**International Journal of Ecology and Environmental Sciences**
**www.ecologyjournal.in**
**Online ISSN: 2664-7133, Print ISSN: 2664-7125**
**Received: 07-10-2024, Accepted: 06-11-2024, Published: 21-11-2024**
**Volume 6, Issue 4, 2024, Page No. 56-61**

# Statistical analysis of rainfall data: A case study of eastern region in Burkina Faso

**Tog-Noma Patricia Emma Bontogho[1*], Michelline Marie Regina Kansole[1], Mamounata Kabore[2], Gaba Olayemi Ursula Charlène[3]**

[1] Department of High Institute of Sustainable Development, University Yembila Abdoulaye TOGUYENI, Fada N'Gourma, Burkina Faso

[2] Department of Environment, Institute of Environmental Engineering and Sustainable Development, University Joseph Ki Zerbo, Ouagadougou, Burkina Faso

[3] Department of Water, National Institute of Water, University of Abomey-Calavi, Benin

## Abstract

Accurate analysis of rainfall characteristics is essential for water resources management practices mainly in water scarce countries. In this study, analysis of historical rainfall time series at four synoptic stations in Burkina Faso was carried out in order to evaluate the statistical characteristics of rainfall data recorded from 1980 to 2021. In each station, skewness and kurtosis were applied to test the annual rainfall distribution. Similarly, the standard normal homogeneity test (SNHT) and the probability of exceedance were computed. The Mann-Kendall test and Sen's slope estimator were then used to detect trends and estimate their magnitude in the time series data. As a result, a negative skewness is observed in Diapaga station indicating that the rainfall data is skewed to the left. Conversely, the station of Fada N'Gourma, Mahadaga and Pama are showing a positive skewness. The kurtosis is positive across all station indicating a peaked distribution. A maximum rainfall magnitude of 1000mm/year in Diapaga and Fada N'Gourma is exceeded with a 25% probability of occurrence. A maximum rainfall magnitude of 730mm /year and 750 mm/year is exceeded with a 87% probability of occurrence respectively in Mahadaga and Pama. In overall, an occurrence of rainfall more than 1000mm is low while the occurrence of rainfall around 700mm is high. Man Kendall test results in Diapaga (2.0754), Mahadaga (0.32614) and Pama (0.69683) exhibit a significant positive trend, while Fada N'Gourma (-0.50403) shows a non-significant decreasing trend.

**Keywords:** Rainfall, distribution, eastern region, Burkina Faso, probability of exceedance

## Introduction

Climate change is severely constraining the socio-economic development of countries worldwide (Abildtrup *et al* 2006; Kriegler *et al* 2012) [10, 13]. Indeed, in many regions, socio-economic development pathways rely on natural resources, which are highly vulnerable to climate variability. According to the World Bank (2000a) [19], agriculture remains the main rural economic activity in Africa. Unfortunately, crop yields have been decreasing over the last several decades. According to the Intergovernmental Panel on Climate Change (IPCC, 1997) [8], agriculture encompasses approximately one-third of Africa's land area and generates thirty percent (30%) of its gross domestic product. Indeed, occurrence of extremes events including floods, droughts and dry spells has huge consequences mainly on countries having an agriculture-based economy. As stated by Fisher *et al*. (2005) potential negative impacts of climate change present significant challenges to nations, affecting trade patterns, resource utilization, regional planning, and societal well-being. In addition, IPCC (2022) [9] reported that rainfall variability can have adverse impacts on the productivity of agricultural speculation. In Burkina Faso, for instance the economy heavily relies on agriculture, with over 90% of the population engaged in the sector. However, the climate of the country is characterized by high rainfall variability and this has a repercussion on crop productivity. Indeed, significant changes in the rainfall patterns have been observed in many watersheds across the country (West, 2008; Lodoun *et al*. 2013; Nouaceur *et al*. 2020) [14, 15, 16]. As highlighted by Cabral *et al*. (2014) [5], rainfall significantly impacts the agricultural growth, GDP, and poverty rates of rain-fed agricultural countries like Burkina Faso and Senegal. In addition, Barrios *et al*. (2008) [4] emphasized that rainfall has been a significant determinant of poor economic growth for African nations. A wide range of methods for analyzing rainfall characteristic can be found in the litterature. Several works have dealt with statistical analysis of rainfall in Burkina Faso. For onstance, Somé *et al* (2022) [17] applied the coefficient of variation (CV), and the standardized index of cumulated rainfall (SPI) to map the spatial and temporal dynamics of rainfall across Burkina Faso. Yanogo *et al*. (2023) [20] applied standardized rainfall indices, rainfall extremes and rainfall concentration to detect breaks in rainfall time series. This study sets out to assess the statistical patterns of rainfall within eastern Burkina Faso for practical and theoretical perspectives.

