World Academics Journal of Engineering Sciences **Research Paper**
Vol.7, Issue.4, pp.28-34, December (2020) **E-ISSN:** 2348-635X

Determine the parameters affecting on the pollution of the Orontes River by statistical analysis methods

Lina Khouri

Dept. Environmental & Sanitary Engineering, Faculty of Civil Engineering, Damascus University, Homs, Syria

Author's Mail id: buterfly.lole@gmail.com

Available online at: www.isroset.org

Received: 10/Dec/2020, Accepted: 20/Dec/2020, Online: 31/Dec/2020

***Abstract*—** This study investigates the quality of the Orontes River by conducting a statistical analysis at certain monitoring points.as 11 monitoring points were distributed along the river from the point of its entry into the Governorate of Homs until its exit from it, and for this purpose, methods of multivariate statistical analysis were applied (cluster analysis, Factor analysis). By using the statistical program /SPSS/ to find out the advantages of the variables, the results indicated the possibility of reducing the number of measured variables and reducing the number of monitoring points along the stream, as the cluster analysis showed the presence of two clusters (low pollution, polluted), which could reduce the number of sampling sites according to similar specifications. By applying the factor analysis to study the changes to compare the compositional patterns between clusters, It gave /4/ basic factors with the degree of measurement quality of 85.7% > 50%, and the best rate for explaining the total variance ratios is 87.86% of the total variance and includes (7) basic variables:
(EC, Na, BOD, COD, NO3, DO, TSS).

Keywords— Orontes River Pollution, Statistical Analysis, Cluster Analysis, Factor Analysis, Multivariate Statistical Methods, River Water Quality.

I. INTRODUCTION

Water resources are considered one of the basics of sustainable development, so there has been a growing interest since ancient times in the quality of water that is used as it constitutes an essential element of the existence of living organisms, as various water sources suffer at present great pollution as a result of the large population increase, industrial and agricultural expansion, and the distribution of many industries Indiscriminately, and the accompanying mismanagement of water resources [1].

Rivers are among the most important sources of surface water pollution from various sources, such as leaks resulting from sewage, industrial networks and excessive use of pesticides and agricultural fertilizers, as these pollutants affect the physical and chemical properties of water, thus leading to a deterioration and reduction of the quality of these sources' water and its suitability for drinking, irrigation or industrial purposes[2].

Hence, the importance of knowing the variables affecting water quality in order to identify the different sources of pollution, which helps in in treating this pollution in the best way.

Most of the studies have been depending on traditional methods such as surveys and expert opinion in determining the variables that affect water quality.

We will review the most important variables that affect the river water quality by conducting a study and statistical analysis of a set of variables to assess water quality using multivariate statistical methods (Cluster Analysis, Factor Analysis).

II. RELATED WORK

Recently, studies dealing with water quality have become one of the tools that used to manage water resources in terms of assessing the current situation of those sources, and determining the treatment required for the water body, so it has become necessary to consider any nonconventional water source to be used economically and effectively [3]. and In the Orontes River case, which is under the influence of pollution resulting from pumping untreated wastewater from random sewage to villages along its course, and industrial drainage resulting from facilities that dump their waste into the river's course without treatment, in addition to pollution resulting from land use in the Orontes Basin area and it was That is why it was chosen to study in this research.

1-Research importance:
Misleading the importance of the study in terms of the large economic area of the Orontes River, where 18% of the total imports in Syria are sent, diverting a water resource and used as a drinking resource for my city, large areas of agricultural land and securing water for industrial

facilities, but no comprehensive study of drinking water has been conducted yet. To the information that has been developed so far in the case of potable water studs in all areas.

2- Research aims:

1- Determining the most influencing factors on water quality by applying standard statistical methods by studying (spatial and temporal changes), to know the relationship with each other and to determine the extent of the contribution of industrial, human and agricultural activities to the changes in water quality.

2- Reducing the number of monitoring areas, which leads to less effort and time spent on sampling and analysing samples and reducing financial and material costs.

3-Providing the simplest methods that reflect the suitability of water for the specific use, based on the results of multivariate statistical analysis.

