36
119	32.39	26.73	318.96	315.17	17.86	2.10	7.45
105	32.03	26.03	372.02	367.59	19.29	2.27	9.71
105	32.03	26.03	372.02	367.59	19.29	2.27	9.71
106	32.25	27.18	376.35	371.87	19.40	1.98	7.28

Table 41: L-Moments for the closest grid data points relative to the stations.

ERA5 Points	L_2 [mm]	L_3 [mm]	L_4 [mm]
152	9.67	4.12	2.78
134	9.35	3.85	2.31
60	9.47	4.51	3.49
98	9.81	3.78	2.64
113	7.97	2.76	2.10
169	11.30	4.87	3.09
114	9.29	3.23	1.87
134	9.35	3.85	2.31
173	9.69	3.45	2.22
152	9.67	4.12	2.78
170	11.09	4.24	2.54
119	8.49	3.36	2.48
105	9.33	3.27	2.32
105	9.33	3.27	2.32
106	9.49	3.35	2.54

4.6.2: Moments' ratios

For the associated couples' numbers see Table 9.

Table 42: Ordinary moments ratios of the associated couples.

Couples' n°	$\bar{\mu}$ ratios [ ]	$\tilde{\mu}$ ratios [ ]	σ^2 ratios [ ]	σ_{ub}^2 ratios [ ]	σ ratios [ ]	S ratios [ ]	K ratios [ ]
1	1.71	1.87	0.70	0.69	0.84	0.53	0.48
2	2.00	2.18	1.57	1.55	1.25	0.67	0.62
3	1.82	2.07	0.93	0.93	0.96	0.48	0.53
4	1.75	1.96	0.99	0.99	1.00	0.55	0.63
5	1.86	1.91	1.28	1.27	1.13	0.56	0.47
6	1.64	1.86	1.43	1.42	1.19	1.71	2.69
7	1.71	1.88	0.99	0.97	0.99	1.12	1.30
8	1.70	1.98	0.86	0.85	0.92	0.24	0.35
9	1.79	1.93	0.55	0.55	0.74	0.40	0.30
10	1.71	1.89	1.38	1.37	1.18	1.02	1.02
11	1.56	1.73	1.17	1.17	1.08	1.21	1.41
12	1.88	1.98	1.66	1.65	1.29	0.56	0.62
13	1.84	2.31	1.09	1.06	1.04	0.67	0.68
14	1.85	2.00	1.69	1.62	1.30	0.65	0.57
15	2.07	2.10	3.80	3.62	1.95	1.17	1.08
Means:	**1.79**	**1.98**	**1.34**	**1.31**	**1.12**	**0.77**	**0.85**

Table 43: L-Moments ratios of the associated couples.

Couples' n°	L_2 ratios [ ]	L_3 ratios [ ]	L_4 ratios [ ]
1	0.98	0.61	0.60
2	1.37	0.98	1.20
3	1.20	0.50	0.54
4	1.13	0.61	0.54
5	1.26	0.95	0.76
6	1.09	0.80	1.15
7	0.98	0.75	1.12
8	1.08	0.32	0.60
9	0.85	0.50	0.69
10	1.14	1.07	1.57
11	1.05	0.98	1.37
12	1.47	0.82	0.66
13	1.13	0.62	0.81
14	1.40	1.14	1.75
15	1.83	2.77	2.57
Means:	**1.20**	**0.89**	**1.06**

4.6.3: Error indicators

Ordinary moments

Table 44: Error indicators of the ordinary moments, between the stations and original grid data points.

Before the correction				**Errors**			
Couples' n°	**$\bar{\mu}$ [mm]**	**$\tilde{\mu}$ [mm]**	**σ^2 [mm^2]**	**σ_{ub}^2 [mm^2]**	**σ [mm]**	**S [ ]**	**K [ ]**
1	25.81	25.67	-443.51	-141.18	-3.52	-1.28	-5.85
2	35.00	34.05	-354.05	205.27	4.91	-0.68	-2.94
3	26.30	28.18	-460.98	-34.76	-0.77	-1.35	-4.68
4	25.36	27.39	-390.96	-6.00	-0.07	-0.83	-2.15
5	28.77	27.13	-257.12	74.02	2.16	-1.02	-5.29
6	25.92	27.72	-558.67	242.37	4.71	1.66	15.32
7	24.60	26.11	-317.50	-8.35	-0.11	0.20	1.74
8	24.67	28.50	-360.42	-55.84	-1.46	-1.58	-5.04
9	28.81	29.62	-405.36	-188.39	-5.26	-1.65	-10.10
10	25.74	26.37	-436.23	169.28	3.76	0.06	0.24
11	22.65	23.74	-479.97	83.78	1.87	0.39	2.60
12	28.51	26.32	-295.96	204.56	5.14	-0.93	-2.83
13	27.05	33.97	-351.89	23.51	0.84	-0.74	-3.13
14	27.34	25.97	-346.97	226.87	5.76	-0.79	-4.13
15	34.56	29.82	-338.54	973.47	18.41	0.33	0.55
Error indicators							
RMSE [ ]	106.92	109.12	1525.69	1115.08	22.60	4.00	22.46
MAE [ ]	411.09	420.56	5798.13	1768.61	36.37	8.21	25.69
MAPE [%]	44.02	49.20	-1713.14	22.58	10.75	-32.31	-23.85
R² [ ]	0.16	0.18	0.05	0.04	0.05	0.00	0.00

Table 45: Error indicators of the ordinary moments, between the stations and corrected grid data points.

After the correction				**Errors**			
Couples' n°	**$\bar{\mu}$ [mm]**	**$\tilde{\mu}$ [mm]**	**σ^2 [mm^2]**	**σ_{ub}^2 [mm^2]**	**σ [mm]**	**S [ ]**	**K [ ]**
1	-2.90	-3.24	-599.76	-284.05	-6.20	-0.66	-4.17
2	7.20	5.80	-482.17	88.11	2.48	-0.20	-1.77
3	0.89	2.50	-624.23	-184.05	-3.51	-0.75	-3.20
4	-1.60	-0.53	-530.18	-133.30	-2.60	-0.41	-1.27
5	2.09	-2.02	-350.53	-11.40	0.08	-0.48	-3.80
6	-6.08	-3.68	-757.63	60.43	1.68	2.20	16.68
7	-3.01	-2.86	-431.17	-112.29	-2.40	0.59	2.60
8	-3.13	0.25	-488.54	-173.00	-3.89	-1.10	-3.87
9	0.03	-1.49	-547.77	-318.61	-7.82	-1.02	-7.94
10	-2.97	-2.54	-592.48	26.41	1.08	0.68	1.92
11	-9.60	-7.99	-650.72	-72.36	-0.94	0.82	3.55
12	2.81	0.24	-403.95	105.80	2.91	-0.45	-1.71
13	1.64	8.57	-477.85	-91.67	-1.57	-0.22	-1.67
14	1.93	0.57	-472.93	111.69	3.35	-0.27	-2.67
15	8.97	3.30	-465.97	856.95	15.99	0.79	1.64
Error indicators							
RMSE [ ]	17.88	15.29	2071.53	1029.70	20.70	3.30	21.02
MAE [ ]	3.75	3.15	7875.87	131.34	1.37	0.48	5.68
MAPE [%]	-0.40	-0.37	-2327.04	-1.68	-0.41	-1.89	-5.27
R² [ ]	0.16	0.18	0.05	0.04	0.05	0.00	0.00

L-Moments

Table 46: Error indicators of the L-Moments, between the stations and original grid data points.

Before the correction	**Errors**		
Couples' n°	**L_2 [mm]**	**L_3 [mm]**	**L_4 [mm]**
1	-0.15	-1.59	-1.10
2	3.44	-0.08	0.46
3	1.94	-2.25	-1.59
4	1.32	-1.47	-1.21
5	2.05	-0.14	-0.51
6	1.00	-0.97	0.46
7	-0.15	-0.82	0.23
8	0.76	-2.61	-0.92
9	-1.43	-1.74	-0.68
10	1.38	0.27	1.58
11	0.51	-0.10	0.93
12	4.00	-0.61	-0.85
13	1.20	-1.24	-0.43
14	3.75	0.46	1.75
15	7.85	5.94	3.98
Error indicators			
RMSE [ ]	10.98	7.66	5.52
MAE [ ]	27.47	6.95	2.10
MAPE [%]	16.09	-14.16	5.26
R^2 [ ]	0.01	0.00	0.02

Table 47: Error indicators of the L-Moments, between the stations and corrected grid data points.

After the correction	**Errors**		
Couples' n°	**L_2 [mm]**	**L_3 [mm]**	**L_4 [mm]**
1	-2.07	-1.15	-1.27
2	1.59	0.33	0.31
3	0.06	-1.77	-1.81
4	-0.62	-1.07	-1.38
5	0.47	0.15	-0.64
6	-1.24	-0.46	0.27
7	-1.99	-0.48	0.11
8	-1.09	-2.20	-1.07
9	-3.35	-1.38	-0.82
10	-0.54	0.71	1.41
11	-1.69	0.35	0.77
12	2.32	-0.25	-1.01
13	-0.65	-0.89	-0.58
14	1.90	0.81	1.60
15	5.97	6.29	3.82
Error indicators			
RMSE [ ]	8.57	7.39	5.50
MAE [ ]	0.93	1.03	0.27
MAPE [%]	-0.54	-2.09	-0.68
R^2 [ ]	0.01	0.00	0.02

4.6.4: Error reductions

Ordinary moments

Table 48: Error reduction for the ordinary moments, between the original grid data and the corrected one.

Error indicators	$\bar{\mu}$	$\tilde{\mu}$	σ^2	σ_{ub}^2	σ	S	K
RMSE	83.28%	85.99%	-35.78%	7.66%	8.40%	17.31%	6.43%
MAE	99.09%	99.25%	-35.83%	92.57%	96.23%	94.15%	77.89%
MAPE	100.91%	100.75%	-35.83%	107.43%	103.77%	94.15%	77.89%
R^2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

L-Moments

Table 49: Error reduction for the L-Moments, between the original grid data and the corrected one.

Error indicators	L_2	L_3	L_4
RMSE	21.90%	3.50%	0.24%
MAE	96.62%	85.22%	87.04%
MAPE	103.38%	85.22%	112.96%
R^2	0.00%	0.00%	0.00%

4.7: Corrected data comparison

Here are included the box-plots for the moments that have the highest R^2, while the others are omitted for brevity.

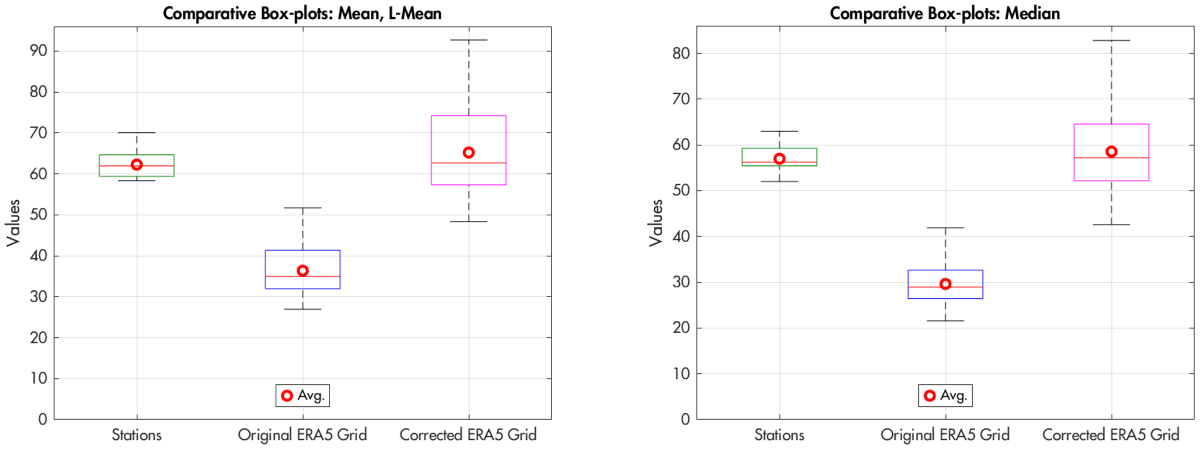


Figure 49: Comparative box-plots, Mean and Median.

As it can be seen, since the grid data is scaled relative to the stations, it shows an increase level of spread on the previous plots.

4.7.1: Trends & QQ plots

Median

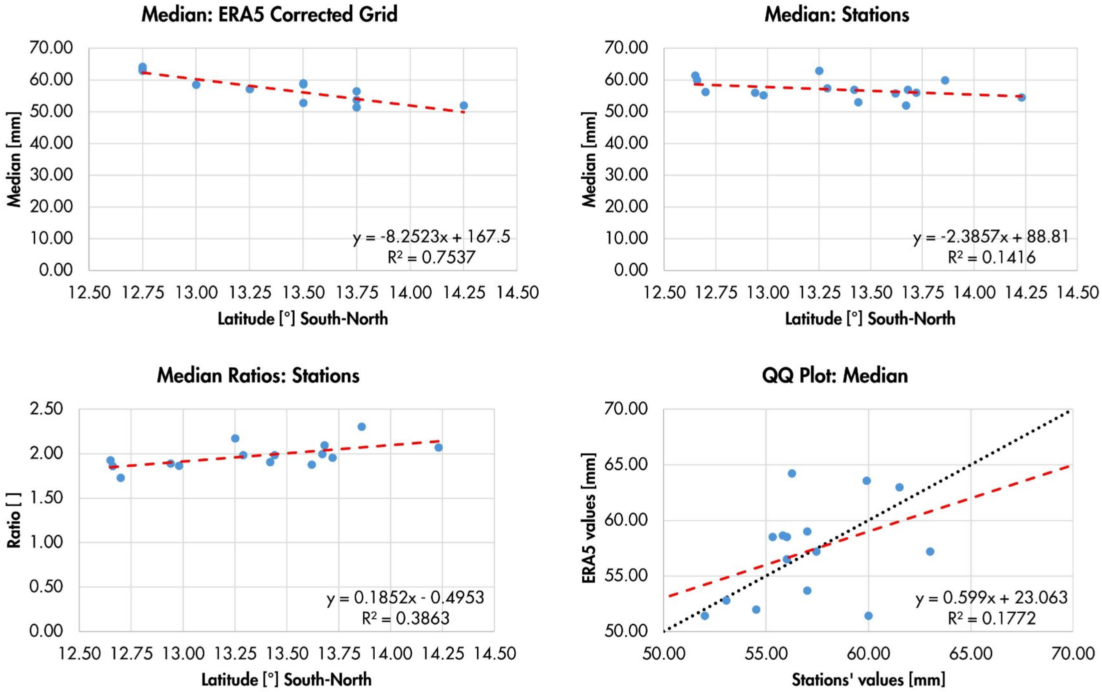


Figure 50: Median trends by latitude and QQ plot.

Mean/L-Mean

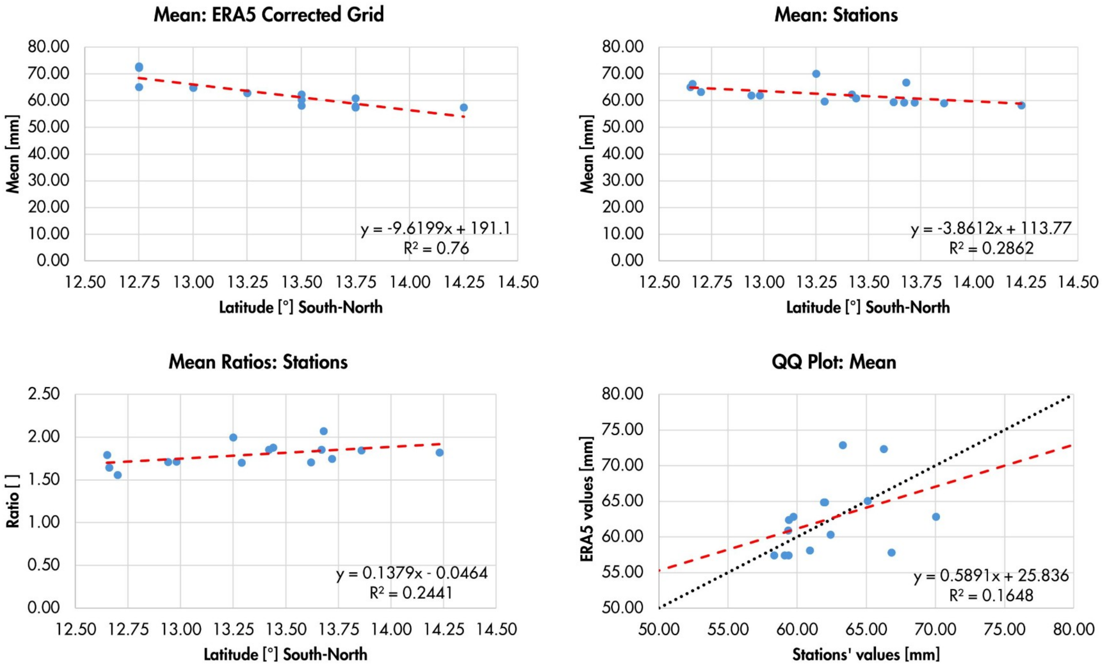


Figure 51: Mean/L-Mean trends by latitude and QQ plot.

4.8: Spatial interpolation: Ordinary Kriging

4.8.1: Moments

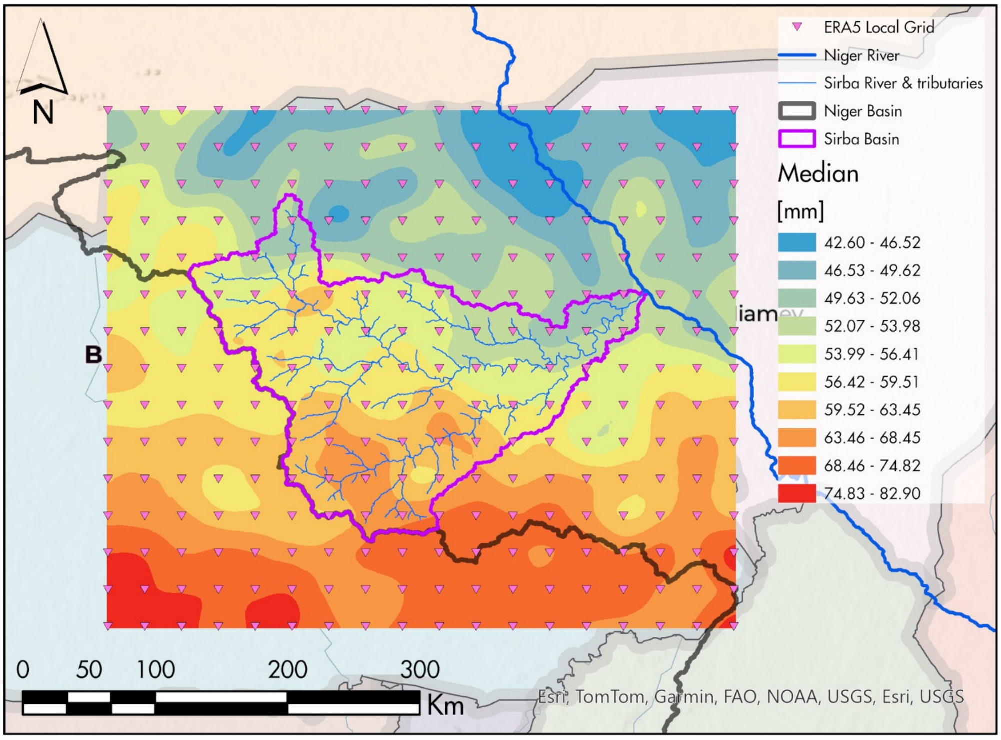


Figure 52: Median at ungauged locations interpolated from the ERA5 Corrected Grid.

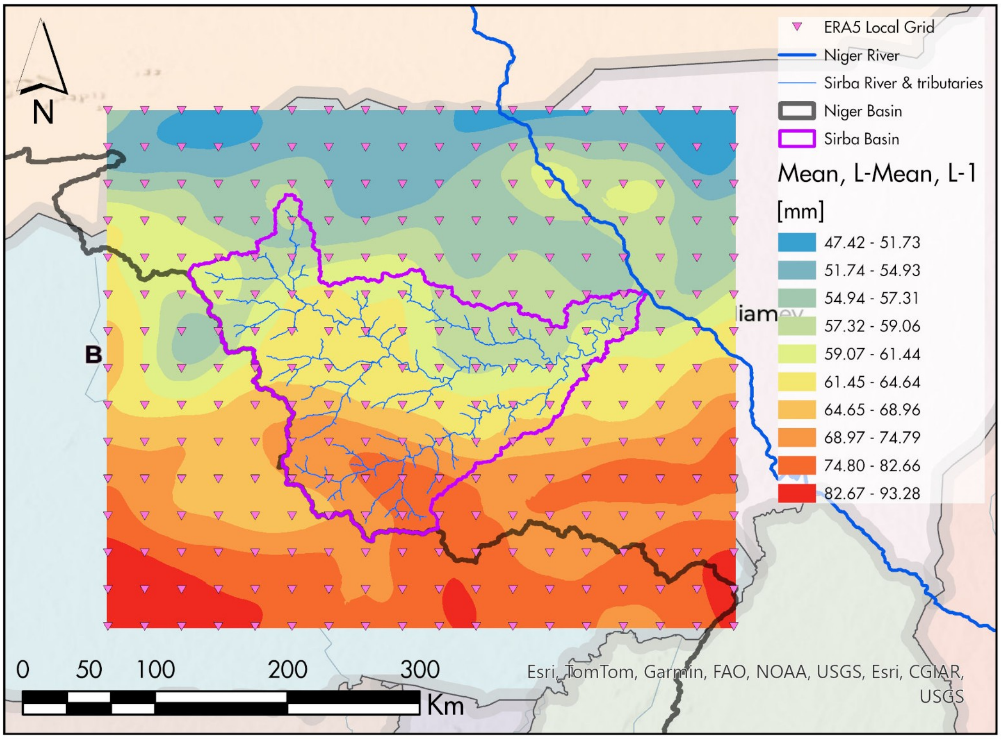


Figure 53: Mean at ungauged locations interpolated from the ERA5 Corrected Grid.

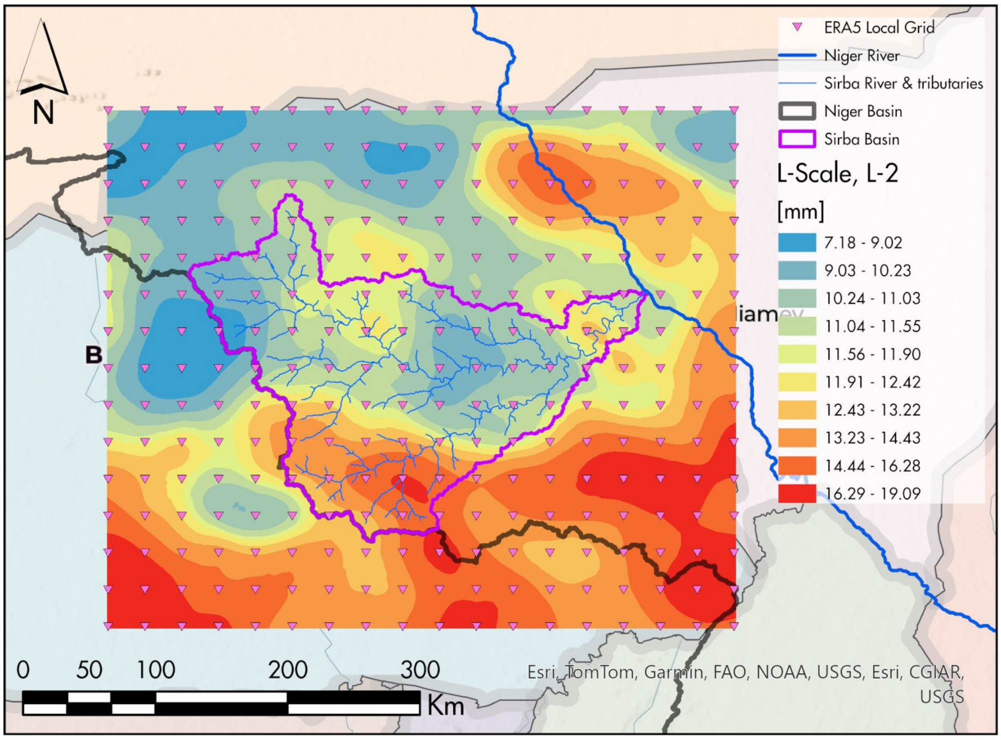


Figure 54: L-Scale at ungauged locations interpolated from the ERA5 Corrected Grid.

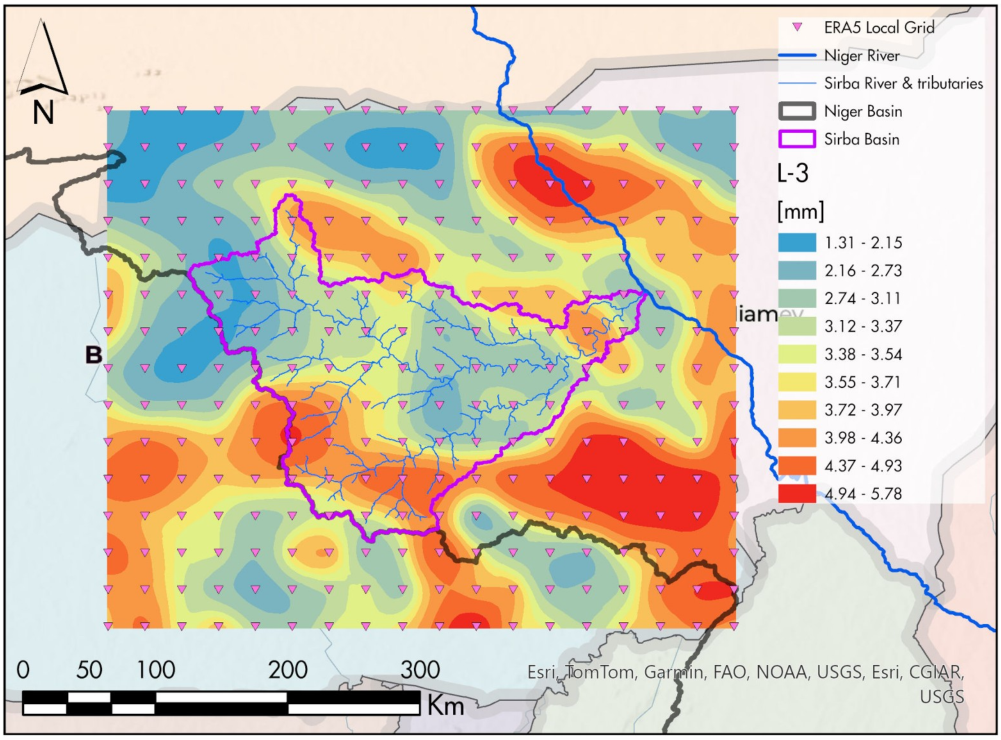


Figure 55: L-3 at ungauged locations interpolated from the ERA5 Corrected Grid.

