ADV MATH SCI JOURNAL
Advances in Mathematics: Scientific Journal **13** (2024), no.4, 479–490
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.13.4.2

STATISTICAL CHANGE POINT ANALYSIS: APPLICATION TO LAKE CHAD WEATHER DATA

M.A. Issaka[1], A.S. Dabye, and D.A. Njamen

ABSTRACT. In this work, we address the problem of parametric and non-parametric statistical tests by proposing a criterion to detect shifts in precipitation, humidity, and maximum temperature data for Lake Chad from 1925 to 2010. The statistical tests employed include the Buishand, Pettitt, and Mann-Kendall tests. Our analysis indicates that the change point occurred between 1960 and 1990, particularly in precipitation and maximum temperature data. For the monthly analysis, shifts were notably observed in April and August, specifically in temperature and humidity.

1. INTRODUCTION

Test statistics for change point detection are crucial tools used to identify points in a data sequence where the statistical properties, such as the mean, variance, or distribution, change significantly. Commonly used tests include the Buishand [3], Pettitt [12], and Mann-Kendall [8] tests, each of which is designed to detect different types of shifts or trends within the data.

Statistical change point analysis has been explored by many statisticians across various application areas such as bioinformatics [13], climatology [14], signal

[1]*corresponding author*
2020 *Mathematics Subject Classification.* 62F03, 62F05, 62J15, 62P12, 62P30.
Key words and phrases. Change point detection, test statistic, Statistical data analysis, Lake Chad.
Submitted: 09.09.2024; *Accepted:* 24.09.2024; *Published:* 19.10.2024.

processing, and neuroscience [9]. Specifically, detecting change points in data sequences using statistical tests presents a significant challenge in fields like time series analysis, regression models, and estimation problems.

In this work, we focus on the problem of detecting shifts in precipitation, humidity, and maximum temperature data for Lake Chad during the period from 1907 to 2010 using various statistical tests. Our motivation stems from the fact that climate changes, driven primarily by meteorological phenomena, are of great concern to public decision-makers and the population, especially in the Sahel region. Indeed, the shrinking surface area of Lake Chad is a critical research topic in this context.

In the current climate and environmental changes, trends in meteorological parameters, particularly rainfall, are frequently debated. Among the studies that have focused on change point detection with applications in meteorology, Ramli et al. [15] examined rainfall and temperature patterns in the Hadejia River Basin (HRB) using ANOVA and the Mann-Kendall trend test.

Asfaw et al. [6] analyzed changes in rainfall and temperature in northcentral Ethiopia using gridded monthly precipitation data from the Global Precipitation and Climate Centre (GPCC V7) and temperature data from the Climate Research Unit (CRU TS 3.23) with 0.5° by 0.5° resolution from 1901 to 2014. The analysis employed coefficients of variation, anomaly indices, precipitation concentration indices, and the Palmer drought severity index, along with the Mann-Kendall test, to detect trends in the time series. Further studies by Suhaila and Yusop [16], and Getahun et al. [17] discussed trend analysis and change point detection of annual and seasonal temperature series in Peninsular Malaysia and Ethiopia, respectively.

Firstly, we will discuss the data sources used in this study and provide an exploratory description of these data to understand their distribution. In the next section, we introduce the statistical tests used for change point detection, along with the proposed detection model. We then present the results for each test statistic, identifying shifts in annual and monthly data for precipitation, maximum temperature, and humidity. Finally, we conclude with a summary of our findings in the last section.

2. Data

The data used in this study were provided by the National Agency of Meteorology of Chad. It includes monthly data on precipitation, humidity, and maximum temperature collected from 1925 to 2010. The rainy season in most regions of Chad begins in April, so we selected data from the months of April to August for analysis.

3. Test Statistics for Change Point Detection

3.1. **Mann-Kendall Test.** The Mann-Kendall test [8], often used in conjunction with Sen's slope calculation, is the most widely applied non-parametric method for trend estimation. It is particularly effective for detecting trends over time in data that may exhibit autocorrelation. Numerous studies have employed the Mann-Kendall test in various fields, including climate change [7], agriculture [10], and environmental sciences [16], [11].