## Study area

The eastern region of Burkina Faso (12°15′ Nord and 1°00′, Fig. 1) has an area of 46 256 km$^2$ representing 17% of the national territory. The region is surrounded by the Sahel region in the north, the Niger in the east and Togo in the south-west. The region's topography is mainly characterized by Precambrian metamorphic bedrock. It can be noticed the presence of sandstone and sandstone-schist formations in the southeastern portion, forming the Gobnangou massif. Elevations range from 160 to 330 meters, with the highest point reaching 408 meters in Dymaci, located in the northern part of Gayéri. The dominant soil types include surface sandy soils, deeper sandy-clayey soils, gravelly soils over compact substrates, and outcrops of bedrock. The eastern region of Burkina Faso falls under Sudanian climate characterized by a hot and dry season lasting from October to June. Annual rainfall differs from 750 to 950 mm, with mean temperatures around 26°C. The onset of the rainy season occurred in late June and the rainy season is characterized by dry spells with repercussion on crops yields. The eastern region is part of the Niger and Oti basins. Consequently, the study area is drained in one hand by the Niger and its tributaries (Faga, Sirba, Goroubi, Dyamougou, Tapoa) and the other hand by thePendjari basin tributaries (Kourtiagou, Arly, Doubodo, Singou, and Koulpelogo). Agriculture is the primary economic activity in the study area involving over 90% of the population. However, floods and droughts frequently disrupt the agricultural activities.

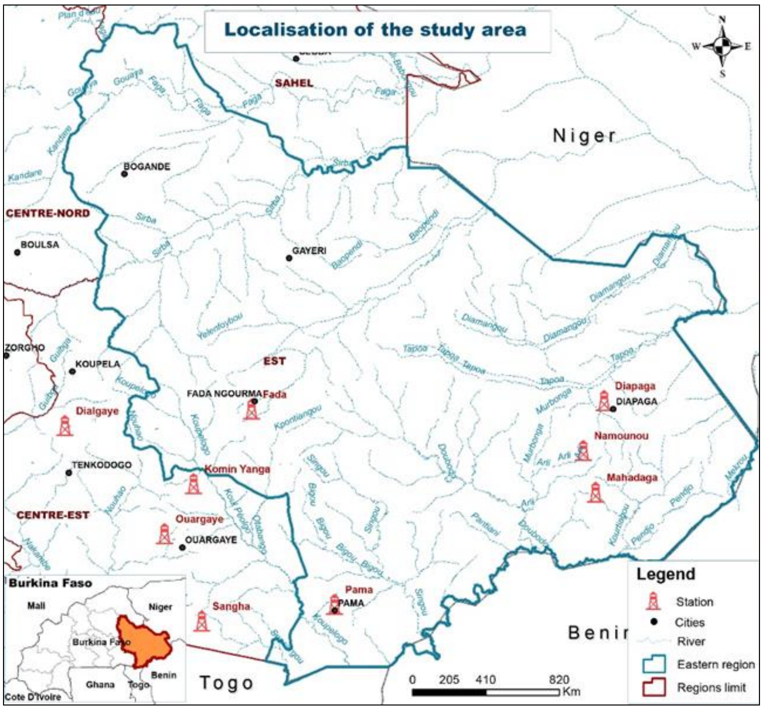



**Fig 1:** Localization of the study area

**Dataset**

This research investigates the statistical analysis of rainfall data in the eastern part of Fada N'Gourma. High quality monthly rainfall records from four synoptic stations (Diapaga, Fada N'Gourma, Mahadaga and Pama) over the period 1980–2021 were obtained from Burkina Faso's National Agency of Meteorology. Basic statistical characteristics of rainfall data has been presented in Table 1.