Although L-Scale and L-3 showed low values for the error indicators between the station data and the corrected grid (see Table 47 and Table 49), they are included here because they were used to calculate the quantiles of the corrected grid.

4.8.2: ERA5 quantiles

EV1, Gumbel

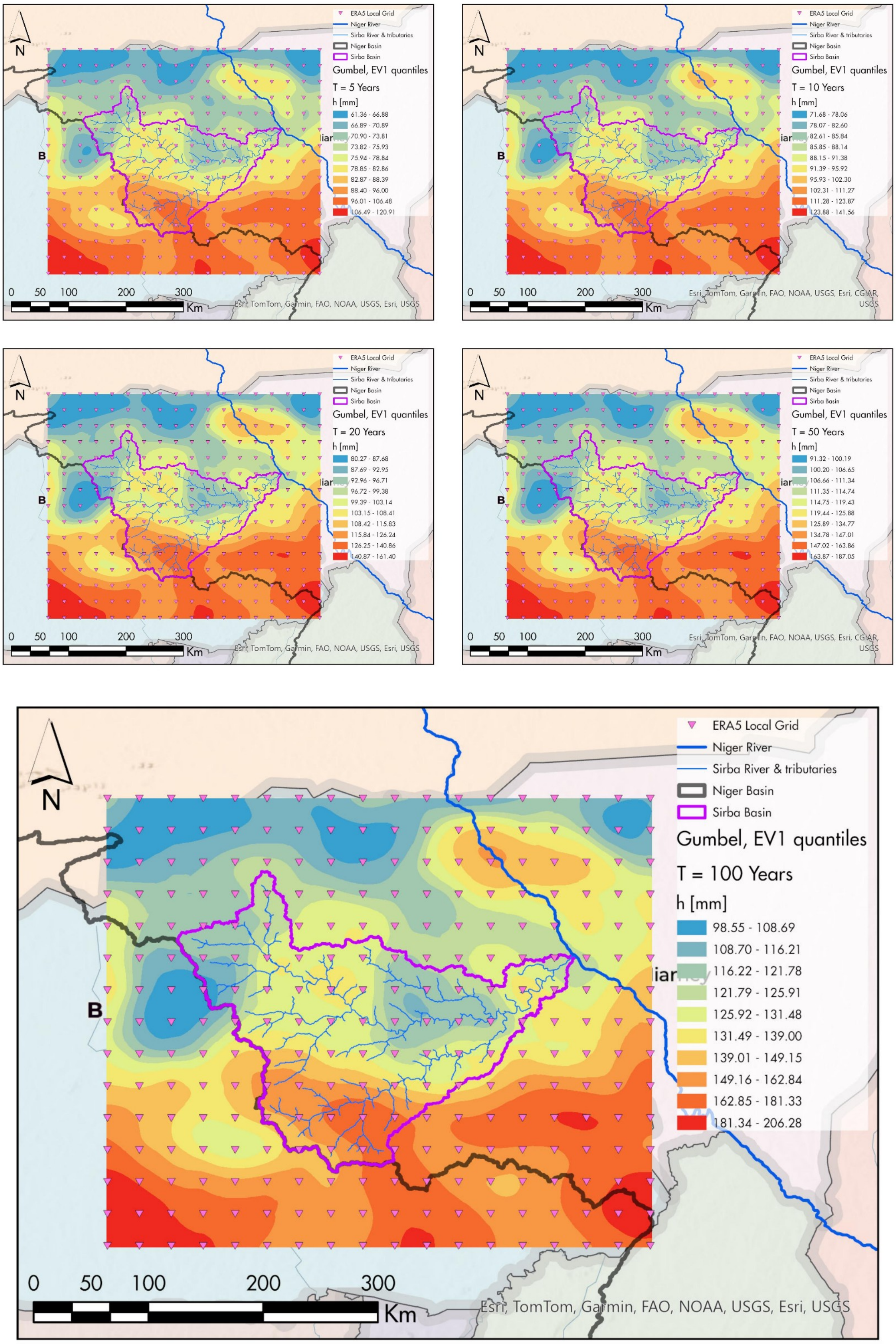


Figure 56: Rainfall heights of the ERA5 Corrected Grid, in terms of Return Period T, for the EV1 distribution.

GEV, Generalized Extreme Values

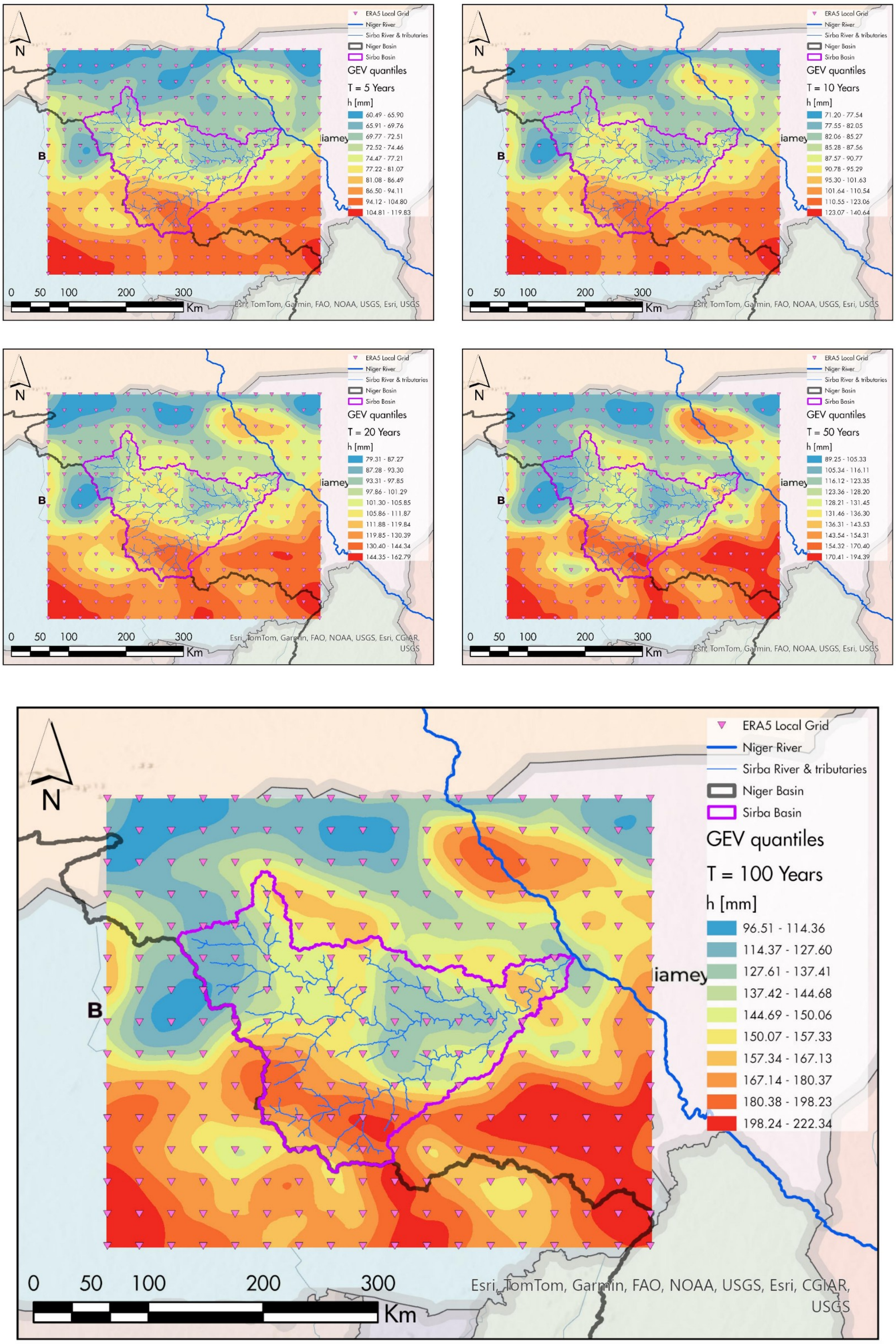


Figure 57: Rainfall heights of the ERA5 Corrected Grid, in terms of Return Period T, for the GEV distribution.

BTXII, Burr Type XII

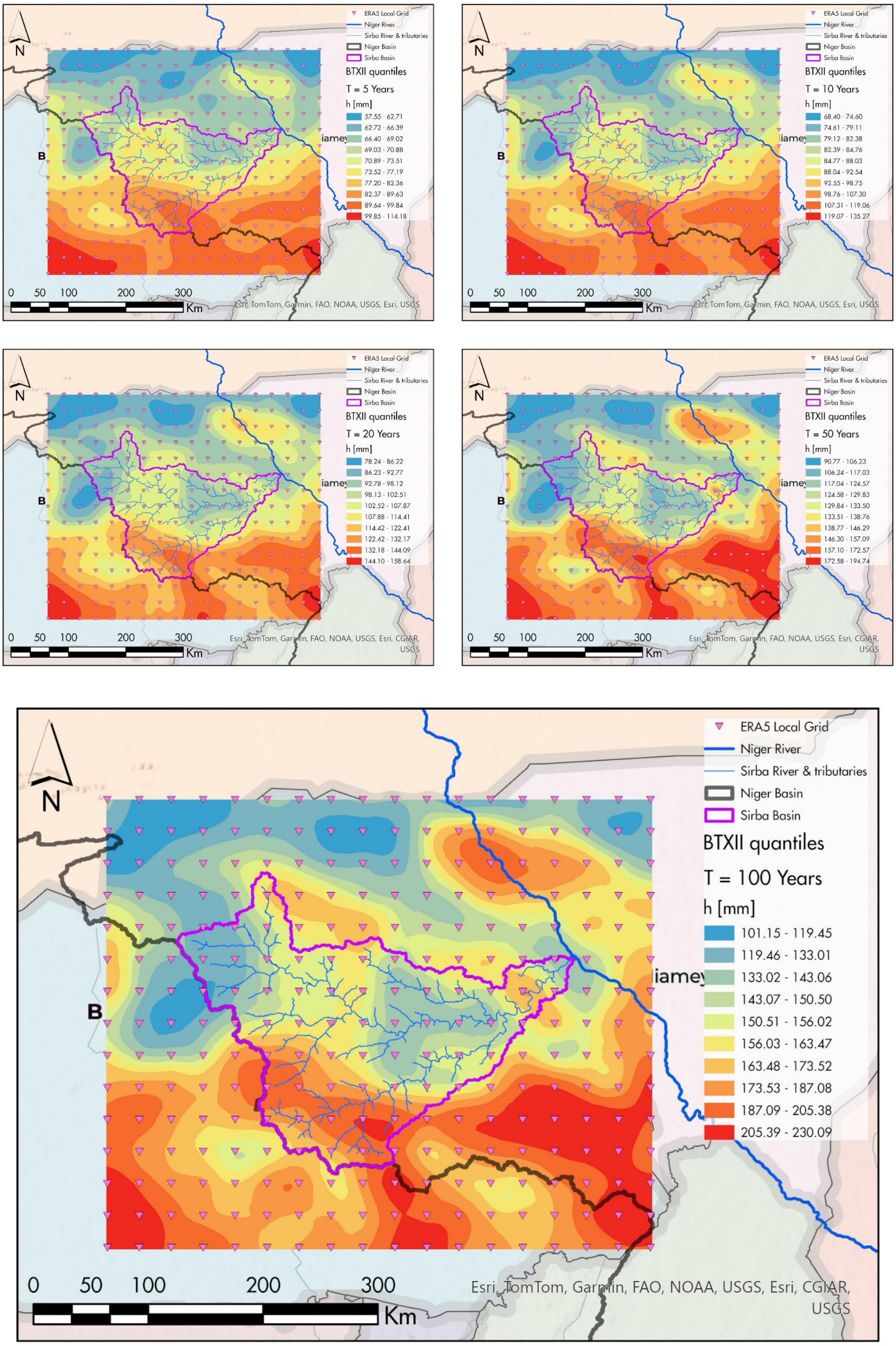
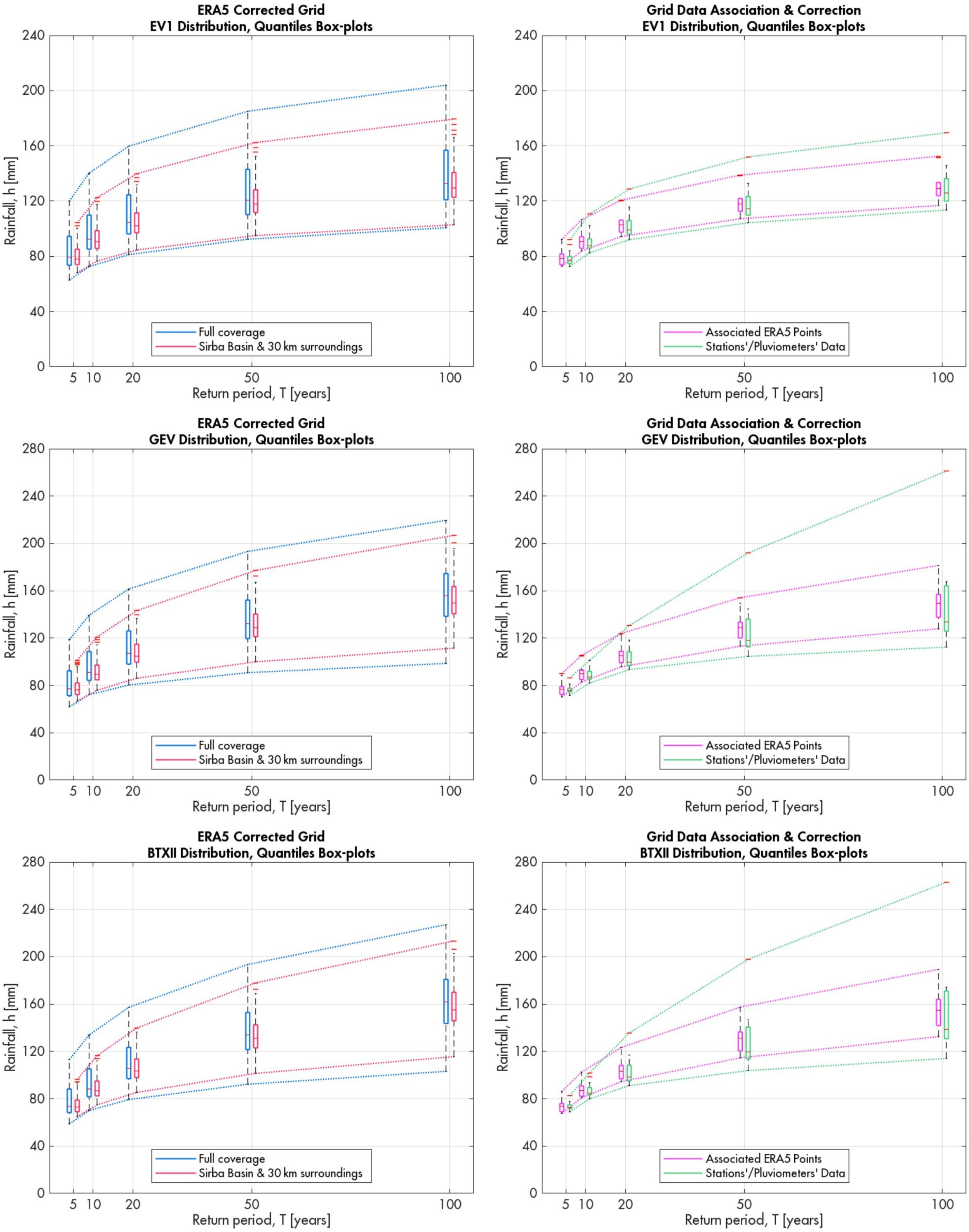


Figure 58: Rainfall heights of the ERA5 Corrected Grid, in terms of Return Period T, for the Burr Type XII distribution.

4.8.3: Quantiles' box-plots



ERA5 Corrected Grid
EV1 Distribution, Quantiles Box-plots

Grid Data Association & Correction
EV1 Distribution, Quantiles Box-plots

ERA5 Corrected Grid
GEV Distribution, Quantiles Box-plots

Grid Data Association & Correction
GEV Distribution, Quantiles Box-plots

ERA5 Corrected Grid
BTXII Distribution, Quantiles Box-plots

Grid Data Association & Correction
BTXII Distribution, Quantiles Box-plots

Figure 59: Quantiles' box-plots between the ERA5 Corrected Grid and the Station Data.

Here are the final comparative box-plots of the computed quantiles, between the ERA5 Corrected Grid, shown in variations of full coverage (see Figure 7), covering the Sirba Basin and its 30 km surroundings, along with the associated grid points (Table 9) and station (or pluviometer) data (see Figure 4 and Table 1).

5: Discussion

5.1: River geometries & discharges

Although this paper focuses solely on rainfall statistics, some remarks on river geometries and discharges can be made, which could serve for future studies.

As discussed in the Methods section, river geometries and basin shapes can be derived from a Digital Terrain Model (DTM). The results depend on the model's resolution, which is also used to compute these features. To achieve a level of detail in the rivers suitable for visual interpretation, certain local slope variations, and by consequence small rivers, are not considered.

In the specific context of the Sirba Basin, where slope variations are generally low (less than 2% in most cases), the complexity of river tributaries poses challenges. The basin contains a high number of small rivers, which are themselves tributaries of larger rivers. These small catchments often contribute to flood risk in areas not directly adjacent to the main rivers. During the dry season, these catchments typically exhibit no surface flow; however, their activation during heavy rainfall events can lead to localized flooding.

Figure 60: Details of small rivers intersecting localities: Barsalogho (left), located in the west, far from major tributaries, and Sebba (right), situated near Sirba's tributary, the Yali river. Source: Maxar Technologies (2024).

Satellite imagery can visually reveal small catchments, many of which intersect with inhabited areas. Future flood simulation studies should account for these local variations to provide a more accurate assessment of flood risks.

Achieving this requires a higher-resolution DTM, as the currently used AfSIS DTM, employed for basic river geometry computations, has a resolution of only 90 meters. Presently, DTMs with resolutions of 30 meters, such as ALOS, Copernicus DEM, NASA's SRTM, and ASTER GDEM, have demonstrated high reliability for hydrological modeling. For even more precise results, LiDAR-based DTMs can be acquired on a smaller scale, offering resolutions as fine as 1 meter. However, this approach requires complex coordination with state authorities.

The basin contains several small earth dams used for irrigation purposes, where discharge data is usually recorded. The most important dams are located in the following localities:

- Bogande;
- Bilanga;
- Dakiri;
- Kossogoudou;
- Liptougou;
- Manni;
- Tougouri;
- Yalgo;
- Zeguedeguen.

In addition to these, there are many other dams that form artificial lakes. Although a statistical analysis of discharges was not included in this study, organizing and analyzing this data would be valuable for establishing a stronger link between rainfall and discharge patterns in the basin, as well as their interactions with these dams.

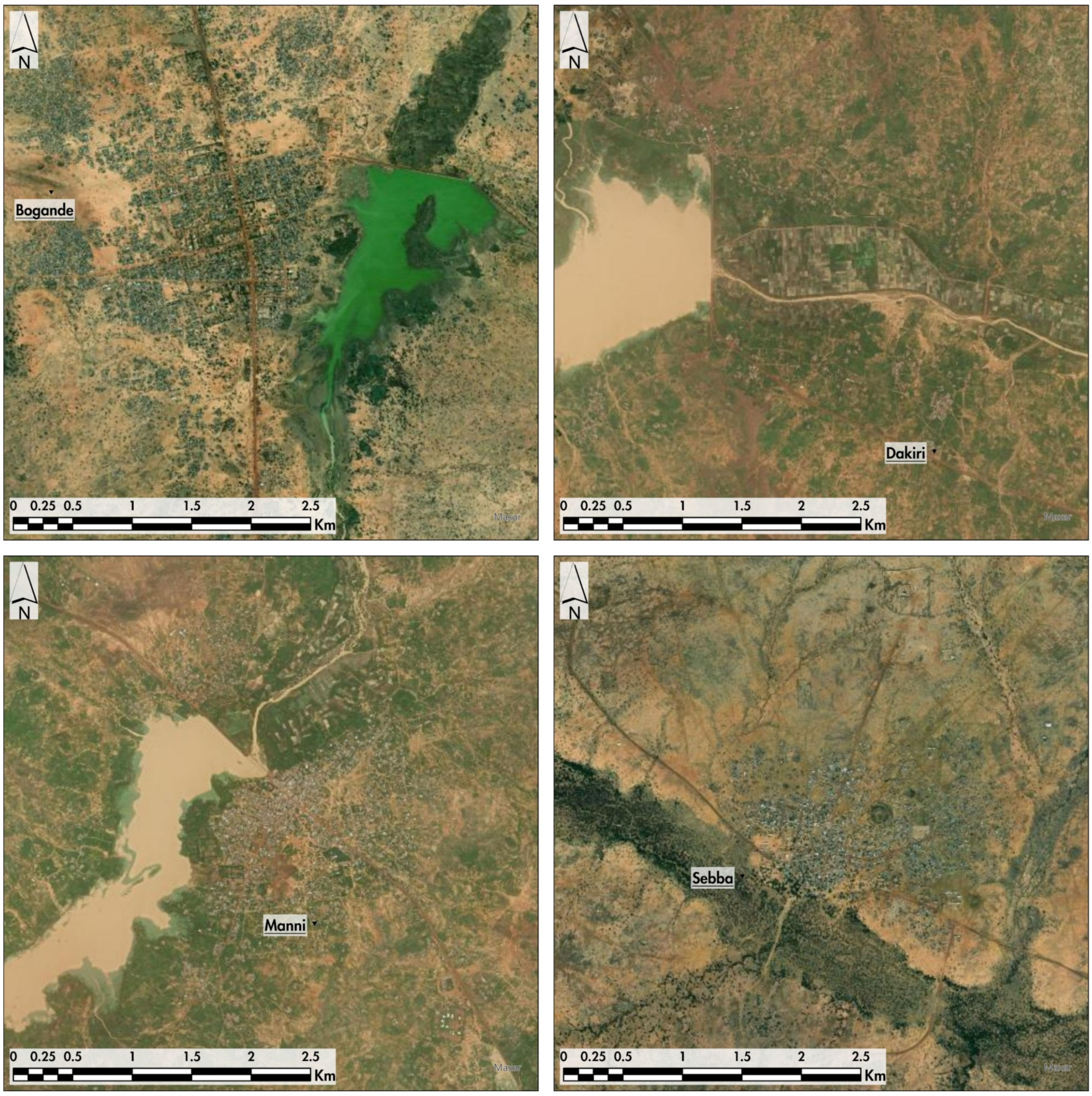


Figure 61: Satellite images of key locations in the Sirba Basin: Bogande, Dakiri, Manni, and Sebba. Note the proximity of rivers or dams to these localities and the small river branches intersecting areas such as Bogande and Sebba. Source: Maxar Technologies (2024).

As a conclusive note to this section, viewing the area through satellite images clearly indicates that some intervention against flood risk is necessary.

Without delving into complex details, the Sirba Basin lacks a levee system, which would require extensive state programs and funding. A viable alternative could be the adoption of drainage systems, such as channels, to redirect water flowing toward localities back into the Sirba's tributaries. The necessary work could be carried out by the local population, which would help reduce the need for extensive government intervention.

Another viable alternative involves techniques of rainwater harvesting as an adaptation strategy for flood mitigation. This was covered by P. Tamagnone, E. Comino, and M. Rosso in their 2020 paper [46]. Without entering into specifics, the techniques consist of digging small holes on large areas to drain water using the surface of the watershed itself. In his doctoral thesis, P. Tamagnone (where he performed HEC-RAS simulations on the watershed of the village of Touré, Niger, in the Sirba Basin) remarked that these methods could improve the existing situation in terms of outgoing flow, as the basin shows clear signs of soil degradation that exacerbate the flood risk [47].

5.2: Distribution fitting

The backbone of this paper is the distribution fitting, which provided some good results, at least for the Burr Type XII function, as it better represents the stations' dataset when considering the adopted error indicators, statistical tests, and information criteria, except for the Anderson-Darling test for the Kossogoudou station (see Table 24).

Obviously, the distribution selection is not arbitrary. Other distributions, such as the Gumbel and GEV, can also represent the stations' dataset, though with less robustness for the Anderson-Darling test (see Table 22), but with similar error indicators and ICs values (see Table 16 and Table 28). This shows how flexible extreme value distributions can be, even in cases like this, where there is a distinct difference between dry and monsoon seasons.

Other distributions, like the Pearson Type III, its log alternative, the GLO, and GPA, which performed well in terms of error indicators, seem unreliable when considering the Anderson-Darling test and, even more so, the information criteria. These distributions showed clear out-of-bounds conditions for some stations. This is due to the fact that for the initial values of the sorted rainfall sample (usually 2 or 3), very low density probabilities were obtained (tending to 0), thus driving the log-Likelihood close to -∞ and the ICs to +∞.

In theory, the criteria for selecting the distribution can be expanded. Other distributions, such as the less commonly used Log-Logistic, Pearson Type V (2 and 3 parameters), Chi, Inverse Chi, and Kappa (K3-Generalized Gumbel and Park), among many others, can be applied. Given their complexities, as demonstrated by C.G. Anghel et al., new approximations for the parameters were applied using L-Moments. These methods were employed to fit some of the previously mentioned distributions to the discharge data from the Siret Basin (Romania). [48] [49].

