

## Effluent quality assessment of sewage treatment plant using principal component analysis and cluster analysis

**Saad Uffi Jebur**

M.Sc. student  
University of Baghdad  
College of Engineering  
Baghdad-Iraq  
[Saad.uffi79@gmail.com](mailto:Saad.uffi79@gmail.com)

**Dr. Basim Hussein Khudair**

Assistant Professor  
University of Baghdad  
College of Engineering  
Baghdad-Iraq  
[dr.basimal-obaidy@coeng.uobaghdad.edu.iq](mailto:dr.basimal-obaidy@coeng.uobaghdad.edu.iq)

### ABSTRACT

Sewage water is a mixture of water and solids added to water for various uses, so it needs to be treated to meet local or global standards for environmentally friendly waste production. The present study aimed to analyze the new Maaymyrh sewage treatment plant's quality parameters statistically at Hilla city. The plant is designed to serve 500,000 populations, and it is operating on a biological treatment method (Activated Sludge Process) with an average wastewater inflow of 107,000m<sup>3</sup>/day. Wastewater data were collected daily by the Mayoralty of Hilla from November 2019 to June 2020 from the influent and effluent in the (STP) new in Maaymyrh for five water quality standards, such as (BOD<sub>5</sub>), (COD), (TSS), (TP) and (TN). The results showed that the removal efficiency was 88%, 75%, 94%, 57%, and 77%, respectively. The results of the cluster analysis (CA) showed the formation of clusters in four stages and then gave the final shape consisting of two groups. At the same time, two influencing factors were extracted in the principal component analysis (PCA). The effluent's final quality (an average of eight consecutive months) complies with the stringent regulations proposed in the Iraqi Quality Requirements.

**Keywords:** Sewage treatment plant, Effluent quality assessment, Statistical analysis, Principal Component Analysis, Cluster analysis.

---

\*Corresponding author

Peer review under the responsibility of University of Baghdad.

<https://doi.org/10.31026/j.eng.2021.04.07>

2520-3339 © 2019 University of Baghdad. Production and hosting by Journal of Engineering.

This is an open access article under the CC BY4 license <http://creativecommons.org/licenses/by/4.0/>.

Article received: 2/10/2020

Article accepted: 30/12/2020

Article published:1/4/2021



## تقييم جودة النفايات السائلة في محطة معالجة مياه الصرف الصحي باستخدام تحليل المكونات الرئيسية والتحليل العنقودي

د. باسم حسين خضير العبيدي  
استاذ مساعد  
جامعة بغداد – كلية الهندسة

سعد عوفي جبر  
طالب ماجستير  
جامعة بغداد – كلية الهندسة

### الخلاصة

تُعرف مياه الصرف الصحي بأنها مزيج من الماء والمواد الصلبة تُضاف إلى الماء لاستخدامات مختلفة ، لذلك تحتاج إلى المعالجة لتلبية المعايير المحلية أو العالمية لإنتاج النفايات الصديقة للبيئة. هدفت الدراسة الحالية إلى التحليل الإحصائي لمعايير الجودة لمحطة معالجة مياه الصرف الصحي الجديدة بالمعمرة في مدينة الحلة. تم تصميم المحطة لخدمة 500000 نسمة وتعمل على طريقة المعالجة البيولوجية (عملية الحمأة النشطة) بمتوسط تدفق مياه الصرف الصحي يصل إلى 107000 متر مكعب / يوم ، تم جمع بيانات مياه الصرف الصحي يوميًا من قبل بلدية الحلة من تشرين الثاني 2019 إلى حزيران 2020 من المؤثر و مياه الصرف الصحي في (STP) الجديدة في المعيميره لخمسة معايير جودة المياه ، مثل (BOD<sub>5</sub>، COD، TSS، TP و TN). أظهرت النتائج أن كفاءة الإزالة كانت 88% ، 75% ، 94% ، 57% و 77% على التوالي. أظهرت نتائج التحليل العنقودي (CA) تكوين العناقيد في أربع مراحل ثم أعطت الشكل النهائي المكون من مجموعتين ، بينما تم استخلاص عاملين مؤثرين في تحليل المكون الرئيسي (PCA). الجودة النهائية للنفايات السائلة (بمتوسط ثمانية أشهر متتالية) تتوافق مع اللوائح الصارمة المقترحة في متطلبات الجودة العراقية.

الكلمات الرئيسية: محطة معالجة مياه الصرف الصحي ، تقييم جودة النفايات السائلة ، التحليل الإحصائي ، تحليل المكونات الرئيسية ، تحليل عنقودي.

### 1. INTRODUCTION

There is a great deal of spatial and temporal variability in the availability of water in Iraq. The rise in population and expansion of economic activity inevitably contributes to an increase in demand for water usage for various purposes. Water supplies in Iraq have also suffered significant stress in terms of water quantity, especially in the last two decades, due to various reasons, such as dams constructed on the Tigris and Euphrates in the riparian countries, global climate changes, and the extreme local decrease in annual precipitation rates and inappropriate planning of water uses within Iraq, (Jones et al., 2008; Trondalen, 2009; Rahi and Halihan, 2010). The amount and consistency of supplies coming from various sources influence the quality of water. Therefore, in order to maintain our knowledge and understanding of our climate, overall national planning and resource management in relation to water with emphasis on the allocation of priorities among different uses is important (Vaux, 2001). It is commonly used in many research publications relating to the needs of sustainable development (Parparov et al., 2006).

