Global Journal of Engineering and Technology [GJET].
ISSN: 2583-3359 (Online)
Frequency: Monthly
Published By GSAR Publishers
Journal Homepage Link- https://gsarpublishers.com/journal-gjet-home/

Statistical Analysis and Homogeneity Test of Annual Rainfall Stations in the Region of Mascara (Algeria)

BY

Mr. Dr. HDR BOUALEM ABDELKADER[1] and Mlle BOUALEM FERIEL OUMNIA[2]

[1,2]Faculty of Natural and Life Sciences University of Mostaganem (Algeria) Option: Hydro-agricultural development
Faculty of Exact Sciences of mascara (Algeria) Computer Science department

Article History

Received: 17/03/2024
Accepted: 25/03/2024
Published: 27/03/2024

Vol – 3 Issue –3

PP: - 01-06

Abstract:
In actuality, it is rather simple to take advantage of the rainfall data in relation to the water balance components, or the amount of water accessible to the primary groundwater reservoirs in the area. The variations in the hydrological system components can also be explained by random variables in time and space, such as rainfall and climate. in the widest sense. Because of this, and in spite of the untrustworthy nature of the data, we are attempting, with this study, to find a reliable statistical analysis of precipitation.

As part of the annual precipitation statistical analysis, we will first concentrate on examining the homogeneity using the linear regression method, as well as by comparing with the Wilcox on data and managing its data with various laws to validate the annual precipitation's reliability and draw conclusions.

Keywords: rainfall, analysis, smoothing, slope, Mascara, Algeria.

Introduction

1. Situations of Climate Stations

The watercourse regime is influenced by conditions as well (Bergaoui, M. (1987). Rain fall actually directly affects Melrir and Hounet feed flow watershed. However, the climate within the equipment is far from ideal; this is explained by the fact that there are only three waterfalls in the basin's extended as ours (see Figure 1. for a map showing the locations of the waterfalls). Another drawback is the difficulty with gaps and the scarcity of long rainfall series.

Several authors have previously examined rainfall in Algeria (Seltzer, P. (1946); (Gaussen, H. and Bagnouuls, M. (1948); Chaumont and Paquin, C. (1971).

Figure 1: Rainfall Stations Situations Map

2. Data Collection

Monthly rainfall collection was carried out with the National Agency for Water Resources Oran, 2012. However, these values represent the monthly rainfall totals as well as the meteorological stations' waterfalls in "Wadi Melrir and Hounet".

3. Choice of Stations Used

We have three stations that can provide an interpolation in our region; we have chosen stations that are inside the boundaries of four watersheds for the purpose of four studies (see Table 1).

Table1:

Presentation of Rainfall Stations

		Coordinates		
Stations	**code**	**X (m)**	**Y(m)**	Period observational
Touhami	111002	245,0	192,2	1990-2004
Fergoug	111506	259,3	250,3	1991-2005
AinFrass	110603	240,1	215,7	1991-2005

It is evident that a thorough analysis of false information is useless. Any analysis needs to go through two steps: gathering the data and treating it afterwards. It is evident that the first operation is crucial. (Abdelkader Boualem ; 2021).

*Corresponding Author: Mr. Dr. HDR BOUALEM ABDELKADER

4. Materials and Methods

4.1 A controlling method

The rear several methods to test the homogeneity:

- Method of line a regression,
- Wilcox on test.
- In our case, we note the method of line a regression and the Wilcox on one

4.1.1 Wilcox on Test

This test is nonparametric; rather than using a set of values, it employs the rankings of the observation series. The following idea is the basis of the Wilcox on test (Wilcox on,. F(1945)): The sample XUY, or union of X and Y, is another option if the sample X originates from the same population as Y. Thus, two samples, X and Y, are derived from a sequence of observations of length N. The sizes of the samples are N1 and N2, respectively, with N=N1N2 and N1≤N2.

The values of our series are then ranked in ascending order. Later, we're only concerned with how each of the two samples in this series ranks. But still, when values are repeated multiple times, we associate the appropriate Average rank.

We then calculate the ranks sum of Wx elements of the first sample in the common set Wx=Σx rank Wilcox on showed that, in case the two samples X and Y is a homogeneous series, the quantity Wx is.