Some authors argue that the potential of extreme value distributions is often not fully utilized. Variants of the GEV, such as the Fréchet and Weibull distributions, could be applied. The same can be said for the TCEV (Two Component Extreme Values) model, can treat the dry and rainy seasons as two distinct components for the Peak Over Threshold approach. For this paper, that approach was omitted, but its inclusion is warranted.

Under specific conditions, an alternative to the Anderson-Darling and Kolmogorov-Smirnov tests could be Zhang's tests [50]. Their values, based on likelihood ratios, are distribution-free and can be applied in a similar manner to the traditional tests, but they are even more sensitive to extreme probabilities, such as 0 or 1. Unlike the Kolmogorov-Smirnov test, which is more sensitive to the maximum deviation, Zhang's tests distribute sensitivity across the entire dataset [50]. Other statistical tests for better comparison could also be applied, such as the simple Maximum Value test [51], Chi-Pearson test [52], or the Cramér-von Mises criterion [53] [54].

All the aforementioned methods can be applied to identify a better candidate distribution, or multiple distributions, to fit the empirical sample from all stations, particularly the Kossogoudou station.

Considering the confidence intervals, they typically decrease for rainfall with a return period greater than 10 years (see from Figure 17 to Figure 21). This is because, even though the datasets span multiple decades for most stations, the Annual Maxima method excludes significant rainfall extremes that cannot be observed within this framework. This limitation underscores the importance of adopting the Peak Over Threshold (POT) method, which allows for the inclusion of multiple extreme events within the same year, thereby enriching the already limited dataset.

An initial attempt was made to isolate different rainfall extremes based on various day windows. However, since the rainfall season does not exhibit a clear delineation, determining appropriate separation periods between extremes is highly complex to implement. This can be clearly seen through the monthly cumulative rainfall plots (see Annex). Additionally, defining the threshold requires careful consideration. The fixed threshold initially adopted was too low, leading to challenges in data quality. A dynamic approach, such as using a moving average, could provide a more robust solution.

For the region, the main challenge lies in the reliability of daily rainfall series from the gauges. The dataset contains evident gaps due to malfunctions or other issues, and some years had to be discarded (see Table 1 and Figure 3), which could have otherwise enhanced the fitting process. For future applications, the data has been better organized; however, further efforts are necessary to install additional rainfall gauges to more accurately quantify rainfall at a regional level.

5.3: DDF & IDF Curves

It was previously demonstrated how to compute average Depth-Duration-Frequency (DDF) and Intensity-Duration-Frequency (IDF) curves across multiple-day windows using the station dataset. The same methodology can be applied to compute these curves for various distributions, different return periods, and multiple rainfall durations measured in days (see from Table 34 to Table 39). While these analyses are typically conducted in terms of hours, the lack of sufficient data from rainfall gauges recording at 15-minute intervals necessitated this approach. The limited availability and insufficient record lengths of these gauges made an Annual Maxima analysis unfeasible.

Nevertheless, IDF curves remain a valuable tool. They can be utilized to create hyetographs, which describe variations in rainfall intensity over time. Using DDF curves and methods like the Chicago method, block rainfall series can be generated with an initial peak, final peak, or an intermediate peak, ensuring that the area under thecon-sidered IDF curve remains constant [55].

These constructed hyetographs can then serve as inputs for hydrograph computation using various methods based on the inflow-outflow of surface water. This allows for flood simulations using software such as HEC-RAS and HEC-HMS.

Moreover, the DDF and IDF curves can act as references for future years, enabling comparisons of cumulative rainfall over 1 to 10 days to determine the corresponding return period (T) to which they align Additionally, they can be integrated into a meteorological warning system.

5.4: Spatial interpolation

As mentioned before, due to the small number of stations and their irregular spacing, traditional interpolation techniques like Kriging or IDW are not viable. The dataset of the used stations shows very low values of spatial correlation. Only a small number of Ordinary Kriging attempts produced acceptable results, and these were based on the cumulative monthly rainfalls of specific years.

Even simple interpolation methods, such as Spline (regularized or tensioned), which could have been used for moments, quantiles, and growth factors, yielded unreliable results, with some values turning negative. The viable alternative to obtain at least a visual understanding of how rainfall-related quantities vary across the Sirba Basin is the use of Thiessen Polygons. While these do not constitute operational results, they were nevertheless implemented (see Annex section: Basic spatial interpolation: Thiessen Polygons).

To work solely with the stations, as mentioned earlier, their number must be increased to achieve acceptable values of spatial correlation. This is especially crucial for applying the procedure, and even so the usage of specialized semi-variance estimators, such as Cressie-Hawkins, Dowd, or Genton, which account for outliers and skewness. Matheron's estimator alone is insufficient for analyzing datasets with high outliers and skewness, such as the one under investigation. The few good results were achieved using the Cressie-Hawkins or Dowd estimators, which at least demonstrate that commonly used semi-variance estimators are not suitable in this context.

A viable alternative is the use of grid data, such as ERA5, which is more regular and offers a certain degree of spatial correlation. However, as a reanalysis dataset, ERA5 differs from station data, showing lower averages for the AM series (even when adopting a smaller time span, e.g., from 1960 or 1980). This discrepancy arises because ERA5 data are computed based on various variables, such as pressure and temperature, and are time-averaged. Max. of the AM series differ, with the grid showing more values in the lower interquartiles and fewer outliers compared to the station data (see Figure 40).

For this reason, the grid data had to be corrected. Although the correction showed poor R^2 values, it produced some comparable results in terms of median and mean rainfall heights when referenced against station data, and the error reductions were significant. Quantiles calculated from the corrected ERA5 grid were comparable to those computed from the stations. As expected, the station-derived quantiles exhibited more outliers due to the dataset's initial characteristics. Nonetheless, the two quantile series displayed a degree of similarity between the lower and higher interquartiles (25% and 75%) in their computed box plots (see Figure 59, right side).

The adopted correction method was straightforward. By associating each station with its nearest grid point and applying a scaling ratio to the grid data, the cumulative distribution function (CDF) of each grid point was adjusted toward higher rainfall values to better align with the corresponding station's distribution.

However, this basic correction alone is insufficient. More refined adjustments can be implemented, such as considering the four grid points surrounding each station and calculating a weighted average, or directly modifying the CDF through quantile correction between the grid and station distributions. This approach allows for targeted corrections, addressing probability differences for middle or tails values of the distribution.

Additionally, as a precaution, the reliability of the station series can be assessed through outlier analysis methods to ensure robust corrections [51].

Other grid datasets, such as CHIRPS and TRMM, could also be utilized.

CHIRPS (Climate Hazards Group Infrared Precipitation with Stations) is a global rainfall dataset designed to monitor rainfall trends and support hydrological studies, particularly in regions with limited ground data. It provides daily, 5-day, or monthly precipitation data from 1981 to the present. CHIRPS offers a finer spatial resolution compared to ERA5, at 0.05° (~5 km at the equator). The dataset combines satellite-derived Cold Cloud Duration (CCD) imagery with station observations, resulting in a comprehensive resource that is particularly valuable for drought monitoring [56].

TRMM (Tropical Rainfall Measuring Mission) was a NASA-JAXA satellite mission aimed at measuring tropical and subtropical rainfall to enhance understanding of climate cycles and patterns. The mission operated from 1997 to 2015, covering latitudes from 50°N to 50°S. Its resolution varied from ~4 km to 25 km depending on the instruments used. TRMM employed various sensors, such as the Precipitation Radar (PR) and the TRMM Microwave Imager (TMI), to measure rainfall intensity and spatial distribution effectively [57].

Both of these datasets are available in .csv or NetCDF formats (.nc) and can be analyzed in the same manner as the ERA5 dataset, with various coding languages and GIS software.

Ultimately, the simple correction demonstrates that adjusting the grid dataset using station data as a reference is a viable approach. This method was employed to generate operational rainfall height maps for different distributions and return periods (see from Figure 56 to Figure 58).

In this paper, spatial interpolation is presented as a method for rainfall regionalization, also referred to as regional frequency analysis. These methods are generally divided into two categories:

- Regional methods:
    - Fixed regions;
    - Cluster regions, based on climatic characteristics in the parameter space rather than the geographic one;
    - Regions of Influence (ROI), where each sample point has its own ROI.
- Regionless methods:
    - Based on spatial interpolation;
    - Based on geostatistical techniques.

Regionless methods are often preferred because they avoid creating artificial borders for regions. As mentioned earlier, spatial interpolation methods encompass various techniques, such as Thiessen polygons, Inverse Distance Weighting (IDW), Spline, Geo-regressions, and different variants of Kriging.

Italy has implemented regionalization methods for discharge analysis, such as VAPI (Valutazione delle Piene) [58] and ARPIEM (Analisi Regionale delle Piene nei bacini montani) [59]. Each of these methods has its strengths and weaknesses, and similar approaches could be adapted to model rainfall across large areas, such as the Sirba Basin.

These approaches are primarily based on the 1997 work of Hosking and Wallis on regional frequency analysis [24], as well as the index flood method initially proposed by Dalrymple [37]. The main objective is to create homogeneous regions where the Growth Factor K(T) is constant. In these regions, the frequency distribution remains invariant, except for a site-dependent scale factor, the index value (which can represent either rainfall or discharge). Ultimately, each region will have its own K(T) curve.

A key challenge, as discussed earlier, is the short duration of records at many stations, leading to very short Annual Maximum (AM) series. In contrast, Italy benefits from a larger network of rainfall gauges, many with extensive historical records dating back to 1900. This allows for more reliable application of the initial Homogeneity Test proposed by Hosking and Wallis. However, in the Sirba Basin, the same test is less dependable due to the high number of discarded years and the brevity of available series at many stations [24].

Further efforts are necessary to enhance regional frequency analysis not only for the Sirba Basin but also for the entire territories of Burkina Faso and Niger, focusing on regional approaches rather than relying solely on spatial interpolation.

Determining a reliable probability curve based solely on station data is only feasible when the observed series length (N) is sufficiently long. For example, Benson suggests that $N>T/2$ [60], where T is the return period. Similarly, the *Flood Estimation Handbook* (FEH) recommends $N>5T$ [61].

In the analyzed context, these conditions are rarely met. For instance, the Namaro station, with only 17 recorded years, would render return periods exceeding 8.5 years unreliable under Benson's criterion. Even relatively better stations, such as Aribinda and Piela, with 60 years of data, struggle to meet these criteria, as return periods beyond 30 years remain unreliable.

This limitation must always be considered when interpreting distribution fitting (from Table 13 to Table 30, Table 31, and from Figure 17 to Figure 21) and quantile maps (from Figure 56 to Figure 58). Discarding stations with short records has been deemed unfeasible, given the already limited number of stations with usable AM series, further compounded by the large number of inactive rainfall gauges in the region.

6: Conclusions

The objective of this paper was to use the rainfall series from the station, together with grid data, to compute operative products that could be applied to assess pluviometric risk in the Sirba Basin.

The computed quantiles of the Burr Type XII distribution for the stations, along with the DDF and IDF curves, represent valid indicators for rainfall representation, at least at a local level near the locality where the station was placed. As mentioned in the Discussion section, some improvements can be made. Obviously, not only one distribution can be applied to the entire basin, but multiple distributions or even a mixed quantile estimation approach may be adopted.

Due to the spatial and numerical configuration of the stations' dataset, more accurate spatial interpolation methods, such as Ordinary Kriging, could not be adopted to represent moments, distribution quantiles, and growth factors. Nevertheless, simpler methods, such as Thiessen Polygons, were computed to provide at least a limited visual understanding of rainfall quantities in the region.

To address the limitations of the station data, the ERA5 grid dataset was adopted and scaled to match the stations' moments for the Annual Maxima series. The initial error indicators showed poor performance, but the error reduction achieved between the initial grid and the corrected/scaled grid (using the stations as a reference) demonstrated improvements, particularly for the Median and Mean.

The corrected grid's moments were used to compute rainfall heights (quantiles) using the best-fitting distributions for the station data, including Gumbel, GEV, and Burr Type XII (the most suitable). These values were comparable to the rainfall heights observed at the stations, as indicated by raw box-plot comparisons, suggesting that scaling the grid data to align with the station data is a viable approach. However, further improvements could be achieved by employing more precise quantile corrections and incorporating spatial averaging techniques.

Further efforts should be made to implement rainfall frequency analysis for the entire countries of Burkina Faso and Niger, utilizing all available rainfall gauges instead of just those in the Sirba Basin, as an alternative to spatial interpolation.

Overall, this paper highlights both essential practices and potential challenges in difficult contexts such as this, where station data is derived from sparse pluviometers installed over different decades and are often characterized by interruptions and unreliable years.

7: Bibliography

[1] Institute of BioEconomy, "ANADIA Project", National Research Council of Italy, link: https://www.ibe.cnr.it/progetti/anadia/;

[2] P. Tamagnone, G. Massazza, A. Pezzoli, and M. Rosso, "Hydrology of the Sirba River: Updating and Analysis of Discharge Time Series", *Water*, vol. 11, no. 1, pp. 1-3, Jan. 2019, doi: https://doi.org/10.3390/w11010156;

[3] G. Massazza, P. Tamagnone, C. Wilcox, E. Belcore, A. Pezzoli, T. Vischel, G. Panthou, M. H. Ibrahim, M. Tiepolo, V. Tarchiani, and M. Rosso, "Flood hazard scenarios of the Sirba River (Niger): Evaluation of the hazard thresholds and flooding areas", *Water*, vol. 11, no. 5, pp. 1-4, May 2019, doi: https://doi.org/10.3390/w11051018;

[4] L. Descroix, I. Bouzou, P. Genthon, D. Sighomnou, G. Mahe, I. Mamadou, J.-P. Vandervaere, E. Gautier, O. Faran, J.-L. Rajot, and others, "Impact of drought and land-use changes on surface-water quality and quantity: The Sahelian paradox", *Curr. Perspect. Contam. Hydrol. Water Resour. Sustain.*, 2013, doi: http://dx.doi.org/10.5772/54536;

[5] L. Descroix, *Processus et enjeux d'eau en Afrique de l'Ouest soudano-sahélienne*, Editions des Archives Contemporaines, France, 2018. ISBN 978-2-8130-0314-0;

[6] L. Descroix, P. Genthon, O. Amogu, J. L. Rajot, D. Sighomnou, and M. Vauclin, "Change in Sahelian rivers hydrograph: The case of recent red floods of the Niger River in the Niamey region", *Glob. Planet. Change*, vol. 98-99, pp. 18-30, 2012, doi: https://doi.org/10.1016/j.gloplacha.2012.07.009;

[7] V. Aich, B. Koné, F. F. Hattermann, and E. N. Müller, "Floods in the Niger basin — analysis and attribution", *Nat. Hazards Earth Syst. Sci. Discuss.*, vol. 2, pp. 5171–5212, 2014, doi: https://doi.org/10.5194/nhessd-2-5171-2014;

[8] Y. S. C. Somé, L. Dango, and G. D. Abdouramane, "Arid Climate in the Sirba Basin in Burkina Faso: Causes for Better Decisions in Land Use Planning", Curr. J. Appl. Sci. Technol., vol. 14, no. 6, pp. 1–8, 2016, doi: https://doi.org/10.9734/BJAST/2016/24098.

[9] T. Kandji, L. Verchot, and J. Mackensen, "Climate Change and Variability in the Sahel Region: Impacts and Adaptation Strategies in the Agricultural Sector", Environment, World Agroforestry Centre, pp. 2–4, 2010, link: https://www.researchgate.net/publication/264626384_Climate_Change_and_Variability_in_the_Sahel_Region_Impacts_and_Adaptation_Strategies_in_the_Agricultural_Sector;

[10] City Population, "Population of Cities in Burkina Faso", link: https://www.citypopulation.de/en/burkinafaso/cities/;

[11] City Population, "Population of Cities in Niger", link: https://www.citypopulation.de/en/niger/;

[12] Africa Soil Information Service (AfSIS): Hydrologically Corrected/Adjusted SRTM DEM (AfrHySRTM), link: https://cmr.earthdata.nasa.gov/search/concepts/C1214155420-SCIOPS;

[13] Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023): ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), doi: https://doi.org/10.24381/cds.adbb2d47;

[14] Instituto di BioEconomia, Consiglio Nazionale delle Ricerche, Niger River link: http://sdicatalog.fi.ibimet.cnr.it/dataset/niger-river;

[15] Instituto di BioEconomia, Consiglio Nazionale delle Ricerche, Niger Basin, link: http://sdicatalog.fi.ibimet.cnr.it/dataset/niger-basin;

[16] "River Niger Tributaries from Burkina Faso", Wikipedia, Feb. 28, 2021, link: https://en.wikipedia.org/wiki/Sirba_River#/media/File:River_Niger_Tributaries_from_Burkina_Faso_OSM.png;

[17] M. Tiepolo, M. Rosso, G. Massazza, E. Belcore, S. Issa, and S. Braccio, "Flood Assessment for Risk-Informed Planning along the Sirba River, Niger", Sustainability, vol. 11, no. 15, pp. 1–18, 2019,. doi: https://doi.org/10.3390/su11154003;

[18] J. R. Hosking, "L-moments: Analysis and estimation of distributions using linear combinations of order statistics", *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 52, no. 1, pp. 105-124, 1990, link: http://www.jstor.org/stable/2345653?origin=JSTOR-pdf;

[19] I. Vladimiriescu, Bazele hidrologiei tehnice, pp. 54–57, Bucharest, Romania: Editura Tehnica, 1984.

[20] C. J. Van Eijk and E. F. J. B. L. van Gelder, "Plotting Positions in Extreme Value Analysis", Journal of Hydrology, vol. 270, no. 1–4, pp. 157–170, 2003, doi: https://doi.org/10.1175/JAM2349.1;

[21] O. B. Adeboye and M. O. Alatise, "Performance of Probability Distributions and Plotting Positions in Estimating the Flood of River Osun at Apoje Sub-basin, Nigeria", Agricultural Engineering International: the CIGR Ejournal, vol. 9, pp. 1-21, July 2007, doi: http://doi.org/10.13031/2013.786;

[22] R. K. Deo, B. L. Sinha, and K. K. Sharma, "Ranking of Plotting Position Formula in Frequency Analysis of Annual and Seasonal Rainfall at Gariyaband, Chhattisgarh", Journal of Water Resources and Ocean Science, vol. 13, no. 1, pp. 1-5, 2024; doi: https://doi.org/10.11648/j.wros.20241301.11;

[23] E. Todini and M. Franchini, Glossario Internazionale di Idrologia, Istituto di Costruzioni Idrauliche, 1986, link: https://hdl.handle.net/11392/519322;

[24] J. Hosking and J. Wallis, Regional Frequency Analysis: An Approach Based on L-Moments, Cambridge University Press, 1997, doi: http://doi.org/10.1017/CBO9780511529443;

[25] V. W. Griffis and J. R. Stedinger, "The log-Pearson type 3 distribution and its application in flood frequency analysis: 1. Distribution characteristics", J. Hydrol. Eng., vol. 12, no. 5, pp. 482–491, 2007, doi: http://doi.org/10.1061/(ASCE)1084-0699(2007)12:5(482);

[26] G. Talbot, "Extreme values statistical analysis library", MATLAB Central File Exchange, 2024, [Online], link: https://www.mathworks.com/matlabcentral/fileexchange/93075-extreme-values-statistical-analysis-library;

[27] D. Ganora and F. Laio, "Hydrological Applications of the Burr Distribution: Practical Method for Parameter Estimation", Journal of Hydrologic Engineering, vol. 20, no. 11, pp. 04015024 (1 to 8), 2015, doi: https://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0001203;

[28] A. Kolmogorov, "Sulla determinazione empirica di una legge di distribuzione", Giornale dell'Istituto Italiano degli Attuari, vol. 4, pp. 83–91, 1933, link: https://www.researchgate.net/post/Kolmogorovs_original_paper_Sulla_determinazione_empirica_di_una_legge_di_distribuzione_from_1933;

[29] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions", Annals of Mathematical Statistics, vol. 19, no. 2, pp. 279–281, June 1948, doi: https://doi.org/10.1214/aoms/1177730256;

[30] T. W. Anderson and D. A. Darling, "A test of goodness of fit", Journal of the American Statistical Association, vol. 49, no. 268, pp. 765–769, 1954, doi: https://doi.org/10.1080/01621459.1954.10501232;

[31] H. Akaike, "A new look at the statistical model identification", IEEE Transactions on Automatic Control, vol. 19, no. 6, pp. 716–723, Dec. 1974, doi: https://doi.org/10.1109/TAC.1974.1100705;

[32] G. Schwarz, "Estimating the dimension of a model", Annals of Statistics, vol. 6, no. 2, pp. 461–464, 1978, doi: https://doi.org/10.1214/aos/1176344136;

[33] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules", IEEE Signal Processing Magazine, vol. 21, no. 4, pp. 36–47, July 2004, doi: https://doi.org/10.1109/MSP.2004.1311138;

[34] K. P. Burnham and D. R. Anderson, Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd ed. New York, NY, USA: Springer, 2002, doi: https://doi.org/10.1007/b97636;

[35] R. E. Horton, Drainage Basin Characteristics, Trans. Am. Geophys. Union, vol. 13, pp. 350–361, 1932, doi: https://doi.org/10.1029/TR013i001p00350;

[36] U.S. Department of Commerce, U.S. Weather Bureau, Rainfall Intensity-Duration-Frequency Curves for the United States for Durations from 5 Minutes to 24 Hours and Return Periods from 1 to 100 Years, Technical Paper no. 25, Washington, D.C.: U.S. Government Printing Office, 1955, link: https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.weather.gov/media/owp/oh/hdsc/docs/TP25.pdf&ved=2ahUKEwi3_o7ck7aJAxX42wIHHVqgMBgQFnoECBkQAQ&usg=AOvVaw1oxXJXM1yqe-SeBoVDWQp1s;

[37] T. Dalrymple, Flood-Frequency Analyses, U.S. Geological Survey Water-Supply Paper 1543-A, Washington, D.C.: U.S. Government Printing Office, 1960, doi: https://doi.org/10.3133/wsp1543A;

[38] A. H. Thiessen, "Precipitation averages for large areas", Monthly Weather Review, American Meteorological Society, vol. 39, no. 7, pp. 1082–1089, July 1911, doi: https://doi.org/10.1175/1520-0493(1911)39%3C1082b:PAFLA%3E2.0.CO;2;

[39] Environmental Systems Research Institute (ESRI), "Create Thiessen Polygons", ArcGIS Pro, [Online], link: https://pro.arcgis.com/en/pro-app/latest/tool-reference/analysis/create-thiessen-polygons.htm;

[40] Environmental Systems Research Institute (ESRI), "How Kriging works", ArcGIS Desktop Help: 3D Analyst Toolbox, [Online], link: https://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/how-kriging-works.htm#ESRI_SECTION1_DD457E27190A4CD0BC1D7A9D7BA5DD5F;

[41] Environmental Systems Research Institute (ESRI), "Get started with Geostatistical Analyst in ArcGIS Pro", ArcGIS Pro Help: Geostatistical Analyst, [Online], link: https://pro.arcgis.com/en/pro-app/latest/help/analysis/geostatistical-analyst/get-started-with-geostatistical-analyst-in-arcgis-pro.htm;

[42] Environmental Systems Research Institute (ESRI), "Understanding ordinary Kriging", ArcGIS Pro Help: Geostatistical Analyst, [Online], link: https://pro.arcgis.com/en/pro-app/latest/help/analysis/geostatistical-analyst/understanding-ordinary-kriging.htm;

[43] M. A. Oliver and R. Webster, "Basic Steps in Geostatistics: The Variogram and Kriging", SpringerBriefs in Agriculture, Springer, Cham, 2015, doi: https://doi.org/10.1007/978-3-319-15865-5;

[44] D. G. Krige, "A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand", Journal of the Chemical, Metallurgical and Mining Society of South Africa, vol. 52, no. 6, pp. 119–139, 1951, link: https://journals.co.za/doi/abs/10.10520/AJA0038223X_4792;

[45] G. Matheron, "Principles of Geostatistics", Economic Geology, vol. 58, no. 8, pp. 1246–1266, 1963, doi: https://doi.org/10.2113/gsecongeo.58.8.1246;

[46] P. Tamagnone, E. Comino, and M. Rosso, "Rainwater harvesting techniques as an adaptation strategy for flood mitigation", Journal of Hydrology, vol. 586, p. 124880, 2020, doi: https://doi.org/10.1016/j.jhydrol.2020.124880;

[47] P. Tamagnone, Progress in planning mitigation and adaptation strategies driven by indigenous knowledge and numerical modelling to face hydrometeorological hazards in the Sahel, Ph.D. dissertation, Politecnico di Torino, Italy, 2021, link: https://hdl.handle.net/11583/2912982;

[48] C. G. Anghel and C. Ilinca, "Parameter estimation for some probability distributions used in Hydrology", Appl. Sci., vol. 12, no. 24, p. 12588, 2022, doi: https://doi.org/10.3390/app122412588;

[49] C. G. Anghel, S. C. Stanca, and C. Ilinca, "Two-parameter probability distributions: Methods, techniques and comparative analysis", Water, vol. 15, no. 19, p. 3435, 2023, doi: https://doi.org/10.3390/w15193435;

[50] J. Zhang, "Powerful Goodness-of-fit Tests Based on the Likelihood Ratio", Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 64, no. 2, pp. 281–294, May 2002. doi: https://doi.org/10.1111/1467-9868.00337;

[51] F. Laio, P. Allamano, and P. Claps, "Exploiting the information content of hydrological 'outliers' for goodness-of-fit testing", Hydrology and Earth System Sciences, vol. 14, no. 10, pp. 1909–1917, 2010, doi: https://doi.org/10.5194/hess-14-1909-2010;

[52] K. Pearson, "X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling", The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, vol. 50, no. 302, pp. 157–175, 1900, doi: https://doi.org/10.1080/14786440009463897;

[53] H. Cramér, "On the composition of elementary errors: First paper: Mathematical deductions", Scandinavian Actuarial Journal, vol. 1928, no. 1, pp. 13–74, 1928, doi: https://doi.org/10.1080/03461238.1928.10416862;

[54] R. E. von Mises, Wahrscheinlichkeit, Statistik und Wahrheit. Berlin, Germany: Julius Springer, 1928;

[55] J. S. Keifer and H. H. Chu, “Synthetic Storm Pattern for Drainage Design”, Journal of the Hydraulics Division, vol. 83, no. 4, pp. 1–25, 1957, doi: https://doi.org/10.1061/JYCEAJ.0000104;

[56] Climate Hazards Center, "CHIRPS: Rainfall Estimates from Rain Gauge and Satellite Observations", University of California, Santa Barbara, link: https://www.chc.ucsb.edu/data/chirps;

[57] NASA, "Tropical Rainfall Measuring Mission (TRMM)", Global Precipitation Measurement, link: https://gpm.nasa.gov/missions/trmm;

[58] E. Ferrari and P. Versace, “La valutazione delle piene in Italia: Rapporto Nazionale di Sintesi”, Gruppo Nazionale per la Difesa dalle Catastrofi Idrogeologiche, Linea 1 – Previsione e prevenzione degli eventi idrologici estremi e loro controllo, C.N.R., u.o. 1.15 Dipartimento di Difesa del Suolo, Università della Calabria, Cosenza, Roma, 1994, link: http://www.idrologia.polito.it/gndci/rapportiPdf/Vapi_Nazionale.pdf;

[59] F. Laio, D. Ganora, P. Claps, and G. Galeati, "Spatially smooth regional estimation of the flood frequency curve (with uncertainty)", Journal of Hydrology, vol. 408, no. 1-2, pp. 67-77, 2011, doi: https://doi.org/10.1016/j.jhydrol.2011.07.022;

[60] M. A. Benson, "Uniform flood-frequency estimating methods for federal agencies", *Water Resources Research*, vol. 4, no. 5, pp. 891-908, 1968;

[61] D. Reed and A. Robson, *Flood Estimation Handbook*, vol. 3, Wallingford: Institute of Hydrology, 1999;

8: Annex

Here the plots follow the order of Table 1, regarding the stations.