The test statistic used in the Mann-Kendall test is defined as follows:

$$S_{t,n} = S = \sum_{i=1}^{t} \sum_{j=t+1}^{n} sgn(x_j - x_i), \tag{3.1}$$

where

$$sgn(x_j - x_i) = \begin{cases} 1, & \text{if} \quad (x_j - x_i) > 0, \\ 0, & \text{if} \quad (x_j - x_i) = 0, \\ -1, & \text{if} \quad (x_j - x_i) < 0, \end{cases}$$

where x_i and x_j are the yearly or seasonal values in years i and j (with $j > i$) accordingly. According to [8] and [**?**], it is shown that $\mathbf{E}(\mathbf{S}) = 0$ when the number of observations $n > 10$.

In this case, the variance is defined

$$\mathbf{Var}(\mathbf{S}) = \frac{n(n-1)(2n+5)}{18}. \tag{3.2}$$

Thus, the distribution can be approximated by the standard normal distribution **Z**

$$\mathbf{Z} = \begin{cases} \frac{S-1}{(Var(S))^{\frac{1}{2}}}, & \text{if} \quad S > 0, \\ 0, & \text{if} \quad S = 0, \\ \frac{S+1}{(Var(S))^{\frac{1}{2}}}, & \text{if} \quad S < 0, \end{cases}$$

Otherwise, the proposed test problem in our study is defined by

$$\mathbf{H_0}: \ \mathbf{Z} \leq \mathbf{Z}_{\alpha/\mathbf{2}}$$

versus

$$\mathbf{H_1}: \ \mathbf{Z} > \mathbf{Z}_{\alpha/\mathbf{2}}$$

where H_0 is the absence of trend in the data and H_1 is the presence of almost one trend. By choosing $\alpha = 0.005$ the significance level, $Z_{\alpha/2}$ is a given value.

3.2. **Pettitt's Test.** The Pettitt's test [12] is a non-parametric test used for detecting change points in a sequence of random variables $X_1, X_2, \ldots, X_T$. The test assumes that the sequence has a change point at τ if the sequence X_t for $t = 1, \ldots, \tau$ follows a common distribution $f_1(x)$, and X_t for $t = \tau+1, \ldots, T$ follows a different common distribution $f_2(x)$, where $f_1(x) \neq f_2(x)$.

Using the Mann-Kendall test statistic $S_{t,n}$ from Equation (3.1), the test statistic K for Pettitt's test and the corresponding confidence level (α) for a sample length n can be defined as follows:

$$K_{t,n} = \max_{1 \leq t \leq n} |S_{t,n}| = \max\left(K^+, K^-\right), \tag{3.3}$$

$$\alpha = \exp\left(\frac{-K_{t,n}}{n^2 + n^3}\right), \tag{3.4}$$

3.3. **Buishand Test.** The Buishand test [3] is a parametric test that assumes normality and independence of the data.

Assuming that the change point is at k, the Buishand test statistic U_k is defined as:

$$U_k = \frac{\sum_{i=1}^{k} \left(\frac{S_i}{\sigma_x}\right)^2}{n(n+1)}, \tag{3.5}$$

where

$$S_k = \sum_{i=1}^{k} (x_i - \overline{x}), \quad \forall k = 1, \ldots, n. \tag{3.6}$$

Here:

- σ_x is the standard deviation of the series,

- $\overline{x}$ is the mean of the time series observations $(x_1, \ldots, x_n)$,
- k is the index at which a change point is suspected.

To detect the location of the change point, the distribution of the Buishand test statistic is given by:

$$B = \max_{0 \leq k \leq n} |U_k| . \tag{3.7}$$

The value of $B/\sqrt{n}$ is compared with the critical value provided by [3]. If the calculated value is less than the critical value, the null hypothesis is not rejected. Otherwise, if the calculated value exceeds the critical value, the null hypothesis is rejected.

4. Results

4.1. **Change Point Detection with the Buishand Test.** The Buishand change-point detection test was applied to identify shifts in the annual and monthly series of maximum temperature, precipitation, and humidity data for the period 1960–2010.

As illustrated numerically in Table 1 and graphically in Fig. 1, the test results indicate a significant increase in precipitation during April and August, which correlates with rising temperatures. For the annual data, the most common month where a shift was detected is August.

