IOSR Journal of Applied Physics (IOSR-JAP)
e-ISSN: 2278-4861.Volume 11, Issue 4 Ser. I (Jul. – Aug. 2019), PP 58-63
www.iosrjournals.org

Analysis of Precipitation Data Using Empirical Orthogonal Functions

M. H. Ali[1] and Sunusi Usman Yerima[2]
[1] Department of Physics, Bayero University, Kano.
[2]Nnamdi Azikwe International Airport Abuja, Nigeria..
Corresponding Author: M. H. Ali

Abstract: *This work presents the analysis of a 90-year (1905-1995) precipitation data of an observatory station. A 90-year precipitation data from Kano State observatory station were acquired and analyzed using Empirical Orthogonal Function and Principal Component techniques. The results show that months of July, August and September contribute about 76% of the annual precipitation for this station. In particular, August was found to contribute about 38% of the total annual rain while the months of July and September each contribute about 23% and 15% of the annual precipitation respectively. These quantitative ratios could help farmers and government in decision making with regards to crop production and sometimes flood management of the area under study.*

Date of Submission: 28-06-2019 Date of acceptance: 15-07-2019

I. Introduction

Understanding the trend and pattern distribution of atmospheric parameters like precipitation is fundamental to many applications in geosciences and climate research. Water is the major controller of the Earth's evolution. Its movement through hydrological cycles has been identified as the largest movement of any substance on Earth. As such, the water influence on climate changes and variations from one region to another cannot be overemphasized. Hydrological cycles give rise to the exchange of moisture and heat between the Earth and the atmosphere. More than 10^{17} Kg of water drops as rain from the terrestrial atmosphere yearly [1]. The temporal and spatial distribution of this huge precipitation over the terrestrial earth varies from one region of the Earth to another.

Essentially, precipitation records would be very useful for researches in global and regional hydrological cycle and climate variability. Analysis of such data could help in verification and calibration of satellite based climate data and also in global circulation and forecast models [1]. As such, investigations on the trend and patterns of precipitation of a given area are very useful for water management, agriculture, electrical power generation and flood and drought controls. However, most weather phenomena are controlled by large number of variables or parameters. Methods of investigations in such phenomena are mostly multivariate [2].

Researchers used different methods of analysis to study certain hydrologic properties of a given terrestrial area. Positive matrix factorization method was used by [3] while cluster analysis (CA) was adopted by [4] in their studies of precipitation data. [5] made use of the non-parametric method to investigate the trend of the precipitation in Turkey. They were able to conduct trend detection analysis on thirteen precipitation variables using their 1929 to 1993 records from96 stations across Turkey. Since rain gauge cannot be placed everywhere, making in-situ data from rain gauge networks unavailable for some places [6], mathematical modeling sometimes is adopted in the study of rainfall characteristics. The uncertainties arising from the use of such techniques are investigated by[7]. Empirical Orthogonal Functions (EOF) are sometime used to estimate future rainfalls of a given location where previous records were used in generating the EOF as done by [8] on 50-year monthly records and in [9].

In this work, principal component analysis was again used to investigate the monthly distribution of the annual precipitation of a meteorological center through the 90 year precipitation records of the center.

Method

A 90-year precipitation records obtained from NIMET office in Kano was used in this study. The data are monthly averaged values for the station. The period of the record is 1905 to 1995 contained in 90 rows by 14 columns in a matrix form. Principal Component Analysis (PCA) was used in analyzing the data.

The trend of the data was removed and the resultant records in matrix formed the Covariance of the records [10]. This last matrix was decomposed using singular value decomposition into eigenvectors (Empirical Orthogonal Functions, EOF) and eigenvalues matrices [11]. The eigenvalue matrix, L was then normalized to identify the most significant EOFs. Three EOFs; EOF12,EOF11 and EOF10 corresponding to the months of August, July and September respectively, were found to explain 76% variance of the data. The corresponding principal components; PC12, PC11 and PC10 were then computed and used in reconstructing the entire records for comparison [12][13].The two sets of data were compared graphically.

II. Results And Discussion

The eigenvector matrix, C is given in table 1.

