Statistical Analysis of the Total Annual (Maximum, Minimum) Temperatures and Total Annual Evaporation for the Period 2011-2017 for the Governorates of Babil and Najaf

Khawla Ali Abd Al-Hameed

Department of Soil Science and Water Resources, College of Agriculture, AL-Qasim Green University,
Babil, Hilla, Iraq

Abstract: The simple linear correlation coefficient (pearson) is a tool for people who work in statistical analysis. The aim of this study is to determine the required sample size (water level) to estimate the Pearson coefficient of correlation in terms of force and direction of total temperature variables (maximum, minimum). The annual total annual evaporation of Babil governorate was marked by the symbol E for the governorates of Babylon and Najaf during the period from 2011-2017. The result of the association of the total annual maximum temperatures and the total annual evaporation of the province of Babylon is strong and positive as a result of the correlation of the total of the minimum temperatures and the total annual evaporation of the province of Babylon is weak and positive, the total temperature (maximum, minimum) and total annual evaporation of Najaf governorate were weak and negative.

Key words: Linear correlation, coefficient, required sample, force and direction, evaporation, association

INTRODUCTION

Often the researcher needs to study the relationship between two or more variables. The correlation coefficient of Pearson is a dimensionless measurement that specifies a simple linear relationship between two variables. Its value varies from -1, called a simple linear negative relationship, to +1, called a fully positive simple linear relationship. When this value approaches zero, the simple linear correlation score is smaller. Of the correlation coefficient of Pearson, a number of other statistics were calculated such as partial correlation, direct and indirect effects of variables in path analysis and canonical association by Hair *et al.* and also in determining variables that could be used indirectly by CRUZ.

As a result, these statistics depend on the accuracy of Pearson's simple linear correlation coefficient. The size of the samples affects the statistical result and its interpretation. In large sample, low-volume coefficients generate a statistical result when the relationship parameters between them are not important in practice. Also, when the sample size is small, the reliability of the estimates is low and does not represent a real relationship between the two variables as by Hair CRUZ, Cargnelutti Filho *et al.* (2010), Cargnelutti Filho *et al.* (2011), Cargnelutti Filho *et al.* (2012), Sari *et al.* (2017), Goktas and Isci (2011), Neter *et al.* (1993) and Keller and Waracck (2001).

We conclude from this that the size of the appropriate sample is accurate. Finally, the aim of this study was to determine the sample size to estimate Pearson's simple linear correlation coefficient between total temperature variables (maximum, minimum) and total annual evaporation of Babil and Najaf governates. In this research, we know that the province of Babylon has a total area of 5116 km^2 and the area of water bodies have 250 km^2, Najaf province has a total area of 28, 824 km^2 and the area of water bodies have 200 km^2. Can we prove that the water bodies of the province of Babylon are greater than the province of Najaf through the correlation coefficient of Pearson?

CORRELATION

Definition of correlation: The correlation means that there is a relationship between two phenomena or two variables, meaning that the change in one of the two variables leads to change in the other variable whether increase or decrease. If the two variables increase together and decrease together, the relationship between them is positive. If either one decreases by increasing the other variable, negative. The simplest way to study the relationship between the two variables or phenomena is the propagation mode. If we have T and E variables and these data are collected from the pairs of values of these two variables, they can be graphically represented in the propagation mode and take different forms (Fig. 1). The points are scattered unconnected to a specific direction indicating that there is no relationship between the two variables (T, E), (Fig. 2). Where the points are spread around a straight line where the values of E decrease with the increase of T values and we conclude that there is an

Fig. 1: The points are scattered unconnected to a specific direction indicating that there is no relationship between the two variables (T, E)
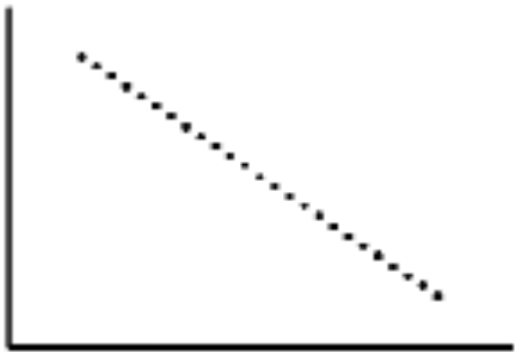