8.1: Cumulative & Max. Daily Rainfalls by month

Here are the two series of plots, for each station:

- Averages of monthly cumulative rainfalls, across all available years;
- Box-plots of the series of daily maximum rainfall for each month, across all available years.

For the second plots, the yearly max. of daily rainfall for each station, clearly coincides with the ones of Table 10.

For the first two stations, Kogho and Salogo, there are only two available years, thus making them unusable for the AM analysis, and show no whiskers or outliers on the box-plots.

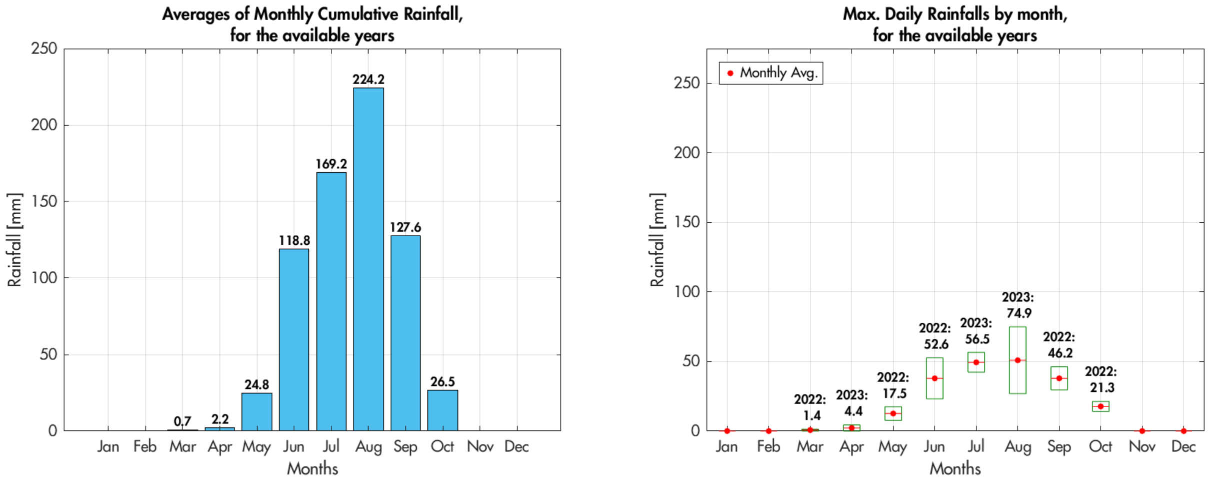


Figure 62: Kogho.

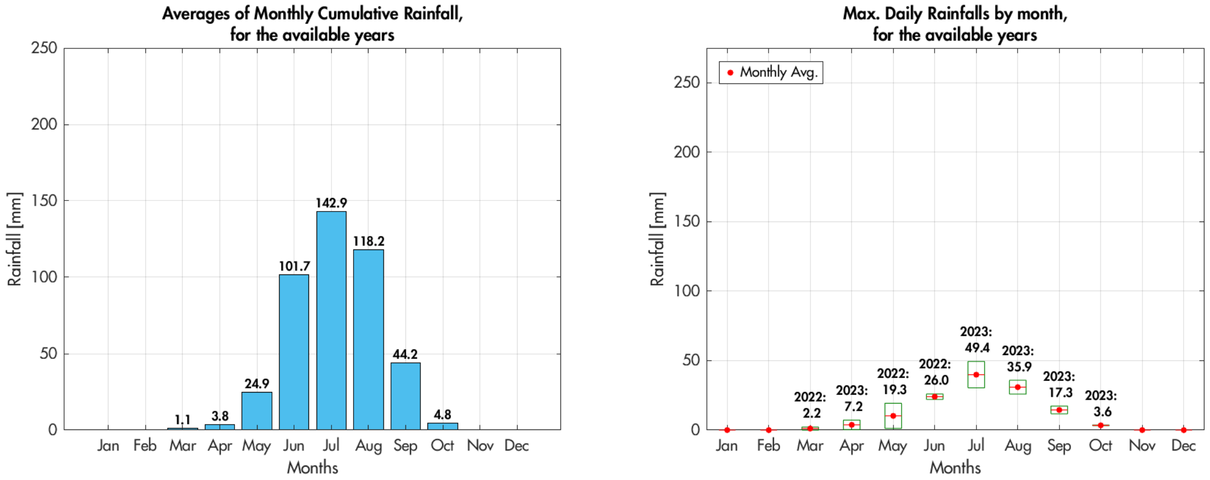


Figure 63: Salogo.

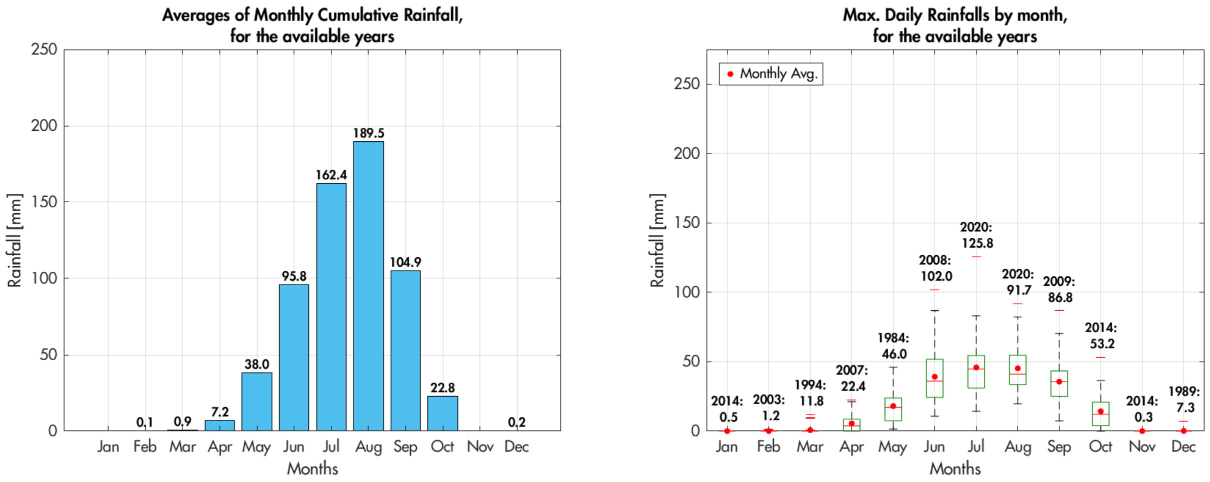


Figure 64: Bogande.

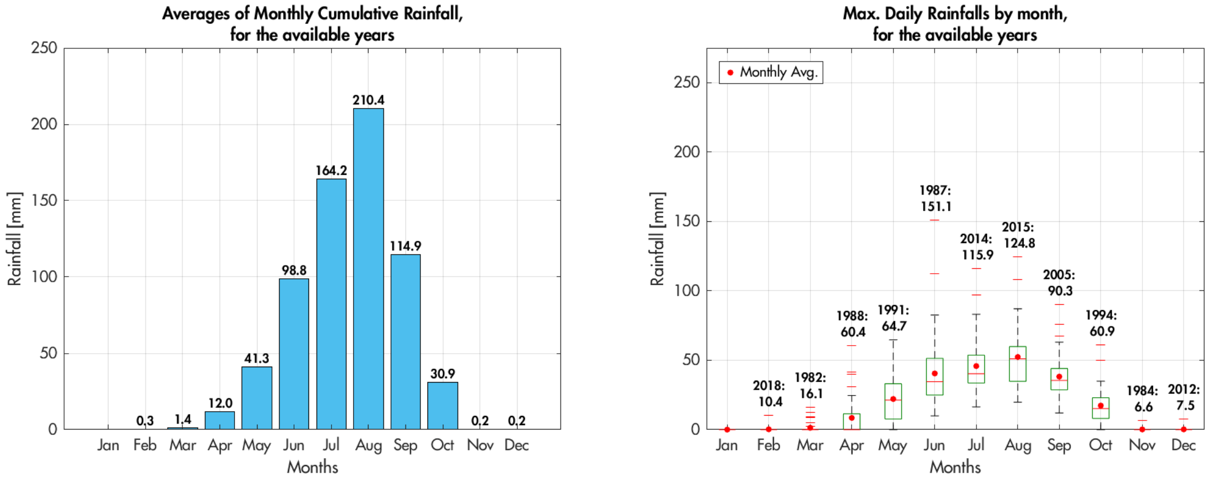


Figure 65: Manni.

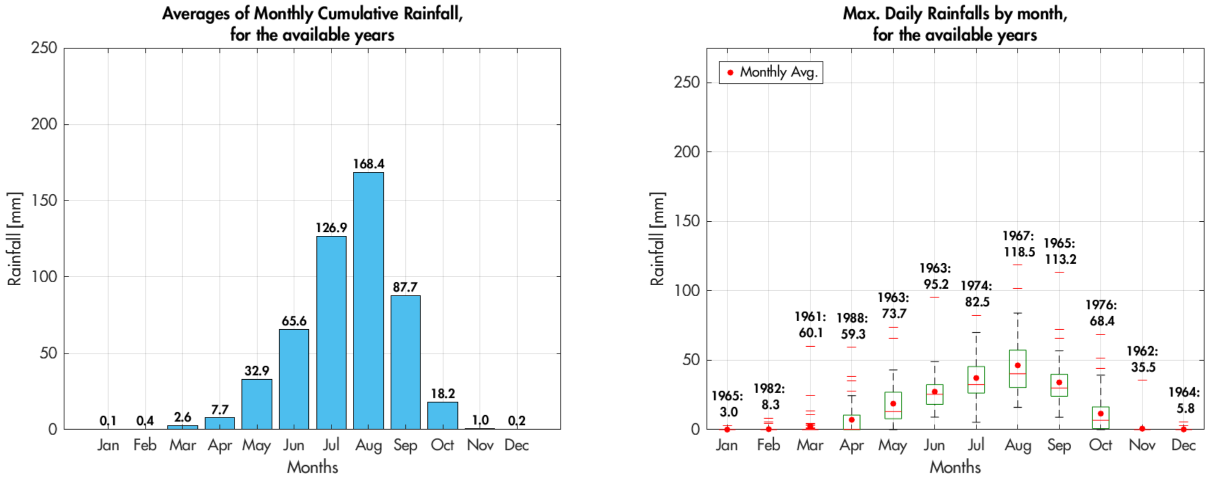


Figure 66: Bani.

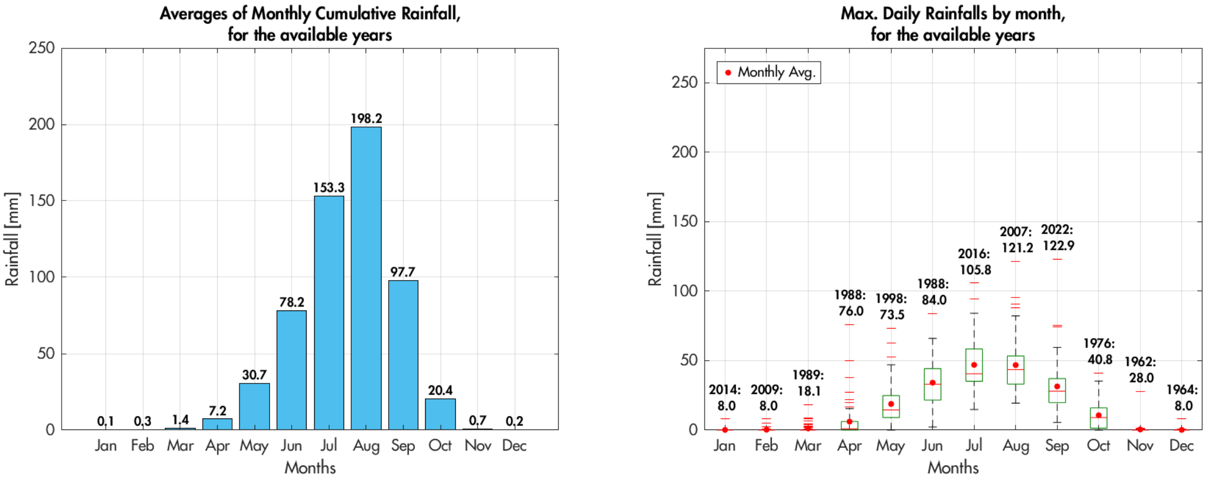


Figure 67: Barsalogho.

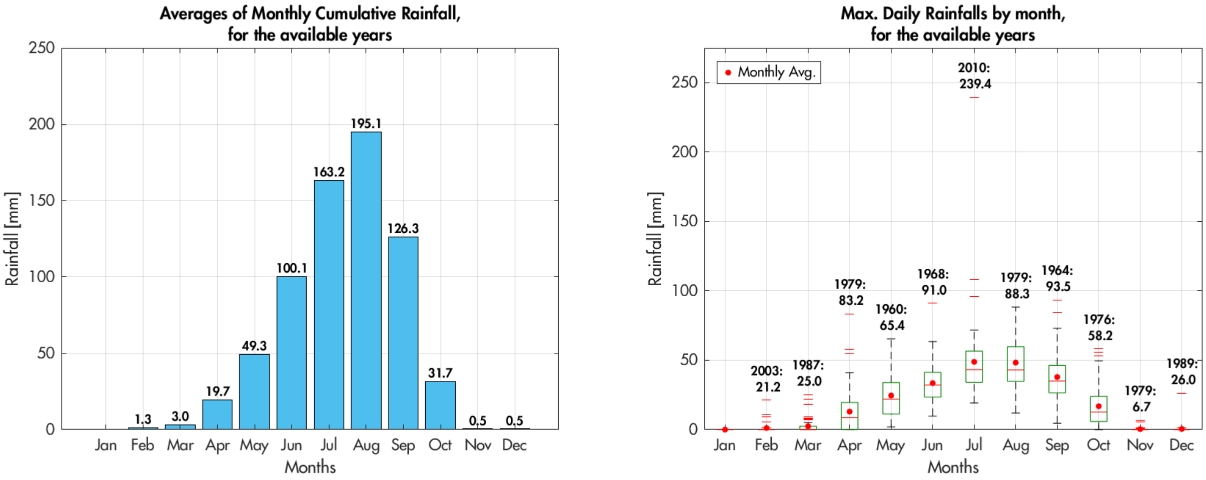


Figure 68: Boulsa.

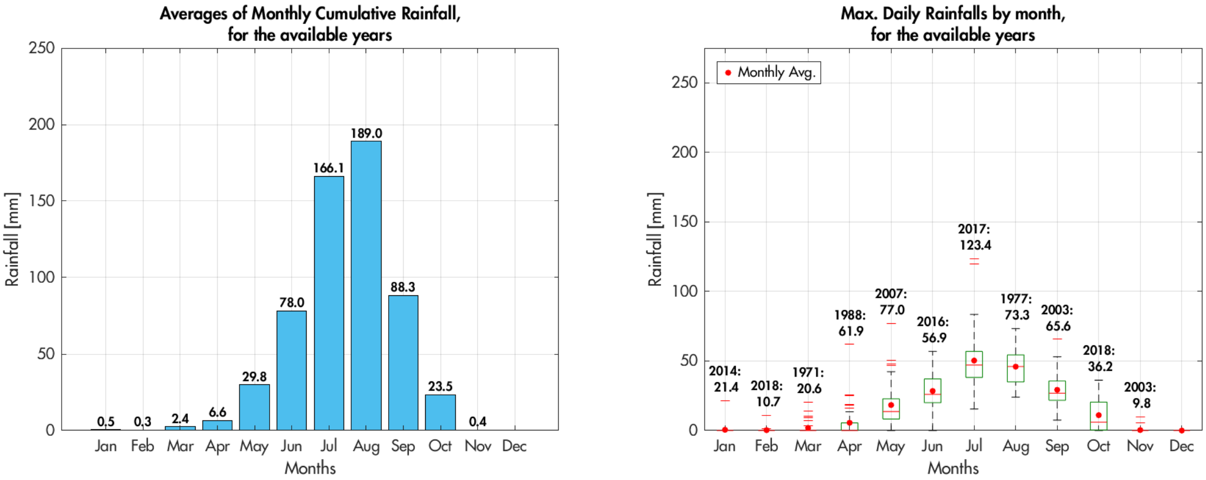


Figure 69: Bouroum.

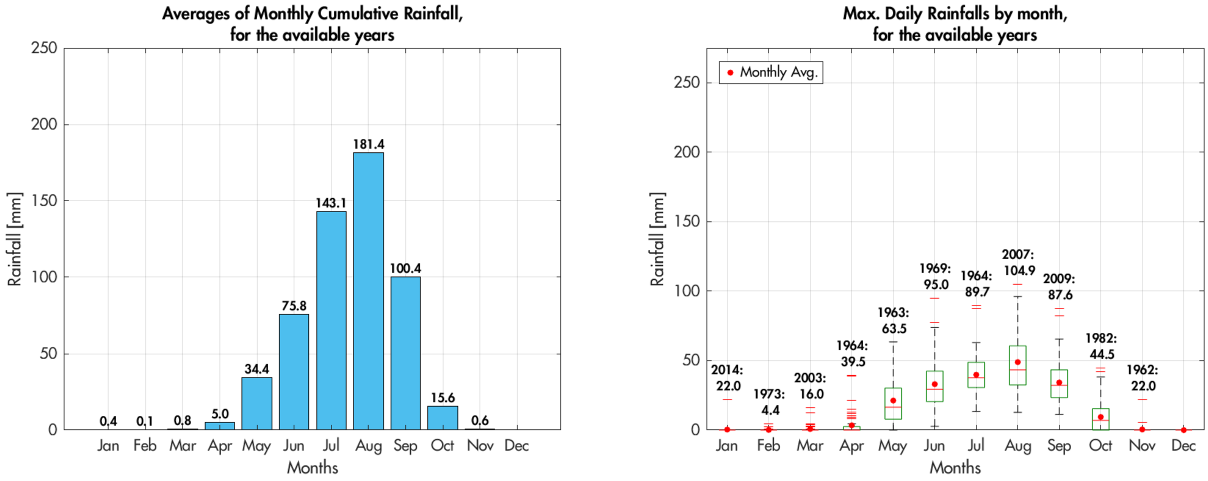


Figure 70: Dakiri.

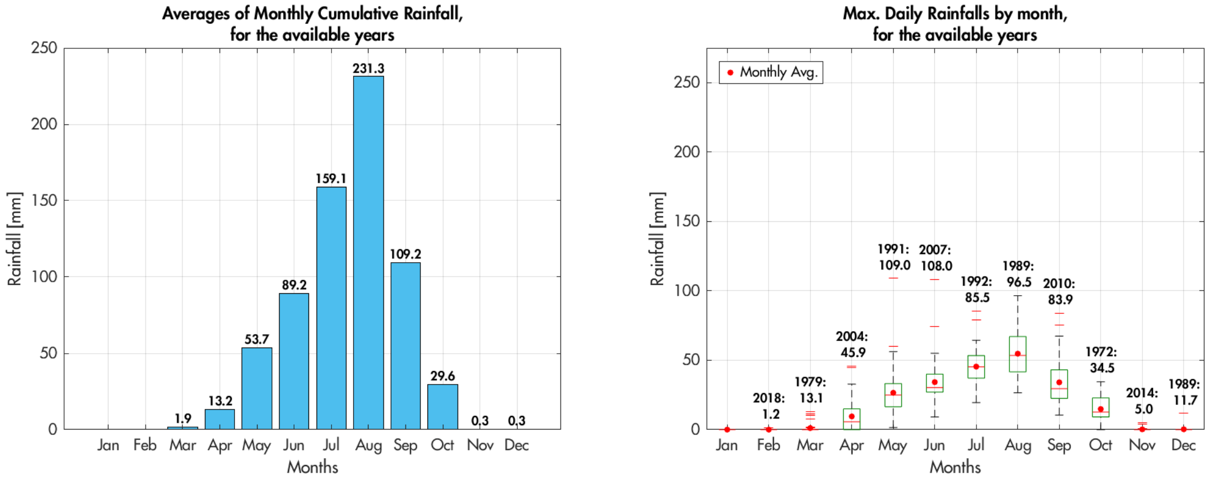


Figure 71: Gayeri.

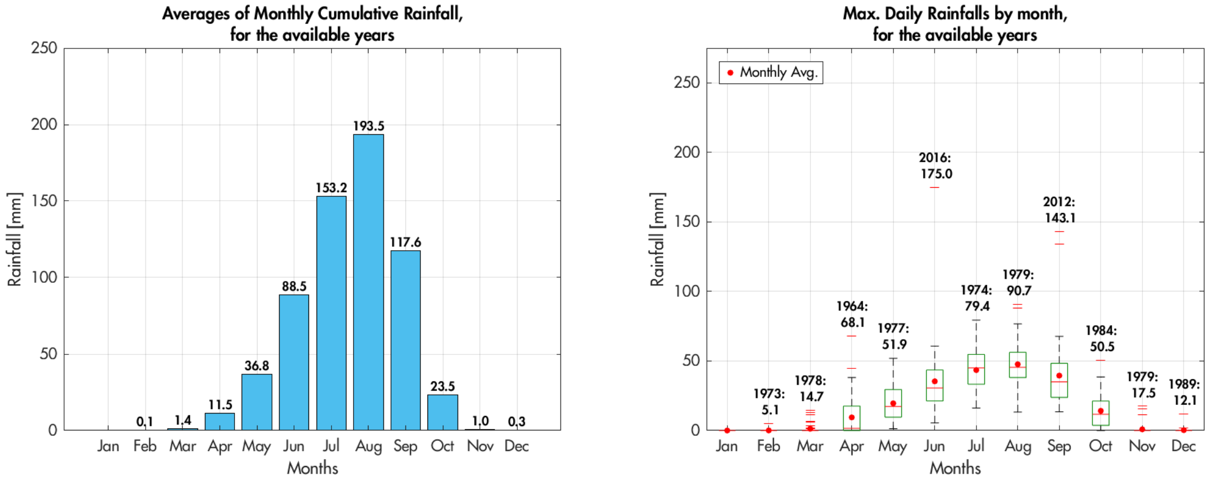


Figure 72: Kossougoudou.

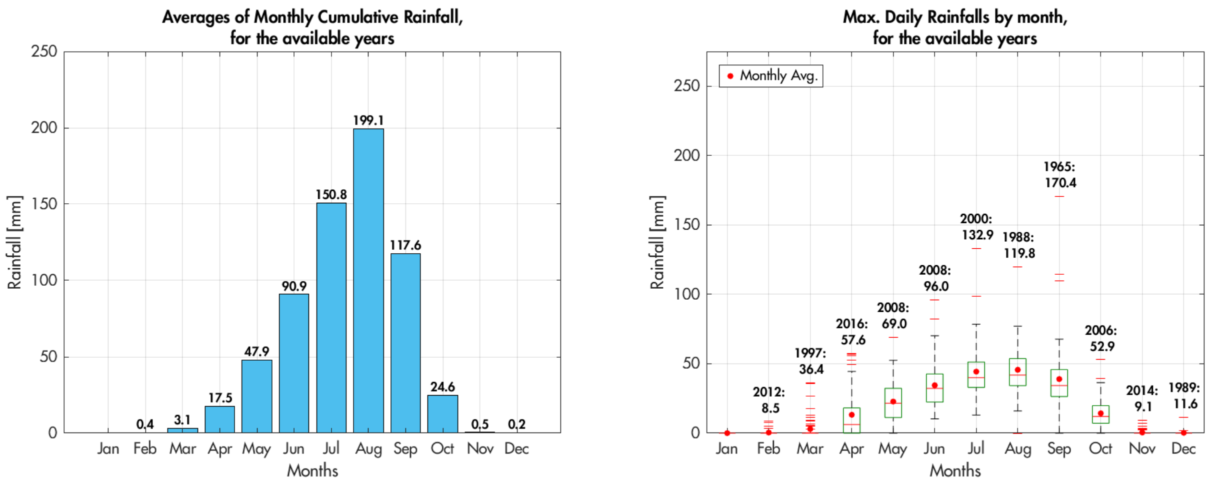


Figure 73: Piela.