**Table 1:** Statistical description of dataset

| | DIAPAGA | FADA | MAHADAGA | PAMA |
|---|---|---|---|---|
| Min | 279.0 | 575.0 | 610.1 | 625.0 |
| 1st Qu | 700.2 | 771.2 | 833.0 | 826.2 |
| Median | 814.5 | 821.1 | 922.0 | 931.0 |
| Mean | 839.8 | 839.9 | 921.8 | 927.9 |
| 3rd Qu | 988.2 | 921.4 | 1007.7 | 1041.5 |
| Max | 1262.0 | 1084.0 | 1279.0 | 1287.0 |
| sd | 213.7 | 117.8 | 153.9 | 162.7 |

**Methodology**

**1. Statistical analysis of yearly rainfall**

A set of statistical analyses were conducted on the annual rainfall data including the skewness, the kurtosis, the standard normal homogeneity test (SNHT) and the probability of exceedance. In the skewness and kurtosis test, the objective was to check whether the annual rainfall follow a normal distribution. All statistics were performed using the statistical software R which was developped by the core team of R development. Equation (1) shows the formula of the skewness. The test is based on the density function partitioned into class intervals. The goodness-of-fit of the model to the sample is assessed by the chi-squared statistic. A lower chi-squared value signifies a better fit. The value of the skewness is computed by:

$$S_k = \frac{\frac{\sum_{i=1}^{N}(k_i-\bar{k})^3}{N}}{\sigma_{\bar{k}}^{3}} \qquad \textbf{(Eq.1)}$$

Similarly, the Kurtosis test allows detecting the normality of the data set. The Kurtosis (Eq.2) is defined as the ratio of the mean square successive difference between the years to the variance. The test statistic is given as follow:

$$K_u = \frac{\frac{\sum_{i=1}^{N}(k_i-\bar{k})^4}{N}}{\sigma_{\bar{k}}^{4}} \qquad \textbf{(Eq.2)}$$

The Standard Normal Homogeneity Test **(**SNHT) was developped in 1986 by Alexanderson. A statistic T(d) compares the mean of the first d years of the record with the last of (n − d) years which can be written as follow:

$$T_d = d\bar{z}_1^{\,2} + (n-d)\,\bar{z}_2^{\,2}, d = 1, 2, \ldots., n \qquad \textbf{(Eq.3)}$$

$$\bar{z}_1 = \frac{1}{2}\sum_{i=1}^{d}(Y_i - \bar{Y})/s \qquad \textbf{(Eq.2)}$$

$$\bar{z}_2 = \frac{1}{n-d}\sum_{i=d+1}^{n}(Y_i - \bar{Y})/s \quad \textbf{(Eq.4)}$$

Are the mean values of i z during the first d years and the last (n - d) years respectively. A high T value in year d implies that a break is located in the year d. The test statistic $T_0$ is defined as:

$$T_0 = \max_{1\le d \le 0} T(d) \quad \textbf{(Eq.5)}$$

The probability of rejecting the null hypothesis when $T_0$ exceeds a certain critical value is depended on the sample size. Then the series would be classified as inhomogeneous at a certain level for instance: 95% level of significance.

**Probability plot of the rainfall**

In order to graphically analyze the probability distribution of the annual rainfall in the watershed, a probability plot is realized. According to Dirk *et al*., (2013) [11], a probability plot refers to a plot of the rainfall depths versus their probabilities of exceedance as determined by one or another method. Gupta *et al* (1975) [7] pointed out that probability analysis of rainfall data allows to compute the expected rainfall at various chances. For instance, a rainfall at 80 percent probability can be safely taken as assured rainfall.

The first step in this analysis consists to rank the historical rainfall time series in a descending order. Then, a serial rank number is attributed to the ranked time series with r ranging from 1 to N (N represents the number of the observations). The next step was to compute the probability of exceedance (F) by the method of Hazen in which F is given by the following formula:

$$F = \frac{r - 0.5}{N} \quad \textbf{(Eq.4)}$$

The probability plot is then created by plotting the probability of exceedance versus the annual rainfall. For the purpose of convenient description and understanding, the probability grades is defined as follows: 0–20% is very low, 20%–40% is low, 40%–60% is moderate, 60%–80% is high, and 80%–100% is very high.