Table 1. The matrix of the eigenvectors (Empirical Orthogonal Functions)

0	0.9995	-0.0293	-0.01	-0.0016	-0.0007	-0.0009	-0.0002	-0.0001	0.0002	0	0
0	0.009	-0.0326	0.9992	-0.0093	-0.0189	-0.003	0.0036	0.0032	-0.0015	0.0011	0.0017
0	0.0019	0.0051	0.0097	0.9998	0.0073	-0.0064	-0.0113	0.0035	0.0023	0.0004	0.0009
0	0.0002	-0.0015	-0.0072	0.0075	-0.5405	0.8242	-0.1529	0.0312	0.017	-0.0595	-0.0156
0	-0.0002	-0.0025	0.0031	-0.0131	0.0896	-0.1217	-0.9049	0.3956	0.011	0.0262	0.0298
0	-0.0001	0.0004	-0.0001	-0.0015	0.0259	0.0091	0.1863	0.4037	0.8826	-0.1392	0.0559
0	-0.0001	-0.0018	0.0009	-0.001	0.0005	0.051	-0.1099	-0.3083	0.3001	0.892	0.0668
0	-0.0001	0	-0.0009	-0.0006	0.0028	0.0112	-0.032	-0.1362	-0.0125	-0.121	0.9826
0	0.0003	0.0007	-0.0043	0.0021	0.0124	0.0808	0.3312	0.7509	-0.3608	0.4052	0.1592
0	-0.0013	-0.0017	-0.0177	0.0027	-0.8358	-0.5445	0.0122	0.0452	0.0124	0.0431	0.0207
0	0.0296	0.999	0.0323	-0.0055	-0.0027	-0.0001	-0.0029	-0.0001	0.0004	0.0014	0.0001
1	0	0	0	0	0	0	0	0	0	0	0

The eigenvalues matrix L is a diagonal matrix with the following elements along its major diagonal; 0, 3.39519792167092, 15.7474308466477, 101.696039243021, 2457.02370079352, 26692.5156217529, 33682.9858061816, 191194.562581702, 30153.484759628, 293080.086369760, 467633.556368927, 769272.472209319.

The transposed matrix of the normalized version of L is given in Table 2.

Table 2. Normalized values of L (Converted to row vector)

0	0	0	0.0001	0.0012	0.0133	0.0167	0.0949	0.1143	0.1455	0.2322	0.3819

The above table shows that three EOFs corresponding to the percentages of the variance of 38.19%, 23.22% and 14.55% contained about 76% of the total variance of the data. The corresponding principal components were computed and compared to identify the months corresponding to the three EOFs. Some of the results of the comparison are shown in the following figures.

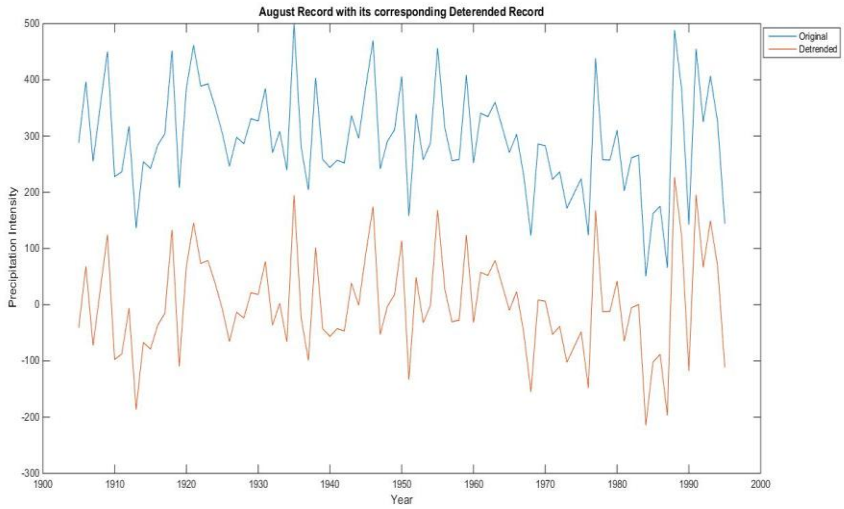


Figure 1.Temporal Variation of the Precipitation for the Month of August Using Original Data Compared with the Corresponding Detrended Data.

