Al-Azhar University Civil Engineering Research Magazine (CERM)
Vol. (39) No. (4) October, 2017

Statistical clustering of water quality in El-Nubaria Canal, Egypt

Anass El-Molla[1], M. Shaban [2], M. A. Abdelall[1], and Ahmed Habash[2]

[1] Prof. of Irrigation Design and Hydraulics, Faculty of Engineering, Al-Azhar University.
[2] Drainage Research Institute, National Water Research Centre, Ministry of Water Resources and Irrigation, El-Kanater El-Khairia, PO 13621/5, Cairo, Egypt.

ملخص البحث :

يعد التدهور في نوعية المياه أحد القضايا الرئيسية التي تواجه إدارة الموارد المائية في مصر . ومن المفترض أن تكون نوعية المياه في القنوات الرئيسية مماثلة لتلك الموجودة في الفروع، وهي تشكل المصدر الرئيسي لمحطات معالجة مياه الشرب.والهدف الرئيسي من هذا البحث هو تحقيق أداة لرصد وإدارة نوعية المياه للري في ترعة النوبارية والتي تعد أحد أهم مصادر المياه في منطقة غرب دلتا النيل بجمهورية مصر العربية .ويتحقق ذلك بتصنيف مستويات تلوث مياه القناة إلى عدة فئات باستخدام نهج تحليل إحصائي جغرافي يمكن أن يضمن معلومات بسيطة ولكنها دقيقة عن مستويات التلوث ومعاييرنوعية المياه في أي موقع من مواقع الرصد الموجودة علي المجري المائي .ويتم إستخدام نظم المعلومات الجغرافية في تحديد المناطق المتشابهة في نوعية المياه علي المجري المائي لترعة النوبارية وعمل خرائط توضح التصنيف الإحصائي لمواقع رصد نوعية المياه علي المجري الرئيسي لترعة النوبارية . وتدعم هذه الخرائط الرقمية صانعي القرار في ادارة نوعية المياه علي المجري المائي الرئيسي لترعة النوبارية . وبالإضافة لذلك فإن التحليل الإحصائي الموضح علي الخريطة الرقمية الخاصة بترعة النوبارية تمكن صناع القرار من العثور على المناطق الساخنة التي تستحق مزيدا من االإهتمام في مراقبة التلوث. ونتيجة لذلك، تسهم نتائج هذا البحث التي تم الحصول عليها في إدارة ورصد نوعية المياه في ترعة النوبارية بطريقة تمكن صناع القرار من إتخاذ قرارت صائبة معتمدة علي رؤية واضحة لنوعية المياه عند مواقع الرصد المختلفة علي المجري المائي لترعة النوبارية .

Abstract

Deterioration of water quality is one of the major issues facing Egypt's water resources and management. Water quality in the main canals is supposed to be similar to those of the branches, and they comprise the main source for downstream drinking water treatment plants. The main objective of this research is to achieve a tool for monitoring and managing of water quality for irrigation in El-Nubaria canal. This is achieved by classifying the pollution levels of the canal water into several classes using a geo-statistical clustering approach that can ensure simple but accurate information about the pollution levels and water criteria at any location through the canal stream. Geographical Information System (GIS) is used to visualize the derived clusters by the mean of digital maps up on the entire system of El-Nubaria canal. The gained maps may support the decision makers to manage and control the water pollution in El-Nubaria canal. In addition, the geo-statistical clustering enables the decision makers to find the hotspots that deserve more monitoring attention to control the pollution. Consequently, the obtained research results contribute essentially in managing and monitoring the water quality of El-Nubaria canal.

1. Introduction

According to the (NWRP-2005), water quality surveys showed a general uniform distribution of parameters from Aswan to Cairo, and also showed that although the Nile receives enormous loads of different matters, it still maintaining its self-purification capacity. However, water quality deteriorates in the Nile branches due to disposal of agricultural drainage as well as decreased flow. Water quality in the canals is supposed

