Journal of the Nigerian Association of Mathematical Physics
Volume 47 (*July, 2018 Issue*), **pp315– 322**
© J. of NAMP

ANALYSIS OF MISSING DATA, QUALITY CONTROL AND HOMOGENEITY TEST OF ANNUAL PRECIPITATION SERIES OF NIGERIAN RAINFALL STATIONS

Uzoma E. K.[1] and Adeniyi M.O.[2]*

[1]**Department of Physics, Hezekiah University, Umudi, Nigeria**
[2]**Department of Physics, University of Ibadan, Ibadan, Nigeria**

Abstract

The devastating consequences of flooding, drought, erosion and desertification have made the investigation of important climate variables such as precipitation a necessity. This study conducted the analysis of missing values in 25 stations across the country, checked the quality and homogeneity of the precipitation series vital for future climate and hydrological studies. Missing values were estimated and completed by polynomial interpolation method while four homogeneity tests – Standard Normal Homogeneity Test, Buish and Range Test, Pettitt Test and von Neumann Ratio Test were applied to the stations' data separately at 95% significance level. Years of missing data were more predominant within the 1960-1969 decade followed by 1970-1979. The quality control results indicated that Calabar and Warri had the highest unique value, P_{out} for replacing any outlier of 4978.1mm and 4206.2mm respectively from southern stations while lowest values of 1293.1mm and 1321.2mm were found in the northern stations of Sokoto and Maiduguri respectively. Moreover, outliers were also detected in the southern stations of Ikeja and Calabar in 2004 and 1975 respectively. Years of break or inhomogeneity detection were more in 1970's as seven stations detected break in 1970 while four stations in 1972 and 1975 as well. The stations were further grouped into "useful", "doubtful", and "suspect" categories depending on the number of homogeneity tests rejecting the null hypothesis. 23(92%) of the stations were further categorized "useful" while 2(8%) of them, Bauchi and Nguru were "doubtful". No "suspect" was found. "Doubtful" stations were therefore recommended to be critically inspected or be subjected to data adjustment for possible correction before being used for any hydrological analyses.

Keywords: Precipitation, homogeneity tests, missing values, quality control, outliers

1.0 Introduction

Precipitation anomalies could result in devastating consequences of flooding, drought, erosion and desertification. Therefore its study requires precipitation data series with high quality and homogeneity to reduce to a barest minimum the biasness in the findings of any study in such area as river basins management study, flood hazards protection, studies related to climate change, erosion modeling and other applications for ecosystem and hydrological impact modeling. As a result of this, it is therefore necessary to test and check for reliability and homogeneity of data recorded at gauging stations before they are being used. For a better representation of an area, it is also important to complete the series of the stations having missing values due to various reasons [1]. Cubic polynomial interpolation method was adopted in the estimation of the missing values in this study. Cubic polynomial interpolation method was used in [2] and [3] to investigate Piecewise cubic interpolation of monotonic data. The problem of shape preserving using piecewise cubic interpolation method was discussed in [4] and [5].Historically, the identification of outliers has been the primary emphasis of quality control work [6, 7]. Non-resistant homogeneity testing methods are used by replacing the outlier values of each annual precipitation series by the unique value, P_{out}[8]. This method, as used in this study, was used in[9] and [8] to undertake the quality control of annual precipitation data in Büyük Menderes Basin, Turkey and the Southwest Europe precipitation data analysis respectively. Consequently, it is an important task to assess the homogeneity of long climate records before they can be reliably used, and as recommended by World Meteorological Organization (WMO), 'it is important, therefore, to remove the inhomogeneities

Corresponding Author**:** Uzoma E.K., Email: uzomaechefulachik@gmail.com, Tel: +2347033814023, +2348033579081 (AMO)

or at least to determine the possible error they may cause' [10]. Inhomogeneity in station data records are often caused by changes in observational routines, among which are station relocations, changes in measuring techniques and changes in observing practices [11]. Various statistical methods for homogeneity testing abound. Four of these tests, namely Standard Normal Homogeneity Test (SNHT), Buishand Range Test (BRT), Pettitt Test (PT) and von Neumann Ratio Test (VNRT) were employed in [11]to the European climate analysis. The results were categorized into three classes, which are useful, doubtful and suspect according to the number of tests rejecting the null hypothesis. Two of these methods – Buishand Range and Pettitt Tests were used in assessing the homogeneity of annual precipitation data in Büyük Menderes Basin, Turkey in [9]. SNHT and PT were used in [12] to check the homogeneity of 212 precipitation records in Turkey for the period 1973-2002. SNHT was also used in[8] to evaluate the homogeneity of precipitation series in Iberian Peninsula, southern France and northern Africa. With the high level of accuracy and reliability of the results of these studies, we considered the methods appropriate for this research. The uniqueness of this study is based on the fact that no recent and comprehensive study has been carried out on this subject in the country.