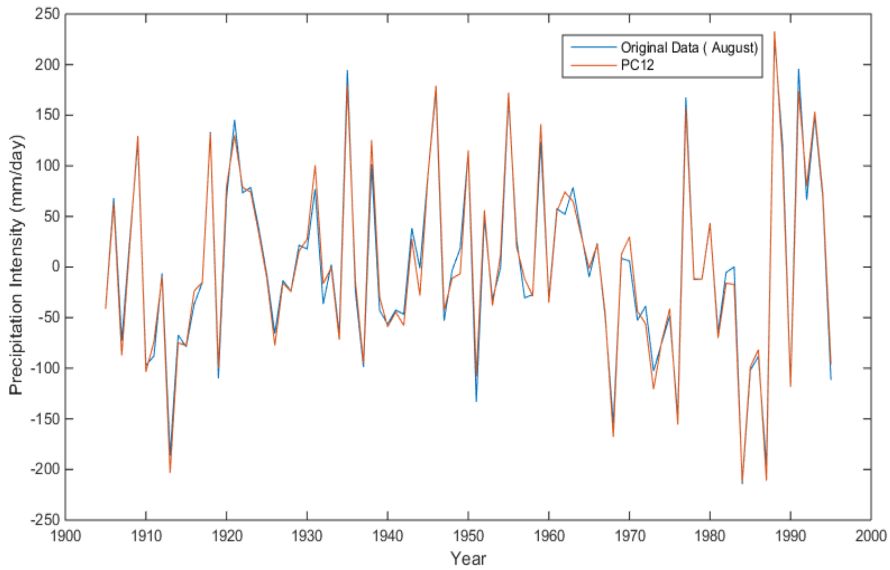


Figure 2.Detrendedrecord for the month of August compared with the Principal Component (pc12) computed from EOF12 (38.19%).

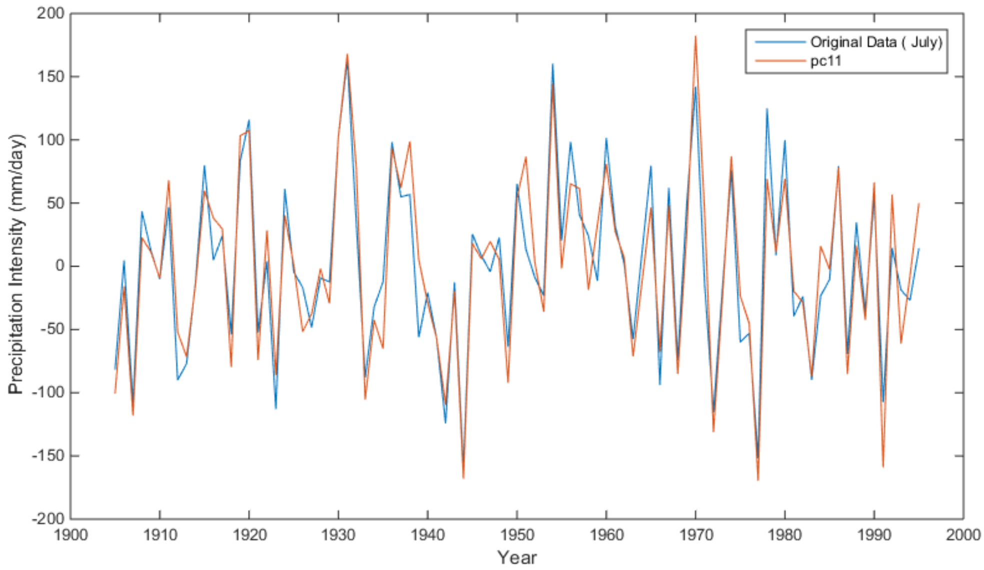


Figure 3.Detrended record for the month of July compared with the Principal Component (pc11) computed from EOF11 (23.22%).