Water quality in an aquatic ecosystem is determined by many physical, chemical, and biological factors (Sargaonkar and Deshpande, 2003). Due to its direct impact on the environment and public health, there has been an increased interest in wastewater treatment in recent decades (Khudair and Awad, 2020). The disposal of wastewater without adequate treatment is an issue

related to health risks and degradation of the freshwater environment (**Khudair and Jasim, 2017**). Sewage is released into the environment characterized as a mixture of waste plus solids from various uses, such as domestic, industrial, commercial, that contains high organic material, several pathogenic microorganisms, and some toxic compounds (**Zhou, 2002**). If the sewage is not treated and the effluent waste is allowed to be disposed into the water body, the environment will be at risk (**Bonmatí and Flotats, 2003**). The bioremediation processes are cost-effective and require no maintenance, easy operation, and eco-friendly (**Metcalf and Eddy, 1991; Beach, 2001**). In fact, the conventional activate sludge (CAS) processing plant is highly effective in removing organic carbon and nutrients (BOD= 96.6%, COD= 94.8%, SS= 77.1%, N= 84.4% and P= 88.0%) (**Colliver and Stephenson, 2000**).

The use of various multivariate statistical techniques such as Cluster Analysis (CA), Principal Component Analysis (PCA) and Factor Analysis (FA) helps to understand better the effluent quality and ecological status of the studied systems in the interpretation of complex data matrices, enabling the identification of possible factors affecting the sources of sewage treatment (**Igbinosa and Okoh, 2009**). The present study aims to assess the Mayoralty STP's sewage quality located in Babil Governorate in Iraq using the principal component analysis and cluster analysis.

## 2. Materials and Methods:

### 2.1 Study area description

A sewage treatment plant was established in Al-Hilla city (Al-Maaymyrh area), the center of Babel governorate, located south of Baghdad governorate, in 1982, serving about 11,097 people with an actual capacity of 12000 m<sup>3</sup>/day as shown in **Fig. (1)**. A new plant was created adjacent to the old plant, and an actual card of 107,000 m<sup>3</sup>/ day showing the details of its units in **Fig. (2)** and entered the service at the end of 2019, which is called the treatment plant in Al-Maaymyrh. The treatment plant in Maaymyrh consists of **Table (1)** (**Babylon Sewerage Directorate, 2020**).



**Figure 1.** Maaymyrh sewage treatment plant, Hilla (**Al-Maaymyrh project office, 2012**).



Table 1. Locations name of Al-Maaymyrh project.

No. Location	Location name
1	units Pool plant liquid spray deodorant
2	Coarse bar screen room
3	lifting pumping station
4	Fine bar screen room
5	Aerated grit sedimentation tank
6	Anaerobic tank
7	Anoxic zone
8	Oxidation ditch
9	Sludge lifting pump tank
10	Final settling tank
11	Distribution water and sludge well
12	Contact tank
13	Water pump room
14	Mechanical thickening Room and Sludge storage tank

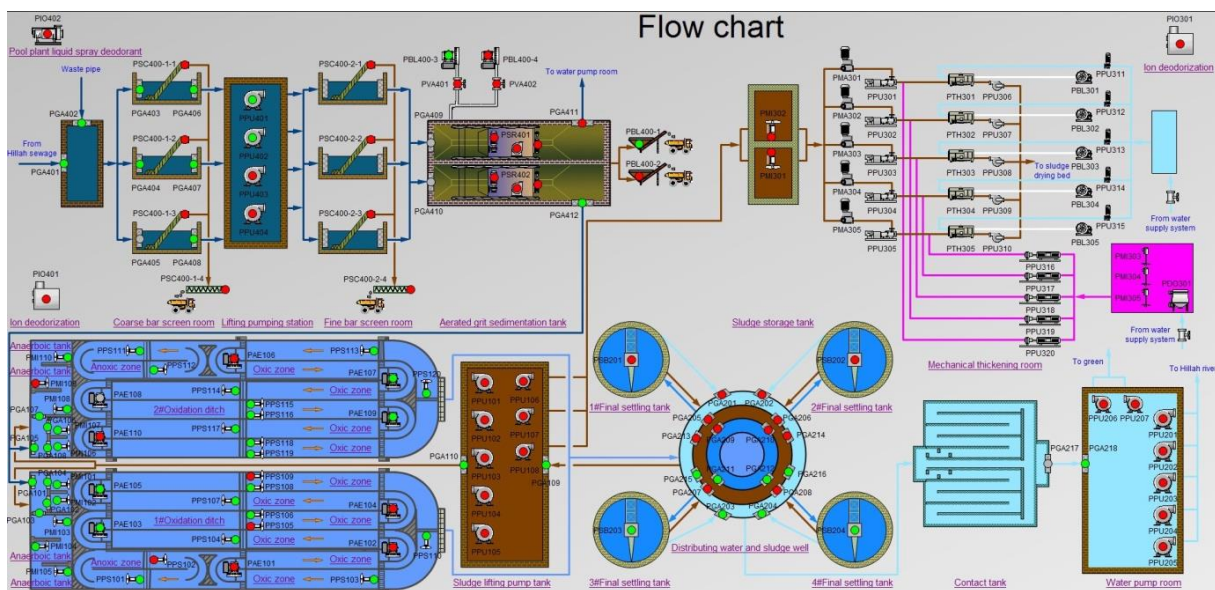


Figure 2. the details in the new AL-Maaymyrh STP.