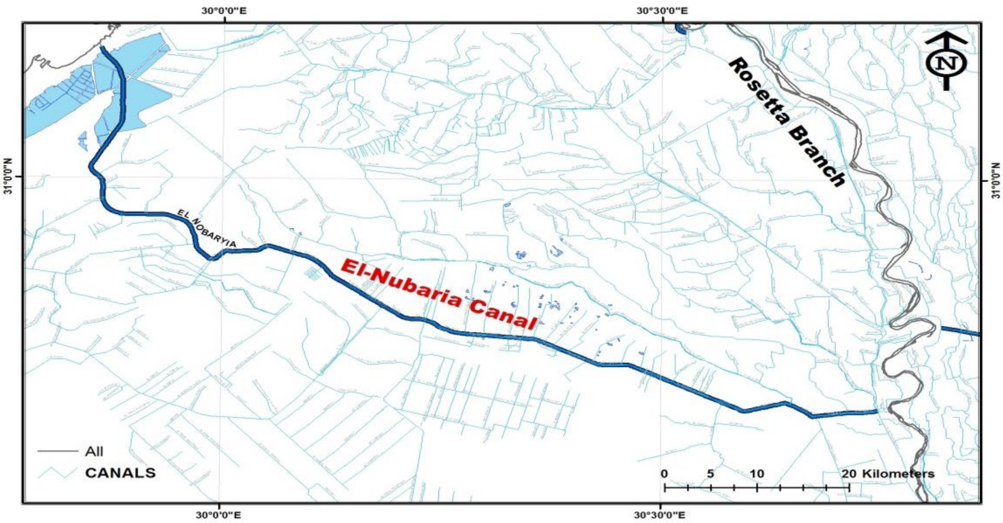
to be similar to those of the branches, and they comprise the main source for downstream drinking water treatment plants. However, most canals suffer from industrial and domestic wastes (liquid & solid). Egypt is considered as one of the most water scarce countries. The only dependable sustainable water resource in Egypt is the river Nile.
The Nile Delta irrigation canals water is deteriorated due to disposal of domestic; industrial and agricultural drainage effluents with decreasing flows (World Bank, 2005). The National Water Resources Plan (NWRP-2017 "Facing the Challenge" issued in 2005) has developed a national policy with three major pillars that are 1) increasing water use efficiency; 2) water quality protection; and 3) pollution control and water supply augmentation. Therefore, optimizing the water quality of irrigation is the main concern of any strategic planning for better water resources management in Egypt. The poor water quality has a direct impact on health and the environmental conditions. Reduction of pollution loads entering the water system will improve the water-related public health conditions, improve the sustainable use of groundwater resources, and contributes to meeting the water quality requirements of the various functions of the water system.
The low level of sanitation service especially in rural areas makes nearby water channels the suitable sites for inhabitant's sewage to be disposed. As the water quality status varies spatially according to the distribution of the human activities (i.e., agricultural and industrial), mapping the water quality parameters using GIS as a decision support system, can provide a useful basis for taking fast and reliable management decisions. GIS is an information system that is specially designed for handling spatial (or geographical) data. It has the advantage of handling attributed data in conjunction with spatial features, which was totally impossible with previous manual cartographic analysis (Anand Kumar, 2004; Veerabhadram, 2003). Statistical analysis was applied in which the median value for the monthly data were calculated for each monitoring site and then compared to the standards of Egyptian legislation for irrigation water uses. The objective of this research is to develop a tool to monitor and control the water pollution through the stream of El-Nubaria canal. This will be achieved by classifying the pollution levels of the canal water into several categories based on a geo-statistical cluster analysis method which provide simple but precise information about pollution level and water quality at any site in the canal monitoring system. This information will be visualized by using GIS to create the digital maps based on the pollution clusters for the monitored sites in El-Nubaria canal.

3. Methods and materials

El-Nubaria canal is located in the western Nile Delta (Fig. 1). El-Nubaria Canal is the major surface water source for the horizontal expansion projects in the West Delta Region. It is mainly fed by fresh water from El-Rayah El-Nassery and El-Rayah El-Beheiry (El-Gammal and Ali, 2008).

3.1. Methods

El-Nubaria Canal is the main canal which serves El Nubaria and El Nasr General Irrigation Directories. Total length of this canal is 100.00 km and cultivated land area on this big canal is 1.150.000 feddan. At present the average discharge of El-Nubaria Canal is 23 million m^3 per day serving a command area of about 1 million acres. Of this flow, 13 million are supplied from the Beheiri Canal. The remaining 10 million m^3 were originally designed to be supplied from El-Nassery Canal. Due to illegal abstractions from El-Nassery Canal, only 7 million are discharged from El-Nassery Canal to El-Nubaria Canal (Donia & Farag, 2010).