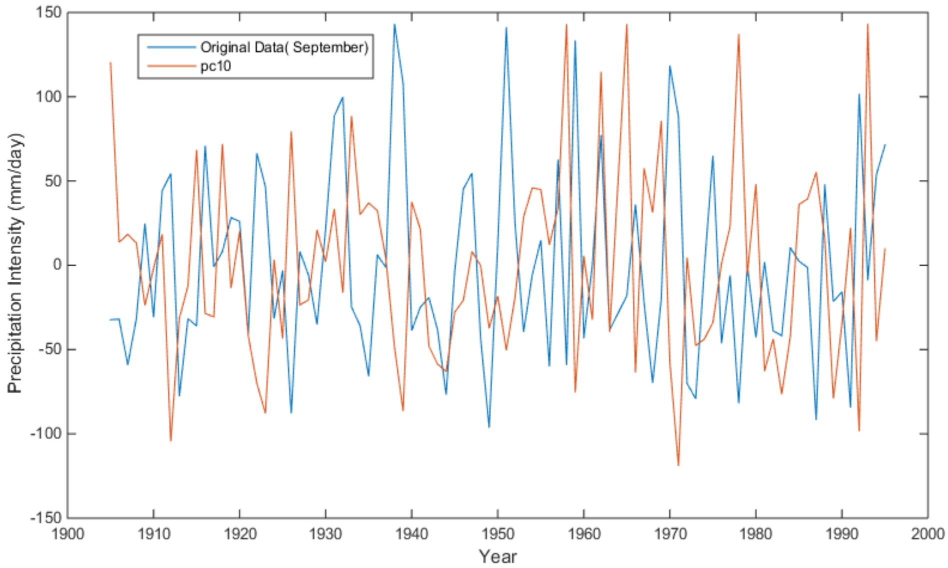


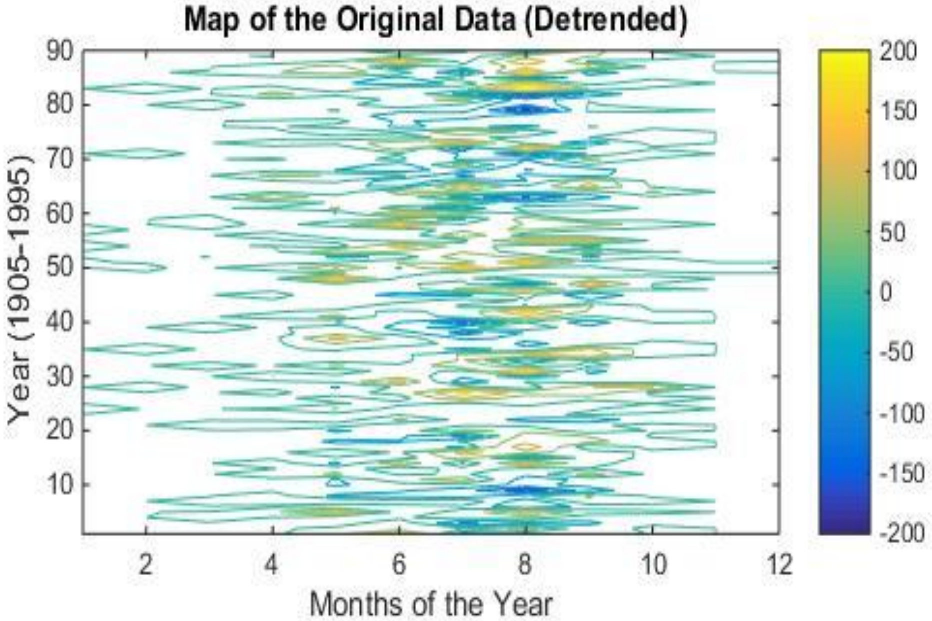
Figure 4. Detrended record for the month of September compared with the Principal Component (pc10) computed from EOF10 (14.55%).

Based on the results in the Figures 2,3 and 4 the contribution of the three identified months are tabulated below.

Table 3. Percent Variance Explained by Different EOF Representing a Month of a Year

EOF	Percent Variance (%)	Month
EOF 12	38.19	August
EOF 11	23.22	July
EOF 10	14.55	September

This table shows that for the 90 year period of this study, the month of August contributed up to about 38% of the total annual precipitation. The months of July and September had up to 23% and 15% of the total annual rain respectively. This make the three months to have a total of about 76% contribution of the annual precipitation of the study area. The original data were the reconstructed using the three EOFs and their corresponding principal components. The two set of data were then compared graphically as seen in Figure 5.