### 2.2 Data collection and analysis

The effluent and impact data at the new wastewater treatment plant (STP) in Maaymyrh were collected in this eight-month study, starting from November 2019 to June 2020, by Al Hilla Sewage Directors. The main water quality criteria were selected; Biological oxygen demand





(BOD), Chemical oxygen demand (COD), Total suspended solids (TSS), Total nitrogen (TN), Total Phosphorous (TP), pH, and Temperature (T). Samples of raw sewage and liquid waste from Maaymyrh STP were taken daily in 1-liter jars and kept at 4°C during transportation to the laboratory. All tests were performed by the personnel working in the laboratories of the new Maaymyrh STP and according to standard sewage analysis methods (APHA et al. 1998)

### 3. STATISTICAL ANALYSIS CONCEPTS

In water quality analysis, the possibility of applying multivariate statistical techniques has been repeatedly demonstrated to classify significant components or variables that explain most of a system's variances. It is intended to minimize the number of variables to a limited number of indicators (i.e., key components or variables) while retaining relationships in the original data and change in water quality characteristics (Milanovic et al. 2010, Varol et al. 2012). In this analysis, mean, maximum, minimum, and standard deviation values were determined using SPSS V25 over eight months. In this research, multivariate data analysis using Principal component analysis (PCA), cluster analysis (CA), and factor analysis (FA) was performed.

#### 3.1 Principal Component Analysis (PCA)

The variance of large data sets of inter-correlated variables can be represented and transformed into smaller sets of independent variables (main components) via a strong pattern recognition technique. (Simeonov et al., 2003). The principal components give information on the most relevant parameters and describe the entire data sets by data reduction with a minimal loss of original information. (Vega et al. 1998, Helena et al., 2000). In brief, linear weighted combinations of the initial variables are the PCA. The main objective of PCA is to check for new latent (hidden) variables of  $n$  samples orthogonal (not correlated) to each other (Jolliffe, 2002). The linear combination of  $p$  variables  $x_i$  is each primary latent variable and represents a specific total variance source.

$$X = TWT + E \quad (1)$$

Where the data matrix is  $X$  ( $n \times m$ ), the matrices of the principal components scores and loadings are  $T$  ( $n \times p$ ) and  $W$  ( $m \times p$ ), respectively, and the residual matrix is  $E$  ( $n \times m$ ) (Hubert and Debruyne, 2009; Rousseeuw and Driessen, 1999)

#### 3.2 Cluster Analysis (CA):

the unsupervised pattern recognition method demonstrates the intrinsic structure of a data set without making a priori assumptions on the data to classify objects of the system into classes or clusters based on their proximity or similarity (Vega et al. 1998). The most common method in which clusters are formed sequentially is hierarchical clustering, beginning with the most similar pair of items and creating higher clusters in a gradual pattern. Within multidimensional space, it is simply the geometrical distance and is calculated as

$$D(X, Y) = \{\sum_i (X_i - Y_i)^2\}^{\frac{1}{2}} \quad (2)$$

Where,

$D(X, Y)$  is the distance between two points  $x$  and  $y$



**3.3 Factor analysis:** is a contribution of less relevant PCA-derived variables and extracts a new group of variables, known as latent variables (LVs), by rotating the PCA need for the defy axis. Hypothetical, not-observable, latent variables can be used in LVs. Each variable can be expressed in factor analysis as a linear mixture of common latent factors and a single unique factor as follows. (Malinowski, 1991):

$$X_{im} = F_{i1}f_{1m} + F_{i2}f_{2m} + \dots + F_{ip}f_{pm} + e_{im} \tag{3}$$

$F_{im}$  is the loading factor where the common and unique (error) variables are  $F_{ip}$  and  $e_{im}$ , respectively.

**4. Results and Discussion**

**Table 2.** Descriptive statistics of the water quality data in Al-Maaymyrh STP.

Parameter	Influent					Effluent				
	N	Min	Max	Mean	St.Dv	N	Min	Max	Mean	St.Dv
BOD	55	53.00	388.0	125.14	66.31	55	4.00	30.00	15.82	5.58
COD	54	98.00	1050.	284.74	185.12	55	18.0	95.00	60.40	22.88
TSS	55	56.00	980.0	222.50	174.90	55	4.00	40.00	13.81	7.55
PH	55	6.91	8.01	7.55	.274	55	7.03	7.94	7.46	.22
T.N	46	.10	46.4	20.25	11.11	48	.10	25.30	9.54	5.55
T.P	51	1.40	57.0	8.02	8.63	51	.01	3.80	1.50	.99

**4.1 Raw sewage assessment**

Characteristics of sewage entering the treatment plant is shown in **Table 3**, which shows the changes in the properties of raw sewage over time, TSS, COD, BOD, NO<sub>2</sub>, NO<sub>3</sub>, TN, TP. The BOD<sub>5</sub>/COD ratio is considered as an indicator of biodegradation capacity (Metcalf and Eddy, 1985). If the BOD<sub>5</sub>/COD ratio is greater than 0.5, biodegradation will be simple if the biodegradation between 0.2 and 0.4 occurs only in an optimal thermal state, and the biodegradation will not proceed if the ratio is less than 0.2 (Contreras et al., 2003). Domestic sewage was found to typically have a BOD<sub>5</sub> / COD ratio of 0.4 to 0.8 (Metcalf and Eddy, 1985). A BOD<sub>5</sub>/COD ratio of 0.4 is commonly known as the distinction point between degradable and non-biodegradable waste (Turak and Afsar, 2004). In the research, the ratio of BOD<sub>5</sub>/COD in raw wastewater was approximately 0.54, suggesting a significant quantity of biodegraded organic matter.



**Table 3.** monthly variation of the average concentration of raw sewage parameter.

Year	Month	BOD	COD	TSS	pH	TN	TP	BOD <sub>5</sub> /COD
2019	November	88.71	354.92	211.28	7.24	20.81	5.49	0.25
	December	164.67	344.66	344.44	7.42	9.51	14.58	0.47
2020	January	102.55	237.88	170.22	7.53	23.45	4.511	0.43
	February	171	293.11	250.88	7.84	15.78	8.82	0.58
	March	140	232.6	256.8	7.64	21.37	11.6	0.6
	April	93.75	144	106.25	7.83	31.42	6.25	0.65
	May	118.33	169	133	7.81	28.03	5.86	0.7
	June	133.5	200	141	7.99	17.5	7.85	0.67
Average		126.56	247.02	201.73	7.66	20.98	8.12	0.54

- Water quality units for all parameters in mg/L

By collecting data in the collected wastewater entering into the plant, noted that its specifications are from weak to moderate concentration levels, as weak, medium, and strong levels of BOD<sub>5</sub> were determined as 110, 220, and 400 mg/L, and dissolved solids concentrations were determined as 100, 220 and 350 mg/L respectively (**Tchobanoglous and Burton, 1991**).