Between two terminals Wmax and Wmin data by the following formulas:

$$\mathbf{W_{min}} = \frac{(N_1+N_2+1)N_1-1}{2} - Z_{1-\alpha/2}\sqrt{\frac{1}{12}N_1N_2(N_1+N_2+1)}$$

$$\mathbf{W_{max}} = (N_1+N_2+1)N_1 - W_{min}$$

$Z1\text{-}\alpha/2_2$: Represents the value of the z-normal distribute on of the corresponding

$Z1\text{-}\alpha/2$:(To a confidence level of 95%, we have=1.96)

We will use the Wilcox on test to check the homogeneity of rainfall data from the Gestations at the significance level of 5%.

a. The Method of Linear Regression

consider a pair of values (xi, yi) on two neighboring rainfall samples, the method of least squares leads to the determination of the line D which has the property to be closest to all coordinate points (xi, yi), the destination being measured by the sum of the square gaps.

The line with equation y=ax+b is the regression line of Y relative to X or X over Y.

- Y: estimated value;
- a: slope of the line;
- b: constant.

Moreover, the set of pairs (xi, yi), is a representative random sample drawn from a population whose character Y is a linear correlation with the character x. if both variables were observed for respective arithmetic means x and Y, for standard deviation Sx and Sy, it help s to know the size of r (correlation coefficient).

- If is equalto1or-1, it has a functional relationship;
- If equals zero, the variable are called separate variables (no correlation).
- If is equalto1, (it was a positive correlation).
- If 0.6 <r< 1, (it was a good correlation).
- If0.3 <r < 0.6, (it was negative correlation).
- If0<r<0.3, (it was a bad correlation).

$$a = \frac{Cov(x,y)}{S_x^2} \quad , \quad b = \overline{Y} - a\,\overline{X}$$

$$\mathbf{Cov\,(x,y)} = \frac{\sum x_i * y_i}{n} \_ \bar{x} * \overline{Y}$$

$$\mathbf{r} = \frac{(\sum_{i=1}^{n} x_i y_i / n \_ \overline{x*y})}{S_x * S_y}, \quad \mathbf{S_x^2} = \frac{\sum_{i=1}^{n} x_i^2}{n} \_ \overline{x^2}, \quad \mathbf{S_y^2} = \frac{\sum_{i=1}^{n} y_i^2}{n} \_ \overline{y^2}$$

S_x^2: Variance of x; Sy^2: variance of y; S_x and S_y are the standard deviation of x and y.

4.2 The Independence Test chi-Square(Reduced Variation)

The expression chi-square covers several statistical tests, three tests mainly (Delhomme, J.P.(1978) **and** Matheron, G.(1970)**.**

The fit test of fitness, which compares the overall observed distribution in a statistical sample to a theoretical distribution, the chi-square.

The test chi-square that controls the separation of character sin given population.

The chi-square test for homogeneity indicates whether samples are from the same population.

The interest test is only the independence test chi-square. This test is used to assess the presence or absence of relationship between two characters in a population where these characters are qualitative, when characters quantitative, and one qualitative, or even when both characters are but quantitative values were pooled.

On the other hand, we note that this test checks for a particular connection but not wholly (except in special cases where it indicates a single causal relationship.

Also, we emphasis that the different chi-square test should not be confused with the theoretical distribution of chi-square, which tabulated values are only used to validate these tests.

So, let's see how this test can be used in the case of a distribution of two characters.

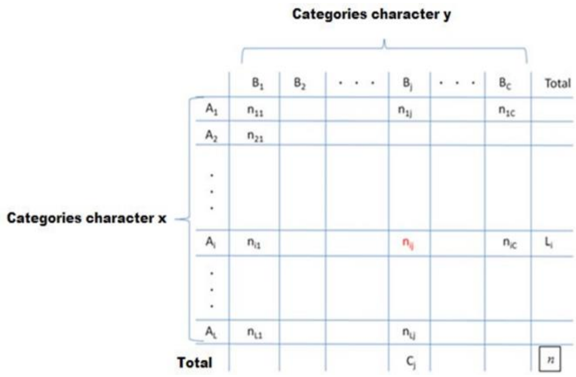
The first character, designated by X, may be quantitative or qualitative in nature, including categories (or classes)(usually from a combination of values of a quantitative trait or the terms of anon-quantitative). This will provide classes A1...AL

The second character, designated by Y, may be quantitative or qualitative in nature, including categories (or classes)(usually from a combination of values of a quantitative trait or the terms of anon-quantitative). This will provide classes B1...BC.

