



# Statistical Change Point Analysis in Air Temperature and Rainfall Time Series for Cocoa Research Institute of Nigeria, Ibadan, Oyo State, Nigeria

N. P. Dibal<sup>1,\*</sup>, M. Mustapha<sup>1</sup>, Adegoke T. M.<sup>2,\*</sup>, A. M. Yahaya<sup>1</sup>

<sup>1</sup>Department of Mathematics and Statistics, University of Maiduguri, Maiduguri, Nigeria

<sup>2</sup>Department of Statistics, University of Ilorin, Ilorin, Nigeria

## Email address:

[pdibal@gmail.com](mailto:pdibal@gmail.com) (N. P. Dibal), [adegoketaiwom@gmail.com](mailto:adegoketaiwom@gmail.com) (T. M. Adegoke)

\*Corresponding author

## To cite this article:

N. P. Dibal, M. Mustapha, Adegoke T. M., A. M. Yahaya. Statistical Change Point Analysis in Air Temperature and Rainfall Time Series for Cocoa Research Institute of Nigeria, Ibadan, Oyo State, Nigeria. *International Journal of Applied Mathematics and Theoretical Physics*. Vol. 3, No. 4, 2017, pp. 92-96. doi: 10.11648/j.ijamtp.20170304.13

Received: August 27, 2016; Accepted: October 20, 2017; Published: December 13, 2017

---

**Abstract:** In this work, a cumsum approach is used to detect change-point in mean of an independent normal random variables. A multiple shift in the mean level was considered and show how such a problem can be straightforwardly addressed through the cumsum approach. Data gotten from Cocoa Research Institute of Nigeria were used and from the result of the analysis, a single change point was detected in the amount of rainfall and a multiple change point was detected in the amount of minimum temperature and no change point was detected in the amount of maximum temperature.

**Keywords:** Change Point, Cumsum

---

## 1. Background

Change point detection is an important part of statistical and economic analysis. Predictions and statistical inference will be invalid if changes in the regimes during the data collection period are not taken into account. The main problems in the change point analysis are to decide whether the statistical model for a series of observations does not change (no change situation) or whether the model changes one or more times and in the latter case to identify when the changes have occurred. Climate variability and their impacts as well as associated vulnerabilities are growing concerns worldwide. Global warming induced changes in temperature and rainfall are already evident in many parts of the world, as well as in India. Hazards likelihoods, droughts, cyclones and others, which may have aggravated due to climate change, are being experienced more frequently in India during the past few decades. Against this background, studies to detect climate change and its various impacts deserve urgent attention. Lowering of agricultural productivity, increased risk of hunger and water scarcity, rapid melting of glaciers and decrease in river flows - all such issues are being

discussed in the context of climate change [1]. In the recent past (1971-2003), the warming trend has accelerated at a rate of 0.22C/10 years [2]. Future projections of climate change using global and regional climate models run by the Indian Institute of Tropical Meteorology (IITM) with different IPCC scenarios indicated a temperature change of about 35C and an increase of 5-10% in summer monsoon rainfall [3, 4]. It is also projected that the number of rainy days may decrease by 20-30%, which implies that the intensity of rainfall is likely to increase. This study is aimed to determining if indeed a change occurs in the annual mean rainfall, annual mean maximum temperature and annual mean minimum temperature time series. The change-point analysis is capable of detecting multiple changes. For each change it provides detailed information including a confidence level indicating the likelihood that a change occurred and a confidence interval in time scale associated with the change. Inhomogeneous climatic data series can bring inaccuracies and make misinterpretation in the investigation of climate change that is visible apparently. The factors causing variations on the long term time series are, locations of the stations, instruments, formulae used to calculate means,

observing practices and station environment [5]. Hence, it is essential for detecting significant change points in air temperature time series with the established method.

## 2. Study Area

Oyo State covers approximately an area of 28,454 square kilometers and is ranked 14<sup>th</sup> by size, it is bounded in the south by Ogun State, in the north by Kwara State, in the west it is partly bounded by Ogun State and partly by the Republic of Benin, while in the East by Osun State.