Furthermore, the analysis reveals that the precipitation in the Lake Chad region experienced a notable increase starting from 1960, peaking around 1980. Following this period, a marked decrease in both monthly and annual precipitation trends was observed between 1981 and 1995, aligning with the well-documented droughts of the 1980s. Additionally, there was a notable rise in humidity from 1990 to 2008.

These results highlight significant temporal variations in precipitation

4.2. **Change Point Detection with Pettitt's Test.** In this analysis, we utilized the Pettitt test statistic to detect changes in precipitation, maximum temperature, and humidity data. As illustrated in Fig. 2 and summarized in Table 2, the results reveal significant shifts in the annual data for precipitation, maximum temperature, and humidity between 1960 and 1990.

TABLE 1. Monthly and annual change point analysis of precipitation, temperature, and humidity using the Buishand test

Period	Precipitation			Maximum temperature			Humidity		
	Statistic	Year	Shift	Statistic	Year	Shift	Statistic	Year	Shift
April	0.087	1961	Yes	0.38	1908	No	0.42	1981	Yes
May	0.068	1959	No	0.56	1965	Yes	0.13	2007	No
June	0.221	1967	No	0.59	1967	Yes	0.43	1980	Yes
July	0.127	1996	No	0.40	1914	No	0.28	1986	No
August	0.682	1960	Yes	0.44	1914	No	0.74	1980	Yes
Annual	0.48	1960	Yes	0.32	1908	No	0.58	1980	Yes

Specifically, the change point analysis for humidity indicates that a change occurred in 1989, affecting the series for April, June, and August. Similarly, significant changes were detected in August and annually in the years 1959 and 1989.

TABLE 2. Monthly and annual change point analysis for precipitation, temperature, and humidity using the Pettitt test

Period	Precipitation			Maximum temperature			Humidity		
	Statistic	Year	Shift	Statistic	Year	Shift	Statistic	Year	Shift
April	329	1997	No	419	1987	Yes	0.42	1989	Yes
May	282	1980	No	387	1980	No	0.13	2007	No
June	409	1967	Yes	556	1959	Yes	0.43	1989	Yes
July	344	1986	No	551	1940	No	0.28	1986	No
August	649	1967	Yes	396	1940	No	0.74	1989	Yes
Annual	558	1960	Yes	432	1987	Yes	0.58	1989	Yes

4.3. **Change Point Detection with the Mann-Kendall Test.** The Mann-Kendall test, a widely utilized non-parametric method for trend analysis, was applied to detect change points in the time series data of precipitation, maximum temperature, and humidity. The results of this analysis are summarized in Table 3.
Analysis of Results:

- **Precipitation:** The Mann-Kendall test did not reveal any significant change points for precipitation across April, May, June, July, August, or the annual data. Although the statistic for August in 1952 suggests a possible change, this is not supported by the results from other months or the annual series.
- **Maximum Temperature:** Significant changes were detected in July in 1952, with a test statistic of 0.13. This suggests a notable shift in the