Fig. 2: Where the points are spread around a straight line where the values of E decrease with the increase of T values and we conclude that there is an inverse linear relationship between the variables (T, E)
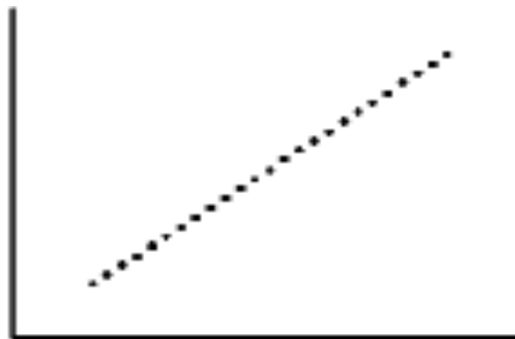

Fig. 3: Where the points are spread around a straight line in which the values of T increase with the values of E

inverse linear relationship between the variables (T, E) (Fig. 3). Where the points are spread around a straight line in which the values of T increase with the values of

Fig. 4: Where the points are spread around the curve and we conclude that there is a non-linear relationship between the variables (T, E)

Table 1: The value of correlation coefficients

Value of correlation coefficient	The meaning
+1	Correlation positive fully
From 0.70-0.99	Correlation positive strong
From 0.50-0.69	Correlation positive average
From 0.01-0.49	Correlation positive weak
0	No correlation positive

E and we conclude that there is a direct linear relationship between the variables (T, E), (Fig. 4) (Neter *et al*., 1993). Where the points are spread around the curve and we conclude that there is a non-linear relationship between the variables (T, E).

Simple linear correlation coefficient (Pearson): This link expresses the strength and direction of the relationship between two variables (phenomena) only, the result of the relationship either positive or negative on the one hand, strong or weak on the other.

Pearson assumes that they change as quantitative and that the relationship takes the form of a straight line Pearson represents the best measure between the two variables of correlation and symbolizes this parameter with the symbol r and takes the following form:

$$r = \frac{\sum_{i=1}^{n} T_i E_i - \left(\sum_{i=1}^{n} T_i\right)\left(\sum_{i=1}^{n} E_i\right)/n}{\sqrt{\sum_{i=1}^{n} T_i^2 - \left(\sum_{i=1}^{n} T_i\right)^2 /n \left(\sum_{i=1}^{n} E_i^2 - \left(\sum_{i=1}^{n} E_i\right)^2 /n\right)}}$$

Correlation properties: What is said about the random correlation applies to the reverse correlation (with a negative signal) (Table 1).

Examples: These examples show how the total temperature (maximum, minimum) and total annual evaporation of the governorates of Babel and Najaf are correlated from 2011-2017.

Table 2: The strength and direction of the relationshipbetween the maximum Temperature (T) and the total annual Evaporation (E)

Years	T			E
2011	366.6			1973.8
2012	387.5			2159.4
2013	371.4			2041.9
2014	385.6			2070.8
2015	387.9			2270
2016	383.5			2078.5
2017	388.7			2097.4
Solution				
T	E	ET	T^2	E^2
366.6	1973.8	723595.08	134395.56	3895886.4
387.5	2159.4	836767.5	150156.25	4663008.36
371.4	2041.9	758361.66	137937.96	4169355.61
385.6	2070.8	798500.48	148687.36	4288212.64
387.9	2270	880533	150466.41	5152900
383.5	2078.5	797104.75	147072.25	4320162.25
388.7	2097.4	815259.38	151087.69	4399086.76
$\sum T_i = $ 2671.2	$\sum E_i = $ 14691.8	$\sum T_i\,E_i = $ 5610121.85	$\sum T^2_i = $ 1019803.48	$\sum E^2_i = $ 30888606.06

Table 3: Relationship between the minimum Temperatureand the total annual Evaporation (E)