Reference study:

Presentation of some studies that dealt with multivariate statistical methods:

- Matejicek [4] studied multivariate statistical methods in managing pollution data (T, EC, pH, DO, COD, TDS, NO_3, DO, PO_3, CL, FC) and water pollution modelling In the aquariums of Czechoslovakia. Singh et. al. [5] In a study of Gomti River in India, 24 variables were chosen to estimate the temporal and spatial changes, where cluster analysis gave 3 clusters, while the factor analysis identified 6 factors responsible for 70% of the variance in the river. Manoj et. al. [6] used multivariate statistical analysis to study the temporal changes of the Fuji River. cluster analysis gave (3) groups and Factor Analysis gave /4/ factors.

III. METODOLOGY

Research materials and methods:

1-Theoretical study:

-The research took into consideration the results of the analysis of water samples taken from the monitoring points along the course of the Orontes River from the year (2007 to 2020), starting from the Amiri area (the point of its entry into the Homs governorate) and ending with the Ghajar Amir area (the point of its exit from the Homs governorate). Choose a set of 21 variables as follows: (EC, COD, BOD, TUR, DO, pH, NH4, Cl, Na, T, TSS, TDS, SS, FC, NO3, PO4, K, Ca, Mg, SAR , TH)

-Eleven observatory points were distributed along the course of the river for sampling and are illustrated in (Fig. 1).

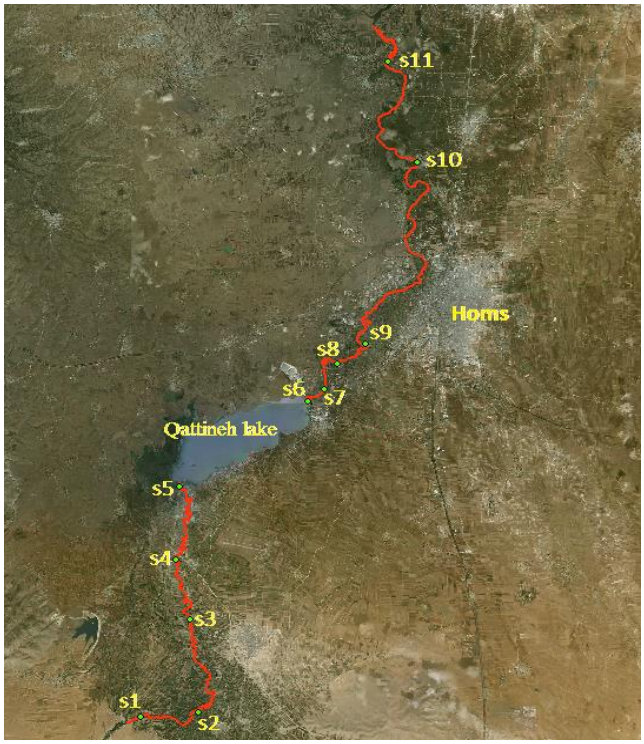

Figure 1. The distribution of observatories along the Orontes River

- **Study area**

The study area is Orontes river which is one of the major rivers in Syria, It is comes in terms of importance after the Euphrates River, as it is the main nerve in the Orontes Basin, and the various activities (agricultural - industrial - tourism) are located near it, and the Orontes River is the main source for securing the water needed for most industrial establishments in the central region, as about /12/ million m^3 for this purpose, and it is the source of drinking water for the city of Hama and the main source of irrigation in the central region, as it irrigates a total area of / 49,000 / hectares and is fed by seasonal rivers, (AL-Sarout- Al Bared – Salhab).

Among the most important projects built on it:

Qattineh Lake: 200 million m^3 per year and irrigates 6000 hectares in Homs Governorate, in addition to securing the water needed for the fertilizer company.

-Al-Rastan Dam lake : 230 m^3 that controls Wadi Harb itself.

-Mharda Dam lake: Its stock is 67 m^3.

-Natural lakes.

- Underground water.-Natural lakes.

- Underground water.