**Trend analysis using the Mann–Kendall test**

The statistical significance of the trends in rainfall records was investigated using the Mann–Kendall test. The Mann-Kendall (MK) test is a non-parametric test that does not require the assumption of normally distributed data (Tabari, 2011) [18]. It tests the null hypothesis ($H_0$) of no trend against the alternative hypothesis ($H_a$) of a trend in the data (Koubahe, 2018).

The Mann Kendall z value (Zs) is computed as followed:

$$Zs = \begin{cases} \frac{S-1}{\sqrt{Var(S)}}, & S > 0 \\ 0, & S = 0 \\ \frac{S+1}{\sqrt{Var(S)}}, & S < 0 \end{cases} \quad \textbf{(Eq.5)}$$

A positive Zs value indicates an increasing trend while a negative Zs shows a decreasing trend.

**2.3.4 Sen's Slope Estimator**.

Sen's estimator is a non-parametric method for detecting trends and estimating their magnitude in time series data. It calculates the change per unit time by determining the slope ($T_i$) of all possible data pairs.

The slope of "n" pairs of data based on Sen'slope test is estimated following the equations:

$$\beta n = \text{Median} \left[\frac{x_j - x_k}{j-k}\right] \quad \forall\,(k < j) \quad \textbf{(Eq.6)}$$

Where:

$X_j$ and $X_k$ represent values data at time j and k, respectively, and time j is after time k ($k \le j$).

The median of "n" values of βn is the Sen's slope estimator test.

Negative βn value depicts a downwarding trend; Positive βn value indicates an upwarding trend.