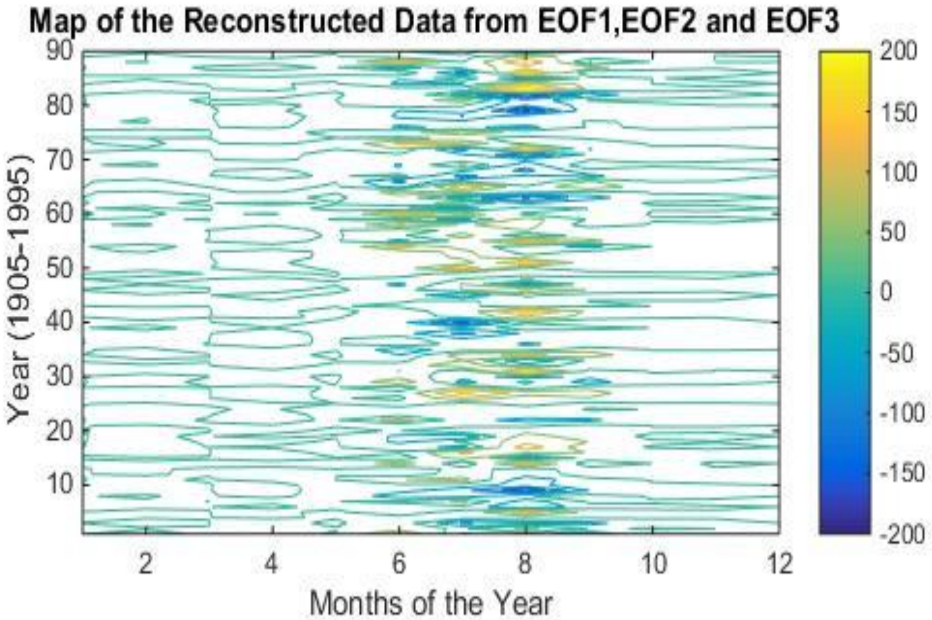


Figure 5 Monthly Mean Precipitation Variation Map over the Period of Study giving by the Original Record and the one Reconstructed from the three leading EOFs.

In Figure 5, the y-axis represent the years (1905-1995); 0- 10 represents 1905-1915, 10-20 represents 1915-1925 etc. while the x-axis represents the number of the months starting from January to December. The color bar on the right hand side of the plot shows the quantitative amount of the precipitation in mm.

Considering the month 8 (August) in Figure 5, all the prominent features in the original data seemed to have been accounted for in the reconstructed data. The less rain in the session of 1915, 1965 and 1985 in the original data are equally indicated in the reconstructed data. Similarly, heavy rain in 1955 and 1982 are also captured in the reconstructed data.

III. Conclusion

In conclusion, we have used the 90-year precipitation data from a meteorological station and identified estimates of the monthly contribution of the annual rainfall of the station. The months of August, July and September were found to contribute up to 76% of the total annual precipitation of the studied area. The method used was the principal component analysis in which empirical orthogonal functions corresponding to each month of a year were used. Month of August accounted for 38.19% of the annual rain while July and September contributed up to 23.22% and 14.55% of the annual precipitation respectively

References

[1] M. T. Chahine, "The hydrological cycle and its influence on climate," *Nature, 359, 6394*, vol. 359, no. 6394. pp. 373–380, 1992.

[2] F. C. G. Gerald, R. N., Thomas, L. B. and Robert, "Sampling Errors in the Estimation of Empirical Orthogonal Functions," *Am. Meteorol. Soc.*, no. July, 1982.

[3] S. Juntto and P. Paatero, "Analysis of daily precipitation data by positive matrix factorization," *Environmetrics*, vol. 5, no. January, pp. 127–144, 1994.

[4] Xiaofeng Gong and M. B. Richman, "On the application of cluster analysis to growing season precipitation data in North America east of the Rockies," *Journal of Climate*, vol. 8, no. 4. pp. 897–931, 1995.

[5] T. Partal and E. Kahya, "Trend analysis in Turkish precipitation data," *Hydrol. Process.*, vol. 20, no. 9, pp. 2011–2026, 2006.

[6] C. Beck, J. Grieser, B. Rudolf, and U. Schneider, "A new monthly precipitation climatology for the global land areas for the period 1951 to 2000," *Geophys. Res. Abstr.*, vol. 7, pp. 181–190, 2005.

[7] M. Strauch, C. Bernhofer, S. Koide, M. Volk, C. Lorz, and F. Makeschin, "Using precipitation data ensemble for uncertainty analysis in SWAT streamflow simulation," *J. Hydrol.*, vol. 414–415, pp. 413–424, 2012.

[8] J. E. Chen, M., Xie, P. T. and Janowick, "Global Land Precipitation : A 50-yr Monthly Analysis Based on Gauge Observations," *Am. Meteorol. Soc.*, pp. 249–266, 2002.

[9] B. Rudolf and U. Schneider, "Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations," *Proc. 2nd Work. Int. Precip. Work. Gr.*, no. January 2005, pp. 231–247, 2004.

[10] L. J. Williams, "Principal component analysis ´," *John Wiley Sons, Inc.*, pp. 433–459, 2010.

[11] P. D. Newman, M. and Sardeshmukh, "A Caveat Concerning Singular Value Decomposition." Iournal of Climate, 1994.

[12] A. Alexander and M. Y. Le, "New Routes from Minimal Approximation Error to Principal Components," pp. 197–207, 2008.

[13] M. B. Cohen and C. Musco, "Dimensionality Reduction for k-Means Clustering and Low," *Proceeding STOC '15 Proc. forty-seventh Annu. ACM Symp. Theory Comput. Pages 163-172*, pp. 163–172, 2015.

M. H. Ali" Analysis of Precipitation Data Using Empirical Orthogonal Functions" IOSR Journal of Applied Physics (IOSR-JAP) , vol. 11, no. 4, 2019, pp. 58-63.