The Orontes River starts flowing from the Baalbek and Hermel regions in Lebanon. It enters into the Syrian territories at rabbla and runs through it for a distance of 325.5 km, so its total length will be about 370 km.

The GPS Coordinates of selected eleven Observatory points / s1-s11/ along the river [7], is shown in table 1:

Table 1.GPS Observatory points

Observatory	Name	Latitude	Longitude
S1	Amiri	36.500	34.458
S2	Rabbla	36.600	34.459
S3	Qanater	36.534	34.521
S4	Al-Nabi Mandu	36.523	34.560
S5	Qattineh Entrance	36.524	34.607
S6	Qattinah Exit	36.617	34.664
S7	Al-Sham Road	36.628	34.672
S8	Tripoli Road	36.628	34.689
S9	Al-Dweir	36.659	34.702
S10	Al-Ghantou	36.693	34.82
S11	Ghajar Amir	36.670	34.884

Sources of pollution on the Orontes River:

1) Sanitation drainage:

In Homs Governorate, there are more than 800 residential communities (villages, municipalities, countries and small communities), most of which are located within the Orontes Basin and part of the Coast Basin.

2) Industrial drainage:

Where the deterioration of the water quality increases after the exit of Qattinah Lake due to the untreated industrial wastewater loaded with pollutants, which is discharged directly into the Orontes River and these facilities (Fig. 2), and these facilities are:

1-The General Company for Nitrogenic Fertilizers: It includes the pollutants resulting from the manufacture of both fertilizers (calantro, urea phosphate, ammonium nitrate).

2-Hessia Industrial City: the industrial drainage that reaches the valley of Rabiah is one of the most important tributaries of the Orontes River.

3-The State Company for Homs Refinery: discharges containing untreated oil, oil and petroleum residues.

4-Military space: includes varying levels of grease and oils, in addition to copper, cadmium, and remnants of coatings.

5-Homs Dairy Company: Due to untreated discharges into the river.

6-Homs Sugar Company's factory: the company's expenses not treated well.

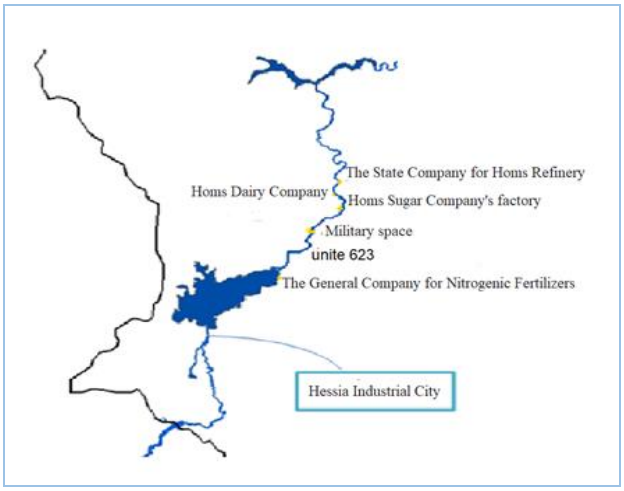

Figure 2. The distribution of Industrial drainage along the Orontes River

2- Statistical study:

Multivariate Analysis contain many Techniques, which used to analyse a set of data. Multivariate data analysis systems provide simultaneous examination of many variables in space and time, and statistics has become essential in the field of various research studies to help it in the interpretation of phenomena, data analysis and numerical application of the results obtained[8] .

Stages of the statistical process:

1. Data collection process.
2. Data classification.
3. The stage of data analysis by applying different statistical methods.
4. Formulating information, as (tables, charts, graphs) [9].
5.Explaining and interpreting the data by statistical methods.

Statistical Analysis:

It is concerned with studying and explaining the behaviour of different phenomena and variables, as it is used in data analysis, correlation matrices, or variance matrices for variables, and the aim of it is to clarify the relationships between those variables and result in a number of new variables called factors, and the multivariate statistical analysis includes a set of methods, the most important of which are:

• Cluster Analysis (CA).
• Factor Analysis (FA).