## 3. Methods

The time series data used in this work includes three variables: annual mean, maximum and annual mean minimum air temperature recorded at Cocoa Research Institute of Nigeria, Ibadan, Oyo State, Nigeria. The period under consideration ranges from 1990 - 2016. The collected monthly data were statistically processed and using software finally annual mean values were calculated for further analysis. The studies of long-term climate change require that data be homogenous. Observed climate abrupt changes in a homogenous climate time series are caused by variations in weather and climate [5]. In recent times, several scientific studies have been conducted on quality control and homogenization of climatologically data for the detection of climate trends [6, 7, 8].

The cumulative sum charts (CUMSUM) and bootstrapping were performed as suggested by [8]. Let  $x_1, x_2, \dots, x_n$ , represents  $n$  data points of a time series, and  $\Sigma_1, \Sigma_2, \dots, \Sigma_n$  are iteratively computed as follows

1. The average  $x_1, x_2, \dots, x_n$  is given by

$$\frac{x_1 + x_2 + \dots + x_n}{n} \tag{1}$$

2. Let  $\Sigma_0$  be equal to zero
3.  $\Sigma_1$  are computed recursively as follows

$$\Sigma_i = \Sigma_{i-1} + (x_i - \bar{x}), I = 1, 2, \dots, n \tag{2}$$

Actually, the cumulative sums are not the cumulative sums of the values. Instead they are the cumulative sums of differences between the values and the average. These differences sum to zero so the cumulative sum always ends at zero,  $\Sigma_{i=0}$ . The confidence level can be determined for the apparent change by performing a bootstrap analysis. Before performing the bootstrap analysis, an estimator of the magnitude of the change is required. One choice, which works well regardless of the distribution and despite multiple change is,  $\Delta$  which is defined as

$$\Delta_j = \max_j \sum - \min_i \sum_{1 \leq i \leq n, 1 \leq j \leq n} \tag{3}$$

Once the estimator of the magnitude of the change has been selected, the bootstrap analysis can be performed. A single bootstrap is performed by:

1. Generating a bootstrap sample of  $n$  data points of time

series, denoted as  $x$  ( $j=1, 2, 3, \dots, n$ ), by randomly reordering the original  $n$  values. This is called sampling without replacement (SWOR).

2. Based on the bootstrap sample, the bootstrap CUSUM is calculated following the same method and denoted as,  $\Sigma_i$ .
3. The maximum, minimum and difference of the bootstrap CUSUM are calculated and the difference between the maximum and minimum bootstrap CUSUM is defined as,

$$\Delta_j = \max_j \sum - \min_i \sum_{1 \leq i \leq n, 1 \leq j \leq n} \tag{4}$$

Determine whether the,  $\Delta_j < \Delta_i$

The bootstrap analysis consists of performing a large number of bootstraps and counting the number of bootstraps for which bootstrap difference  $\Delta_j$  is less than the original difference  $\Delta_i$ . Let  $N$  be the number of bootstrap samples performed and let  $K$  be the number of  $j$  bootstraps for,  $\Delta_j < \Delta_i$ . Then the confidence level that a change occurred as a percentage is calculated as follows:

$$\text{Confidence Level } CL = \frac{k}{N} 100 \tag{5}$$

Bootstrapping results is a distribution free approach with only a single assumption, that of an independent error structure. Once a change has been detected, an estimate of when the change occurred can be made. One such estimator is the CUSUM estimator. Let  $i = m$ , such that:

$$|\Sigma_m| = \max_j |\Sigma_j| \tag{6}$$

Then  $m$  is the point furthest from zero in the CUSUM chart. The point  $m$  estimates last point before the change occurred. The point  $m+1$  estimate the first point after the change. The second estimator of when the change occurred is the mean square error (MSE) estimator. Let MSE ( $m$ ) be defined as:

$$MSE(m) = \sum_{i=1}^m (x_i - \bar{x}_1)(x_i - \bar{x}_2) \tag{7}$$

Where

$$\bar{x}_1 = \frac{\sum_{i=1}^m}{m} \text{ and } \bar{x}_2 = \frac{\sum_{i=m+1}^n}{n - m} \tag{8}$$