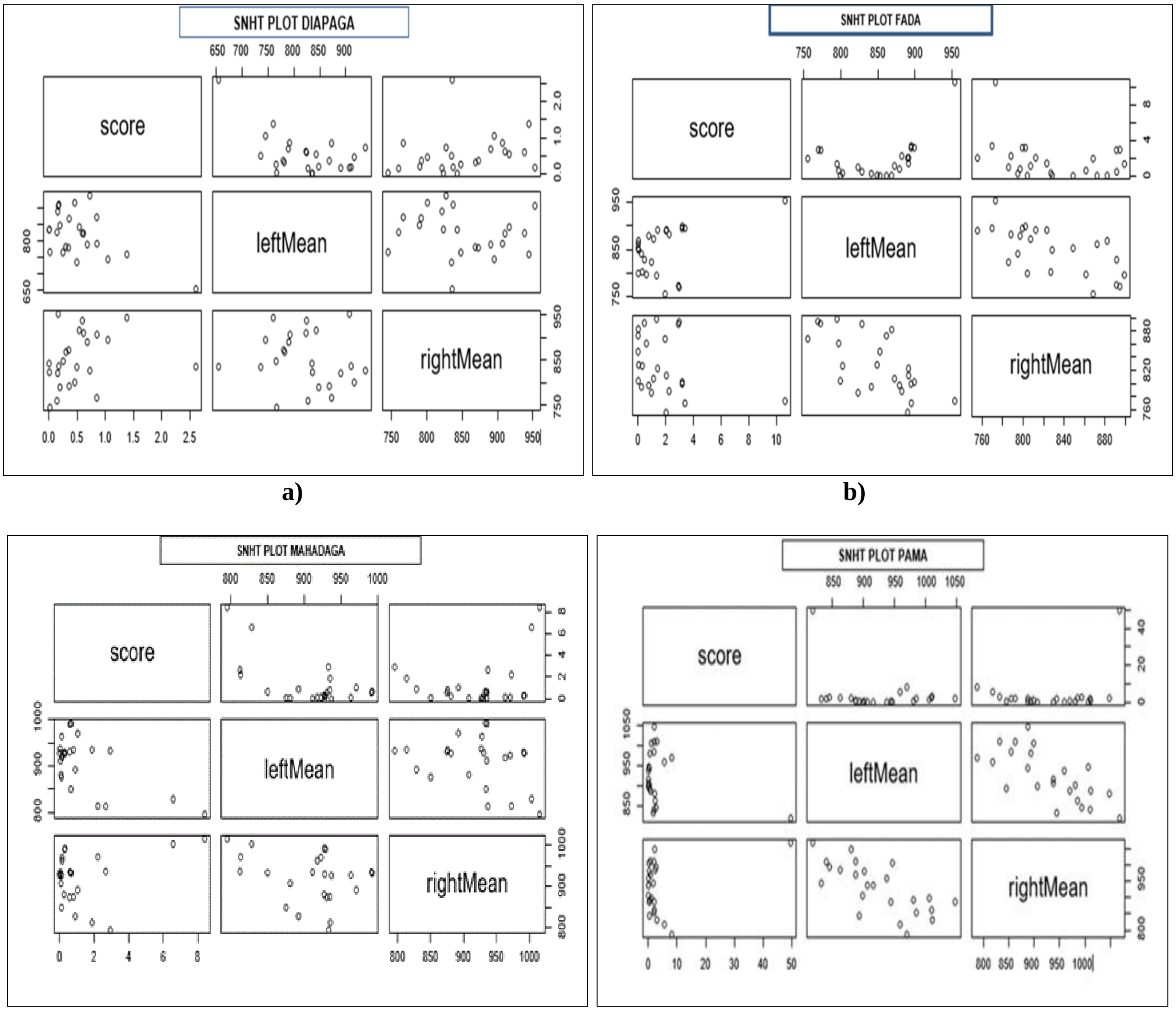
**Results and discussion**

In the present study, rainfall data of eastern Burkina Faso during the period 1980–2021 was analyzed to derive the statistical characteristics of rainfall over the region. The Standard deviation varies from 117.85 mm to 213.68 mm. The coefficient of skewness ranges from -0.019 to 0.203 while the coefficient of kurtosis varies from 2.222 to 3.051. As stated by Azzalini, *et al*. (2005) [3] the skewness for a normal distribution is zero, and any symmetric data should have skewness near zero. Negative values for the skewness indicate that data are skewed to the left and positive values for the skewness indicate that data are skewed to the right (Alaa *et al*; 2014) [1]. A negative skewness is observed in Diapaga station indicating that the rainfall data is skewed to the left. The remaining stations is showing positive a skewness. A dataset with a high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak. The kurtosis is positive in all station. This is indicating a peaked distribution. The rainfall feature analysis especially the distribution and probability of occurrence are necessary for hydro related schemes designing and sustainable water resources management.

**Table 2:** Statistical description of data sets

| | DIAPAGA | FADA | MAHADAGA | PAMA |
|---|---|---|---|---|
| sd | 213.7 | 117.8 | 153.9 | 162.7 |
| Skewness | -0.019 | 0.068 | 0.203 | 0.084 |
| Kurtosis | 3.051 | 2.689 | 2.831 | 2.222 |

Standard Normal Homogeneity Test (SNHT) results for each station are given in figure 2. The mean score is 0.541722 in Diapaga. The value of the maximum leftMean in Diapaga station is 938.2 while the value of the maximum rightMean is around 953.0. The value of the minimum leftMean in Diapaga station is 654.4 while the value of the minimun rightmean is around 745.1 mm. The Standard Normal Homogeneity Test (SNHT) results for Fada N' Gourma station indicated a value of 1.742508 as mean score. The value of the maximun leftMean in this station is 953.6 while the value of the maximun rightmean is around 899.0. The value of the minimun leftMean in Fada N'Gourma station is 755.4 mm while the value of the minimun rightmean is around 755.4 mm. At the station of Mahadaga 1.310859 is recorded as the mean score. The value of the maximun leftMean in is 992.4 while the value of the maximun rightmean is around 1015.7. In addition, the value of the minimun leftMean in Mahadaga station is 795.2 mm while the value of the minimun rightmean is around 795.2 mm. As far as Pama station is concerned, the mean score is 3.67636. The value of the maximun leftMean in Pama is 1048.2 while the value of the maximun rightMean is around 1068.2. The value of the minimun leftMean in Pama station is 818.0 mm while the value of the minimun rightMean is around 787.4 mm.