A) Clustering

Many researchers have used cluster analysis in their water quality publications. The Hierarchical Cluster Analysis was described by Güler et al. (2002) that it is an efficient means to recognize groups of samples that have the same chemical and physical characteristics. Various approaches of cluster analysis have been used to view water quality data for both surface and sub-surface waters (Santos-Roman et al., 2003; Odom, 2003; Dobbie et al., 2003; Zeng and Rasmussen, 2005; Ryberg, 2006). Cluster analysis classifies homogeneous sub-groups of monitoring sites in a statistical population by minimizing the Within-group and maximizing Between-group variations. Hierarchical and Relocation are the broad categories for conventional cluster analysis methods.

Relocation cluster analysis includes methods of expectation-maximization and k-means. These conventional methods are effective and precise on little datasets. Therefore, breaking these sets into smaller ones will be suitable for the small scale of dataset of El-Nubaria canal monitoring sites (Fig. 2).

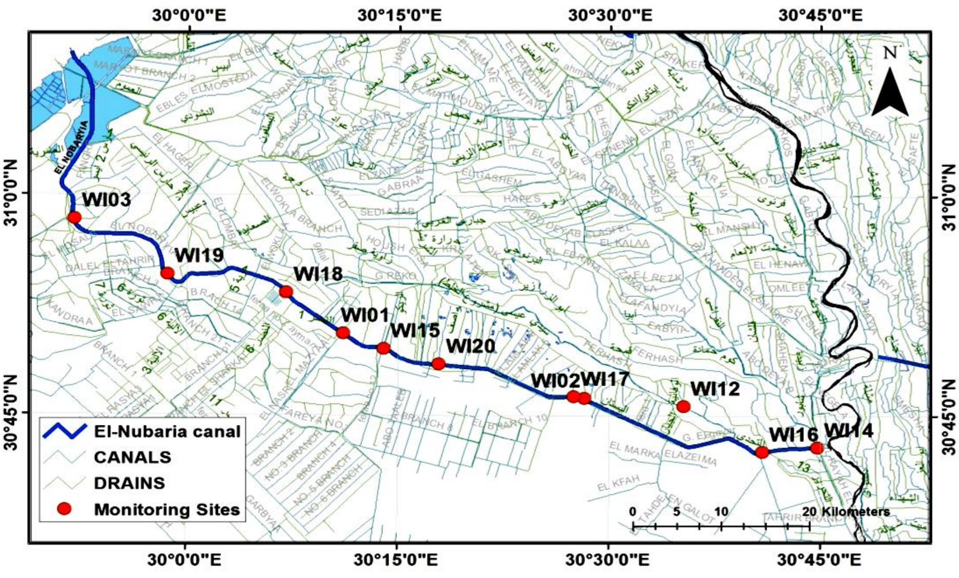


Fig. 2: Water quality monitoring sites cover El-Nubaria canal

The sequential clustering methodology (Theodoridis and Koutroumbas, 1999) examines the values one by one and decides if the current value should merge with the previously formed clusters or start a new one based on the distance criterion. In the next step, a cluster method is employed on the sub-clusters to check the desired number of clusters. The Two-Step Multivariate Cluster Analysis method depends on the likelihood distance measure which assumes that the variables in the cluster model are independent, each continuous variable is normally distributed and each categorical variable is assumed to have a multinomial distribution. However, empirical internal testing tells that the procedure is fairly robust to violations of these assumptions.

To cluster data of measured water quality, nine water quality parameters were selected. These were Total Suspended Solids (TSS mg/l), Chemical Oxygen Demand (COD mg/l), Biological Oxygen Demand (BOD mg/l), Dissolved Oxygen (DO mg/l), Nitrate (NO_3 mg/l), Total Nitrogen (TN mg/l), Negative Logarithm of Hydrogen Ion Concentration (pH), Total Dissolved Salts (TDS mg/l) and Ammonia (NH4 mg/l). The selected parameters represent the biological, physical, and chemical characteristics of El-Nubaria canal water. The analysis started with the application of a Two-Step Multivariate Cluster Analysis method procedure for the parameters monthly measured from August 2007 to July 2016. The clustering results then were illustrated using GIS the techniques.

B) GIS

A Geographic Information System (GIS) is a computer system that collects, stores, manages, calculates, analyzes and displays geographical or spatial data. Combined with data monitoring technology of Remote Sensing (RS) and data collecting technology of Global Positioning System (GPS), GIS has been widely used in environmental protection projects (Dabberdt et al., 2004; Warda and Johnson, 2007), including water quality monitoring of piped drinking water (Samuels et al., 2006) and spill management in inland waterways (Camp et al., 2010).