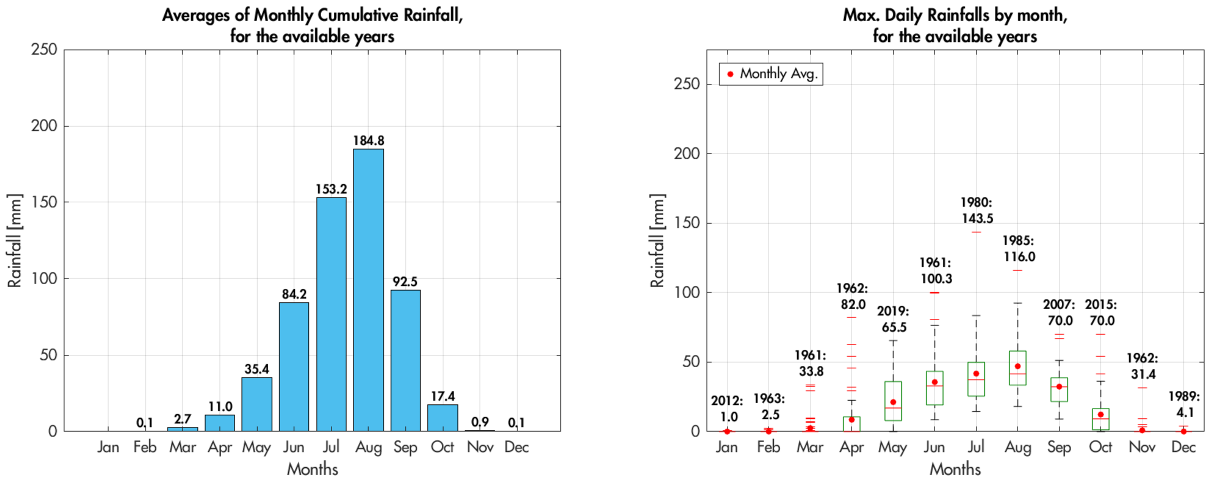


Figure 74: Sebba.

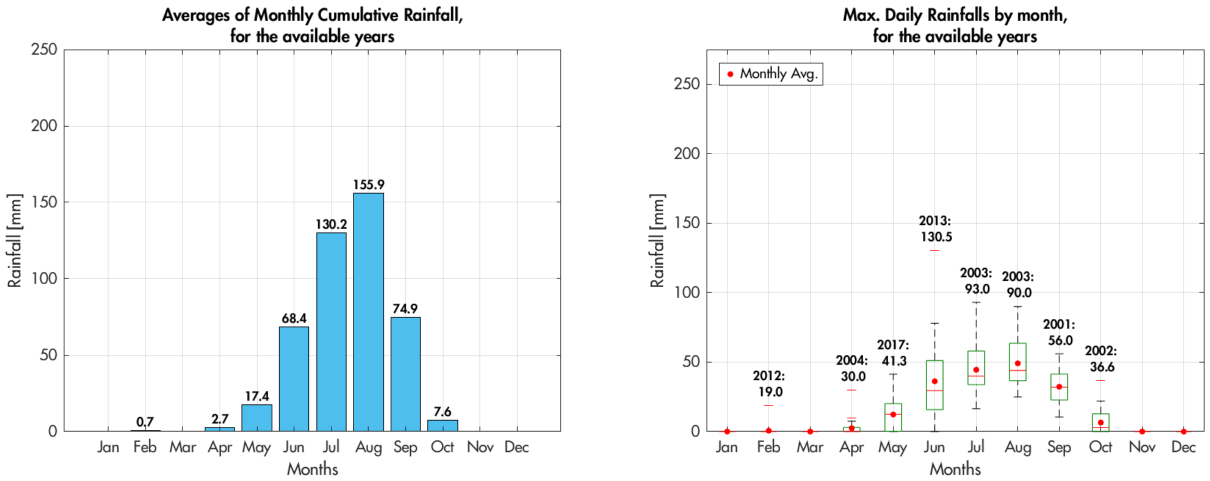


Figure 75: Gotheye.

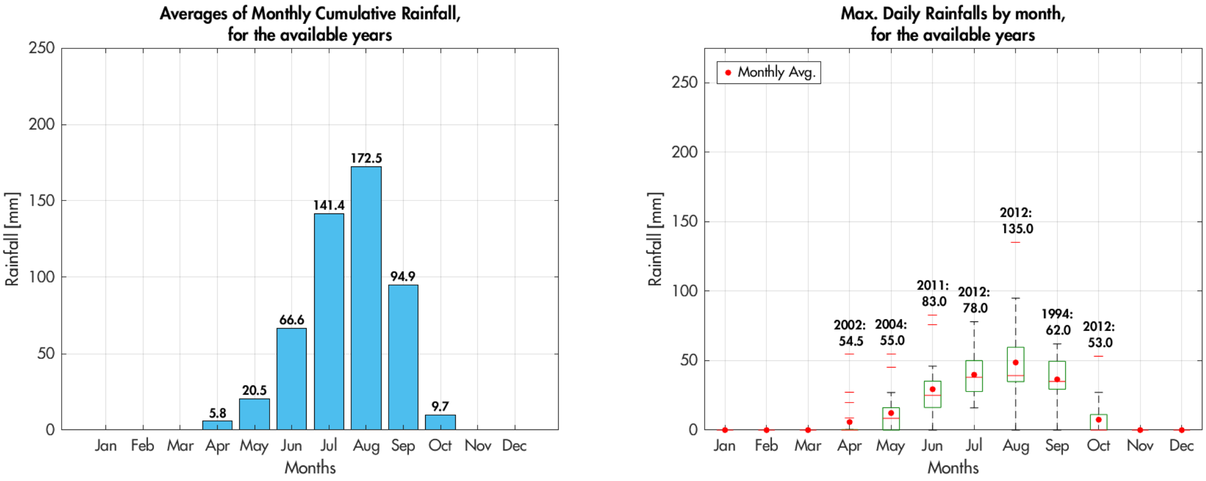


Figure 76: Karma.

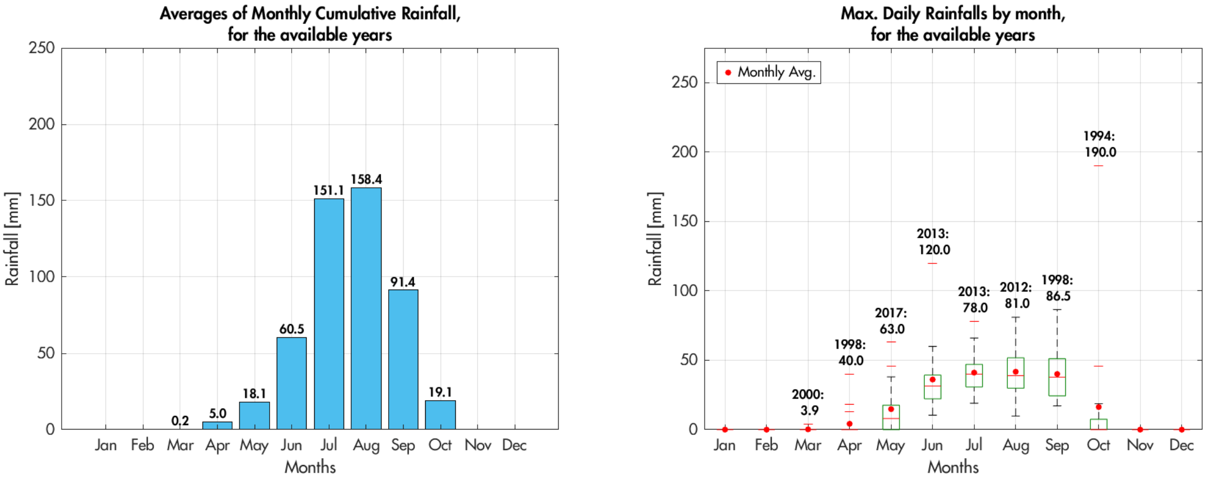


Figure 77: Namaro.

8.2: Yearly Cumulative Rainfalls

The Kogho and Salogo station's plots are omitted since they each contain only two years. For each station the discarded years are indicated in Table 1 and Figure 3.

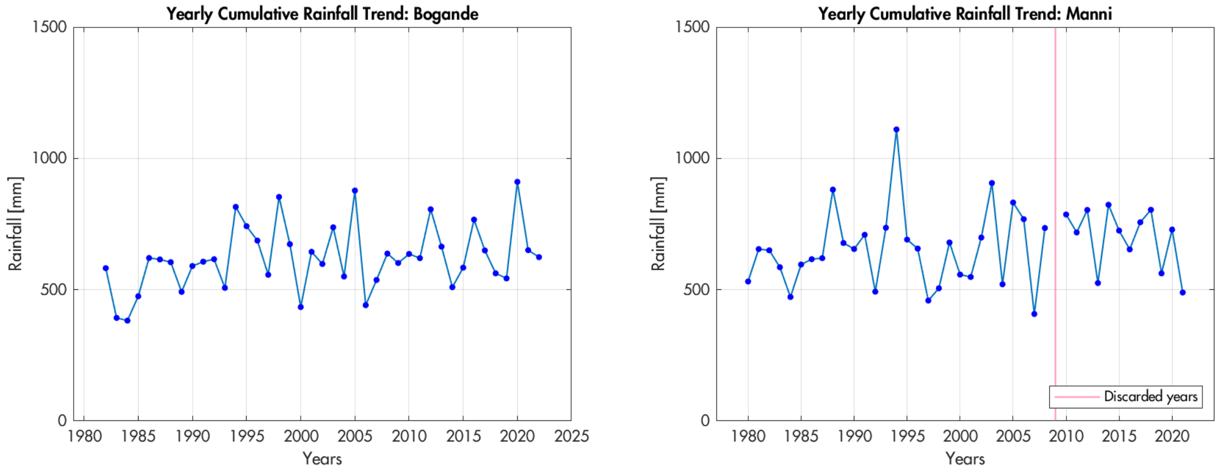


Figure 78: *Yearly cumulative rainfall trends, Bogande and Manni.*

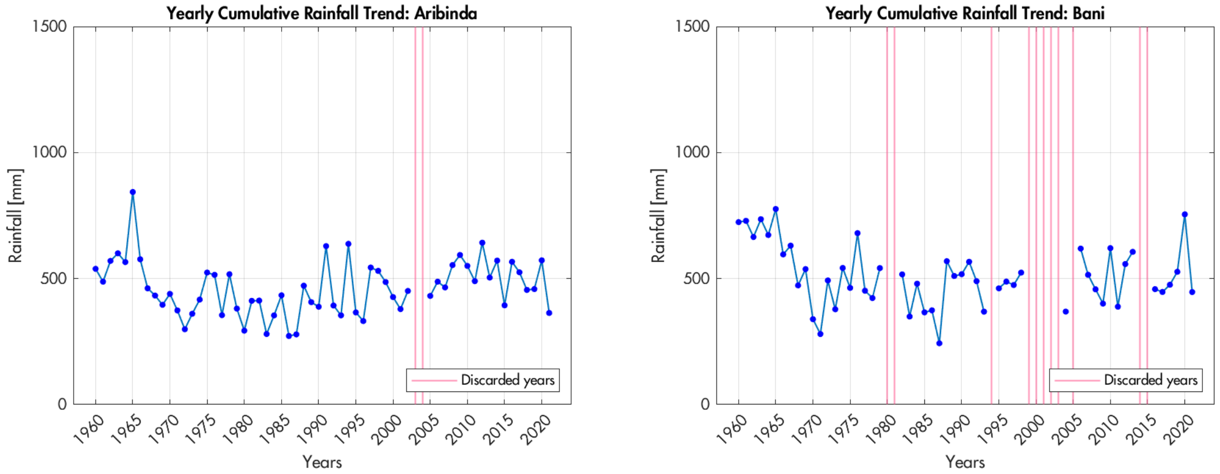


Figure 79: Yearly cumulative rainfall trends, Aribinda and Bani.

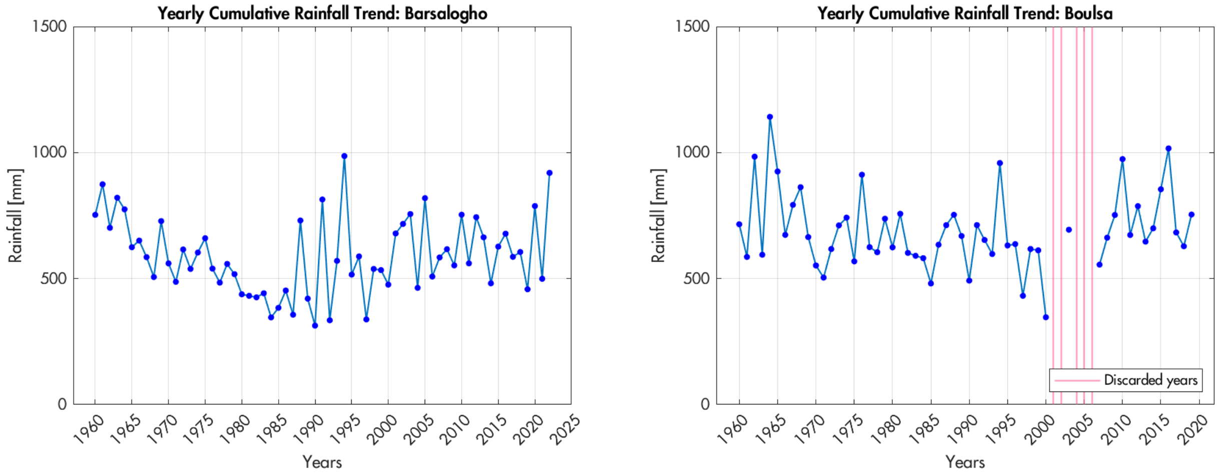


Figure 80: Yearly cumulative rainfall trends, Barsalogho and Boulsa.

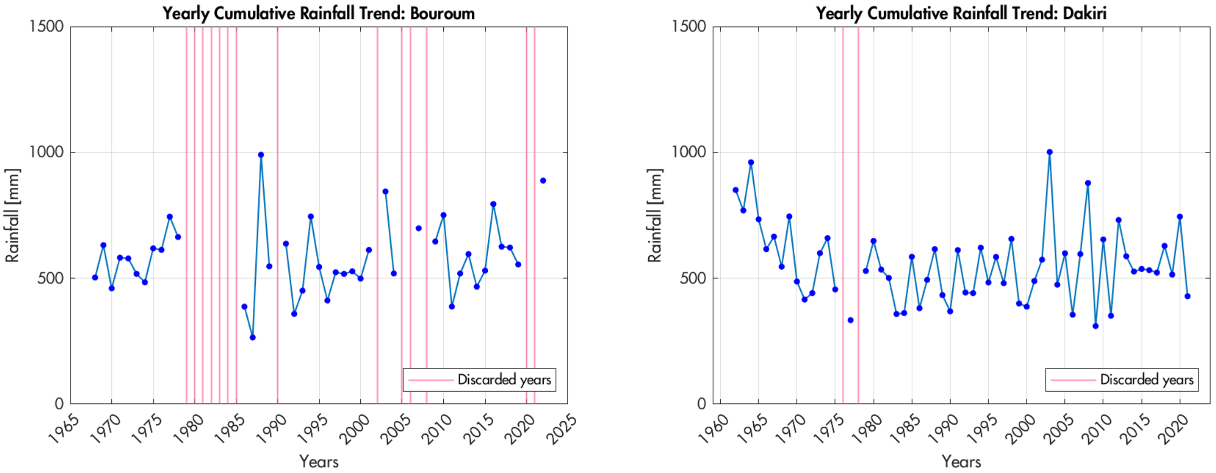


Figure 81: Yearly cumulative rainfall trends, Bouroum and Dakiri.

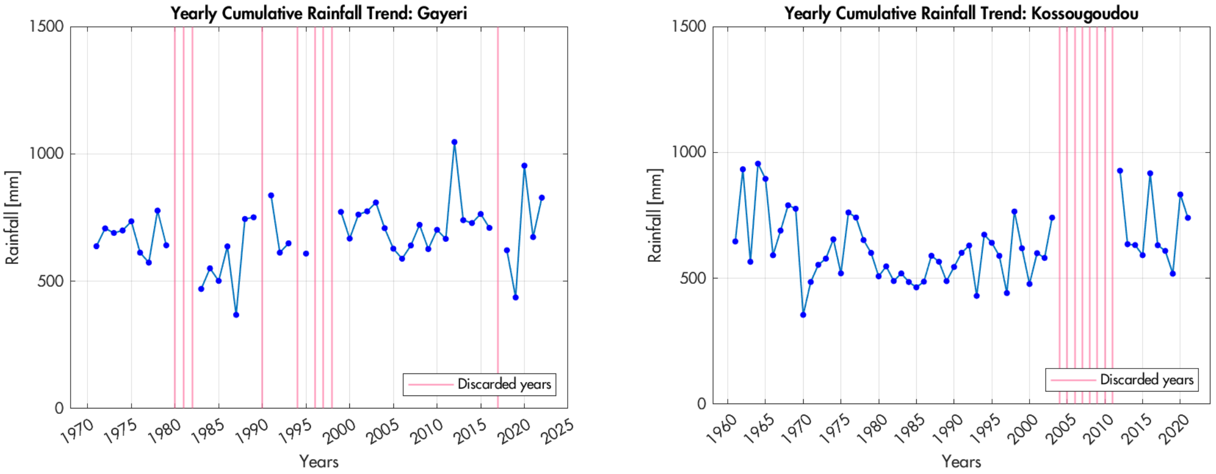


Figure 82: Yearly cumulative rainfall trends, Gayeri and Kossougoudou.

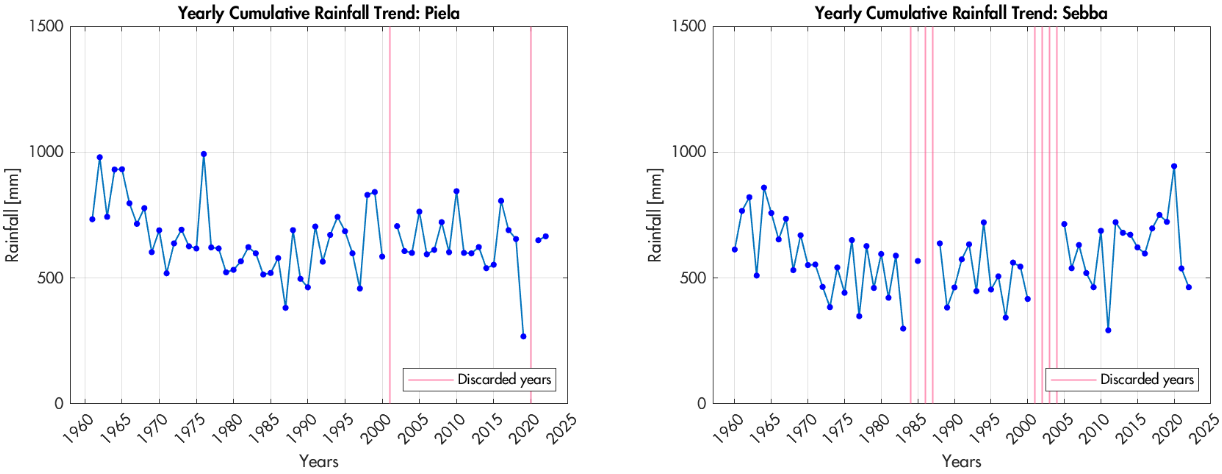


Figure 83: Yearly cumulative rainfall trends, Piela and Sebba.

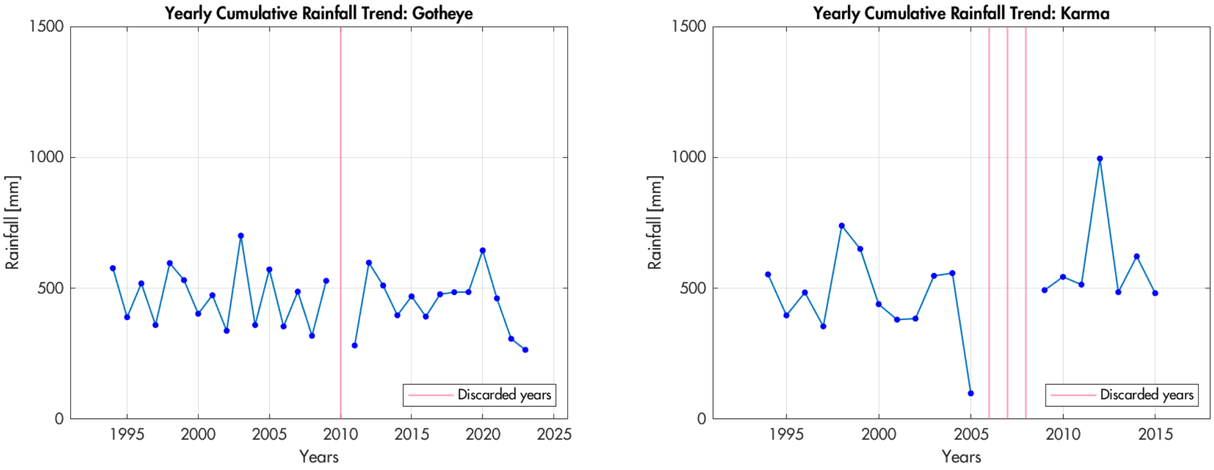


Figure 84: Yearly cumulative rainfall trends, Gotheye and Karma.

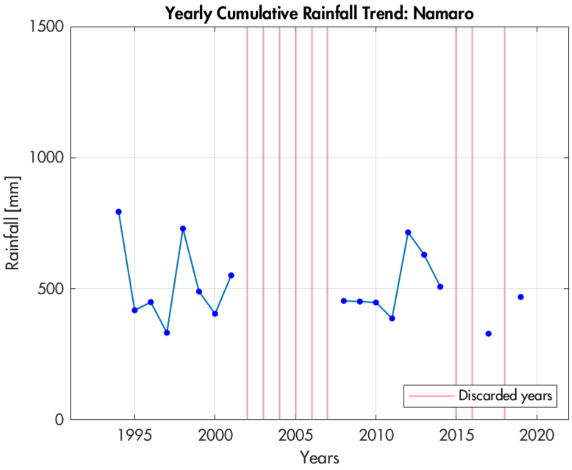


Figure 85: Yearly cumulative rainfall trend, Namaro.

8.3: L-Moments trends

8.3.1: By years

For each station, the yearly trends obviously converge to the values of Table 11. Discarded years are in Table 1.

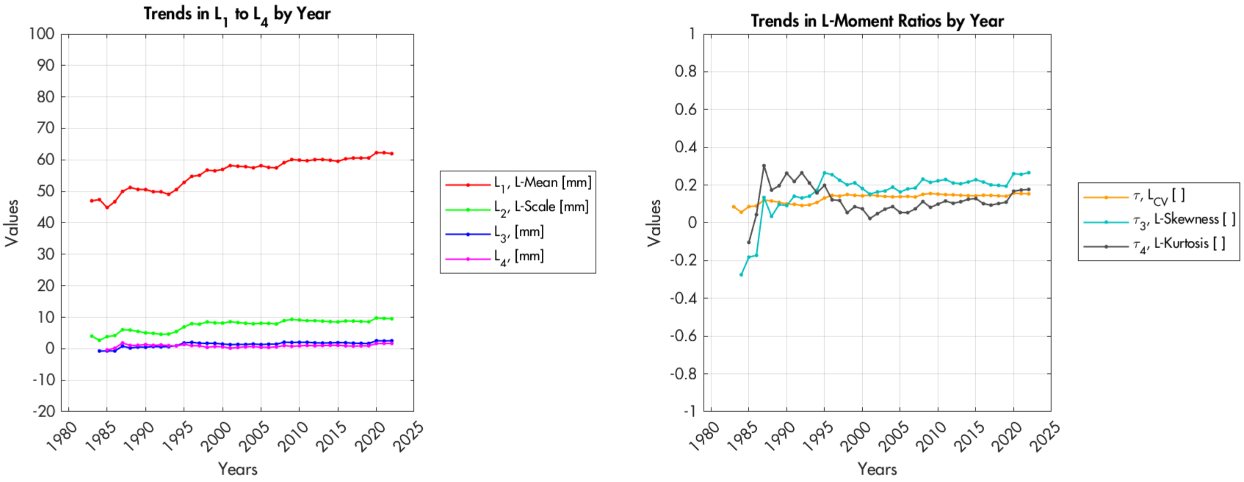


Figure 86: Trends in L-Moments by year, Bogande.

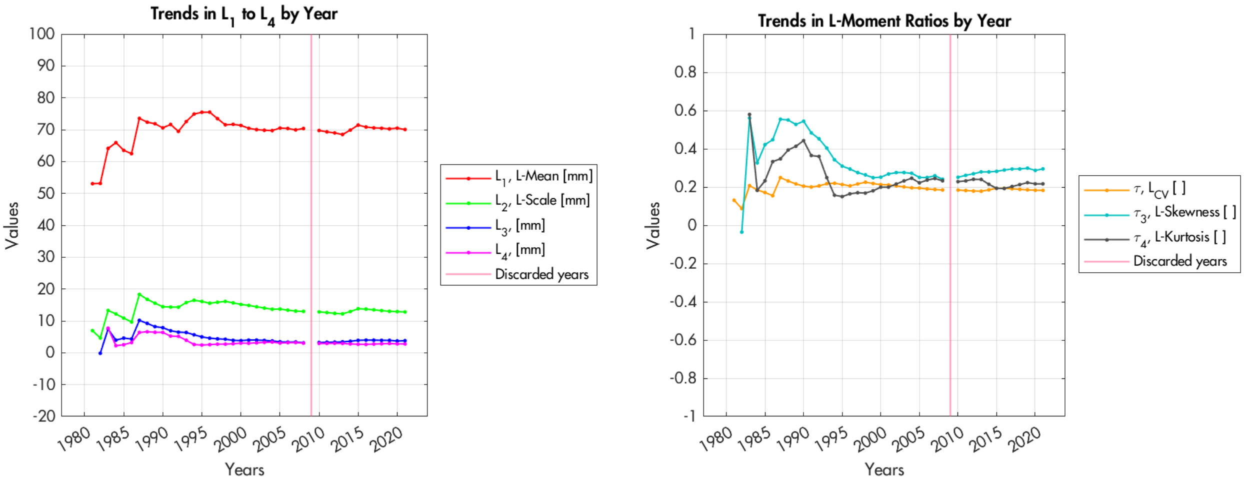


Figure 87: Trends in L-Moments by year, Manni.