2.0 Materials and Methods

Data

Monthly precipitation data for the stations were obtained from the archives of the Nigerian Meteorological Agency (NIMET) Oshodi. The data length of the series varied between 30 and 108 years. Fig.1 and Table 1 below show the distribution of the stations over the country and their exact location.

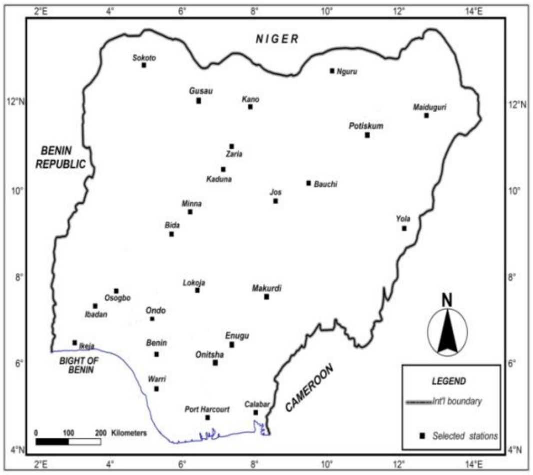

Fig. 1, Distribution of 25 precipitation stations over Nigeria used in the study

Table 1. Location of stations used in the study

S/N	Stations	Latitudes (^oN)	Longitudes (^oE)
1	Kano	12.05	8.50
2	Bauchi	10.50	10.00
3	Gusau	12.17	6.7
4	Jos	9.93	8.88
5	Bida	9.08	6.01
6	Kaduna	10.52	7.43
7	Lokoja	7.82	6.75
8	Maiduguri	11.83	13.15
9	Makurdi	7.73	8.53
10	Nguru	12.88	10.46
11	Zaria	11.07	7.7
12	Potiskum	11.71	11.07
13	Sokoto	13.08	5.25
14	Yola	9.23	12.46
15	Minna	9.61	6.56
16	Portharcourt	4.85	7.02
17	Ikeja	6.58	3.33
18	Calabar	4.97	8.35
19	Ibadan	7.43	3.90
20	Ondo	7.10	4.83
21	Osogbo	7.78	4.48
22	Enugu	6.47	7.55
23	Benin	6.32	5.60
24	Warri	5.52	5.73
25	Onitsha	6.15	6.78

Missing Data Analysis

The missing data analysis was done using polynomial interpolation by finding a formula whose graph passed through a given set of points(x, y). To fit N+1 points with an Nth degree polynomial, an exact function of which only discrete values are known was used to establish an interpolating or approximating function which passed through all specified *interpolation points* (also referred to as *data points*) and by so doing, members of this approximating function that agree with the known discrete values of the exact functions were estimated. There exists only one Nth degree polynomial that passes through a given set of points and expressed as a power series thus:

$$g(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_Nx^N . \qquad (1)$$

where g(x) is interpolating or the approximating function, a_i= unknown coefficients, i = 0,N (N+1, coefficients).

Quality Control

The data for each of the 25 stations were subjected to quality control test. The identification of outliers has been the primary emphasis of quality control work [8]. Outliers were identified as those values trespassing a maximum threshold for each time series [13] defined by

$$P_{out} = q_{0.75} + 3IQR. \qquad (2)$$

Where P_{out} is the unique value for replacing any outlier found in any series, $q_{0.75}$ is the third quartile and IQR is the interquartile range. The IQR has been used in quality control of climate data [14] because it is resistant to outliers. Values over P_{out} were substituted by this limit. This way of proceeding reduces the bias caused by outliers and yet keeps the information of extreme events [15].