In MSE estimation, the data series is splitted into two segments, 1 to  $m$ , and  $m+1$  to  $n$ , then it is estimated that how well the data in each segment fits their corresponding averages. The value of  $m$ , for which MSE ( $m$ ) is minimized, gives the best estimate of the last point before change, while the point  $m+1$  denote the first point after change. In the same way data of each segment can be passed through the above method to find level 2 change points that divides corresponding segments into sub-segments. Repetition of the procedure mentioned above helps finding significant change points at subsequent levels for each of which associated confidence limits and levels can be determined by bootstrapping. In this manner

multiple change points can be detected by incorporating additional change points each at successive passes that will continue to split the segments into two. Variations and trends of air temperature and amount of rainfall in Oyo State were examined following the method mentioned above. The cumulative sum charts (CUSUM) and bootstrapping were used

for the detection of abrupt changes. Section of the CUSUM chart with an ascending trend indicates a period when the values remaining above the overall average. Likewise, a section with a descending trend indicates a period of time where the values lie below the overall average. The confidence level can be determined by performing bootstrap analysis.

### 4. Data Analysis

**CUSUM Chart of minimum temperature**

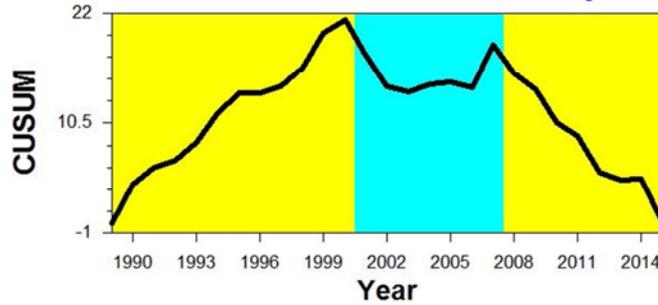


Figure 1. Cumsum chart of minimum temperature.

**CUSUM Chart of maximum temperature**

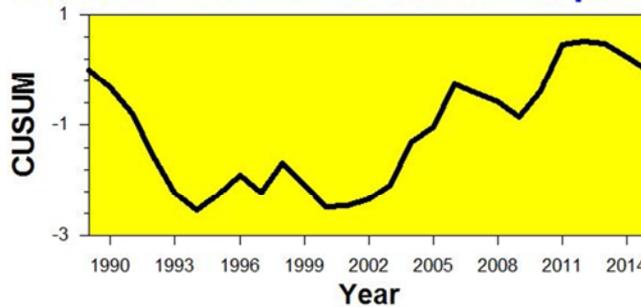


Figure 2. Cumsum chart of maximum temperature.

**CUSUM Chart of rainfall**

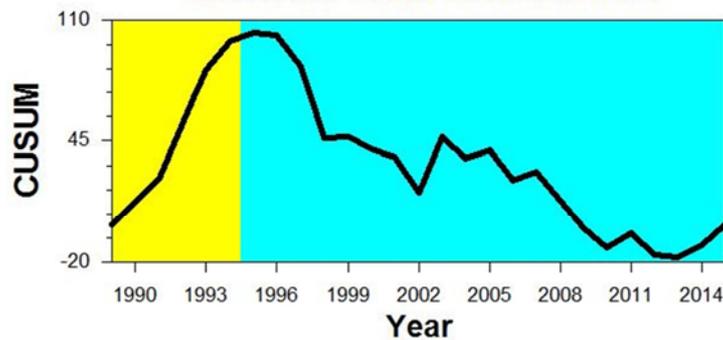


Figure 3. Cusum chart of rainfall.

**Table of Significant Changes for minimum temperature**

Confidence Level for Candidate Changes = 50%, Confidence Level for Inclusion in Table = 90%, Confidence Interval = 95%, Bootstraps = 1000, Without Replacement, MSE Estimates

Year	Confidence Interval	Conf. Level	From	To	Level
2001	(1998, 2007)	98%	19.967	17.643	1
2008	(2002, 2013)	93%	17.643	15.696	2

Figure 4. Significant changes in mean of minimum temperature.