Years	T			E
2011	194.1			1973.8
2012	205.7			2159.4
2013	198.2			2041.9
2014	206.7			2070.8
2015	211.4			2270
2016	203.3			2078.5
2017	198.2			2097.4
Solution				
T	E	ET	T^2	E^2
194.1	1973.8	383114.58	37674.81	3895886.4
205.7	2159.4	444188.58	42312.49	4663008.36
198.2	2041.9	404704.58	39283.24	4169355.61
206.7	2070.8	428034.36	42724.89	4288212.64
211.4	2270	479878	44689.96	5152900
203.3	2078.5	422559.05	41330.89	4320162.25
198.2	2097.4	415704.68	39283.24	4399086.76
$\sum T_i = $ 1417.6	$\sum E_i = $ 14691.8	$\sum T_i\,E_i = $ 2978183.83	$\sum T^2_i = $ 287299.52	$\sum E^2_i = $ 30888606.06

Example: Between the strength and direction of the relationship between the maximum Temperature (T) and the total annual Evaporation (E) of Babil governorate during the period from 2011-2017 (Table 2):

$$r = \frac{\sum_{i=1}^{n} T_i E_i - \left(\sum_{i=1}^{n} T_i\right)\left(\sum_{i=1}^{n} E_i\right)/n}{\sqrt{\sum_{i=1}^{n} T_i^2 - \left(\sum_{i=1}^{n} T_i\right)^2 /n \left(\sum_{i=1}^{n} E_i^2 - \left(\sum_{i=1}^{n} E_i\right)^2 /n\right)}}$$

$$r = \frac{5610121.85 - 5606390.88}{\sqrt{(473.56)(53036.454286)}}$$

$$r = \frac{3730.97}{5011.5809174028} = 0.744$$

Note that the relationship is positive and strong

Example: Between the strength and direction of the relationship between the minimum Temperature (T) and the total annual Evaporation (E) of Babil governorate during the period from 2011-2017 (Table 3):

Table 4: Between the maximum Temperature (T) total annual Evaporation (E) of Najaf governorate

Year	T	E
2011	376.4	2724.3
2012	389.9	2685.8
2013	382	2575.2
2014	389.4	2464.2
2015	395.4	2598.7
2016	406.2	2595.3
2017	399	2686.1
Solution		
T	E	ET	T^2	E^2
376.4	2724.3	1025426.52	141676.96	7421810.49
389.9	2685.8	1047193.42	152022.01	7213521.64
382	2575.2	983726.4	145924	6631655.04
389.4	2464.2	959559.48	151632.36	6072281.64
395.4	2598.7	1027525.98	156341.16	6753241.69
406.2	2595.3	1054210.86	164998.44	6735582.09
399	2686.1	1071753.9	159201	7215133.21
$\sum T_i = $ 2738.3	$\sum E_i = $ 18329.6	$\sum T_i\,E_i = $ 7169396.56	$\sum T^2_i = $ 1071795.93	$\sum E^2_i = $ 48043225.8

$$r = \frac{\sum_{i=1}^{n} T_i E_i - \left(\sum_{i=1}^{n} T_i\right)\left(\sum_{i=1}^{n} E_i\right)/n}{\sqrt{\sum_{i=1}^{n} T_i^2 - \left(\sum_{i=1}^{n} T_i\right)^2 /n \left(\sum_{i=1}^{n} E_i^2 - \left(\sum_{i=1}^{n} E_i\right)^2 /n\right)}}$$

$$r = \frac{2978183.83 - 2975299.3828571}{\sqrt{(259215.26857143)(53036.454286)}}$$

$$r = \frac{2884.4471429}{117251.26328454} = 0.024$$

Note that the relationship is positive weak.

Example: Between the strength and direction of the relationship between the maximum Temperature (T) and the total annual Evaporation (E) of Najaf governorate during the period from 2011-2017 (Table 4):

$$r = \frac{\sum_{i=1}^{n} T_i E_i - \left(\sum_{i=1}^{n} T_i\right)\left(\sum_{i=1}^{n} E_i\right)/n}{\sqrt{\sum_{i=1}^{n} T_i^2 - \left(\sum_{i=1}^{n} T_i\right)^2 /n \left(\sum_{i=1}^{n} E_i^2 - \left(\sum_{i=1}^{n} E_i\right)^2 /n\right)}}$$

$$r = \frac{7169396.56 - 7170277.6685714}{\sqrt{(612.0885715)(46906.348572)}}$$

$$r = \frac{-881.1085714}{5358.2497041213} = -0.164$$

Note that the relationship is negative weak.