Statistical Analysis of Homogeneity

Four homogeneity tests were used to test the homogeneity of the precipitation data in this study. They are Standard normal homogeneity Test (SNHT) [16], Buish and Range Test (BRT) [17], Pettitt Test (PT) [18], and von Neumann Ratio Test (VNRT) [19]. Under null hypothesis, the annual values *Yi* of the testing variables *Y* are independent and identically distributed and the series are considered as homogeneous. Meanwhile under alternative hypothesis, SNHT, BRtest and Pettitt test assume the series consisted of break in the mean and considered as inhomogeneous. These three tests are capable to detect the year where break occurs.VNR test is not able to give information on the year of break because the test assumes the series is not randomly distributed under alternative hypothesis. There are some differences between SNHT, BR test and Pettitt test. SNHT is sensitive in detecting the breaks near the beginning and the end of the series. BRT and PT are easier to identify the break in the middle of the series [20]. Besides, the SNHT and BR test assumed *Yi* is normally distributed, whereas Pettitt test does not need this assumption because it is a non-parametric rank test. In their mathematical formulations adopted from [11], Y_i (*i* is the year from 1 to n) is the annual series to be tested, $\overline{Y}$ is the mean and *s* the standard deviation.

Standard Normal Homogeneity Test

The SNH test is based on the T(k) statistic that compares the mean of the first k observations with the mean of the remaining n-k observations:

$$T(k) = k\overline{z}_1^2 + (n-k)\bar{z}_2^2 \quad k = 1, \dots, n. \qquad (3)$$

where

$$\bar{z}_1 = \frac{1}{k}\sum_{i=1}^{k}\frac{(Y_i - \overline{Y})}{s} \quad and \ \bar{z}_2 = \frac{1}{n-k}\sum_{i=k+1}^{n}\frac{(Y_i - \overline{Y})}{s}$$

The year *k* consisted of break if value of *T(k)*is maximum. To reject null hypothesis, the test statistic,

$$T_0 = \max_{1 \le k < n} T(k). \qquad (4)$$

Is greater than the critical value, which depends on the sample size, n.

This study used 95% critical values for a single shift SNHT as a function of n.

Buishand Range Test

In this test, the adjusted partial sums are defined as:

$$S_0^* = 0 \ and \ S_k^* = \sum_{i=1}^{k}(Y_i - \overline{Y}) \qquad k = 1, \dots, n\,. \qquad (5)$$

When the series is homogeneous, then the value of S_k^*will rise and fall around zero.The year *k*has break whenS_k^*has reached a maximum (negative shift) or minimum (positive shift). Rescaled adjusted range, *R* is obtained by

$$R = \frac{\max\limits_{0 \le k \le n} S_k^* - \min\limits_{0 \le k \le n} S_k^*}{s}\,. \qquad (6)$$

The $R/\sqrt{n}$is then compared with the critical values given by [17].This study used 95% critical values which depend on the sample size, n.

Pettitt Test

This is a non-parametric rank test. The ranks $r_1,\ldots,r_n$ of the $Y_1,\ldots,Y_n$ are used to calculate the statistic:

$$X_k = 2\sum_{i}^{k} r_i - k(n+1) \qquad k = 1, \ldots, n \ . \qquad (7)$$

If a break occurs in year E, then the statistics is maximal or minimal near the year $k = E$:

$$X_E = \max_{1 \le k \le n} |X_k| \ . \qquad (8)$$

The value is then compared with the critical value at 95% significant level depending on the sample size, n.

Von Neumann Ratio

The von Neumann ration N is defined as the ratio of the mean square successive (year to year) difference to the variance [19]:

$$N = \frac{\sum_{i=1}^{n-1}(Y_i - Y_{i+1})^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \ . \qquad (9)$$

When the sample is homogeneous the expected value is N = 2. If the sample contains a break, then the value of N tends to be lower than this expected [21]. If the sample has rapid variations in the mean, then values of N may rise above two [22].

Results and Discussion

Missing Value Analysis

The missing values in the precipitation series were completed by polynomial interpolation Method on a monthly scale. The stations and their respective year(s) of missing values are contained in Table 2 below.

Table 2:The list of precipitation stations and the results of the outlier trimming process.