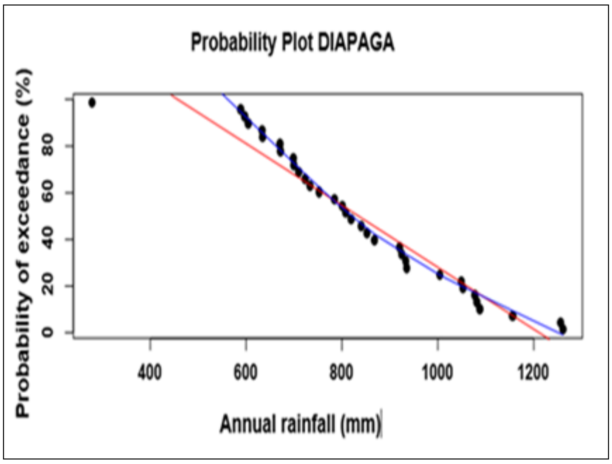

a) b)

c) d)

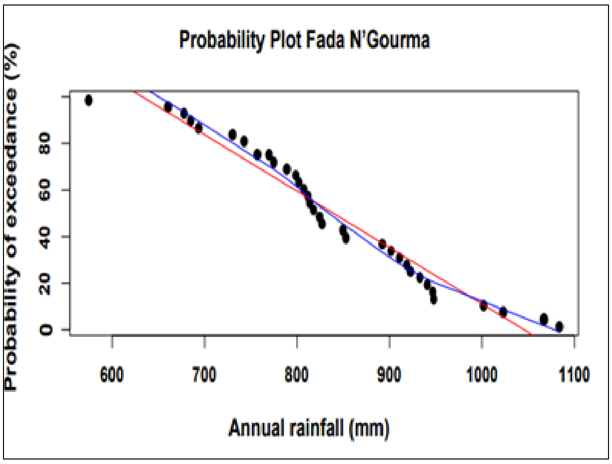
**Fig 2:** SNHT results for Diapaga, Fada N'Gourma, Mahadaga and Pama from 1980 to 2021

**Probability plot**

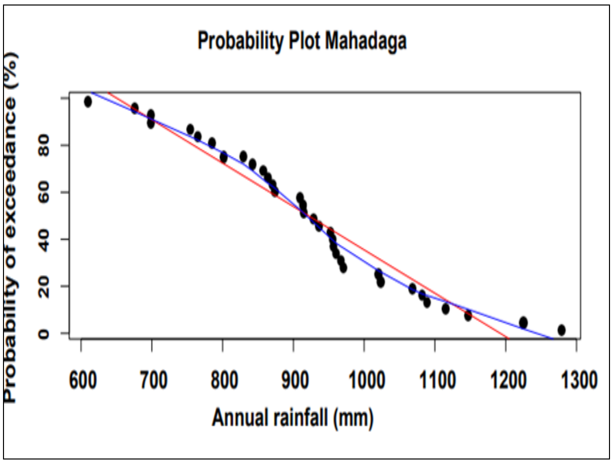
Figure 3a) shows the graphical representation of the probability distribution in Diapaga station. A maximum rainfall magnitude of 1000mm /year in Diapaga is exceeded with a 25 percent probability of occurrence. A maximum rainfall magnitude of 800 mm/year is exceeded with a 75 percent probability of occurrence. Figure 3b) shows the graphical representation of the probability distribution in Fada N'Gourma station. A maximum rainfall magnitude of 1000mm /year in Fada N'Gourma is exceeded with a 25 percent probability of occurrence. A maximum rainfall magnitude of 800 mm/year is exceeded with a 80 percent probability of occurrence. Figure 3c) shows the graphical representation of the probability distribution in Mahadaga station. A maximum rainfall magnitude of 1000mm /year in Mahadaga is exceeded with a 25 percent probability of occurrence. A maximum rainfall magnitude of 750 mm/year is exceeded with a 87 percent probability of occurrence. Figure 3d) shows the graphical representation of the probability distribution in Pama station. A maximum rainfall magnitude of 1000mm /year in Pama is exceeded with a 25 percent probability of occurrence. A maximum rainfall magnitude of 730 mm/year is exceeded with a 87 percent probability of occurrence. A frequency grades for the description of probability is suggested: a frequency grade ranging from 0–20% is very low, 20%–40% is low, 40%–60% is moderate, 60%–80% is high, and 80%–100% is very high. Based on this classification, it come out this study that in the study area, the occurrence of rainfall more than 1000mm is low while the occurrence of rainfall around 700mm is high.