*Corresponding Author: Mr. Dr. HDR BOUALEM ABDELKADER

Under these conditions ,the size n of the population is distributed in an array cross:

$$e_{ij} = \frac{L_i C_j}{n} = \frac{C_j L_i}{n} \quad \chi^2_{Calculated} = \sum_{i=1}^{l} \sum_{j=1}^{c} \frac{(n_{i,j} - e_{i,j})^2}{e_{i,j}}$$



5. Results and Discussion

5.1 Test of Homogeneity between Stations

5.1 .1 Homogeneity test between the station of Fergoug and Touhami

The equation of the linear regression line is of the form:
Y = **18,04x - 37,005**the correlation coefficient is equal: r=0.865**(see Figure2)**

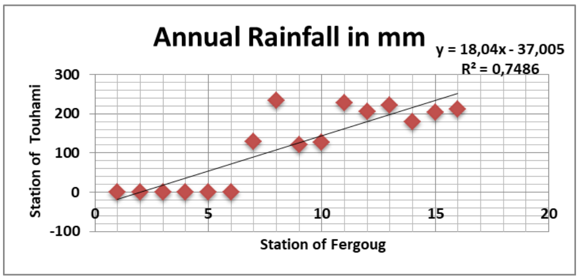


Figure 2: Line is regression between the station Fergoug and Touhami

5.1.2 Homogeneity test between the station Fegoug and AinFrass

The equation of the line regression line is of the form: Y=23,108 x – 70.49 the correlation efficient is equal: r= 0.884 **(seeFigure3).**

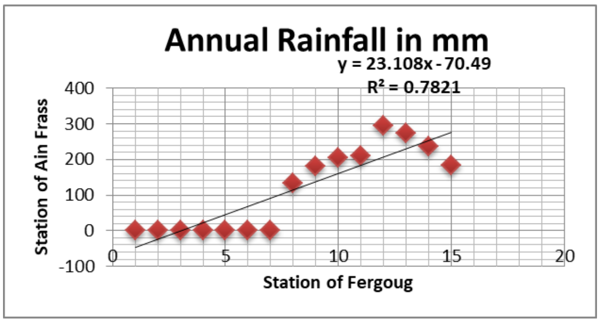


Figure.3: Line is regression between the station Fergoug and Frass

5.2 Applying the method of Wilcox on for the three study sites

5.2.1 Application of Wilcox on Station method of Fergoug (seeTable2).

Table 2: Application of Wilcox on Station method of Fergoug.

X	Y	Ranks	XUY	Original	Ranks	XUY	Original
380,7	161,9	1	161,90	Y	11	299,40	**Y**
298,7	225,7	2	175,90	X	12	335,60	**Y**
175,9	206,5	3	206,50	Y	13	345,10	**Y**
216,4	236,4	4	216,40	X	14	355,70	**Y**
260,6	280,8	5	225,70	Y	15	364,20	**X**
364,2	345,1	6	236,40	Y	16	380,70	**Y**
	295,1	7	260,60	X			
	299,4	8	280,80	Y			
	335,6	9	295,10	Y			
	355,7	**10**	**298,70**	**X**			

*Corresponding Author: Mr. Dr. HDR BOUALEM ABDELKADER

Was : $N_1 = 6$ and $N_2=10$

$W_X = 38$

$W_{max}=69.57$ It satisfies the inequality:

$W_{min}<W_X<W_{max}$ $32.43<38<69.57$

$W_{min}=3$

- As a result, the inequality is confirmed, and our homogeneous series

5.2.2 The Application of Wilcox on Station Method of Touhami (seeTable3).

Table.3:The Application of Wilcox on Station Method of Touhami

X	Y	Ranks	XUY	Original	Ranks	XUY	Original
250.2	132,9	1	132,9	Y	9	209,3	**Y**
152.0	180,2	2	152,0	X	10	215,8	**X**
163.6	203,9	3	152,4	X	11	235,2	**Y**
159.9	209,3	4	159,9	X	12	250,2	**X**
411.6	295,8	5	163,6	X	13	274,1	**Y**
152.4	274,1	6	180,2	Y	14	295,8	**Y**
215.8	235,2	7	184,2	Y	15	411,6	**X**
	184,2	**8**	**203,9**	**Y**			