S/N	Stations	Data Period	Missing Data Years	P_{out} (mm)	Extreme year(s) replaced with P_{out}(mm)
1	Kano	1905 – 2012	1964, 2011	1930.8	
2	Bauchi	1918 – 2012	1978	1844.2	
3	Gusau	1961 – 2012	-	1655.1	
4	Jos	1922 – 2012	-	2207.8	
5	Bida	1950 – 2012	1966, 1974	1874.5	
6	Kaduna	1931 – 2012	-	2044.9	
7	Lokoja	1931 – 2012	-	2287.2	
8	Maiduguri	1926 – 2012	-	1321.2	
9	Makurdi	1927 – 2012	1952, 1978	2172.4	
10	Nguru	1916 – 2012	2003	2353.8	
11	Zaria	1943 – 2012	-	1973.3	
12	Potiskum	1936 – 2012	1939	1531.8	
13	Sokoto	1950 – 2012	-	1293.1	
14	Yola	1931 – 2012	1966, 1967,1968	1617.6	
15	Minna	1916 – 2012	-	2142.3	
16	Portharcourt	1931 – 2010	1967, 1968, 1979	4084.1	
17	Ikeja	1944 – 2012	1966	3034.6	2004
18	Calabar	1905 – 2012	1964, 1967, 1968	4978.1	1975
19	Ibadan	1905 – 2012	-	2355.8	
20	Ondo	1906 – 2012	-	2843.2	
21	Osogbo	1958 – 2012	-	2341.9	
22	Enugu	1980 – 2012	-	2978.2	
23	Benin	1906 – 2012	1979	3558.9	
24	Warri	1907 – 2012	-	4206.2	
25	Onitsha	1931 – 1960	-	3673.8	

The result also indicated that years of missing values were more predominant within the 1960-1969 decade followed by 1970-1979 which likely could be as a result of the civil war (Biafran war) between 1966 and 1970 in the country which could have led to inadequate attention to maintenance of instruments.

Quality Control Test

The results of quality control process are given in Table 2 above in which P_{out} values and extreme years corrected for each station are tabulated. It is obvious that the variation of the data reaches maximum towards the southern stations while the lowest values occurred towards the northern stations. Calabar and Warri have the highest P_{out} of 4978.1mm and 4206.2mm respectively from southern stations while lowest values of 1293.1mm and 1321.2mm were found in the northern stations of Sokoto and Maiduguri respectively. There were two corrected values both of which were located in the southern part of the country. The stations are Ikeja and Calabar. Ikeja has a total annual precipitation of 3388.6mm in 2004 which is higher than the P_{out}value of 3034.9mm and was replaced by the P_{out} while Calabar has a total annual precipitation of 5116.4mm in 1975 which is higher than the P_{out} value of 4978.1mm with which it was replaced. These outliers could be assumed to be error in measurement rather than natural variation since other stations or neighbouring stations within the same geographical location have total annual precipitation values that are not more than their P_{out} values in the same years.

Homogeneity Test

The homogeneity of the annual total precipitation time series of the stations were tested using Standard Normal Homogeneity Test (SNHT), Buishand Range Test (BRT), PettittTest (PT) and von Neumann Ratio Test (VNRT). In the application of these methods, observation series of each station were considered separately. The results of each method were evaluated at 95% significance level and the inhomogeneities or breaks were detected following [11]. Table 3 below shows the list of stations having an inhomogeneity and the year(s) of break calculated by the three methods. Von Neumann is not location-specific and hence does not detect year of break.

Table 3:The results of the homogeneity test, the significant change points at 95% level

S/N	Stations	Data period	SNHT	BRT	PT	VN
1	Kano	1905 – 2012	1995 2004 2011	1981	1981	
				1990	1990	0.91
				1995	1995	
2	Bauchi	1918 – 2012	2004		1960	
			2007	1990	1964	1.41
			2010		1970	
					1972	
					1990	
3	Gusau	1961 – 2012	-	-	-	1.72
4	Jos	1922 – 2012	1947	1959	1959	
			1949	1964	1964	1.65
				1970	1970	
5	Bida	1950 – 2012	-	-	1974	
					1975	2.33
6	Kaduna	1931 – 2012	-	-	1979	1.97
7	Lokoja	1931 – 2012	1946	1983	1983	
			1947	1990	1988	1.89
			1952			
8	Maiduguri	1926 – 2012	-	-	1970	
					1972	1.59
					1975	
9	Makurdi	1927 – 2012	-	-	1971	
					1972	1.76
					1975	
					1976	
10	Nguru	1916 – 2012	1918	1963	1963	
			1920	1965	1965	0.41
			1925	1970	1970	
11	Zaria	1943 – 2012	1995	1975	1975	
			1997	1977	1977	1.70
12	Potiskum	1936 – 2012	1936	1963	1963	
			1958		1974	1.62
					1979	
13	Sokoto	1950 – 2012	1965	1972	1970	
					1971	1.69
14	Yola	1931 – 2012	-	1970	-	
				1971		1.74
15	Minna	1916 – 2012	1916	1962	1962	
			1918	1964	1964	1.78
			1925	1969	1969	
				1975	1975	
16	Portharcourt	1931 – 2010	-	-	1967	
					1968	1.75
					1971	
					1972	
17	Ikeja	1944 – 2012	-	-	1979	
					1980	1.65
					1981	
18	Calabar	1905 – 2012	1906	-	-	
			1909			1.81
19	Ibadan	1905 – 2012	-	1961		
				1977	1977	1.82
				1981		
				1982		
20	Ondo	1906 – 2012	-	1984	1983	
					1984	1.64
21	Osogbo	1958 – 2012	-	1983		
				1984	-	2.33
				1989		
22	Enugu	1980 – 2012	-	2011	-	1.53
23	Benin	1906 – 2012	1990	1960	1960	
			1999	1967	1967	1.49
			2005	1974	1974	
			2009	1988	1988	
24	Warri	1907 – 2012	-	1936		
				1938		1.94
				1939		
25	Onitsha	1931 – 1960	-	1959	-	1.46