Example: Between the strength and direction of the relationship between the minimum Temperature (T) and the total annual Evaporation (E) of Najaf Governorate during the period from 2011-2017 (Table 5):

Table 5: Relationship between the minimum Temperature (T) and the total annual Evaporation (E) of Najaf Governorate

Year	T	E
2011	214.7	2724.3
2012	228.2	2685.8
2013	227.3	2575.2
2014	234.2	2464.2
2015	233.9	2598.7
2016	233.1	2595.3
2017	237	2686.1

Solution

T	E	ET	T^2	E^2
214.7	2724.3	584907.21	46096.09	7421810.49
228.2	2685.8	612899.56	52075.24	7213521.64
227.3	2575.2	585342.96	51665.29	6631655.04
234.2	2464.2	577115.64	54849.64	6072281.64
233.9	2598.7	607835.93	54709.21	6753241.69
233.1	2595.3	604964.43	54335.61	6735582.09
237	2686.1	636605.7	56169	7215133.21
$\sum T_i =$ 1608.4	$\sum E_i =$ 18329.6	$\sum T_i E_i =$ 4209671.43	$\sum T_i^2 =$ 369900.08	$\sum E_i^2 =$ 48043225.8

$$r = \frac{\sum_{i=1}^{n} T_i E_i - \left(\sum_{i=1}^{n} T_i\right)\left(\sum_{i=1}^{n} E_i\right)/n}{\sqrt{\sum_{i=1}^{n} T_i^2 - \left(\sum_{i=1}^{n} T_i\right)^2/n \left(\sum_{i=1}^{n} E_i^2 - \left(\sum_{i=1}^{n} E_i\right)^2/n\right)}}$$

$$r = \frac{4209671.43 - 4211618.3771428}{\sqrt{(335.71428572)(46906.348572)}}$$

$$r = \frac{-1946.9471428}{3968.2655287395} = -0.490$$

Note that the relationship is negative weak.

CONCLUSION

After studying the simple linear correlation coefficient (Pearson) for the quantitative data of the total temperature (maximum, minimum) and total annual evaporation from 2011-2017 for the districts of Babylon and Najaf, it became clear to us that the water level of Babil province is greater than Najaf.

REFERENCES

Cargnelutti Filho, A., M. Toebe, C. Burin, T.R.D. Silveira and G. Casarotto, 2010. Sample size for estimating the Pearson correlation coefficient among corn characters. Braz. Agric. Res., 45: 1363-1371.

Cargnelutti Filho, A., S.J. Lopes, B. Brum, M. Toebe and T.R. da Silveira *et al.*, 2012. Sample size to estimate the Pearson correlation coefficient among characters of castor bean. Semina Agric. Sci., 33: 953-962.

Cargnelutti Filho, A., S.J. Lopes, M. Toebe, T.R.D. Silveira and I.A. Schwantes, 2011. [Sample size to estimate the Pearson correlation coefficient among characters of *Crambe abyssinica* (In Portuguese)]. Agronomic Sci. Mag., 42: 149-158.

Goktas, A. and O. Isci, 2011. A comparison of the most commonly used measures of association for doubly ordered square contingency tables via. simulation. Methodol. Notes, 8: 17-37.

Keller, G. and B. Waracck, 2001. Statistics for Management and Economics. 6th Edn., Duxbury Press, California, USA.,.

Neter, J., W. Wasserman and G.A. Whitmore, 1993. Applied Statistics. 4th Edn., Allyn and Bacon, Boston, Massachusetts, USA., Pages: 989.

Sari, B.G., A.D.C. Lucio, C.S. Santana, D.K. Krysczun and A.L. Tischler *et al.*, 2017. Sample size for estimation of the Pearson correlation coefficient in cherry tomato tests. Rural Sci., 47: 1-6.