#### 4.2 Effluent Quality Assessment

The treatment plant's efficiency was evaluated by the characteristics of the treated sewage from the plant, as shown in table 4. To characterize the quality of the sanitation averages Based on the results, the maximum and minimum values for the selected parameters were calculated, as shown in **Table (2)**, which illustrates the descriptive values of the data.

**Table 4.** monthly variation of the average concentration of treated sewage parameter.

Year	Month	BOD	COD	TSS	pH	TN	TP	BOD <sub>5</sub> /COD
	IQS 2012	40	100	60	6-8.5	<10	<1	0.1-0.3
2019	November	18.15	60.29	19.21	7.24	12.64	1.17	0.3
	December	19.44	67.55	14.88	7.37	5.86	1.1	0.29
2020	January	15.55	79.889	13.11	7.41	11.45	1.81	0.19
	February	15.4	50.44	13.11	7.62	9.16	1.3	0.31
	March	11.2	52.4	9.2	7.5700	0.475	1.92	0.21
	April	13	43	7.75	7.6525	10.65	2.6	0.3
	May	10.66	58.66	7.33	7.69	12.33	0.93	0.18
	June	11	43.5	11	7.93	6.7	2	0.25
Average		14.3	56.97	201.73	7.56	8.66	1.6	0.25

- Water quality units for all parameters in mg/L

The composition of effluent treated sewage varies from the facility by treatment stage, a form of households, enterprises, industries, and discharging public facilities into the system, and this may be a significant contributing factor to the observed variations in pH. A water system's pH level determines its utility for a number of uses. Quality standard for water sources according to regulations issued by Iraqi government Regulation 25. In general, influencer and effluent pH



values were slightly alkaline (mean value), ranging from 6.91 to 8.01 and 7.03 to 7.94, respectively, with a mean value of 7.55 and 7.46 during the study period, as shown in Table 2. This refers to the decrease in the concentration of dissolved CO<sub>2</sub> by reducing the concentration of organic matter. (Colmenarejoa et al., 2006)

TSS is a critical element in regulating sewage discharge, which represents a part of being the root of aesthetic disturbance along the banks of the river; TSS causes irrigation systems harm in the form of irrigation Tubing, sprinklers, emitters, and narrow water channels may be blocked by algae. TSS can adsorb heavy metals on their surfaces and thus encourage complexes of heavy metals (Nkegbe, E. et al., 2005). The means of raw influential TSS concentration was 222.50, while treated effluent TSS means were found to decline sharply to 13.81. In contrast, the raw impact concentration of TSS varied greatly (between 56.00 and 980.00). A slight variation in the TSS concentration of treated effluent (between (4.00 and 40.00) was observed over the eight-month. This means that the efficiency of the Maaymyrh STP is independent of the characteristics affecting it. The processing plant discharges its effluent into the Shatt Al-Hilla River; high TSS can decrease the intensity of sunlight in bodies of water and decrease primary productivity, particularly on green algae, and less light can also have a negative impact on the primary and secondary development of aquatic life and the stratification of the temperature system in aquatic environments (Nkegbe et al., 2005).

The most important factors used to verify the quality of sewage are chemical oxygen demand (COD) and biological oxygen demand (BOD). They reflect the amount of organic matter in sewage (Huertasaet al., 2008). COD is a measure of the amount of oxygen required by a strong oxidant (for example, H<sub>2</sub>SO<sub>4</sub>) to analyze both organic and inorganic substances in an aqueous system (Akan et al., 2008). High COD levels in water lead to severe oxygen depletion, which adversely affects aquatic organisms (Fatoki et al., 2003). The agonist's chemical oxygen COD demand ranged from an average value of 247.02 mg/L to an average value of 56.97 mg/L for the remaining COD value in the final flow (Table 2). At an average value of 126.56 mg/L, the corresponding biological demand for BOD in raw water varied. At an average value of 14.3 mg/L, the BOD concentration in the handled effluent complied with Iraqi quality standards.

### 4.3 Removal efficiency assessments

Treated sewage is one of the alternatives to be used for watering crops and others. Many physical, chemical and biological processes are designed and operated to simulate natural treatment processes to reduce the pollutant load to the level that nature can deal with. In this regard, it is necessary to pay special attention to assessing the environmental impacts of existing wastewater treatment facilities and achieving efficiency in removing pollutants at high rates (Jamra, 1999).





Table 5. Monthly variance in the removal efficiency of parameters for water quality.

Year	Month	BOD%	COD%	TSS%	TN%	TP%
2019	November	79.54	83.01	90.91	39.25	78.69
	December	88.19	80.4	95.68	38.39	92.45
2020	January	84.84	66.42	92.3	51.17	59.88
	February	90.99	82.79	94.77	41.95	85.26
	March	92	77.47	96.42	97.78	83.45
	April	86.13	70.14	92.71	66.1	58.4
	May	90.99	65.29	94.49	56.01	84.13
	June	91.76	78.25	92.2	61.71	74.52
	Average	88.06	75.47	93.69	56.55	77.10

The removal efficiency results in Maaymyrh STP during November 2019 to June 2020, achieved good efficiency over eight months excluding total nitrogen removal and total phosphorous removal, it was noted there is a decrease in removal rates for a some months, and this may be due to insufficient process control as shown in **Table (5) and Fig. (3)**.