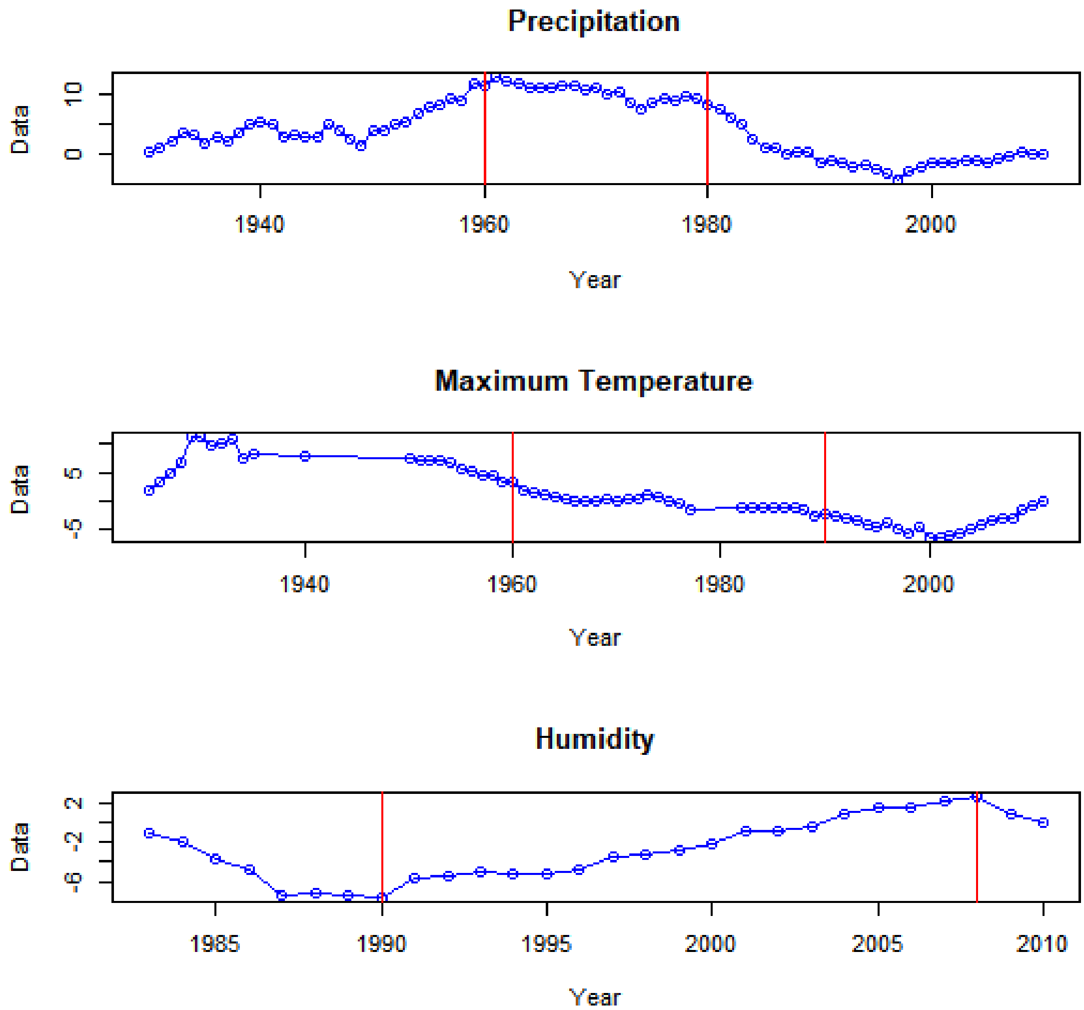

FIGURE 1. Buishand shift detection

trend of maximum temperatures during this period. However, no other significant changes were observed in the remaining months or the annual data.

- **Humidity:** The Mann-Kendall test identified significant changes in humidity in August 1984 and for the annual data in 1991. These results highlight shifts in humidity trends during these years. No significant changes were detected for April, May, June, or July.