From the analysis, the number of stations passing the critical test value by SNHT at a 95% significance level was 24 while Nguru was inhomogeneous. In the evaluation of the SNHT results, the stations with a test statistic higher than the critical value as given in [11] were considered to be inhomogeneous depending on the sample size.The results show that the years of break detected by using SNHT method were mostly at the beginning or towards the end of the series. In Calabar (1906, 1909), Potiskum (1936, 1958), Nguru (1918, 1920, 1925) and Minna (1916, 1918, 1925) stations the inhomogeneity detections were at the beginning of the series while Kano (1995, 2004, 2011), Bauchi (2004, 2007, 2010), and Benin (1990, 1999, 2005, 2009) stations experienced it towards the end. With BRT only one station (Sokoto) was found to be inhomogeneous and the years of break or inhomogeneity in the stations were majorly in the middle of the series as reflected in Kano (1981, 1990, 1995), Bauchi (1990), Lokoja (1983, 1990), and Osogbo (1983, 1984, 1989) stations. The results of the PT show that one station (Bauchi) was inhomogeneous. It is obvious that years of break or inhomogeneity were detected more between 1970 and 1975 as can be seen in Table 3. Seven stations detected break in 1970, four in 1972 and also in 1975.The stations with inhomogeneity in 1970 are Bauchi, Jos, Maiduguri, Nguru, Sokoto and Yola and one can suggest that the inhomogeneity detected at the stations could be caused by the variations in the natural climate conditions due to their relationship in terms of regional and geographical location (Fig. 1).Inhomogeneity was detected in 1990 for Kano and Bauchi stationsby both BRT and PTmethods. Since the stations are from the same region as shown in the map (Fig. 1), one can infer that the inhomogeneity might be related to the variations of natural meteorological conditions. This result is comparable with [12], in which 43 out of the 212 stations studied were inhomogeneous, with SNHT having higher percentage of homogeneous stations than PT.On a general note, the homogeneity analysis was carried out further by grouping the stations into “useful”, “doubtful”, and “suspect” categories depending on the number of tests rejecting the null hypothesis. The test results were classified according to [11] stated at the appendix. Following the classifications, 23(92%) of the stations were categorized “useful”, 2(8%) of the stations was categorized “doubtful”. The “doubtful” stations are Nguru and Bauchi which could be as a result of some variations of natural meteorological conditions. None of the stations was found to be “suspect” in the analysis of the study. The homogeneity results have shown that the precipitation data series of Nigerian stations are to a large extent homogeneous and reliable and are fit for use in any climate and hydrological analyses. The “doubtful” ones should be critically inspected before being used for any hydrological study in the country.

Conclusion

The study has shown that missing values were more predominant within the 1960-1969 decade followed by 1970-1979. The variation of the data reaches maximum towards the southern stations while the lowest values occurred towards the northern stations. Outliers were also detected in the two southern stations of Ikeja and Calabar. Each of SNHT, BRT and PT detected one inhomogeneous station while the stations’ data are considered reliable as they recorded high percentage of homogeneity generally. The “doubtful” stations should be critically inspected before they are being used for further analysis of trend and variability while adequate data adjustment of the inhomogeneous series could help to improve the series.

Appendix

Category 1: Useful

The series that rejects one or none null hypothesis under the four tests at 95% significance level are considered here. Under this category, the series is grouped as homogeneous and can be used for further analysis of trend and variability.