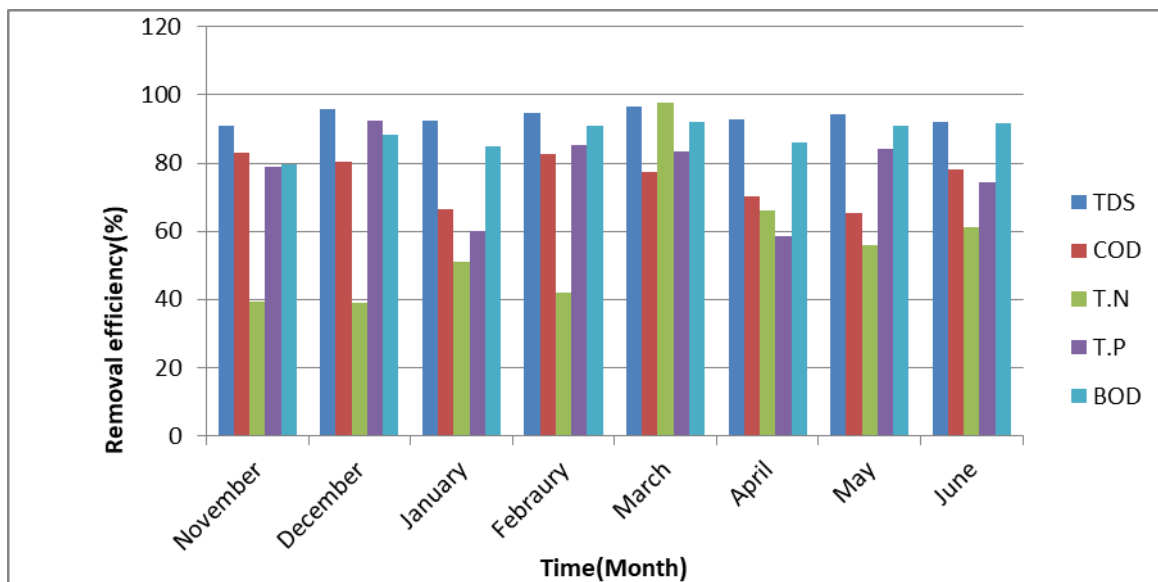


Figure 3. Removal efficiency of selected effluent parameters in the Maaymyrh STP.

### 4.3 Cluster analysis

The cluster analysis findings were seen as a genogram in which the gap between two months corresponds to the similarity and dissimilarity between two months (in terms of treatment effectiveness, the similarity between two months). **Fig. (4)** showed the process of forming clusters with four stages. It then gave the final form consisting of two clusters, one of which is made up of several branches, where in the first stage, the amount of similarity in terms of its efficiency was shown in Cluster 1 for months (November, December, February) as well as similarity the efficiency of the second cluster of months (June, April, and July) and the third cluster represented the efficiency of the plant in March. The fourth cluster represents the efficiency of the plant in May. The second phase consists of three clusters; the first cluster represents the plant's efficiency



for the months (November, December, February, and May). The second cluster represents the plant's efficiency for the months (January, April, and June). The third cluster represents the plant's efficiency for the month (March). The fourth stage, in its final form, consisted of two clusters, where the first cluster gathered the efficacy of the plant for months (November, December, January, February, April, May, and June), as it was close to efficiency, while the second cluster represented the efficiency of the station for March, which was higher than the other months, where we observe the high efficiency of the removal of large nitrogen and total phosphorous (Singh et al., 2004, Pejman et al. 2009, Varol et al. 2012).

The method of cluster analysis is considered one of the most important methods used in the statistical analysis of the components in order to easily know the efficiency of the treatment plants to diagnose the problems that occur in the plant and access a high-efficiency operation on a regular and continuous basis