The statistical study was carried out using a program /SPSS/ statistical package for social science for the analysis of complex data [10]**.**

2-1. Cluster Analysis (CA):

Cluster analysis is based on revealing the behaviour of a group of data and working on compiling it into clusters, depending on the extent of the similarity in its characteristics, so that one cluster includes all variables that are similar in their behaviour with each other and different of others data in the second cluster[11].

2-2. Factor Analysis (FA):
Factor analysis aims to reduce the number of variables and simplify the correlations between the various variables involved in the analysis, as it gives a set of basic factors to explain the phenomenon, and clarifies the relationship between the variables, where factor analysis is used to extract the most influential factors, leading to finding a set of factors responsible for Contrast generation[12].

IV. RESULTS AND DISCUSSION

Clustering of Sites Using Cluster Analysis:
Hierarchical cluster analysis was performed on the dataset resulting from all the observatories along the Orontes River from its entry point to Homs until its exit from it for the years (2007-2020) in order to group the monitoring stations with similar characteristics in clusters using the SPSS program to study the possibility of reducing the number of sampling sites. Obtaining two clusters: / Cluster 1, Cluster 2/ where each cluster includes a group of observatories with similar characteristics according to the Dendogram (Fig. 3):

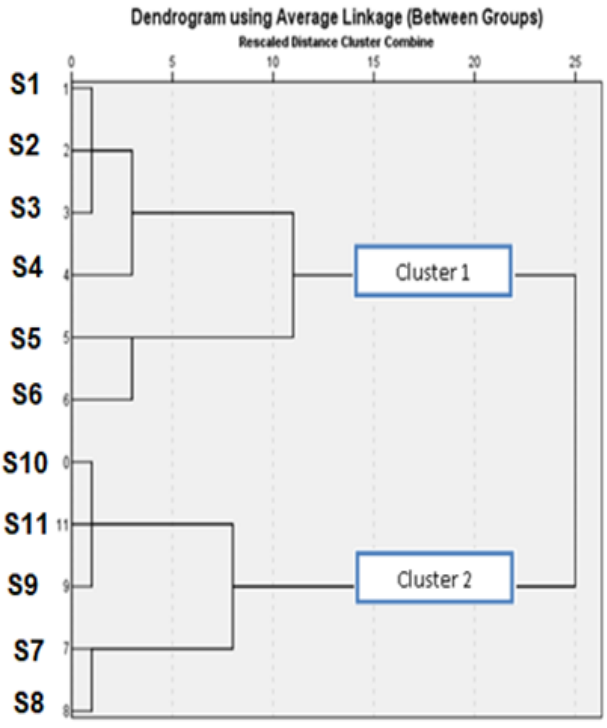

Figure 3. The cluster analysis tree diagram Dendogram

Apply factor analysis to variables:
Factor analysis was applied to the set of variables taken along the Orontes River from the period (2007–2020) To study the changes to compare the structural patterns between two Clusters, the results showed the value of the Kaiser-Mayer test according to the table / KMO / which shows the quality of the measurement with a degree of = 85.7%> 50%, meaning that the test quality is considered good, as shown in the following table / 2 /:

Table 2. Measurement Quality Score

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure	.857
Approx. Chi-Square	4008.200
df	136
Sig.	.000

Study of structural patterns according to clusters:

1) (Cluster1): By applying the factor analysis to the first cluster, which includes 6 observatories (Fig. 4):