Category 2: Doubtful

The series that reject two null hypotheses of the four tests at 95% significance level is placed in this category. In this category, the series have the inhomogeneous signal and should be critically inspected before further analysis.

Category 3: Suspect

When there are three or all tests are rejecting the null hypothesis at 95% significance level, then the series is classified into this category. In this category, the series can be deleted or ignored before further analysis. They lack credibility.

References

[1] Mahmut F., Fatih D., ACem K.¸ and Mahmud G., (2010).Missing data analysis and homogeneity test for Turkish precipitation series. *Journal of the Indian Academy of Sciences 35: doi:707-720. 10.1007/s12046-010-0051-8*

[2] Fritsch F and Carlson R., (1980). Monotone Piecewise Cubic Interpolation. *SIAM J. Numer. Anal.,**17**(2): 238-246.*

[3] Fritsch F., and Butland J., (1984).A Method for Constructing Local Monotone Piecewise Cubic Interpolants. *SIAM J. of Sci. Stat. Comput. **5**:300-304.*

[4] Butt S., and Brodlie K. (1993). Preserving Positivity Using Piecewise Cubic interpolation. *Computers and Graphics. **17**(1):55-64.*

[5] Brodlie K, and Butt S. (1991). Preserving Convexity Using Piecewise Cubic Interpolation. *Computers and Graphics, **15**(1):15-23.*

[6] Filippov V., (1968). Quality control procedures for meterorological data. *Tech. Rep. 26, WMO, Geneva, Switzerland, 38pp.*

[7] Grant E and Leavenworth R., (1972). Statistical quality control, *McGraw Hill, 764pp.*

[8] Gonzalez-Rouco J., Luis J., Vicente Q., and Francisco V., (2001). Quality Control and Homogeneity of Precipitation data in the Southwest Europe. *Journal of Climate **14**: 964-978.*

[9] Ercan Y., Selin A., Necdet D., Talih G., and Fuat S. (2009). Quality Control and Homogeneity of Annual Precipitation data in Büyük Menderes Basin, Western Turkey. *Fresenius Environmental Bulletin 18,1748-1757.*

[10] Aguilar E., Auer I., Brunnet M, Peterson T and WieringaJ. (2003). Guidelines on Climatemetadata and homogenization. *WMO-TD No 1186, WCDMP No 53, World Meteorological Organization, Geneva.*

[11] Wijngaard J., Kleink Tank A., Konnen G., (2003).Homogeneity of 20th Century European Daily Temperature and Precipitation Series. *Int. J. Climatol, 23, 679-692.*

[12] Karabörk M¸ Kahya E, Komuscu AU. (2007). Analysis of Turkish precipitation data: Homogeneity and the Southern Oscillation forcings on frequency distributions. *Hydrological Processes 21: 3203–3210.*

[13] Peterson T., Vose R., Schwoyer R., and Razuveav V., (1998a). Global Historical climatology Network (GHCN) quality control of monthly temperature data. *Int. J. Climatol, 18,1169-1179.*

[14] Eischeid T., Baker C., Karl T., and Diaz H., (1995).The quality of long term climatological data using objective data analysis. *J. Appl. Meteor, 34, 2787-2795.*

[15] Barnett, V and Lewis T. (1994). Outliers in statistical data.3d ed. *J. Wiley and Sons, 604 pp.*

[16] Alexandersson H., (1986). A Homogeneity Test Applied to Precipitation Test. *J. Climatol., 6, 661-675.*

[17] Buishand T., (1982).Some Methods for Testing the Homogeneity of Rainfall Records. *J. Hydrol., 58 , 11-27.*

[18] Pettitt A., (1979). A Non-Parametric Approach to the Change-Point Detection. *Appl. Stat., 28(1979), 126-135.*

[19] Von Neumann J., (1941).Distribution of the Ratio of the Mean Square Successive Difference to the Variance. *Ann. Math. Stat., 13(1941), 367-395.*

[20] Hawkins M. 1977. Testing a sequence of observations for a shift in location. *Journal of the American Statistical Association* **72**: 180–186.

[21] Buishand T. (1981). The analysis of homogeneity of long-term rainfall records in the Netherlands. *KNMI Scientific Report WR 81-7, De Bilt, The Netherlands.*

[22] Bingham C. and Nelson L. (1981).An approximation for the distrivution of the Von Neumann ratio. *Technometrics 23: 285-288.*