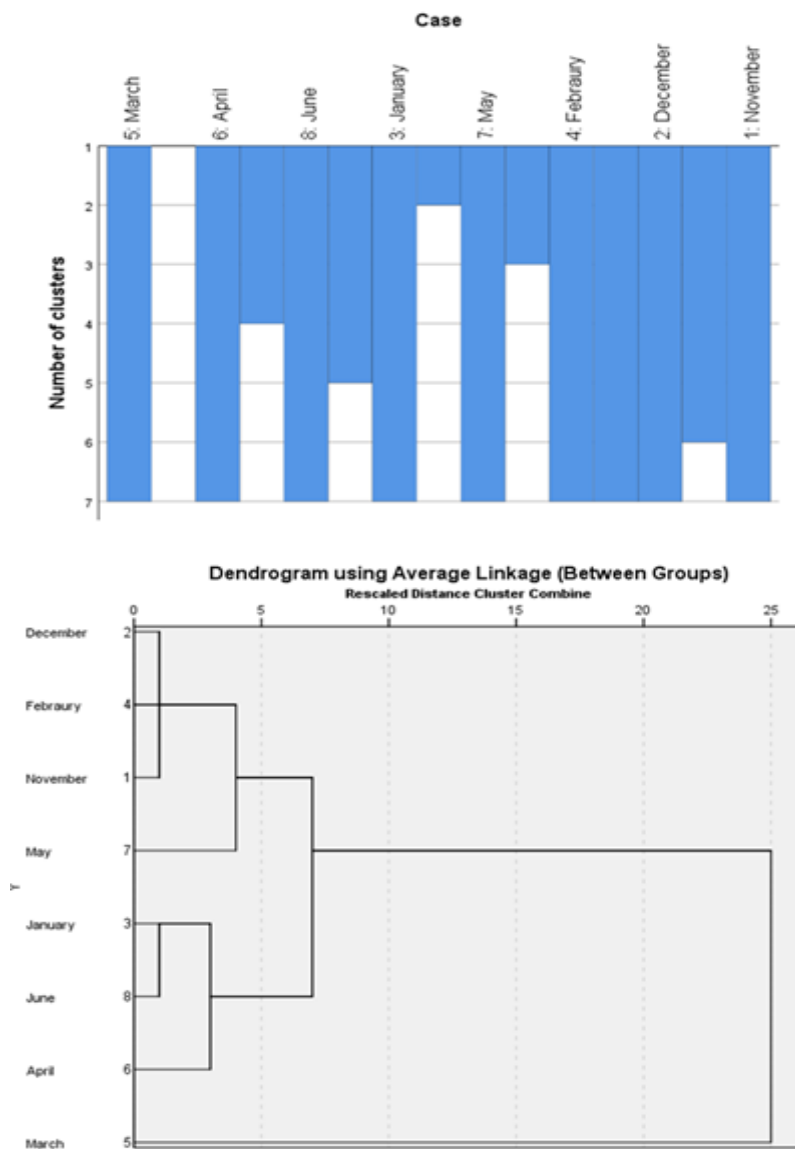


Figure 4. Dendrogram using Average Linkage (Between Groups).



#### 4.4 Principal component analysis/factor analysis

The data obtained from the laboratory analysis was used as PCA/FA inputs and output variable of the sewage samples from the Maaymyrh STP. It is described by six physical and chemical parameters, before the data analysis was standardized to generate the normal distribution of all variables since water quality standards have different metrics and scales (Davis, 1973). The loading factor value was categorized as strong (LF > 0.75), moderate (0.5-0.75) and weak (0.4-0.5) (Vega et al.,1998). In analyzing the components of raw sewage entering the STP, two factors were extracted with eigenvalues > 1, representing more than 75.7% of the total variance in the dataset. As shown in Table (6) and Fig. (5). The first factor constitutes 54.192 % of the overall variance and includes TSS, BOD, TP, and COD - with strong positive loading. This indicates a rise in the total suspended solids, as well as organic materials in the sewage entering the treatment plant. The second factor explains 21.598% of the total variance and is strongly loaded with the pH, with a moderate TN positive loading. The high pH in the raw sewage entering the plant and the moderate increase in total nitrogen explains.

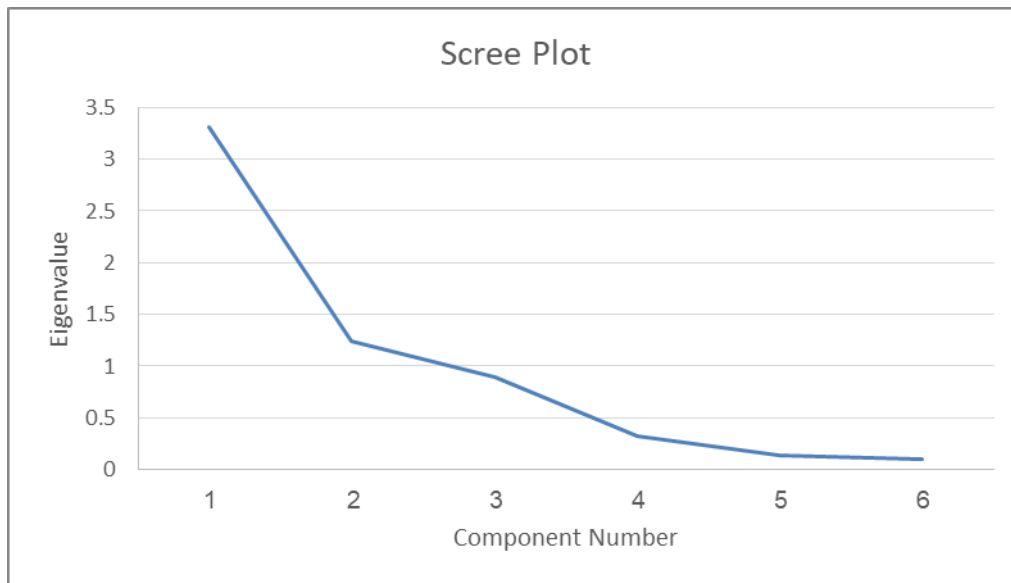


Figure 5. Scree plots of the eigenvalues of principal components.

**Table 6.** Eigenvalue and percentage of variance for the influent treatment plant.

Parameter	Factor 1	Factor 2
Eigenvalues	3.252	1.296
Total Variance Explained%	54.192	21.598
Cumulative %	54.192	75.79
Rotated Component Matrix		
BOD	0.926	0.127
COD	0.824	-0.314
TSS	0.939	-0.194
T.N	-0.169	0.564
T.P	0.892	0.002
PH	0.095	0.908

In this analysis of treated wastewater components from the water treatment plant, two factors were extracted with eigenvalues greater than 1, representing more than 54,896 of the total variance in the dataset, as shown in **Table (7)** and **Fig. (6)**. The first factor constitutes 34.638% of the total variance and includes the materials. TDS, with a strong positive load, indicates an increase in the total suspended solids in the treated wastewater leaving the treatment plant. Still, it is within the standard parameters (**Vega et al. 1998**), while BOD and COD are associated with a moderate positive load, this indicates The organic matter has been processed in a manner that meets the requirements of the standard but requires further treatment. The second factor explains 20.258% of the total contrast and is heavily loaded with TP, with a moderate positive loading of TN. This can be explained by the presence of total phosphorous and total nitrogen despite treatment due to poor management of the operation process. (**Vega et al., 1998**).