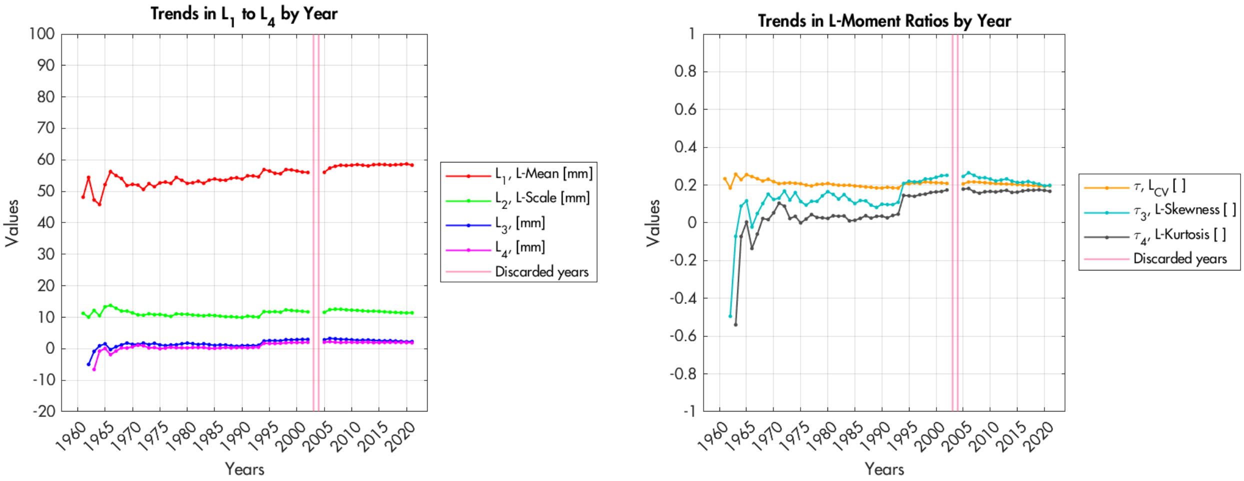


Figure 88: Trends in L-Moments by year, Aribinda

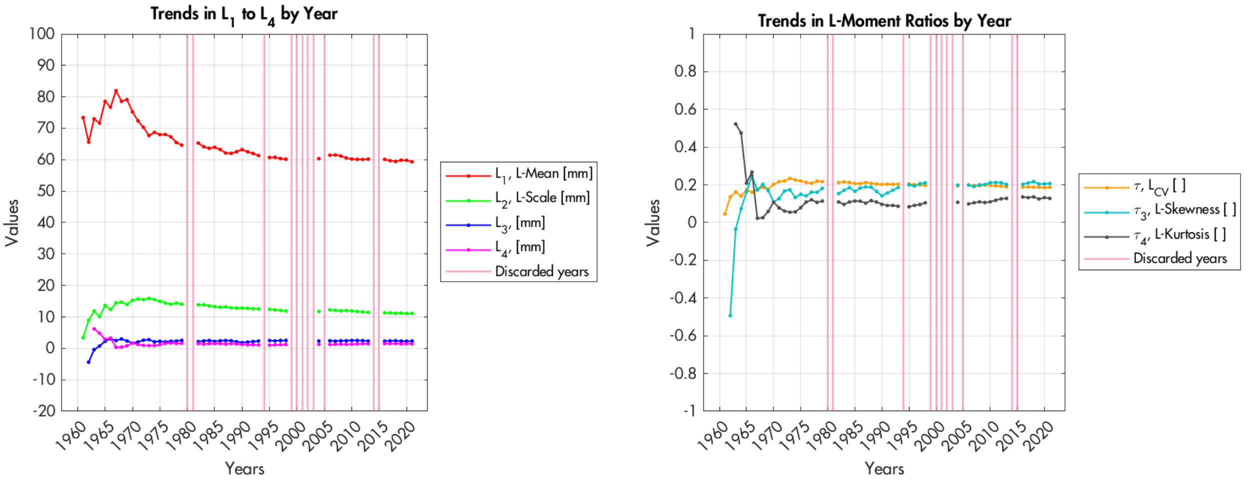


Figure 89: Trends in L-Moments by year, Bani.

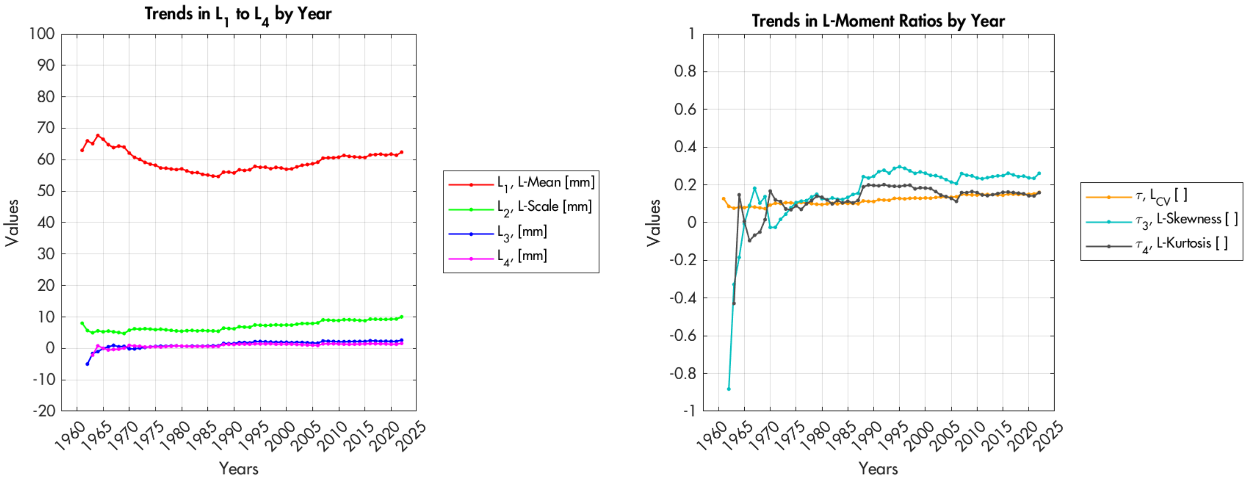


Figure 90: Trends in L-Moments by year, Barsalogho.

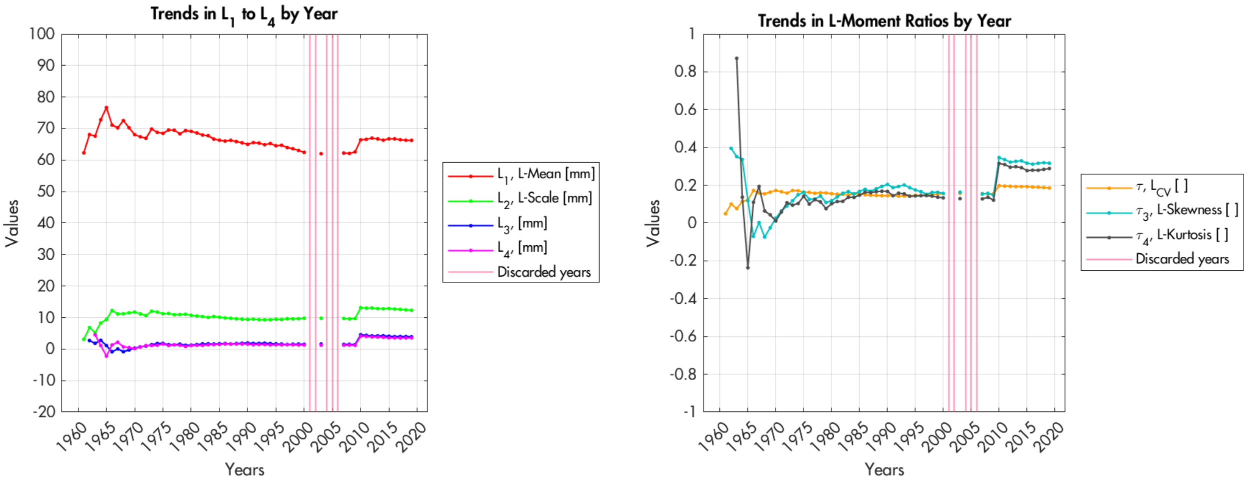


Figure 91: Trends in L-Moments by year, Boulsa.

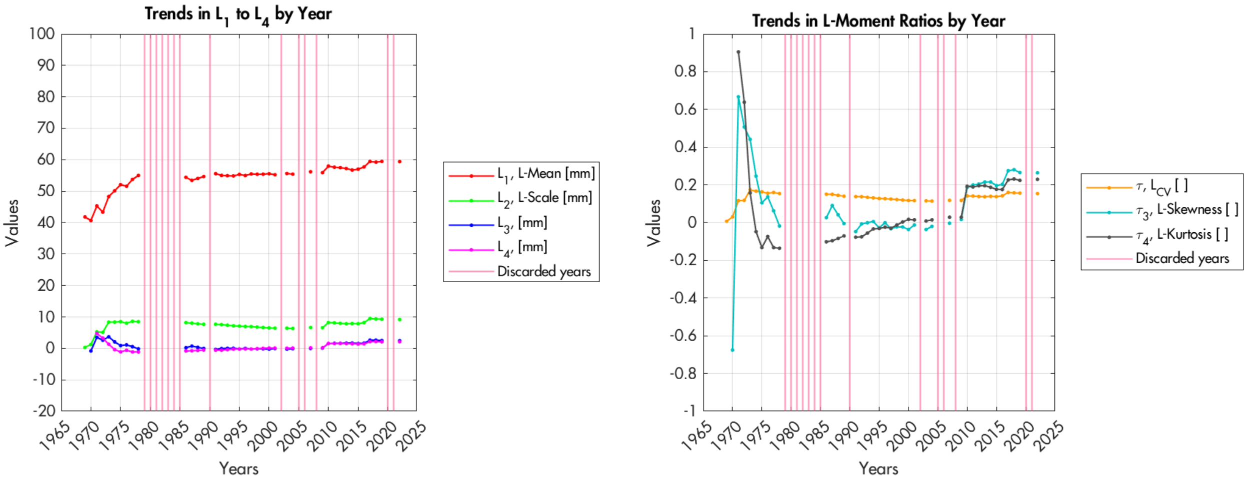


Figure 92: Trends in L-Moments by year, Bouroum.

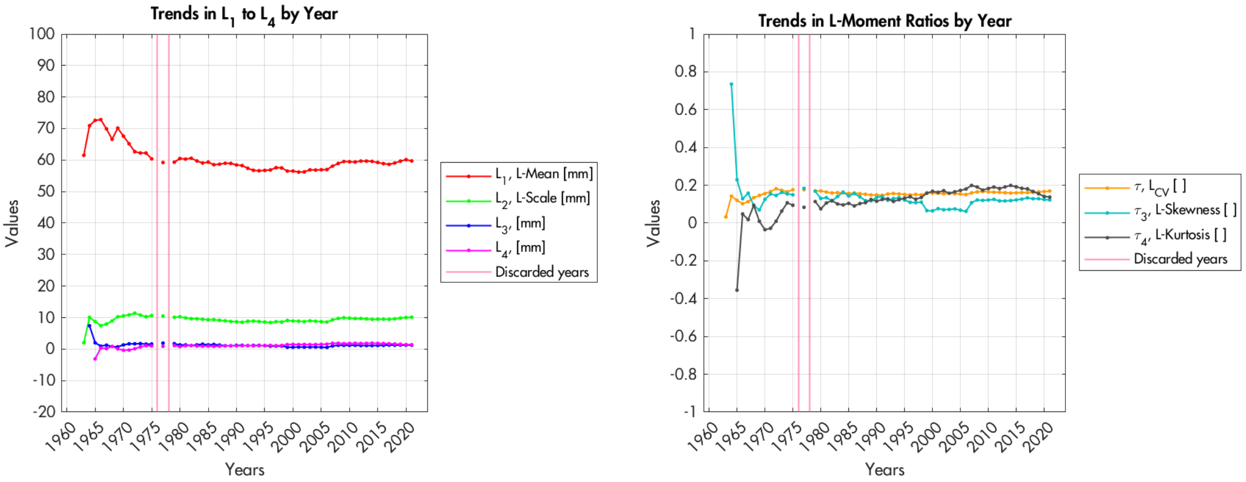


Figure 93: Trends in L-Moments by year, Dakiri.

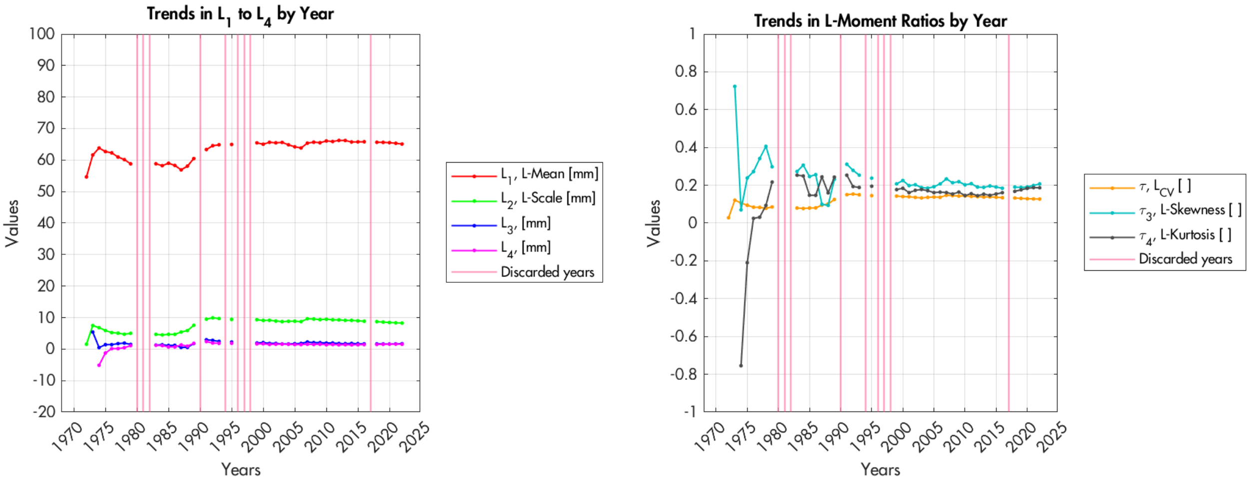


Figure 94: Trends in L-Moments by year, Gayeri.

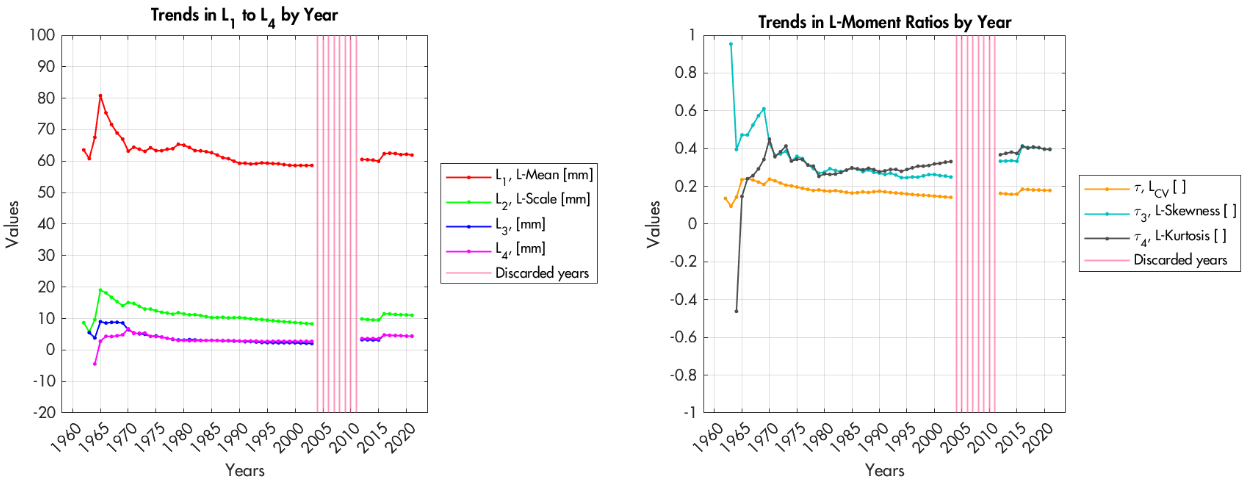


Figure 95: Trends in L-Moments by year, Kossogoudou.

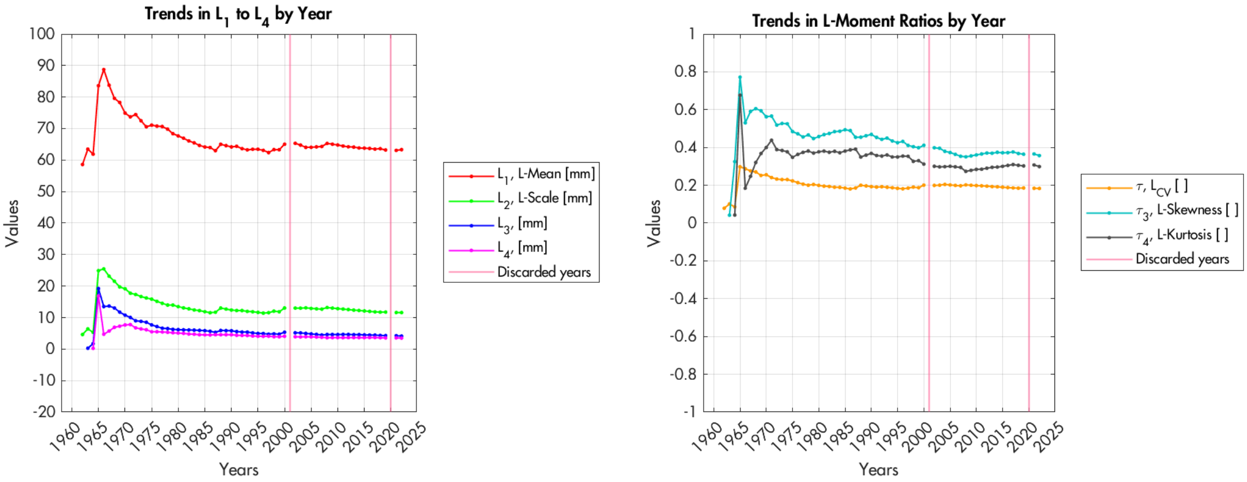


Figure 96: Trends in L-Moments by year, Piela.

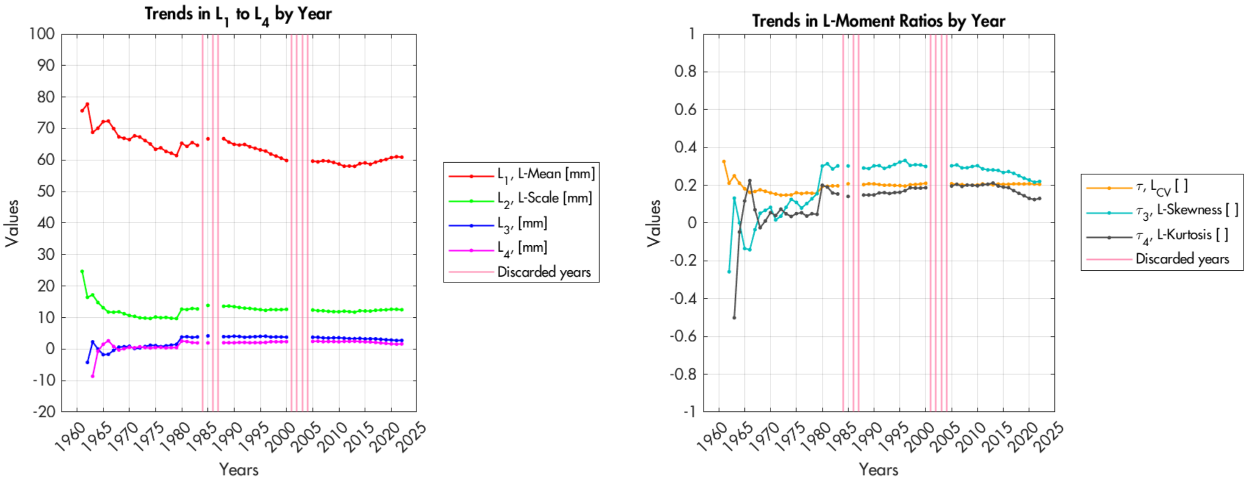


Figure 97: Trends in L-Moments by year, Sebba.

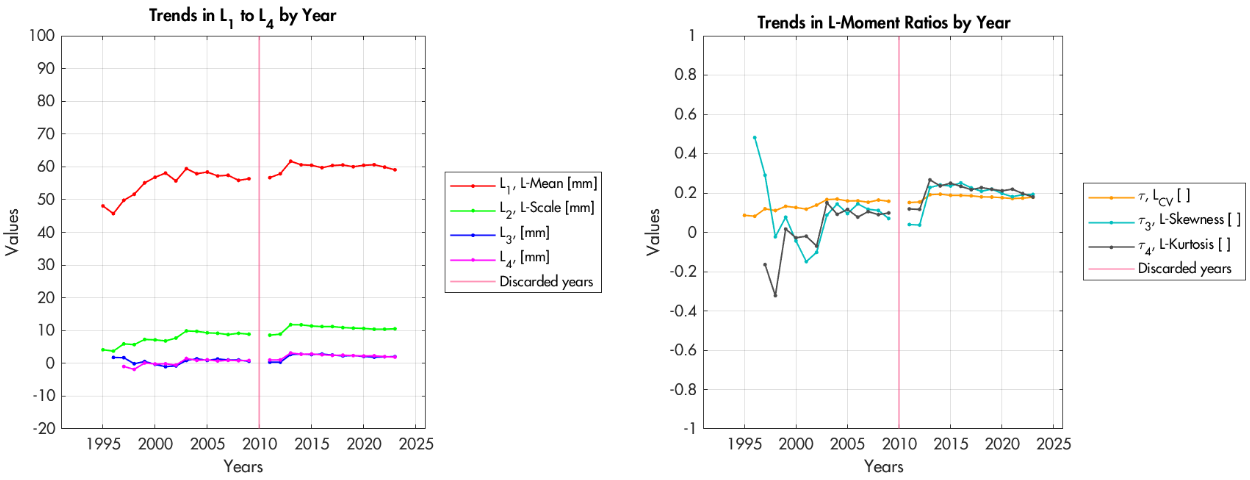


Figure 98: Trends in L-Moments by year, Gotheye.

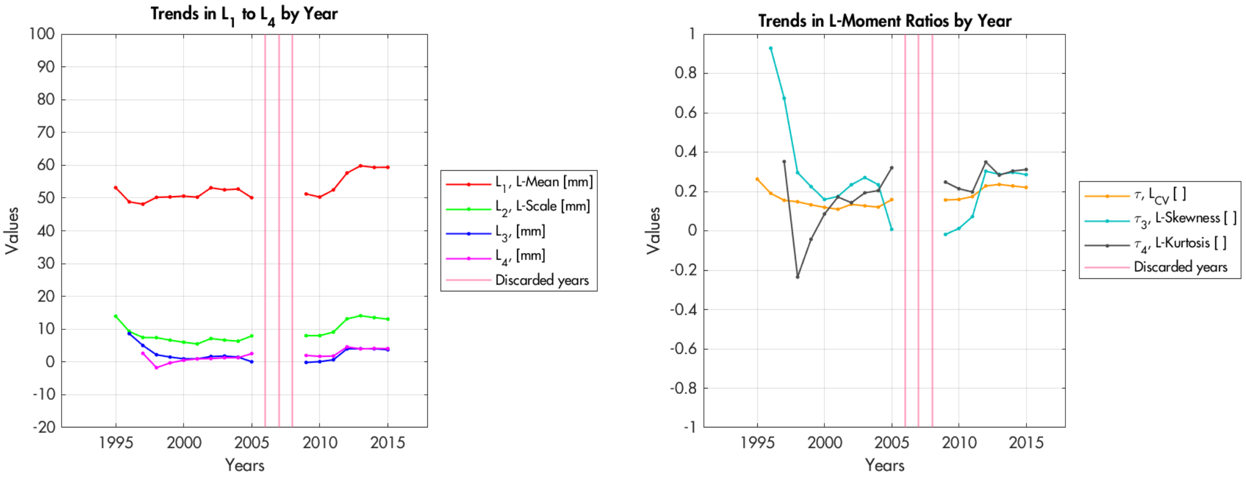


Figure 99: Trends in L-Moments by year, Karma.

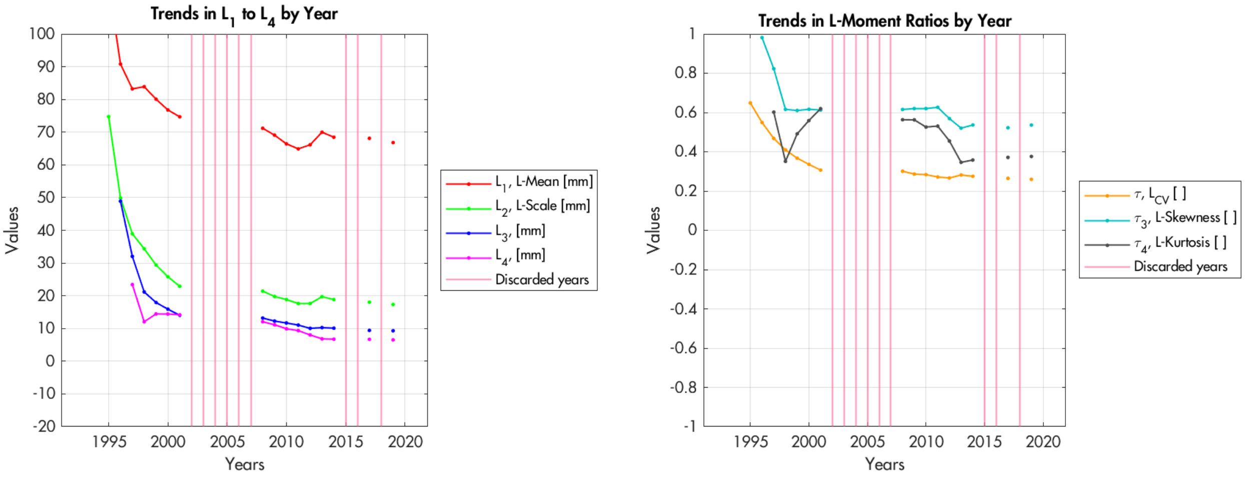


Figure 100: Trends in L-Moments by year, Namaro.