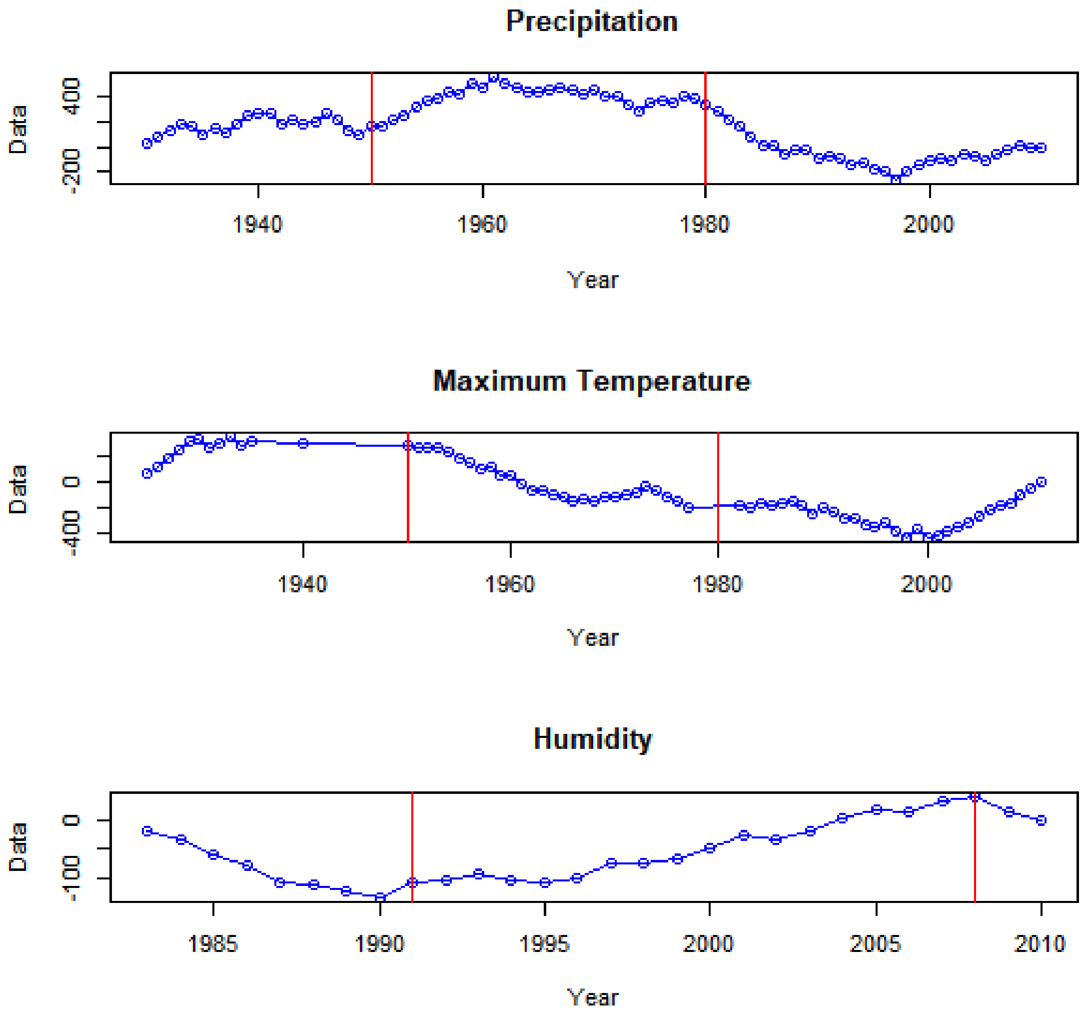

FIGURE 2. Pettitt shift detection

The Mann-Kendall test primarily identified change points in the maximum temperature and humidity series, while the precipitation series did not show significant changes. These findings indicate that notable variations in temperature and humidity occurred during specific periods, offering valuable insights into climatic trends over the decades. Such insights are crucial for further investigations into climate change and its effects on various meteorological parameters.

TABLE 3. Monthly and annual change point analysis for the rain precipitation, temperature and humidity with Mann-Kendall test

Period	Precipitation			Maximum temperature			Humidity		
	Statistic	Year	Shift	Statistic	Year	Shift	Statistic	Year	Shift
April	-0.41	-	No	1.27	-	No	1.29	-	No
May	-0.85	-	No	2.05	-	No	0.63	-	No
June	-0.64	-	No	2.57	-	No	1.85	-	No
July	0.64	-	No	0.13	1952	Yes	0.92	-	No
August	-2.11	1952	Yes	-	1914	No	2.26	1984	Yes
Annual	-1.35	-	No	0.32	-	No	2.21	1991	Yes

Additional Observations: As illustrated in Figure 3 and detailed in Table 3, the Mann-Kendall test results show that significant shifts in humidity and maximum temperature occurred in August during the period from 1950 to 1991. For precipitation, the analysis indicated a change point only for August, as demonstrated by the test statistic.

5. CONCLUSION

In this study, we conducted a comprehensive analysis of change points in annual and monthly series of maximum temperature, precipitation, and humidity data spanning from 1925 to 2010. We employed a range of parametric and non-parametric statistical tests, including the Buishand, Pettitt, and Mann-Kendall tests, to identify significant shifts and trends within these meteorological variables.

Our analysis revealed several key findings:

(1) **Precipitation and Maximum Temperature:** The Buishand and Pettitt tests identified notable change points primarily between 1960 and 1990. This period is characterized by significant shifts in both precipitation and maximum temperature data. Specifically, the shifts were detected in April and August for precipitation, and in August for temperature, indicating substantial variations in these variables during this timeframe.

(2) **Humidity:** The Pettitt and Mann-Kendall tests indicated changes in humidity starting from August, with significant shifts occurring between 1960 and 1990. This period saw notable variations in humidity levels, which are crucial for understanding climatic trends and their impact on the environment.