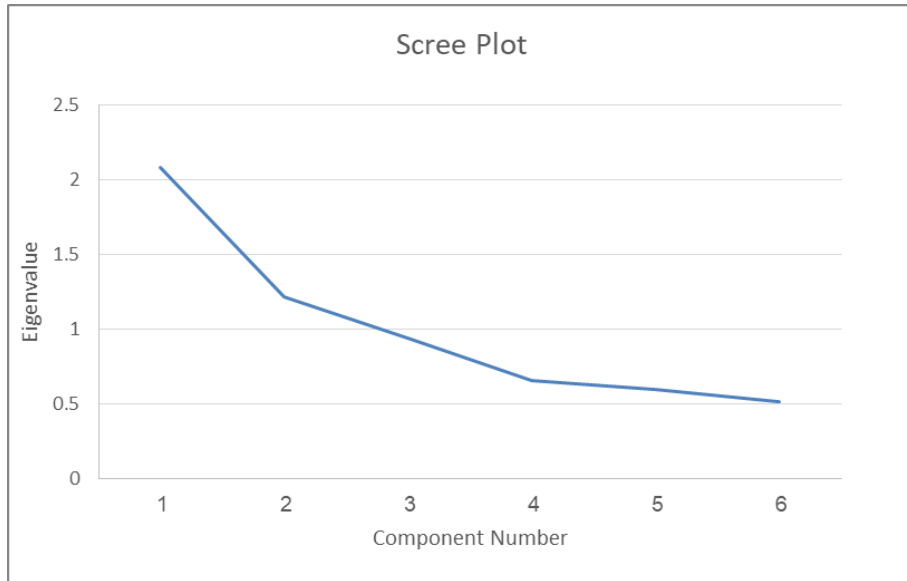


Figure 6. Scree plots of the eigenvalues of principal components.

Table 7. Eigenvalue and percentage of variance for the effluent treatment plant.

	Factor 1	Factor 2
Eigenvalues		
Total variance Explained %	34.638	20.258
Cumulative %	34.638	54.896
Rotated Component Matrix		
TSS	0.782	-0.07
COD	0.609	0.377
T.N	0.195	0.609
T.P	-0.281	0.799
BOD	0.728	0.143
PH	-0.67	0.193

### 5. CONCLUSIONS

The first step in determining the general characterization of domestic sewage in Hilla will be this research, and the findings are as shown below:

- 1- The concentrations of wastewater entering the plant varied from weak to medium. It was also observed that the ratio of BOD<sub>5</sub> /COD in raw water was about 0.54, indicating the presence of a large amount of organic matter subject to biodegradation.
- 2- The effluent's final content (in a mean value of eight consecutive months) complies with the regulations proposed in the Iraqi Quality Requirements.





- 3- The efficiency of TSS removal was found to be over 94 percent, and sewage treatment units emphasize that during sewage operations, the removal of suspended solids (TSS) is important. It was found that the BOD removal efficiency was over 88%, COD was 75%, TN was 57%, and TP was 77%.
- 4- Cluster analysis results showed that the months (November, December, January, February, March, April, and July) indicate an average level of treatment compared to the level of the month (March) when the efficiency was good. The sewage component analysis results by means of the analysis of the main components, the extraction of two influencing factors.
- 5- The overall performance of the current sewage treatment plant was satisfactory. A difference in efficiency was observed during the months, especially in the nitrogen removal process. Still, in general, the level of performance increased in recent months, as the plant was newly established and it needed a certain period for the growth of bacteria.

## 6. Acknowledgment

The authors would like to thank the Babil Sewer Directorate for allowing us to obtain data and provide facilities to communicate with the Station Manager, and the Laboratory Staff. The authors also thank the staff of the Sanitary Engineering Lab and the Civil Engineering Department/College of Engineering-University of Baghdad for their valuable support to complete this work.

## REFERENCES

- A. Parparov, K. D. Hambright, L. Hakanson, and A. Ostapenia., 2006. Water Quality Quantification: Basics and Implementation, *Hydrobiologia*, Vol. 560, No. 1, pp. 227-237.
- A. Sargaonkar and V. Deshpande., 2003. Development of an Overall Index of Pollution for Surface Water Based on a General Classification Scheme in Indian Context, *Environmental Monitoring and Assessment*, Vol. 89, No. 1, pp. 43-67.
- Akan J C, Abdulrahman F.I., Dimari G.A. and Ogugbuaja V.O., 2008. Physicochemical determination of pollutants in sewage and vegetable samples along the Jakara sewage channel in Kano Metropolis, Kano State, Nigeria, *European Journal of Scientific Research*, 23(1), 122–133
- APHA, AWWA, WEF, 1998 Standard methods for examination of water and wastewater, 19th end, American Public Health Association, Washington.
- Beach, M., 2001. Water, pollution, and public health in China. *Lancet*, 32(2): 358-735.
- Bonmatí, A. and Flotats, X., 2003. Air stripping of ammonia from pig slurry: characterization and feasibility as a pre-or post-treatment to mesophilic anaerobic digestion. *Waste Manag.* 23(2): 261-272.