Was: $N_1=7$et $N_2=8$

$W_X = 51$

$W_{max}=73.44$: It satisfies the inequality

$W_{min}<W_X<W_{max}$ $38.56<51<73.44$

$W_{min}=38.56$

- As a result, the inequality is confirmed, and our homogeneous series

5.2.3 The Application of Wilcox on Station Method of Ain Frass(see Table 4).

Table 4 The Application of Wilcox on Station Method of AinFrass

X	Y	Ranks	XUY	Original	Ranks	XUY	Original
280.0	128,9	1	120,5	Y	11	228,7	Y
284.1	234,0	2	126,5	Y	12	234,0	Y
166.0	120,5	3	128,9	Y	13	263,0	X
172.8	126,5	4	166,0	X	14	280,0	X
163.0	228,7	5	172,8	X	15	284,1	X
357.2	205,0	6	179,9	Y	16	357,2	X
	221,5	7	203,9	Y			
	179,9	8	205,0	Y			
	203,9	9	212,4	Y			
	212,4	**10**	**221,5**	**Y**			

*Corresponding Author: Mr. Dr. HDR BOUALEM ABDELKADER

Was: $N_1 = 6$ and $N2 = 10$

$W_x = 67$

$W_{max} = 69.58$ It the inequality:

$W_{min} < W_x < W_{max}$ 32.42<67<69.58

$W_{min} = 32.42$

- As a result, the inequality is confirmed, and our homogeneous series

6.Statistical Analysis by Chi-square test (χ^2)

6.1 Employment adjustment methods in the resort of Fergoug

Statistical analysis of rainfall data and hydrometrics (Lubes, H. Masson, and J.M. 1991) and Morlat, G. (1969) aim at defining climate regions around certain values and characteristics that are representative. The character's tic values are of two types:

- Core Values
- Extreme Values.

The core values characterize the abundance of regions and their irregularities (mean, median). The extreme values represent the value at which precipitation may take a certain probability chosen in advance. For statistical analysis of annual rainfall (Boualem Abdelkader; 2022), we will adjust to the Gauss law (laws normal, lognormal, and root). The series of observed station rainfall is 1991-2005 (See Tables 5, 6, 7 and Figures 4, 5,6).

Table 5: Suitability test χ2(Gauss law)

Classes K	Class boundaries	Observed numbers(n_i)	The theoretical(np_i)	$\frac{(ni-npi)^2}{npi}$
1	185.1-226.4	4	3.75	0.0166
2	227,2-299,9	4	3.75	0.0166
3	312,4-350,9	4	3.75	0.0166
4	355-384,7	3	3.75	0.15

Table 6: Suitability test χ2(log Normal)

Classes K	Class boundaries	Observed numbers(n_i)	The theoretical(np_i)	$\frac{(ni-npi)^2}{npi}$
1	2,27-2,35	4	3.75	0.0166
2	2,36-2,48	4	3.75	0.0166
3	2,48-2,55	4	3.75	0.0166
4	2,55-2,59	3	3.75	0.15

Table.7: Suitability test χ2(root law)

Classes K	Class boundaries	Observed number s(n_i)	The theoretical (np_i)	$\frac{(ni-npi)^2}{npi}$
1	13,61-15,05	4	3.75	0.0166
2	15,07-17,32	4	3.75	0.0166
3	17,67-18,73	4	3.75	0.0166
4	18,84-19,61	3	3.75	0.15

- χ^2 calculated =0.1998:

The number of degrees is ddl=k-3 =4-3 =1; k: class number (sample size means the standard deviation).

For risk α=5% χ^2tabulated=3.841 >χ^2calculated=0.1998; so the fitness for the normal distribution is acceptable.

- χ^2calculated=0.1998

The number of degrees ddl=k-3=4-3=1; χ^2calculated=0.1998<χ^2tabulated=3.841; so the fitness for the log-normal distribution is acceptable.

- χ2calculated=0.1998; the number of degrees ddl=k-3=4-3=1;
- χ^2calculated=0, 1998<χ^2tabulated=3.841;

6.2 Calculation of the return period (when the module is the strongest)

With: T: return period; P=384.7andF=0,938; therefore T=16ans

T= 1/ 1-F

For this purpose, the maximum annual

Precipitation is repeated every 16 years and this was justified by the statistical analysis