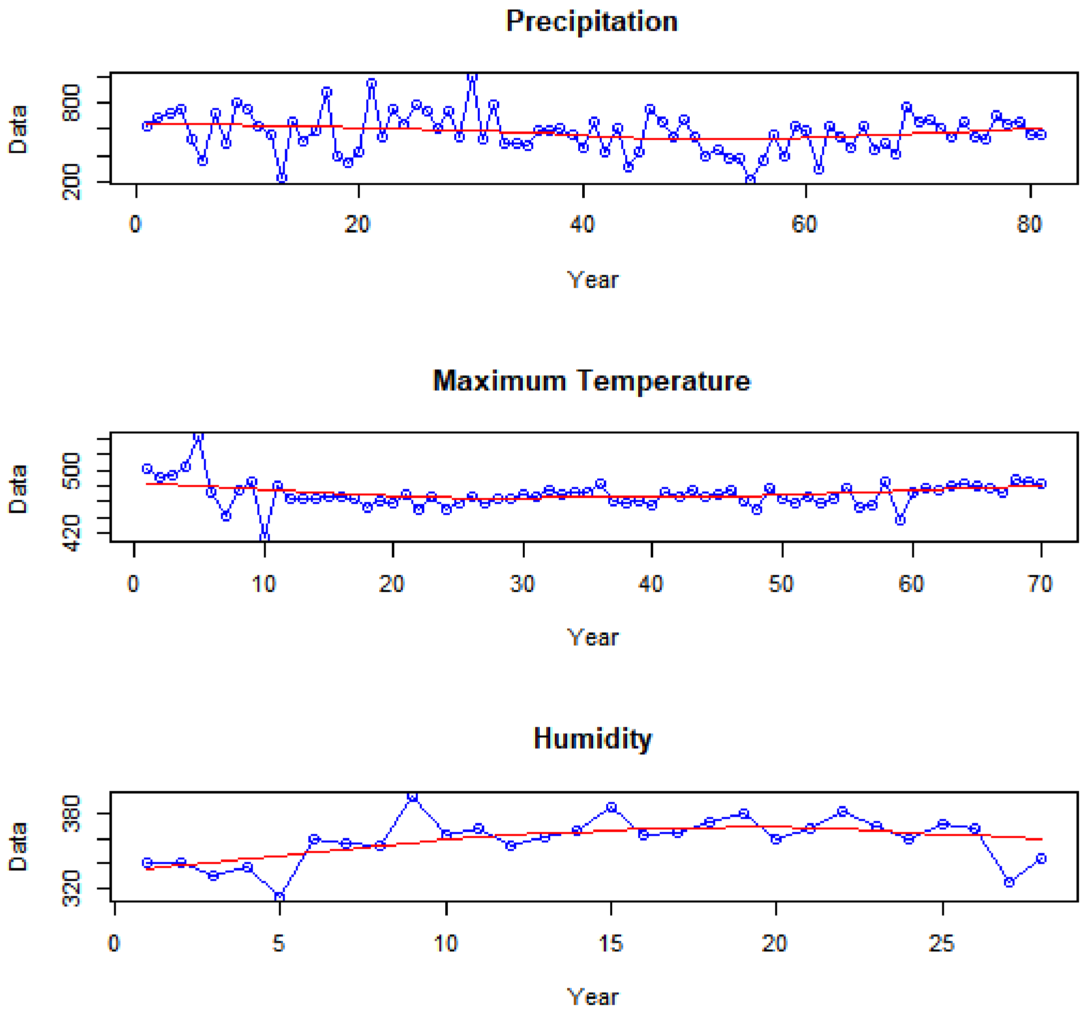

FIGURE 3. Mann-Kendall Shift detection

The results underscore the importance of these change points in interpreting historical climate data. The identification of shifts in maximum temperature, precipitation, and humidity provides valuable insights into the evolution of climatic conditions over the decades. These findings are essential for understanding past climate variability and for making informed predictions about future climate trends.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the National Agency of Meteorology for providing the data essential for this study. Their support and access to the meteorological data have been invaluable in conducting the research and achieving the results presented in this work. We also acknowledge the contribution of all individuals and organizations who have supported and facilitated this research endeavor.