- C. Jones, M. Sultan, E. Yan, A. Milewski, M. Hussein, A. Al-Dousari, S. Al-Kaisy, and R. Becker, 2008. Hydrologic Impacts of Engineering Projects on the Tigris-Euphrates System and its Marshlands, *Journal of Hydrology*, Vol. 553, No. 1-2, pp. 59-75.
- Chapman D., 1996. *Water quality assessments: A guide to the use of biota, sediments and water in environmental monitoring* (2nd ed.), London: UNESCO, World Health Organization, United Nations Environment Programme.
- Colliver, B., Stephenson, T., 2000. Production of nitrogen oxide and dinitrogen oxide by autotrophic nitrifiers. *Biotechnology Advances* 18, 219–232.
- Colmenarejoa M.F., Rubioa A., Sanchez E., Vicenteb J. , M.G. Garcya, and R. Borja, 2006. Evaluation of municipal sewage treatment plants with different technologies at Las Rozas, Madrid (Spain) *J. Environ. Manage.* 81:399-404.
- Contreras S., Rodriguez M., Al-Momani F., Sans C., and Esplugas S., 2003. Contribution of the ozonation pre-treatment to the biodegradation of aqueous solutions of 2,4-dichlorophenol, *Water Res.*, 37: 3164-3171.
- Davis, J.C., 1973. *Statistics and data analysis in geology*. John Wiley and Sons, Inc., New York. 550p.
- E.O. Igbinosa, and A.I. Okoh, 2009. Impact of discharge sewage effluent on the physico-chemical qualities of a receiving watershed in a typical rural community, *International Journal of Environment Science and Technology*, 6(2) 175- 182
- Fatoki S.O., Gogwana P., and Ogunfowokan A.O., 2003. Pollution assessment in the Keiskamma River and in the impoundment downstream, *Water South Africa*, 29(3), 183–187.
- H. J. Vaux, 200. *Water Quality (Book Review)*, *Environment*, Vol. 43, No. 3, p. 39.
- Hammer, M. J., 1986. *Water and sewage technology*
- Helena, B., Pardo, R., Vega, M., Barrado, E., Fernandez, J.M., and Fernandez, L., 2000. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis, *Water Research*, 34, pp. 807–816.
- Hubert, M., 2009. Debruyne, M. Minimum covariance determinant. *Comput. Stat.* 2, 36–43.
- Huertasa E., Salgota M., Hollenderb J., Weberb S. and Dottb W. et al., 2008. Key objectives for water reuse concepts, *Desalination*, 218: 120-131.
- J. M. Trondalen, 2009. *Climate Changes, Water Security and Possible Remedies for the Middle East*, United Nations Educational, Scientific and Cultural Organization.
- Jamrah, A.I., 1999. Assessment of characteristics and biological treatment technologies of Jordanian wastewater, *Bioprocess Engineering*, 21: 331-340
- Jolliffe IT, 2002. *Principal component analysis*, 2nd eds. Springer, New York
- K. A. Rahi and T. Halihan, 2010. Changes in the Salinity of the Euphrates River System in Iraq," *Regional Environment Change*, Vol. 10, No. 1, pp. 27-35.



- Khudair, B., and Awad, S., 2020. Performance Evaluation of Trickling Filter and Extended Aeration of Wastewater Treatment Plants, *Journal of Engineering*, Number 3 Volume 26
- Khudair, B., and Jasim, S., 2017. Performance Evaluation of Sequencing Batch Reactor and Conventional Wastewater Treatment Plant based on Reliability assessment. *Journal of Engineering*, 23(11), pp. 97, 112.
- Liu, C.W.; Lin, K.H. and Kuo, Y.M., 2003. Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan. *Science Total Environment*, 313(1-3): 77-89
- Malinowski ER (1991) *Factor analysis in chemistry*, 2nd eds. Wiley, New York
- Metcalf and Eddy INC., 1985. *Sewage Engineering Treatment, Disposal and Reuse*, 3rd Edn. McGraw-Hill, New York, USA.
- Metcalf and Eddy, Inc., 1991. *Sewage Engineering: Treatment Disposal and Reuse*, third edition. New York: McGraw-Hill.
- Negev E., Emongor V. and Koorapetsi I., 2005. Assessment of Effluent quality at Glen Valley Sewage Treatment Plant, *Journal of Applied Sciences* 5 (4): 647-650.
- Pejman, A.H., Nabi Bidhendi, G.R., Karbassi, A.R., Mehrdadi, N., and Esmaeili Bidhendi, M., 2009. Evaluation of spatial and seasonal variations in surface water quality using multivariate statistical techniques, *International Journal of Environmental Science and Technology*, 6, 3, pp. 467–476
- Reghunath, R., Murthy, T.R., and Raghavan, B.R., 2002. The utility of multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India, *Water Research*, 36, pp. 2437–2442.
- Rousseeuw, P.J.; Driessen, K.V., 1999. A Fast Algorithm for the Minimum Covariance Determinant Estimator. *Technometrics* 41, 212–223.
- Simeonova, P., Simeonov, V., and Andreev, G., 2003. Water quality study of the Struma River Basin, Bulgaria, *Central European Journal of Chemistry*, 2, pp. 121–136
- Singh, K.P., Malik, A., Mohan, D., and Sinha, S., 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) – a case study, *Water Research*, 38, pp. 3980–3992.
- Tchobanoglous, G, Burton, F., 1991. *Wastewater Engineering, Treatment, Disposal, Reuse*, Third Edition,
- Uz G., Turak F., and Afsar H., 2004. Research of BOD and COD values of sewages that contain certain organic materials, Adnan Menderes University, 4th AACD Congress, Sept. 29-Oct. 3, Kupadasy-AYDIN, Turkey Proceedings Book 177.
- Varol, M., Gökot, B., Bekleyen, A., and Şen, B., 2012. Water quality assessment and apportionment of pollution sources of Tigris River (Turkey) using multivariate statistical techniques – a case study, *River Research and Applications*, 28, pp. 1428–1438.



- Vega, M., Pardo, R., Barrado, E., and Deban, L., 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis, *Water Research*, 32, pp. 3581–3592.
- Wunderlin, D.A., Diaz, M.P., Ame, M.V., Pesce, S.F., Hued, A.C., and Bistoni, M.A., 2001. Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba, Argentina), *Water Research*, 35, pp. 2881–2894
- Zhou and Smith Zhou, H., and Smith, D. W., 2002. Advanced technologies in water and sewage treatment