### Table of Significant Changes for rainfall

Confidence Level for Candidate Changes = 50%, Confidence Level for Inclusion in Table = 90%, Confidence Interval = 95%, Bootstraps = 1000, Without Replacement, MSE Estimates

Year	Confidence Interval	Conf. Level	From	To	Level
1995	(1995, 2000)	95%	123.85	99.32	2

Figure 5. Significant changes in mean of rainfall.

### No Significant Changes for maximum temperature

Confidence Level for Candidate Changes = 50%, Confidence Level for Inclusion in Table = 90%, Confidence Interval = 95%, Bootstraps = 1000, Without Replacement, MSE Estimates

Estimated Average = 27.394346

Figure 6. No significant changes for maximum temperature.

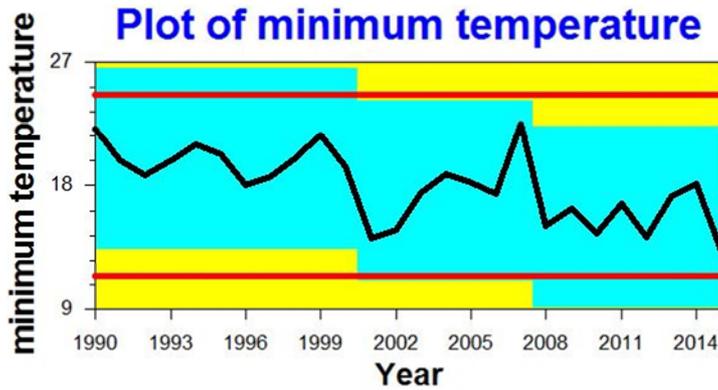


Figure 7. Trend plot for minimum temperature.

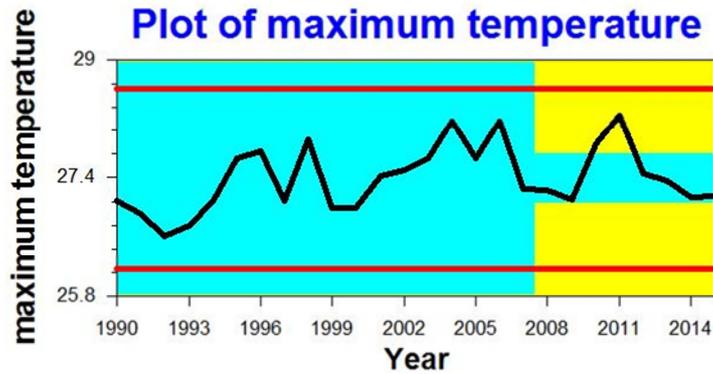


Figure 8. Trend plot for maximum temperature.

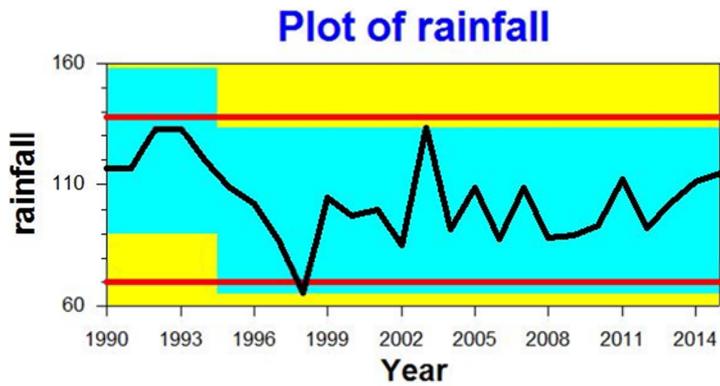


Figure 9. Trend plot for rainfall.