8.3.2: By rainfall durations

Here are the plots of the L-Moments by rainfall duration, throughout the considered moving windows (1-10 days), computed by using the initial rainfall series, not the normalized ones.

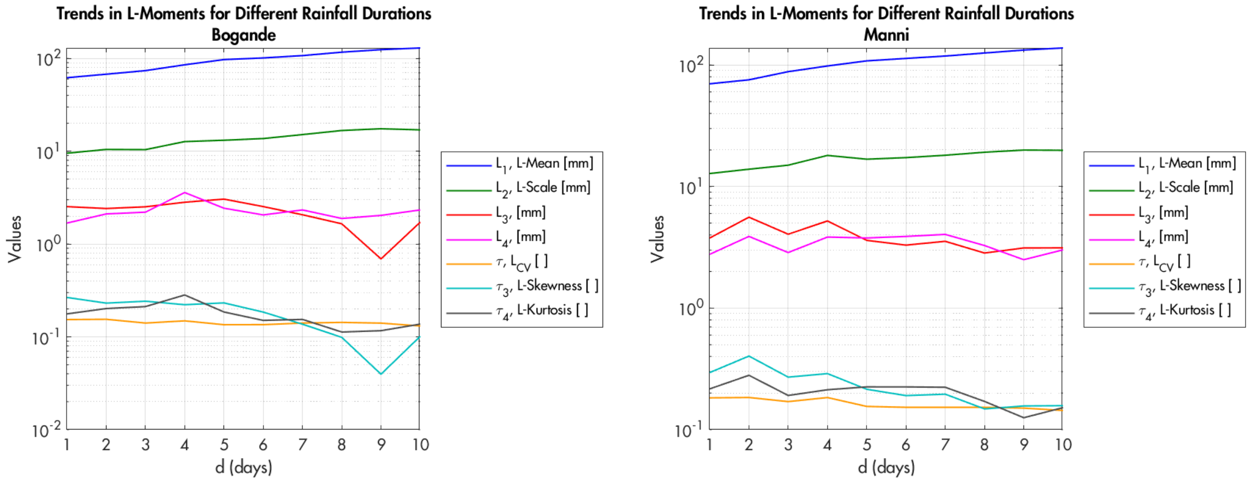


Figure 101: Trends in L-Moments for rainfall durations, Bogande and Manni.

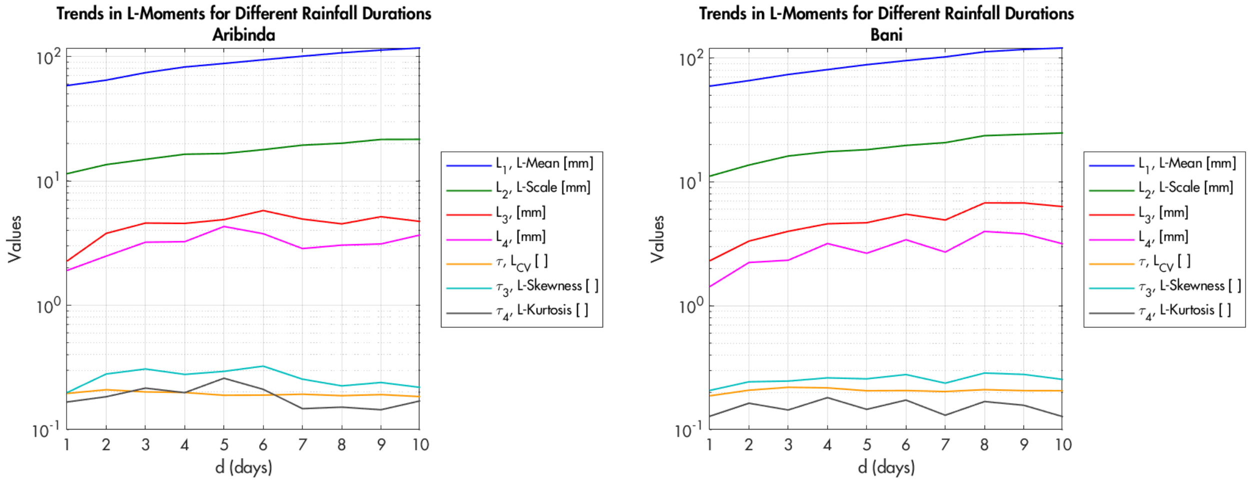


Figure 102: Trends in L-Moments for rainfall durations, Aribinda and Bani.

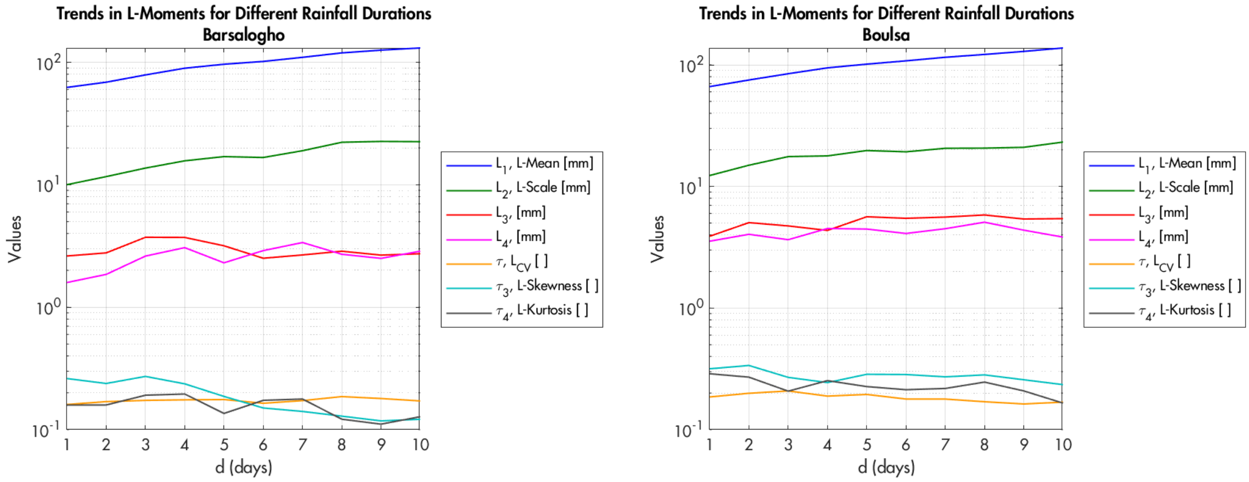


Figure 103: Trends in L-Moments for rainfall durations, Barsalogho and Boulsa.

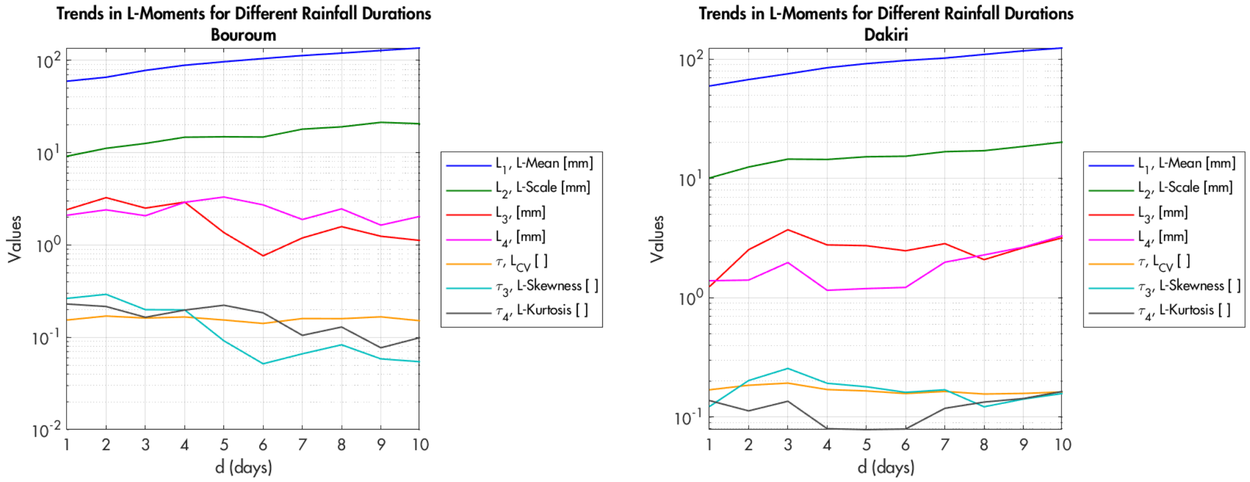


Figure 104: Trends in L-Moments for rainfall durations, Bouroum and Dakiri.

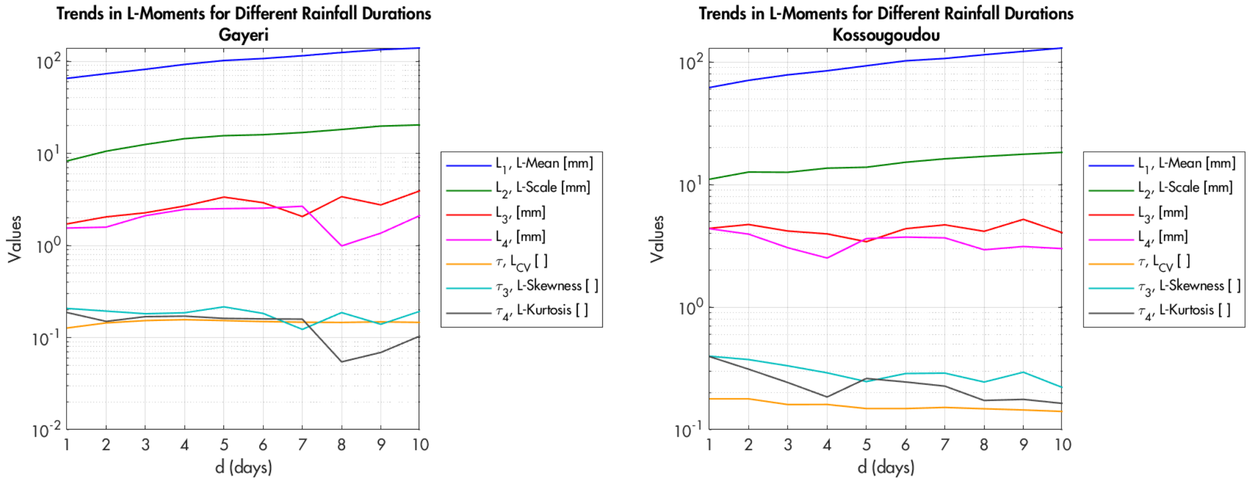


Figure 105: Trends in L-Moments for rainfall durations, Gayeri and Kossougoudou.

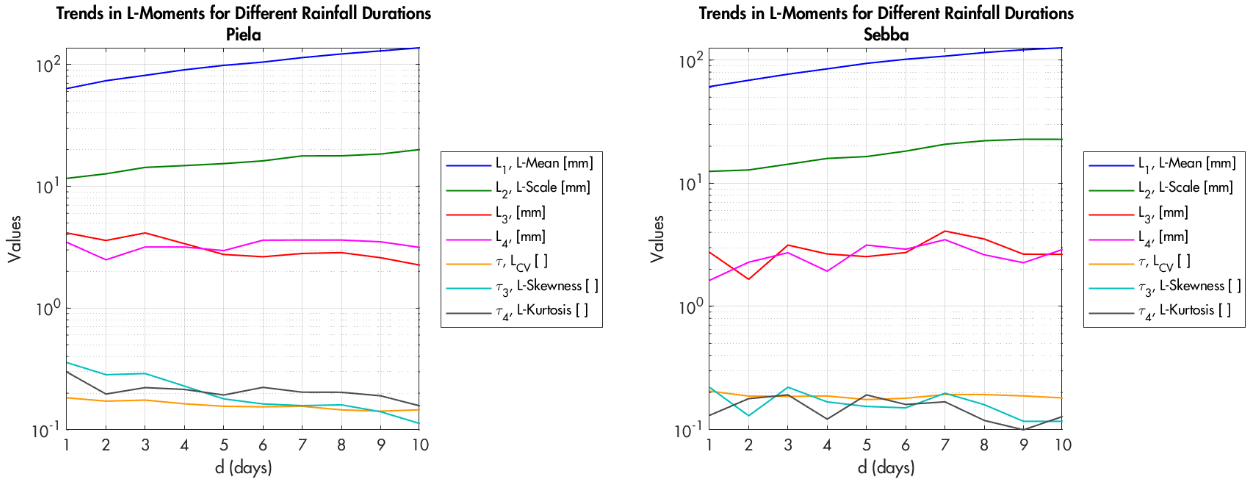


Figure 106: Trends in L-Moments for rainfall durations, Piela and Sebba.

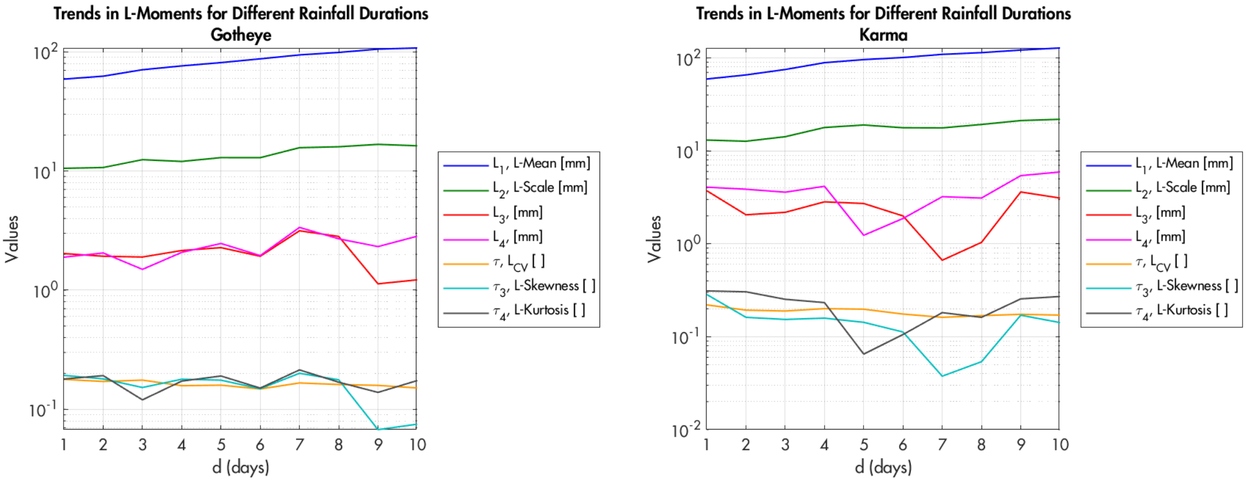


Figure 107: Trends in L-Moments for rainfall durations, Gotheye and Karma.

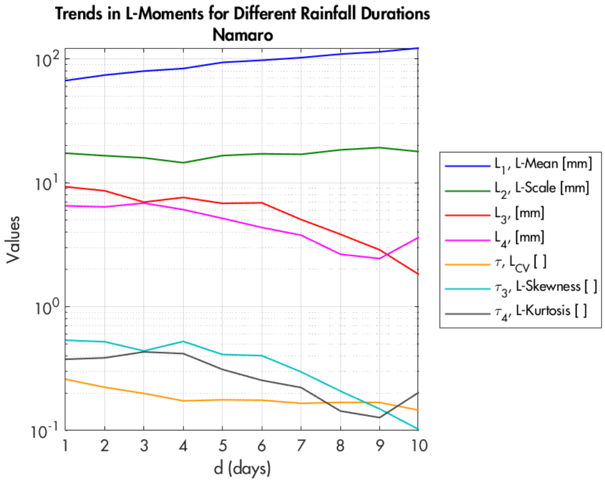


Figure 108: Trends in L-Moments for rainfall durations, Namaro.

8.4: Distributions

Distributions with 'log' in the prefix require the L-moments to be calculated from log(x) instead of the raw value x. When the probability density function (PDF) shows imaginary parts, the value is set to zero, thus rendering the distribution out of bounds in terms of Information Criteria. Consequently, CDF values that are imaginary numbers are set to NaN

8.4.1: 2 Par.

Normal/Gaussian:

Parameters: location $\mu = L_1$ and scale $\sigma = \sqrt{\pi} L_2$

CDF: $F = \frac{1}{2}[1 + erfc(\frac{u}{\sqrt{(2)}})] = \Phi(u)$ where $u = \frac{x - \mu}{\sigma}$ and erfc is the error function

CDF^{-1}: $\hat{x} = \sigma(-\sqrt{2}\, erfc^{-1}(2\phi)) + \mu$

PDF: $f = \frac{\exp(\frac{-1}{2} \cdot u^2)}{2 \cdot \pi \cdot \sigma}$

Log-Normal 2 Par.:

Here the L-moments are in function of log(x).

Parameters: location $\mu = L_1(\log(x))$ and scale $\sigma = \sqrt{\pi} L_2(\log(x))$

CDF: $F = \frac{1}{2}[1 + erfc(\frac{u}{\sqrt{2}})] = \Phi(u)$ where $u = \frac{x - \mu}{\sigma}$

CDF^{-1}: $\hat{x} = \sigma(-\sqrt{2}\, erfc^{-1}(2\phi)) + \mu$

PDF: $f = \frac{\exp(\frac{-1}{2} \cdot u^2)}{2 \cdot \pi \cdot \sigma \cdot \exp(\log(x))}$

Exponential:

Parameters: scale $\alpha = 2 L_2$ and location $\xi = L_1 - \alpha$

CDF: $F = 1 - \exp(\frac{\xi - x}{\alpha})$

CDF^{-1}: $\hat{x} = \xi - \alpha \log(1 - \phi)$

PDF: $f = \frac{\exp(\frac{\xi - x}{\alpha})}{\alpha}$

$f(x < \xi) = 0$

Gamma:

Parameters:

ζ and α represent the shape

if tau > 0 tau < 0.5, then $\zeta = \pi \tau^2$ and $\alpha = \frac{1 - 0.3080 \zeta}{\zeta - 0.05812 \zeta^2 + 0.01765 \zeta^3}$;

if tau > 0 tau < 0.5, then $\zeta = 1 - \tau$ and $\alpha = \frac{0.7213 \zeta - 0.5947 \zeta^2}{102.1817 \zeta + 1.2113 \zeta^2}$;

rate: $\beta = \frac{L_1}{\alpha}$.

$$\text{CDF:}\quad F=gammainc\left(\frac{x}{\beta},\alpha\right)\quad \text{where gammainc is the gamma incomplete function}$$

$$\text{CDF}^{-1}\text{:}\quad \hat{x}=gammainc^{-1}(\phi,\alpha)\cdot\beta$$

$$\text{PDF:}\quad f=\frac{x^{(\alpha-1)}\cdot\dfrac{\exp\left(\frac{-x}{\beta}\right)}{\beta^{\alpha}}}{\Gamma(\alpha)}$$

Logistic:

$$\text{Parameters: scale}\quad \alpha=L_2\quad \text{and location}\quad \xi=L_1\quad ;$$

$$\text{CDF:}\quad F=\frac{1}{1+\exp(-u)}\quad \text{where}\quad u=\frac{x-\xi}{\alpha}$$

$$\text{CDF}^{-1}\text{:}\quad \hat{x}=\xi-\alpha\log(y)\quad \text{where}\quad y=\frac{1-\phi}{\phi}$$

$$\text{PDF:}\quad f=\frac{\dfrac{\exp(-u)}{\alpha}}{(1+\exp(-u))^2}$$

Gumbel (EV1):

$$\text{Parameters: scale}\quad \alpha=\frac{L_2}{\log 2}\quad \text{and location}\quad \xi=L_1-0.577\,\alpha\quad ;$$

$$\text{CDF:}\quad F=\exp\left(-\exp\left(\frac{\xi-x}{\alpha}\right)\right)$$

$$\text{CDF}^{-1}\text{:}\quad \hat{x}=\xi-\alpha\log(\log(\phi))$$

$$\text{PDF:}\quad f=\exp\left(\frac{\xi-x}{\alpha}\right)\frac{\exp\left(-\exp\left(\frac{\xi-x}{\alpha}\right)\right)}{\alpha}$$

8.4.2: 3 Par.

Log-Normal 3 Par.

As for the Log-Norm. 2 Par., here also the L-moments are in function of log(x).

Parameters:

Coefficients for the approximations taken from Table A.1 from [24].

$$E_0=2.0466534;E_1=-3.6544371;E_2=1.8396733;E_3=-0.20360244;$$
$$F_1=-2.0182173;F_2=1.2420401;F_3=-0.21741801$$

$$\text{shape:}\quad \kappa=-\tau_3\frac{E_0+E_1\tau_3^2+E_2\tau_3^4+E_3\tau_3^6}{1+F_1\tau_3^2+F_2\tau_3^4+F_3\tau_3^6}\quad \text{; scale:}\quad \alpha=\frac{L_2\cdot\kappa\cdot\exp\left(\frac{-1}{2}\kappa^2\right)}{1-2\Phi\left(\frac{-\kappa}{\sqrt{2}}\right)}\quad \text{; location:}\quad \xi=L_1-\alpha\frac{1-\exp\left(\frac{1}{2}\kappa^2\right)}{\kappa}$$

CDF:

$$\text{if }\kappa=0\text{, then}\quad y=\frac{x-\xi}{\alpha}\quad \text{; if }\kappa\neq 0\text{, then}\quad y=\frac{-\log\left(\frac{1-\kappa\cdot(x-\xi)}{\alpha}\right)}{\kappa}$$

$$F=\Phi(y)\quad .$$

$$\text{CDF}^{-1}\text{:}\quad y=\Phi^{-1}(\phi)$$

$$\text{if }\kappa=0\text{, then}\quad \hat{x}=\alpha\cdot y+\xi\quad \text{; if }\kappa\neq 0\text{, then}\quad \hat{x}=\xi+\alpha\frac{1-\exp(-\kappa y)}{\kappa}$$

$$\text{PDF:}\quad f=\frac{\frac{\exp(\kappa\cdot y-\frac{1}{2}\cdot y^{2})}{\alpha}}{\sqrt{(2\cdot\pi)}}$$

Pearson Type III

Parameters:

If tau-3 ≥ 1/3 and tau-3 < 1, then $\zeta=1-|(\tau_3)|$ and $\alpha=\frac{0.36067\cdot\zeta-0.59567\cdot\zeta^{2}+0.25361\cdot\zeta^{3}}{1-2.78861\cdot\zeta+2.56096\cdot\zeta^{2}-0.77045\cdot\zeta^{3}}$;

if else, then $\zeta=3\cdot\pi\cdot\tau_3^{2}$ and $\alpha=\frac{1+0.2906\cdot\zeta}{\zeta+0.1882\cdot\zeta^{2}+0.0442\cdot\zeta^{3}}$

shape: $\gamma=\frac{2}{\sqrt{(\alpha)}}\cdot sign(\tau_3)$; location: $\mu=L_1$; scale: $\sigma=L_2\cdot\sqrt{(\pi)}\cdot\sqrt{(\alpha)}\cdot\frac{\Gamma(\alpha)}{\Gamma(\alpha+\frac{1}{2})}$;

$$\beta=\frac{1}{2}\cdot\sigma\cdot|(\gamma)| \quad ; \quad \xi=\mu-\frac{2\cdot\sigma}{\gamma}$$

If α is bigger than approximately 171, β is infinite (overflow) and Γ is very small. So, it must be chosen a case where Γ = 0 and the distribution becomes Normal, with α = µ and β = σ.

$$\alpha=L_1 \quad ; \quad \beta=\sqrt{(\pi)}\cdot L_2 \quad ; \quad \xi=n/a \quad ; \quad \gamma=0$$

CDF:

if γ > 0, then $u=\frac{x-\xi}{\beta}$ and $F=gammainc(u,\alpha)$ where gammainc is the incomplete gamma function;

if γ < 0, then $u=\frac{\xi-x}{\beta}$ and $F=1-gammainc(u,\alpha)$;

if γ = 0, than it is Normal distribution, so $u=\frac{\log(x)-\alpha}{\beta}$ and $F=\frac{1}{2}\cdot(1+erf(\frac{u}{\sqrt{(2)}}))$

CDF^{-1}:

if γ > 0, then $\chi=gammainc(\phi,\alpha)$ and $\hat{x}=(\chi\cdot\beta)+\xi$

if γ < 0, then $\chi=gammainc(1-\phi,\alpha)$ and $\hat{x}=(-\chi\cdot\beta)+\xi$

if γ = 0, than it is Normal distribution, so $\hat{x}=\Phi^{-1}(\phi,\alpha,\beta)$

PDF:

if γ > 0, then $f=\frac{(x-\xi)^{\alpha-1}\cdot\exp\frac{\xi-x}{\beta}}{\beta^{\alpha}\cdot\Gamma(\alpha)}$

if γ < 0, then $f=\frac{(\xi-x)^{\alpha-1}\cdot\exp\frac{x-\xi}{\beta}}{\beta^{\alpha}\cdot\Gamma(\alpha)}$

if γ = 0, than it is Normal distribution, so $u=\frac{x-\alpha}{\beta}$ and $f=\frac{\frac{\exp(\frac{-1}{2}\cdot u^{2})}{2\cdot\pi}}{\beta}$

Log-Pearson Type III

Here the L-moments are computed relative to log(x) not just x. Without repeating, the expressions for the parameters, CDF and CDF^{-1}, are the same ones as for the Pearson Type III, the only difference is that x must be substituted with log(x). Thus when γ = 0, the CDF and CDF^{-1} are relative to a Log-Normal distribution.