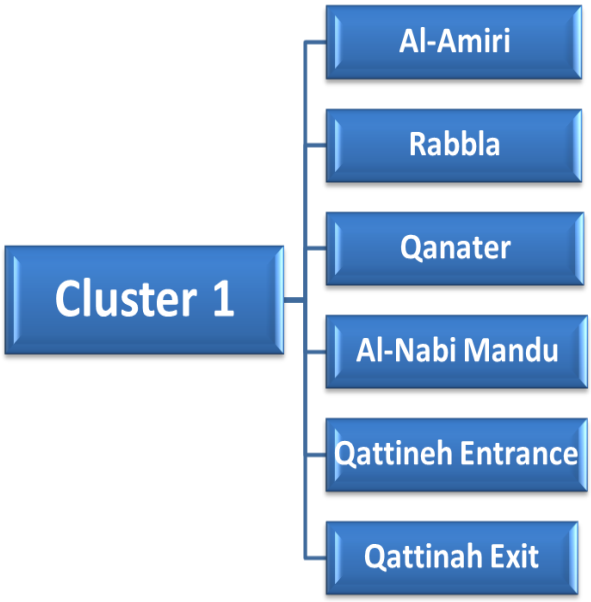

Figure 4. The Cluster 1 Observatories

The results of "Cluster 1" for low-pollution areas showed the presence of / 3 / basic factors including each factor has a set of basic variables that affect it in addition to the variance ratio for each factor. Table /3/ shows the variance ratio for each factor from the total variance:

Table 3. Total factor variance ratio

Total Variance Explained

Component	Total	% of Variance	Cumulative %
1	9.639	48.197	48.197
2	3.778	18.891	67.088
3	1.817	9.085	76.173

The analyses showed that there are three main factors responsible for 76.2% of the variation in water quality, where the first and second factors constitute 48.2% and 18.9%, respectively, and they belong to human activities (sanitation, agricultural drainage), and the remaining 9.1% is due to the remaining factors To industrial activities, as shown in Table /4/:

Table 4. Percentage of influencing factors

		component		
		48.2%	18.9%	9.1%
	Parametric	1	2	3
1	NO_3	.980		
2	COD	.935		
3	PO_4	.932		
4	BOD_5	.924		
5	K	.911		
6	NH_4	.858		
7	tur	.847		
8	TSS	.838		
9	EC		.957	
10	TDS		.940	
11	DO			.926

Thus, by applying the factor analysis for the sum of the variables for the first cluster, it gave /3/ basic factors and includes 11 basic variables, where the best interpretation ratio for the variance ratios for the total factors was 76.173% of the total variance, and a good interpretation ratio is considered for the total of /21/ variables :
(NO3, COD, PO4, BOD, K, NH4, TUR, TSS, EC, TDS, DO).

2) (Cluster2): By applying the factor analysis to the second cluster, which includes /5/ observatories (Fig. 5):

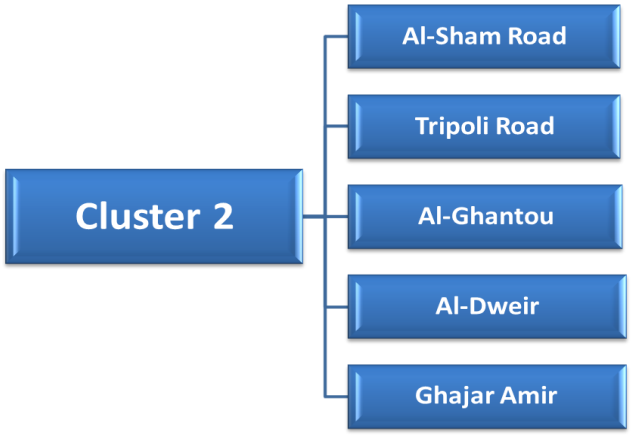

Figure 5. The Cluster 2 Observatories

The results of "Cluster 2" for the polluted areas showed the presence of /4/ basic factors and each factor includes a set of basic variables that affect it, In addition to the variance ratio for each factor, table /5/ shows the variance ratio for each factor out of the total variance:

Table 5. Total factor variance ratio

Total Variance Explained			
Component	Total	% of Variance	Cumulative %
1	8.589	42.945	42.945
2	3.703	18.517	61.462
3	3.397	16.987	78.449
4	1.883	9.415	87.864

The analyzes showed that there are four main factors responsible for 87.86% of the variation in water quality, where the first factor constitutes 42.95% and it belongs to industrial activities (industrial drainage), and the second factor is 18.52% which is related to human activities (estuaries of sewage), and the two workers The third and fourth, at a rate of 16.99% and 9.42%, respectively, indicate pollution by discharges (agricultural drainage and nitrogen fertilizer plant), as shown in Table /6/:

Table 6. Percentage of influencing factors

		Component			
		42.95%	18.52%	16.99%	9.42%
	parametric	1	2	3	4
1	EC	.965			
2	Na	.957			
3	BOD_5		.949		
4	COD		.944		
5	DO		.920		
6	NO_3			.962	
7	TSS				.982

Thus, by applying the factor analysis for the sum of the variables for the second cluster, it gave /4/ basic factors and includes 7 basic variables, as the best interpretation ratio for the variance ratios for the total factors was 87.864% of the total variance, and a good interpretation ratio for the total of /21/ variables is considered.