REFERENCES

[1] A. ASFAW, B. SIMANE, A. HASSEN, A. BANTIDER: *Variability and time series trend analysis of rainfall and temperature in northcentral Ethiopia: a case study in Woleka sub-basin*, Weather and Climate Extremes, **19** (2018), 29–41.

[2] A. ALEMAYEHU, W. BEWKET: *Local spatiotemporal variability and trends in rainfall and temperature in the central highlands of Ethiopia*, Geografiska Annaler: Series A, Physical Geography, **99**(2) (2017), 85–101.

[3] T.A. BUISHAND: *Some methods for testing the homogeneity of rainfall records*, Journal of Hydrology, **58**(1-2) (1982), 11–27.

[4] J. CHEN, A.K. GUPTA: *Parametric Statistical Change Point Analysis: With Applications to Genetics*, Medicine and Finance, Boston: Birkhäuser, 2012.

[5] C. LEE, M. CSÖRGŐ, L. HORVÁTH: *Limit Theorems in Change-Point Analysis*. Wiley, Chichester, 1997.

[6] A. ASFAW, B. SIMANE, A. HASSEN, A. BANTIDER: *Variability and time series trend analysis of rainfall and temperature in northcentral Ethiopia: a case study in Woleka sub-basin*, Weather and Climate Extremes, **19** (2018), 29–41.

[7] A. ALEMAYEHU, W. BEWKET: *Local spatiotemporal variability and trends in rainfall and temperature in the central highlands of Ethiopia*, Geografiska Annaler: Series A, Physical Geography, **99**(2) (2017), 85–101.

[8] H.B. MANN: *Nonparametric tests against trend*, Econometrica, **13** (1945), 245–259.

[9] K.J. OH, M.S. MOON, T.Y. KIM: *Variance change-point detection via artificial neural networks for data separation*, Neurocomputing, **68** (2005), 239–250.

[10] O.I. OJO, M.F. ILUNGA: *Application of nonparametric trend technique for estimation of onset and cessation of rainfall*, Air, Soil and Water Research, **11**(2018), 1–4.

[11] C. ONYUTHA: *Graphical-statistical method to explore variability of hydrological time series*, Nordic Hydrology, **52**(1) (2021), 266–283.

[12] A.N. PETTITT: *A Nonparametric Approach to the Change-Point Problem*, Applied Statistics, **28**(1979), 126–135.

[13] F. Picard, S. Robin, M. Lavielle, C. Vaisse, J. Daudin: *A statistical approach for array CGH data analysis*, BMC Bioinformatics, **6**(1) (2005).

[14] J. Reeves, J. Chen, X.L. Wang, R. Lund: *A review and comparison of changepoint detection techniques for climate data*, Journal of Applied Meteorology and Climatology, **46**(6) (2007), 900–915.

[15] M.F. Ramli, A.Z. Aris, N.R. Jamil, A.A. Aderemi: *Evidence of climate variability from rainfall and temperature fluctuations in semi-arid region of the tropics*, Atmospheric Research, **224**(2019), 52–64.

[16] J. Suhaila, Z. Yusop: *Trend analysis and change-point detection of annual and seasonal temperature series in Peninsular Malaysia*, Meteorology and Atmospheric Physics, **130**(5) (2018), 565–581.

[17] Y.S. Getahun, et al.: *Trend and change-point detection analysis of rainfall and temperature over the Awash River basin of Ethiopia*, Heliyon, 2021.

[18] T.A. Buishand: *Some methods for testing the homogeneity of rainfall records*, Journal of Hydrology, **58**(1-2) (1982), 11–27.

[19] A.N. Pettitt: *A Nonparametric Approach to the Change-Point Problem*, Applied Statistics, **28** (1979), 126–135.

[20] H.B. Mann: *Nonparametric tests against trend*, Econometrica, **13**(1945), 245–259.

Laboratoire de Modelisation Mathematique, Informatique, Applications et Simulations (L2MIAS), Université de N'Djamena, B.P. 1027, N'Djamena, Chad.

Université Roi Fayçal du Tchad, N'Djamena, Chad.
Email address: mhtali1@gmail.com

LERSTAD UFR SAT, Université Gaston Berger, Saint Louis, Senegal.
Email address: dabye−ali@yahoo.fr

Department of Mathematics, University of Maroua, Maroua,, Cameroon.
Email address: didiernjamen1@gmail.com