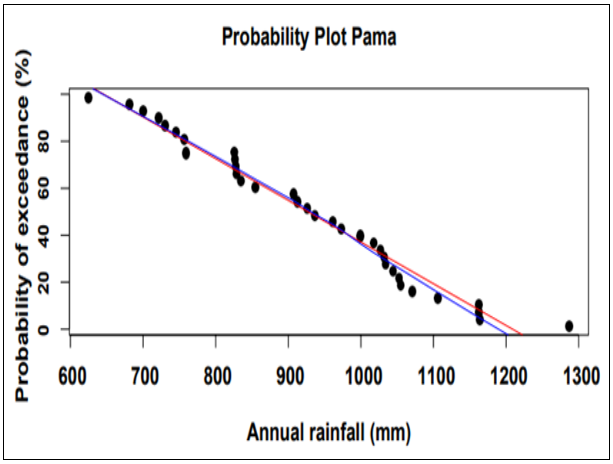

a)




b)




c)




d)

**Fig 3:** Probability plot of Diapaga, Fada N' Gourma, Mahadaga and Pama station

Table 3 provide the results of the non-parametric Mann-Kendall test. The trend of rainfall from 1980 to 2021 has been calculated for each station. In the Mann-Kendall test the Z statistics revealed positive trend of the annual rainfall for Diapaga (2.0754), Mahadaga (0.32614) and Pama (0.69683). However, for Fada N'Gourma Station (-0.50403), there is an evidence of decreasing trend over the period of analysis with a Z value of -0.50403. Thus, Z values for three station show a positive trend and only one station shows a negative trend. Based on associated p-values, non-significant condition is depicted. Diapaga, Mahadaga, and Pama stations exhibit a significant positive trend in annual rainfall, as indicated by their positive Z-statistic values. This suggests a consistent increase in rainfall over the study period.

**Table 3:** Values of the Mann Kendall test for annual rainfall (1980-2021) over eastern region of Burkina Fada

| Mann Kendall | | | |
|---|---|---|---|
| | **z** | **P-value** | **MK-tau** |
| DIAPAGA | 2.0754 | 0.03795 | 0.2513369 |
| FADA | -0.50403 | 0.6142 | -0.06238859 |
| MAHADAGA | 0.32614 | 0.7443 | 4.099822e-02 |
| PAMA | 0.69683 | 0.4859 | 8.563786e-02 |

Table 4 provide the results of the estimation of the rainfall magnitude. The estimated Sen's slope from 1980 to 2021 has been calculated for the station Of Diapaga, Fada N'Gourma, Mahadaga and Pama. Only the station of Fada N'Gourma shows a decreasing trend in the annual rainfall with a sen's slope value of -1.088889. The Sen's slope revealed positive trend of the annual rainfall for Diapaga(2.0754), Mahadaga(0.32614) and Pama(0.69683). Consequently, Fada N'Gourma is experiencing a significant decline in annual rainfall. This decreasing trend could have severe implications for the region, including water scarcity, reduced agricultural productivity, and potential environmental degradation. Diapga, Mahadaga and Pama have experienced a gradual rise in rainfall amounts over the observed period. Diapaga with the highest Sen's slope value (2.0754) shows the most significant positive trend. This could lead to increased water availability, potentially benefiting agriculture and water resources. Within Mahadaga and Pama the trends are less pronounced compared to Diapaga. This could have positive implications for water resources and vegetation in these areas.

**Table 4**: Values of the Mann Kendall test for annual rainfall (1980-2021) over eastern region of Burkina Fada

| Sen's slope | | | |
|---|---|---|---|
| | **z** | **P-value** | **Sen's slope** |
| DIAPAGA | 2.0754 | 0.03795 | 7.772414 |
| FADA | -0.50403 | 0.6142 | -1.088889 |
| MAHADAGA | 0.32614 | 0.7443 | 0.95 |
| PAMA | 0.69683 | 0.4859 | 1.6375 |

**Conclusions**

Statistical characteristics of rainfall are highly significant in hydro related schemes designating. The aim of this paper was to investigate the statistical patterns of rainfall in eastern region of Burkina Faso. Statistical analysis of annual rainfall series for the duration of 1980–2021 were performed using skewness, kurtosis, the standard normal homogeneity test (SNHT) test and the probability of exceedance based on Hazen method. The Mann-Kendall test and Sen's slope estimator were used to detect and quantify trends in annual rainfall.. The analysis has allowed to determine the distribution and the probability of exceedance of the rainfall on a yearly basis. A peaked distribution is identified in all station. The result indicated also that the occurrence of yearly rainfall exceeding 1000mm is low while the occurrence of rainfall less than 750 mm is very high. Further analysis is required to be carried out to in order to establish other rainfall characteristics for sustainable water management perspectives.