From these results, a specific set of variables were selected which are of greatest importance in controlling the quality of the Orontes River water. Which:
(EC, Na, BOD, COD, DO, NO3, TSS).

The results show the importance of the cluster analysis in classifying the sampling sites over the entire area being studied, which helps in reducing the number of sites scheduled for sampling according to the similarity of their specifications, thus reducing the costs of the sampling process, in addition to saving the effort and time required. Ordinary clusters (interpretation of the tree diagram /Fig. 3/ Dendogram of cluster analysis):

Cluster 1: This cluster is considered to be less polluted and includes (Al-Amiri Observatory, Rablah, Al-Qanater, Al-Nabi Mandu, Lake Qattineh entrance and Lake Qattinah exit). It is located at the beginning of the river's entry into Homs Governorate, where it was classified as low pollution and is almost pure due to the small amount of pollutants that you pour out.

Cluster 2: This cluster can be classified as polluted and includes observatories (Al Sham Road, Tripoli Road, Al Ghantou, Al Dweir, Ghajar Amir) and it was considered

polluted due to receiving large quantities of pollutants from the industrial areas, the fertilizer plant, Homs refinery, Wadi Al Rabiah, the mouth of the industrial city (Hassia) As industrial water is thrown untreated into the river, in addition to the sewage treatment of some communities located near the Orontes River.

For factor analysis results:

- **/Cluster 1/** for low-pollution areas showed the presence of / 3 / main factors, and each factor includes a set of basic variables that affect it in addition to the variance ratio for each factor, as:

- ✓ **The first factor:** it constitutes about (48.2)% of the total variance of the basic factors, and it has a strong relationship with 8 variables out of 21 which are (NO3, COD, PO4, BOD5, K, NH4, TUR, TSS). **These include a clear effect of seasonal changes on chemical and biochemical oxygen consumption.**
- ✓ **The second factor:** it constitutes about (18.9)% of the total variance and has a strong relationship with 2 out of 21 variables, namely (EC, TDS).
- ✓ **The third factor:** it constitutes about (9.1)% of the total variance, and it has a strong relationship with / 1 / variable out of 21, which is (DO).

- **/ Cluster 2/** showed the existence of / 4 / basic factors, and each factor includes the main variables that affect it, as:

- ✓ **The first factor:** it constitutes about (42.95)% of the total variance of the basic factors, and it has a strong relationship with 2 of the 21 variables which are (EC, Na).
- ✓ **The second factor:** it constitutes about (18.52)% of the total variance, and it has a strong relationship with 3 of the 21 variables, namely (COD, BOD, DO).
- ✓ **The third factor:** it constitutes about (16.99)% of the total variance, and it has a strong relationship with 1 variable out of 21 which is (NO3).
- ✓ **The fourth factor:** it constitutes about (9.42)% of the total variance, and it has a strong relationship with 1 out of 21 variable, which is (TSS).

V. CONCLUSION

By comparing Factor Analysis results between two clusters, it showed that by using variables of cluster 2, it gave /4/ basic factors, and the best rate for explaining the variance ratios was 87.86% of the total variance and included (7) basic variables, it is considered good for a total of (21) variables. This reflects that the most important variables that affect the quality of the Orontes River are (7) variables and they are classified according to importance as follows: (EC, Na, BOD, COD, NO3, DO, TSS).

From this result, we can see that the areas of the second cluster are greatly affected by industrial water drainage, while the areas of the first cluster have the effect of some sanitation and agricultural. Thus, we were able to determine the source of pollution according to the type of resulting variables.