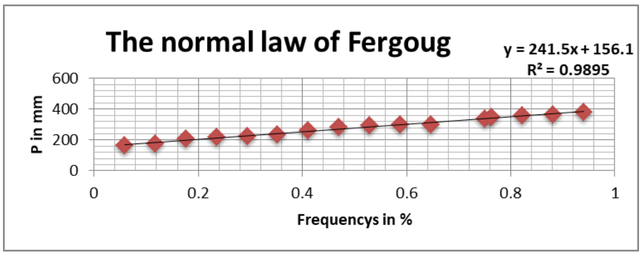


Figure 4: Graphical representation of the normal distribution of station Fergoug

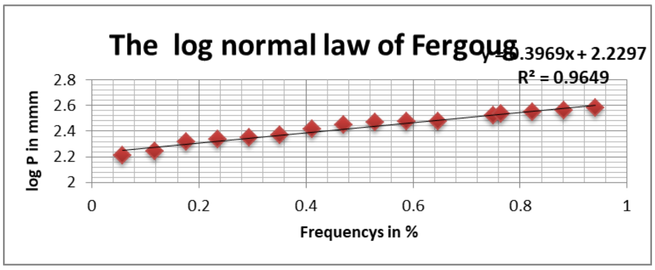


Figure 5: Graphical representation of the lognormal law of Station Fergoug

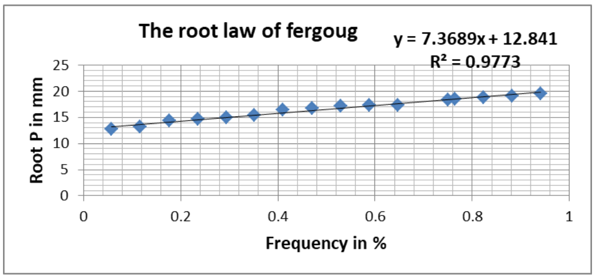


Figure 6: Graphical representation of the root law Station Fergoug

7. Conclusion:

Finally, it is better to comprehend the basins' climatic traits. Therefore, the hydrometric network needs to be reorganized.

It is evident from our study that there is a significant amount of variation in the rainfall data using all the methods employed, including the Wilcoxon on method and linear regression. However, in the end, we confirm the validity of our data using the normal distribution and the confirmation of its laws, which are supported in the earlier curves in a number of different ways

Acknowledgements:

I express my gratitude to Mademoiselle Boualem Fériel Oummia for her tireless efforts in reformulating the text's paragraphs into clear, concise English.

6. References

1. Boualem Abdelkader ; (2021): Comparative Study of Annual Precipitation between Oran Aerodrome (Algeria) and Marseille Marignane (France) Stations Located in the Mediterranean Basin; Journal of Basic and Applied Scientific Research www.textroad.com; ISSN 2090-4304
2. 2-Boualem Abdelkader; 2022: STATISTICAL STUDY AND ANALYSIS OF ANNUAL RAINFALL IN THE NAAMA REGION (ALGERIA) Journal of Fundamental and Applied Sciences ISSN 1112-9867 Available online at http://www.jfas.info
3. 3-Bergaoui, M. (1987).Study of the spatio-temporal distribution of rainfall fine scales of time and space. PhD thesis MechanicsUM2.
4. 4-Delhomme,J.P.(1978).Application of the theory of regionalized variable sin Water Sciences. Bulletin BRGM(2)III,4.In1978.
5. 5- Lubes, H. Masson and J.M. (1991). Evaluation study of software adjustments Statistical distributions of hydrological variables. Laboratory of Hydrology.ORSTOM-L.H.M.
6. 6-Matheron, G.(1970). The theory of regionalized variables and its applications. The Journal of the CMM, Issue 5, SchooloftheMines.Paris.
7. 7-Morlat, G.(1969). Statistical problem sin hydrology. Session organized by International Association for Statistics in the Physical Sciences, London, September 9, 1969.464
8. Seltzer;P.(1946):The climate of Algeria.Univ.From Algiers, IMPG Carbonnel ,Algiers.
9. Gaussen, H and Bagnouls,M.(1948): Rainfall Maps of Algeria(1913-1947average). Scale. 1:500000, IGN,Paris
10. Chaumont and Paquin. C(1971). Carte rainfall of Algeria, ex ch.1/50000. Soc. Hist. Nat. North Africa, Algiers. 4 sheets. Record.
11. Wilcoxon,.F(1945): ividual comparisons byrankingmethods.Biometrics,1,80-83.
12. National water resources agency: rainfall database, 2012, Oran, Algeria .

*Corresponding Author: Mr. Dr. HDR BOUALEM ABDELKADER