## 5. Discussion and Conclusion

The results of the analysis for detection of change points in annual rainfall, annual mean maximum temperature and annual mean minimum temperature of Oyo state are presented in Figures 1 - 3. The shaded background in this figure indicates the event of change over the considered period and maximum range of temperature fluctuation indicated by red lines under the situation of no change in trend. For rainfall and minimum temperature, a change(s) do actually occurred while no change point detection was detected in maximum temperature. The confidence levels of those changes are presented in Figures 3, 4 and 5. According to this method, the level 1 change signifies the first change that admittedly present in the CUSUM chart. The year in which level 1 change in mean of minimum temperature occurred is 2001 (Figures 4 and 7). Prior to the level 1 change, there is a change found at level 2. By excuse of the change point analysis from the independent error structure, no outlier's assumptions were made in maximum temperature, minimum temperature and the rainfall. For the level 1 change, the confidence interval is restricted within one year only. Prior to this change in 2001, the minimum temperature was 19.967°C; while after the change, the average minimum temperature became 17.643°C with a confident interval of 98%. The amount of change of minimum temperature is 2.324°C. It should be noted that, it is the minimum amount of change found in the whole analysis. The widest interval is associated with the change that occurred in 2008 (confidence interval 2002-2013) at level 2. For the rainfall time series there are also distinct change points at level 2. The rainfall time series exhibits a level 2 change in 2001 (confidence interval 1998, 2001) at a confidence level of 95%. Prior to the level 1 change of rainfall time series in 1995, the mean rainfall was 23.85mm; while after the change, average rainfall became 99.23mm. The amount of change of mean rainfall is 75.38mm. In case of maximum temperature, Figures 2, 4, and 8 shows that indeed no change point was detected occurred.

## Acknowledgements

The authors are grateful to the Editors, Associate Editor, and referees whose comments and suggestions were valuable

to improve the exposition of the paper.

---

## References

- [1] IPCC. Summary for policymakers. climate change. 2007.
- [2] D. R. Kothawale and K. K. Rupa. On the recent changes in surface temperature trends over india. *Geophys Res Lett*, 32, 2005.
- [3] NATCOM. Indias initial national communication to the united nations framework convention on climate change. Technical report, National Communication Project, Ministry of Environment and Forests, Government of India., 2004.
- [4] Nebojsa Nakicenovic, Joseph Alcamo, Gerald Davis, Bert de Vries, Joergen Fenhann, Stuart Gaffin, Kenneth Gregory, Arnulf Grübler, Tae Yong Jung, Tom Kram, Emilio Lebre La Rovere, Laurie Michaelis, Shunsuke Mori, Tsuneyuki Morita, William Pepper, Hugh Pitcher, Lynn Price, Keywan Riahi, Alexander Roehrl, Hans-Holger Rogner, Alexei Sankovski, Michael Schlesinger, Priyadarshi Shukla, Steven Smith, Robert Swart, Sascha van Rooijen, Nadejda Victor and Zhou Dadi (2000), Special report on emissions scenarios: A special report of working group iii of the intergovernmental panel on climate change. Technical report, Cambridge University Press, Cambridge, UK, 2000.
- [5] C. Serra, A. Burgueno, and X. Lana (2014), Analysis of maximum and minimum daily temperatures recorded at fabran observatory in the period 1971-1998. *International Journal of Climatology*, 21: 617(636).
- [6] T. Szentimrey, J. Salinger, E. J. Førland, I. Hanssen-Bauer, H. Alexandersson, P. Jones and D. Parker (1998), Homogeneity adjustments of in situ atmospheric climate data: A review. *International journal of climatology*, 18: 1493(1517).
- [7] T. R. Karl and C. N. Williams (1987), An approach to adjusting climatological time series for discontinuous inhomogeneities. *Journal of Climate & Applied Meteorology*, 26: 1744(1763).
- [8] Taylor, Wayne (2000), Change-point Analyzer 2.0 shareware program, Taylor Enterprises, Libertyville, Illinois. Web: <http://www.variation.com/cpa>.
- [9] Shrestha AB, Wake CP, Dibb JE, Mayewski PA (2000) Precipitation fluctuations in the Nepal Himalaya and its vicinity and relationship with some large scale climatological parameters. *International Journal of Climatology* 20: 317-327.