VI. RECOMMENDATIONS AND FUTURE SCOPE

1- The results indicate the positive advantages of applying multivariate statistical methods in reducing the number of sampling.
2- Application of statistical analysis methods reduces sampling costs and reduces effort and time.
3- There is a possibility to conduct studies on more spacious areas with greater accuracy.
4- The possibility of reducing the number of measured parameters to know the quality of river water pollutants.
5- Statistical analysis methods can be used to monitor water quality and pollutant transport.
6- The importance of statistical analysis in selecting the most influential set of variables for designing the Water Quality Index /WQI/.

ACKNOWLEDGMENT

This work was supported by environmental directorate in Homs provided by the Department of environmental observatory

GLOSSARY OF TERMS:

FA	**factor analysis**
CA	**Cluster Analysis**
DO	**Dissolved Oxygen**
BOD	**Biological Oxygen Demand**
COD	**Chemical Oxygen Demand**
TSS	**Total Suspended Solids**
NO3	**Nitrate**
EC	**Electrical Conductivity**

REFERENCES

[1] W. Flint, "The Sustainable Development of Water Resources," Water Resources Update, Issue. 127, **pp. 41-51, Feb 2004.**

[2] M.Denchack, "Water Pollution: Everything You Need to Know", Networks R (https://www.nrdc.org/stories/water-pollution-everything-you-need-know).

[3] W. J. Cosgrove, D. P. Loucks, " Current and future challenges and research directions," Vol. 51, Issue. 6, pp. 4823-4839, 2015.

[4] L. Matejicek, " Modeling of Water Pollution in Urban Areas with GIS and Multivariate Statistical Methods," Paper presented at the 1 st Biennial Meeting of iEMSs, Lugano, Switzerland, **June 24-27, pp.60-65, 2002.**

[5] K. P. Singh, A. Malik, S. Sinha, D. Mohan, "Multivariate statistical techniques for the evaluation of spatial and temporal variations in water qualityof Gomti River (India)-a case study," Water Research,**Vol. 38, pp. 3980–3992, 2004.**

[6] K. Manoj, S. Ghosh, P. K. Padhy," haracterization and Classification of Hydrochemistry Using Multivariate raphical and Hydro Statistical Techniques," Research journal of chemical sciences**Vol 5, No 3, pp.32-42, 2013.**

[7] K. Rahman, S. Barua, S. Alam, A. Aalm, "Ecological Risk Assessment of Heavy Metals concentration in Sediment of the Shitalakhya River," International Journal of scientific in Multidisciplinary studies, **Vol. 6 Issue. 6, pp. 15-20, june 2020.**

[8] M.A.K.Shiker, " Multivariate Statistical Analysis," British jornal of scince,**vol.6 No 1, pp 55, july 2012.**

[9] A.Mousa, "Statistical Data Analysis," Center for Advancement of Postgraduate Studies and Research in Engineering sciences,Faculty of Engineering - Cairo University, **Vol. ,No.2372, 2005.**

[10] C.O. Adejoro, F. N. Ogwueleka, "Users Perception of Social Media Opportunities and Challenges," International Journal of Computer Sciences and Engineering, **Vol.8, Issue.4, pp.93, 2020.**

[11] Y. Qian, K. W. Migliaccio, Y. Wan, Y. Li, "Surface Water Quality Evaluation sing Multivariate Methodes and a New Water Quality Index in THe India River Lagoon, Florida," Sharing Science, **Vol 43, No 8 , pp.21, 2007.**

[12] A. Field, " Discovering Statistics Using SPSS third edition," (3 ed.). London: SAGE Publication, **2009**.

AUTHORS PROFILE

Ms.Lina Khouri, She is currently working as Head of the Environmental Observatory in the "Environment Directorate" in Homs Governorate, affiliated to the Ministry of Local Administration and Environment.

A member of the trustees of the Environmental Development Foundation and a member of the founders of the National Centre for Peace. She is studying Ph.D. degree in Environmental & Sanitary Engineering at Damascus University, She received master degree in Civil Engineering- Environmental Department. She has 11 years research Environmental experience.