Generalized Extreme Value (GEV):

$$b_0=\frac{1}{n}\sum_{i=1}^{n}x_i \quad ; \quad b_1=\frac{1}{n}\frac{\sum_{i=1}^{n}(i-1)}{(n-1)}x_i \quad ; \quad b_2=\frac{1}{n}\frac{\sum_{i=1}^{n}(i-1)(i-2)}{(n-1)(n-2)}x_i \quad ;$$

Parameters:

$$c=\frac{2}{3+\tau_3}-\frac{\log 2}{\log 3} \quad \text{; shape:} \quad \kappa=7.8590\,c+2.9554\,c^2 \quad ;$$

$$\text{scale:} \quad \alpha=\frac{\frac{L_2\kappa}{1-(2^{(-\kappa)})}}{\Gamma(\kappa+1)} \quad \text{; location:} \quad \xi=L_1-\alpha\frac{1-\Gamma(1+\kappa)}{\kappa}$$

CDF:

$$\text{if } \kappa = 0\text{, then} \quad F=\exp(-\exp(\frac{\xi-x}{\alpha})) \quad \text{; if } \kappa \neq 0\text{, then} \quad y=\frac{-\log(\frac{1-\kappa\cdot(x-\xi)}{\alpha})}{\kappa} \quad ;$$

$$F=\exp(-\exp(y))$$

CDF^{-1}:

$$\text{if } \kappa = 0\text{, then} \quad \hat{x}=\xi-\alpha\log(-\log(\phi)) \quad \text{; if } \kappa \neq 0\text{, then} \quad \hat{x}=\xi+\alpha\frac{1-(-\log(\phi))^{\kappa}}{\kappa}$$

PDF:

$$\text{if } \kappa = 0\text{, Gumbel distribution then} \quad f=\exp(\frac{\xi-x}{\alpha})\frac{\exp(-\exp(\frac{\xi-x}{\alpha}))}{\alpha}$$

$$\text{if } \kappa \neq 0\text{, then} \quad f=\frac{\exp((\kappa-1)\cdot y-\exp(-y))}{\alpha}$$

Generalized Logistic (GLO):

Parameters:

$$\text{shape:} \quad \kappa=-\tau_3 \quad \text{; scale:} \quad \alpha=\frac{L_2\cdot\sin(\pi\cdot\kappa)}{\frac{\pi}{\kappa}} \quad \text{; location:} \quad \xi=L_1-\alpha(\frac{1}{\kappa}-\frac{\pi}{\kappa\cdot\pi})$$

CDF:

$$\text{if } \kappa = 0\text{, then} \quad u=\frac{x-\xi}{\alpha} \quad \text{; if } \kappa \neq 0\text{, then} \quad u=\frac{-\log(1-\kappa\frac{x-\xi}{\alpha})}{\kappa} \quad ;$$

$$F=\frac{1}{1+\exp(-u)}$$

$$\text{CDF}^{-1}\text{:} \quad y=\frac{1-\phi}{\phi}$$

$$\text{if } \kappa = 0 \text{ then} \quad \hat{x}=\xi-\alpha\log(y) \quad \text{; if } \kappa \neq 0\text{, then} \quad \hat{x}=\xi+\alpha\frac{1-y^{\kappa}}{\kappa}$$

$$\text{PDF:} \quad f=\frac{\frac{\exp(u\cdot(\kappa-1))}{\alpha}}{(1+\exp(-u))^2}$$

Generalized Pareto (GPA):

Parameters:

$$\text{shape:} \quad \kappa=\frac{1-3\tau_3}{1+\tau_3} \quad \text{; scale:} \quad \alpha=(1+\kappa)(2+\kappa)L_2 \quad \text{; location:} \quad \xi=L_1-L_2(2+\kappa)$$

CDF:

if κ = 0, then $u=\frac{x-\xi}{\alpha}$; if κ ≠ 0, then $u=\frac{-\log(1-\kappa\frac{x-\xi}{\alpha})}{\kappa}$; $F=1+\exp(-u)$

CDF^{-1}:

if κ = 0, then $\hat{x}=\xi-\alpha\log(1-\phi)$; if κ ≠ 0, then $\hat{x}=\xi+\alpha\frac{1-(1-\phi)^{\kappa}}{\kappa}$

PDF:

$$f=\frac{\exp((\kappa-1)\cdot u)}{\alpha}$$

if κ = 0, case Exponential distribution then $f(x<\xi)=0$

if κ =1, then $f(x<\xi\vee x\geq\frac{\xi+\alpha}{\kappa})=0$

if κ > 0, then $f(x<\xi\vee x>\frac{\xi+\alpha}{\kappa})=0$

if κ < 0, then $f(x<\xi)=0$

Burr Type XII:

if: $\tau_3>\frac{1+3\tau}{3+\tau}$ and $\tau_3<\frac{1}{\tau}(-2+2\cdot 3^{(\frac{\log(1-\tau)}{\log 2})}+3\tau)$: tau and tau-3 are out of bounds;

Parameters:

λ is for scale, c represents the shape of the distribution's tail and κ is the shape for the peak & tail

$$t_r=24-33\tau+9\tau^3-8(1-\tau)^{(\frac{\log(3)}{\log(2)})}(3+\tau)$$

$$\kappa=\frac{\log\frac{-(t_r+\sqrt{t_r^2-24\tau\cdot(-4(1-\tau)^{\frac{\log(3)}{\log(2)}}(3+\tau)+3(4-5\tau+\tau^3))\cdot(-1+\tau(-3+\tau^3)+3\tau^3))}}{24(1-\tau)^{\frac{\log(3)}{\log(2)}}(3+\tau)-18(4-5\tau+\tau^3)}}{\log(3/2)}$$

if κ < -1, then $c=\frac{-(\kappa+\frac{5}{2}^{1+\kappa}\cdot\kappa+\kappa\cdot\tau+\kappa\cdot\sqrt{(\frac{2}{5}^{-2\cdot(1+\kappa)}+2^{-\kappa}\cdot 5^{1+\kappa}\cdot(1-3\cdot\tau)+(1+\tau)^2)})}{4\cdot\tau}$;

if κ ≥ -1, then $c=-(\frac{2}{3})\cdot(\frac{5}{2}-(\frac{2}{5})^{(-1-\kappa)})\cdot(1+\kappa-\frac{1}{\tau})+\frac{2}{3}\cdot(-1+(\frac{5}{2})^{(1+\kappa)})\cdot(-\kappa-\frac{\log(2)}{\log(1-\tau)})$;

$$\lambda=\frac{L_1\cdot(-\kappa)^{(1+\frac{1}{c})}\cdot\Gamma(1-\frac{1}{\kappa})}{-\Gamma(1+\frac{1}{c})\cdot\Gamma(\frac{1}{c}+\frac{1}{\kappa})}$$

CDF: $F=1-(1-\kappa\cdot(\frac{x}{\lambda})^{c})^{(\frac{1}{\kappa})}$

CDF^{-1}: $\hat{x}=\lambda(\frac{1}{\kappa}(1-(1-\phi)^{\kappa}))^{(\frac{1}{c})}$

PDF:

$$f=\frac{c\cdot(1-\kappa\cdot(\frac{x}{\lambda})^{c})^{-1+\frac{1}{\kappa}}\cdot(\frac{x}{\lambda})^{-1+c}}{\lambda}$$

8.5: Growth factors

The parameters calculated with the average normalized L-moments across all durations, will be indicated with an over-line, and they are dimensionless, thus correctly rendering K(T) dimensionless. Here are the expressions for the K(T):

Normal: $K(T)=\overline{\sigma}(-\sqrt{2}\,erfc^{-1}(2\overline{F}))+\overline{\mu}$

Log-Norm. 2 Par.: $K(T)=\overline{\sigma}(-\sqrt{2}\,erfc^{-1}(2\overline{F}))+\overline{\mu}$

Log-Norm. 3 Par.: $y=\Phi^{-1}(\overline{F})$;

if $\overline{\kappa}=0$, then $K(T)=\overline{\alpha}\cdot y+\overline{\xi}$; if $\overline{\kappa}\neq 0$, then $K(T)=\overline{\xi}+\overline{\alpha}\dfrac{1-\exp(-\overline{\kappa}\,y)}{\overline{\kappa}}$.

Exponential: $K(T)=\overline{\xi}-\overline{\alpha}\log(1-\overline{\phi})$

Gamma: $K(T)=gammainc^{-1}(\overline{F},\overline{\alpha})\cdot\overline{\beta}$

Logistic: $K(T)=\overline{\xi}-\overline{\alpha}\log(y)$ where $y=\dfrac{1-\overline{F}}{\overline{F}}$

Pearson Type III:

if $\overline{\gamma}>0$, then $\overline{\chi}=gammainc(\overline{F},\overline{\alpha})$ and $K(T)=(\overline{\chi}\cdot\overline{\beta})+\overline{\xi}$

if $\overline{\gamma}<0$, then $\overline{\chi}=gammainc(1-\overline{F},\overline{\alpha})$ and $K(T)=(-\overline{\chi}\cdot\overline{\beta})+\overline{\xi}$

if $\overline{\gamma}=0$, than it is Normal distribution, so $K(T)=\Phi^{-1}(\overline{F},\overline{\alpha},\overline{\beta})$.

Log-Pearson Type III: Same expressions as for Pearson Type III.

Gumbel: $K(T)=\overline{\xi}-\overline{\alpha}\log(\log(\overline{F}))$

GEV: if $\overline{\kappa}=0$, then $K(T)=\overline{\xi}-\overline{\alpha}\log(-\log(\overline{F}))$; if $\overline{\kappa}\neq 0$, then $K(T)=\overline{\xi}+\overline{\alpha}\dfrac{1-(-\log(\overline{F}))^{\overline{\kappa}}}{\overline{\kappa}}$

GLO: $y=\dfrac{1-\overline{F}}{\overline{F}}$;if $\overline{\kappa}=0$ then $K(T)=\overline{\xi}-\overline{\alpha}\log(y)$; if $\overline{\kappa}\neq 0$, then $K(T)=\overline{\xi}+\overline{\alpha}\dfrac{1-y^{\overline{\kappa}}}{\overline{\kappa}}$

GPA:

if $\overline{\kappa}=0$, then $K(T)=\overline{\xi}-\overline{\alpha}\log(1-\overline{F})$; if $\overline{\kappa}\neq 0$, then $K(T)=\overline{\xi}+\overline{\alpha}\dfrac{1-(1-\overline{F})^{\overline{\kappa}}}{\overline{\kappa}}$.

Burr Type XIII: $K(T)=\overline{\lambda}\left(\frac{1}{\overline{\kappa}}(1-(1-F)^{\overline{\kappa}})\right)^{\left(\frac{1}{\overline{c}}\right)}$

8.6: Basic spatial interpolation: Thiessen Polygons

8.6.1: Stations/Rainfall Gauges

Max. & L-Moments

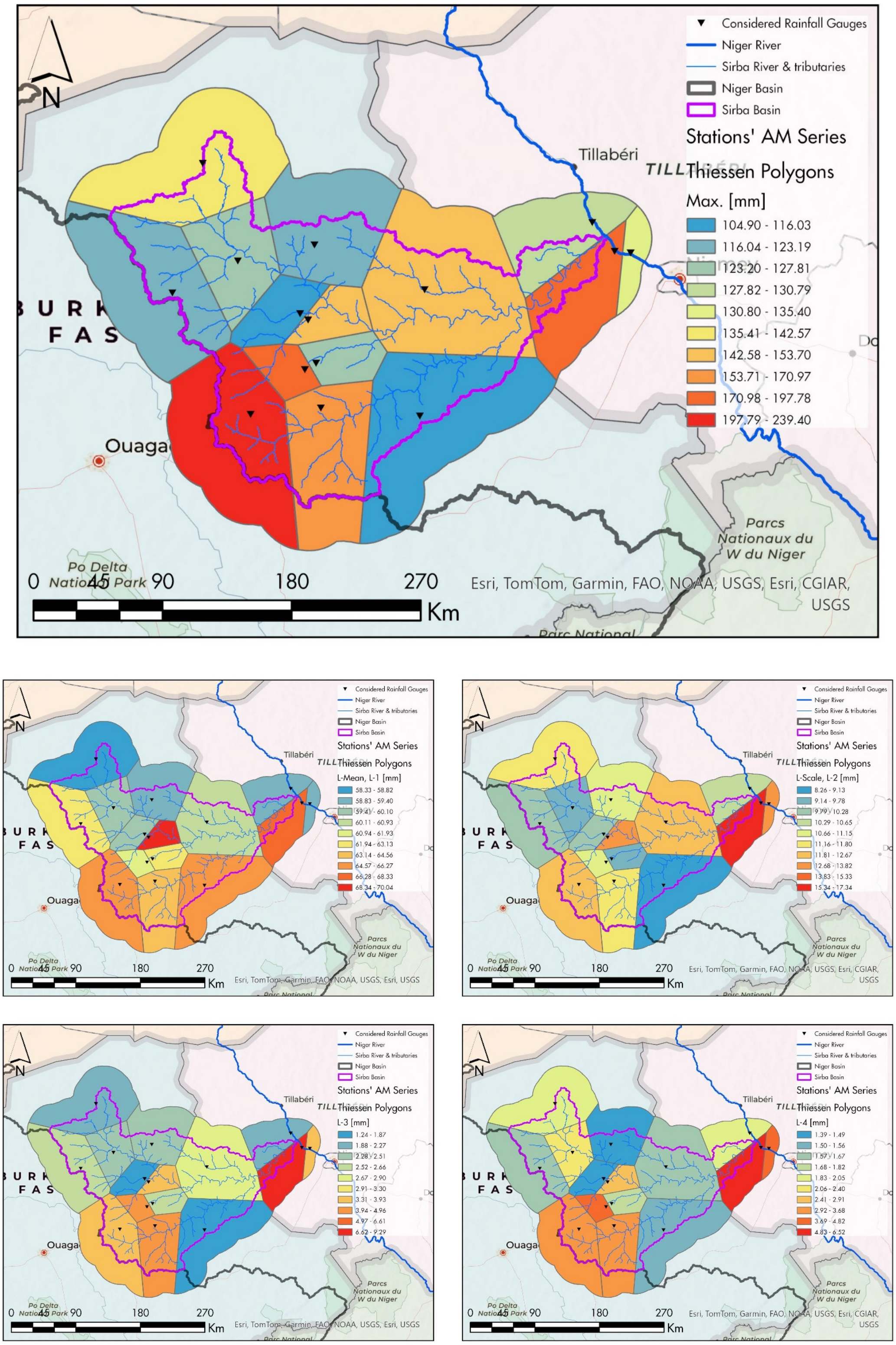


Figure 109: Thiessen polygons of the considered rainfall gauges (stations), representing Max. and L-Moments of the AM Series of Daily Rainfall.

BTXII Quantiles

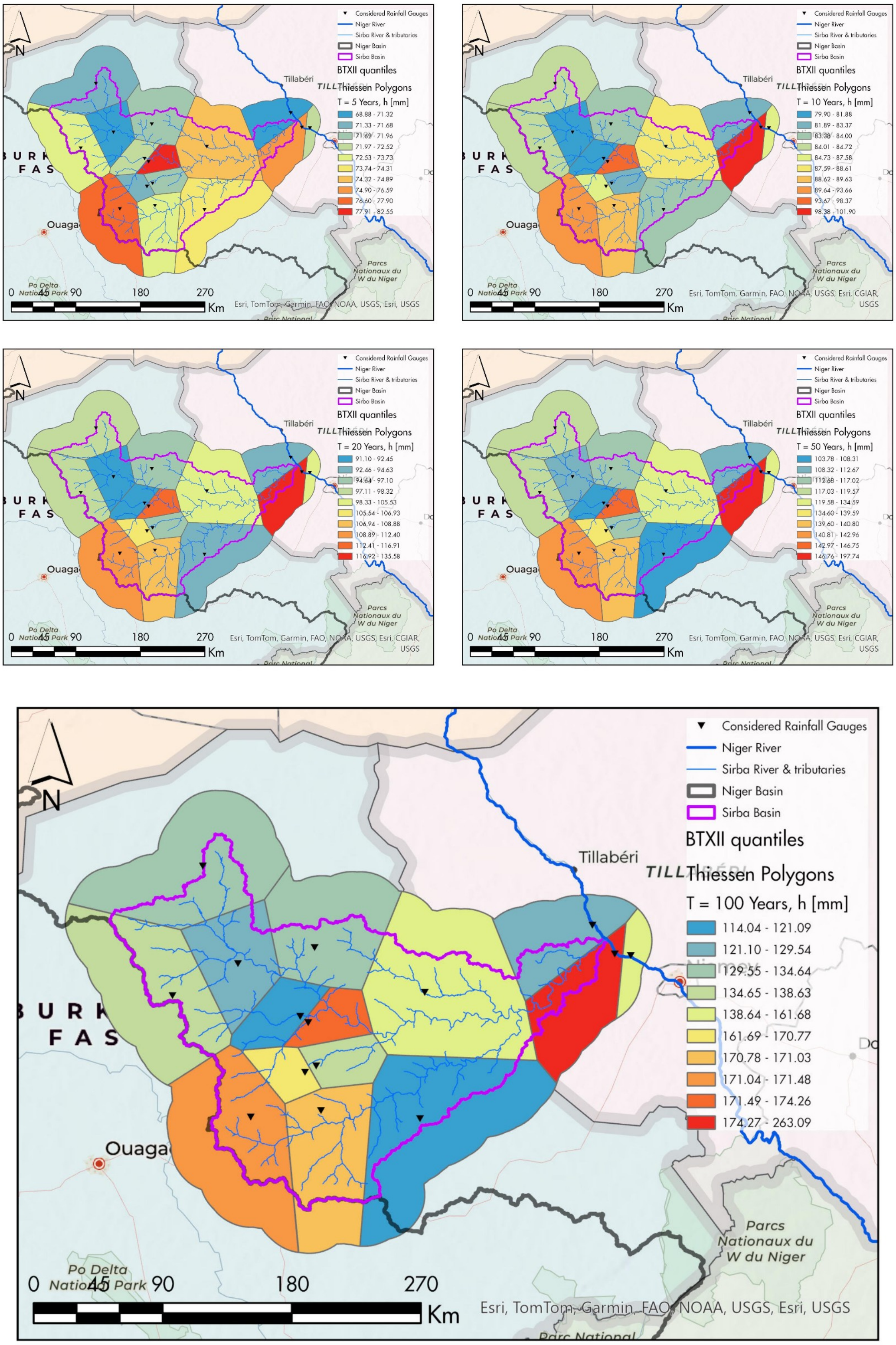


Figure 110: Thiessen polygons of the considered rainfall gauges (stations), representing quantiles of the BTXII distribution, computed using the L-Moments.

BTXII Growth Factors, K(T)

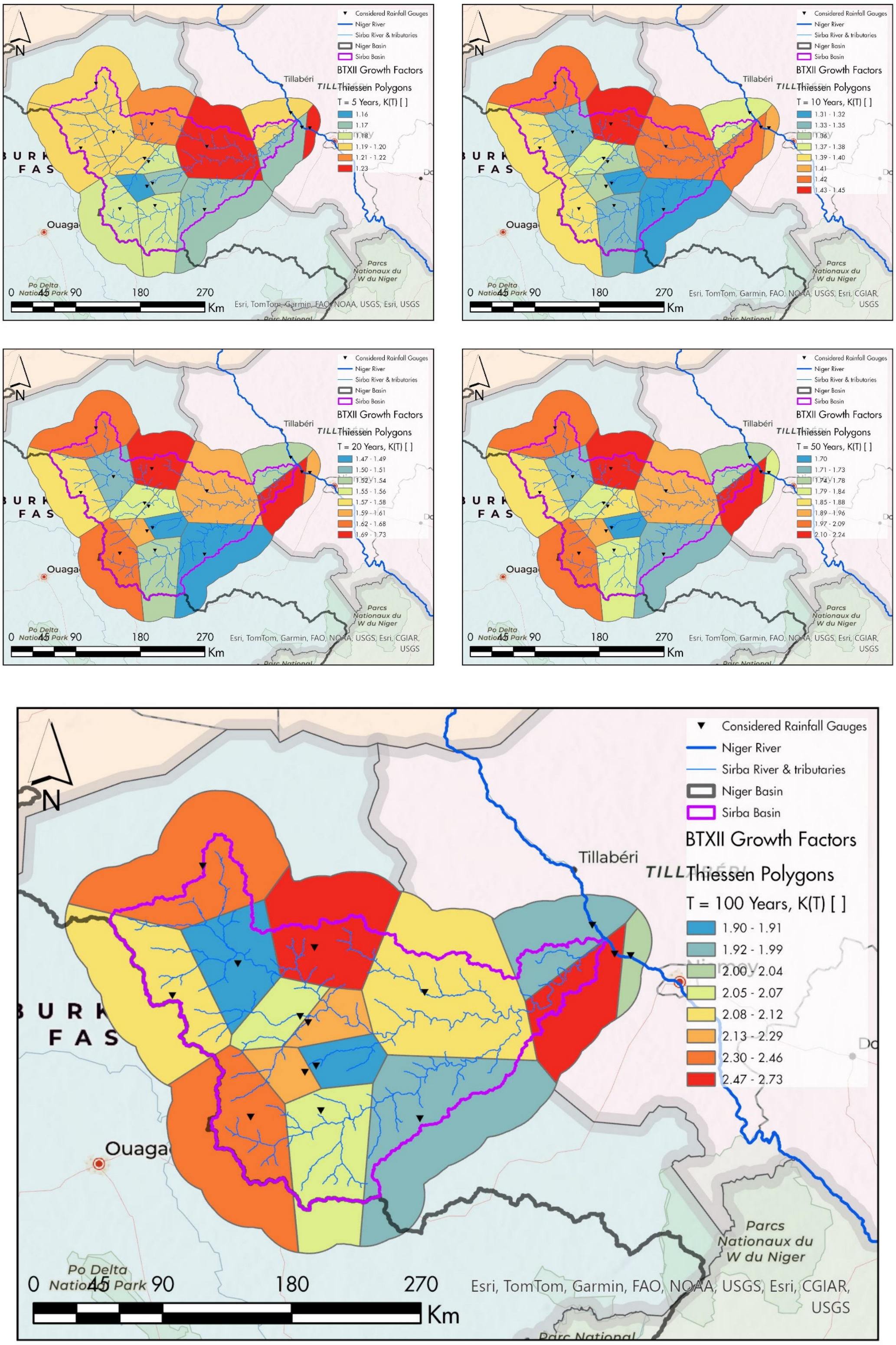


Figure 111: Thiessen polygons of the considered rainfall gauges (stations), representing the growth factors of the BTXII distribution, computed using the L-Moments, and used for the DDF and IDF curves.

8.6.2: ERA5 Local Grid

Max. & L-Moments

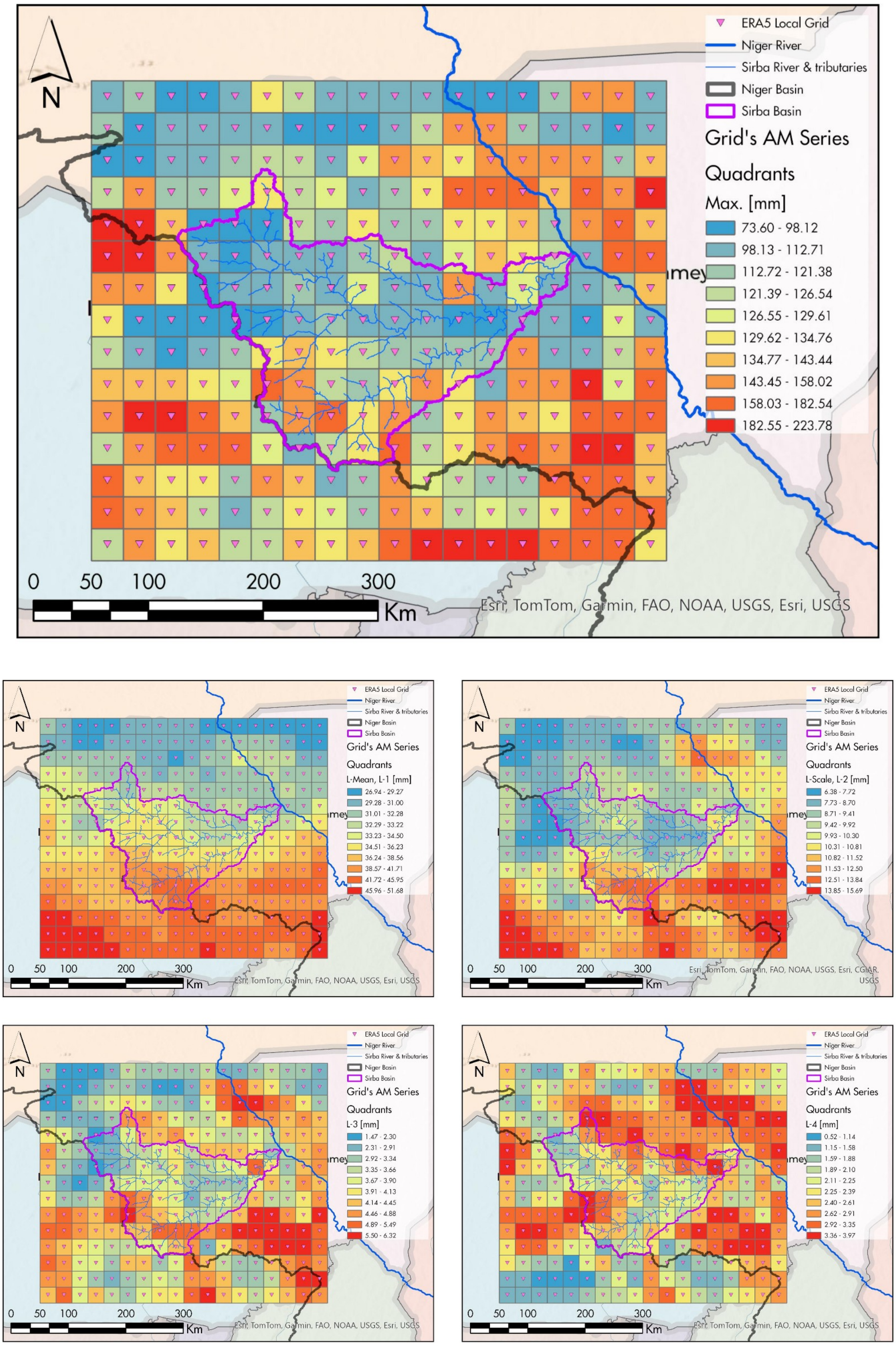


Figure 112: Thiessen polygons (Quadrants) of the initial ERA5 Local Grid, representing Max. and L-Moments of the AM Series of Total Precipitation.