**References**


1. Alaa M AL-Lami, Yaseen K AL-Timimi, Ali M AL-Salihi. The Homogeneity Analysis of Rainfall Time Series for Selected Meteorological Stations in Iraq,2014:60:10(2). ISSN: 2222-8373
2. Alexandersson H. A homogeneity test applied to precipitation data, Int. J. Climatol.,1986:6:661-675.
3. Azzalini A. The skew-normal distribution and related multivariate families. Scandinavian journal of statistics,2005:32(2):159-188.
4. Barrios S, Ouattara B, Strobl E. The impact of climatic change on agricultural production: Is it different for Africa?. Food policy,2008:33(4):287-298.
5. Cabral JF. Rainfall and economic growth and poverty: Evidence from Senegal and Burkina Faso, 2014.
6. Fischer G, Shah M, Tubiello NF, Van Velhuizen H. Socio-economic and climate change impacts on agriculture: an integrated assessment, 1990–2080. Philosophical Transactions of the Royal Society B: Biological Sciences,2005:360(1463):2067-2083.
7. Gupta VK, Duckstein L. A stochastic analysis of extreme droughts. Water Resources Research,1975:11(2):221-228.
8. IPCC. Intergovernmental Panel on Climate Change/Organization for Economic Cooperation and Development) Guidelines for National Greenhouse Gas Inventories. OECD/OCDE, Paris, 1997.
9. IPCC. Climate Change 2022: impacts, adaptation and vulnerability contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2022.
10. Abildtrup J, Audsley E, Fekete-Farkas M, Giupponi C, Gylling M, Rosato P, *et al*. Socio-economic scenario development for the assessment of climate change impacts on agricultural land use: a pairwise comparison approach. Environmental science & policy,2006:9(2):101-115.
11. Dirk R. Frequency Analysis of Rainfall Data. In: College on Soil Physics 30th Anniversary (1983-2013), The Abdus Salam International Centre for Theoretical Physica, 2013, 244-10.
12. Koudahe K, Djaman K, Kayode JA, Awokola SO, Adebola AA. Impact of Climate Variability on Crop Yields in Southern Togo. Environ. Pollut. Clim. Chang,2018:2:148. [CrossRef]
13. Kriegler E, O'Neill BC, Hallegatte S, Kram T, Lempert RJ, Moss RH, *et al*. The need for and use of socio-economic scenarios for climate change analysis: a new approach based on shared socio-economic pathways. Global Environmental Change,2012:22(4):807-822.
14. West CT, Roncoli C, Ouattara F. Local perceptions and regional climate trends on the Central Plateau of Burkina Faso. Land Degradation & Development,2008:19(3):289-304.
15. Nouaceur Z, Murarescu O. Rainfall variability and trend analysis of rainfall in West Africa (Senegal, Mauritania, Burkina Faso). Water,2020:12(6):1754.
16. Lodoun T, Giannini A, Traoré PS, Somé L, Sanon M, Vaksmann M, *et al*. Changes in seasonal descriptors of precipitation in Burkina Faso associated with late 20th century drought and recovery in West Africa. Environmental Development,2013:5:96-108.
17. SOMÉ, Yélézouomin Stéphane Corentin, ZOROM, Alimata, SOMÉ, Wièmè, *et al*. Analysis Of The Vegetation Dynamics Of Burkina Faso Using Time Series Of FAPAR Images, 2022
18. Tabari H, Marofi S, Aeini A, Talaee PH, Mohammadi K. Trend analysis of reference evapotranspiration in the western half of Iran. Agric. For. Meteorol,2011:151:128–136. CrossReF
19. World Bank. MAKE Award Winner. Internal Memo, World Bank, 2000a.
20. Yanogo IP, Yaméogo J. Recent rainfall trends between 1990 and 2020: contrasting characteristics between two climate zones in Burkina Faso (West Africa). Glasnik Srpskog Geografskog Drustva,2023:103(1):87-106.