The GIS layers of the canals networks were extracted from digitized and scanned surveillance topographic maps and remotely sensed images for western Nile delta. Those layers include the irrigation system and monitoring sites. For instance as illustrated in Table (1), the site (WI14) represents El-Nubaria canal intake from El-Rayah El-Beheary; The site (WI15) represents El-Bustan canal intake from El-Nubaria canal and (WI03) represents Mariut canal intake and the end of El-Nubaria canal.

Table 1: Monitoring sites in the main drains in El-Nubaria canal

Site Code	Cluster	Site Name	Site Description
WI14	**1**	El-Rayah El-Beheary	El-Nubaria canal Intake
WI16	**2**	El-Nubaria	After Dalengat Extension
WI12	**3**	El-Hager	El-Hager Canal
WI02	**3**	El-Nubaria	El-Nubaria Before Bostan

Site Code	Cluster	Site Name	Site Description
WI17	**4**	El-Nubaria	Before Bostane
WI20	**2**	El-Nubaria	Before Umoum 3
WI15	**2**	El-Bustan	El-Bustan canal Intake
WI01	**2**	El-Nubaria	El-Nasr Canal Intake
WI18	**3**	El-Nubaria	El-Nubaria canal at km 61.00
WI19	**3**	El-Nubaria	Mariut PS No. 1
WI03	**3**	Mariut	Mariut canal Intake & End of El-Nubaria canal

3.2. Materials

In order to perform the statistical clustering, geo-referencing of topographic maps scaled 1: 50,000 for the western Nile delta were digitized to extract the features of main canals, branch canals and pump stations. The GIS data mapping software used in this research is the ArcGIS 10.5 Desktop. Utilizing the geographic position of monitoring sites on the irrigation system in the western delta were geographically located and referenced to their positions on the canal. Then they were overlain on the drains layers. The stream of El-Nubaria canal was divided into segments each actually represented by the water quality measurements of a monitoring site.

4. Results and Discussions

This section presents the detailed output of the statistical package SPSS 20 for the Two-Step Multivariate Cluster Analysis method that was carried out to classify the monitoring sites in the canals water quality monitoring network which cover El-Nubaria canal. Table (2) shows the clusters' distribution for the 11 monitoring sites. These sites were classified into 4 clusters. Table (3) presents the clusters centroids (median and standard deviations) for the examined parameters. The centroids showed that the clusters were well separated by the continuous variables.

As shown in figures (3) to (11) highest DO, TN and pH levels were recorded at the monitoring sites in cluster 1. These sites also showed lowest BOD, COD and TSS levels. However, TDS was generally moderate.

As illustrated in Figure (7) and table (3) no significant differences in the pH median in-between different clusters were recorded. The standard deviation of pH for all clusters is in a small domain.

As presented in figure (7) and table (3) the values of DO are in the same range of records for clusters (2, 3 and 4). The same case is also observed in figure (5); the TSS records are almost in the same domain for clusters (2, 3 and 4).

Table 2: Clusters distribution for water quality monitoring sites in El-Nubaria canal

Cluster	N	% of Combined	% of Total
1	1	9.09	9.09
2	4	36.36	36.36
3	5	45.45	45.45
4	1	9.09	9.09
Combined	**11**	**100.00**	**100.00**
Total	**11**	**100.00**	**100.00**

Table (3): The clusters centroids (median and standard deviations) calculated for the examined parameters

Cluster	1		2		3		4	
Statistics	**Median**	**Std. Dev.**	**Median**	**Std. Dev.**	**Median**	**Std. Dev.**	**Median**	**Std. Dev.**
BOD (mg/l)	6.0	2.2	8.0	4.4	7.0	7.3	12.0	5.2
COD (mg/l)	10.0	3.3	11.0	7.7	12.0	8.1	18.0	5.6
TSS (mg/l)	14.0	13.9	17.3	14.6	17.0	15.1	17.0	15.0
NO3 (mg/l)	0.3	0.3	0.6	1.4	0.2	0.3	1.3	1.1
NH4 (mg/l)	0.1	0.2	0.4	2.3	0.1	0.3	0.6	1.7
TN (mg/l)	4.9	3.7	4.0	4.6	3.9	4.8	2.6	2.2
TDS (mg/l)	306	148	319	283	333	281	303	298
DO (mg/l)	7.0	0.9	6.0	0.6	6.0	0.8	6.0	0.7
pH	7.9	0.2	7.9	0.2	7.9	0.2	7.8	0.3

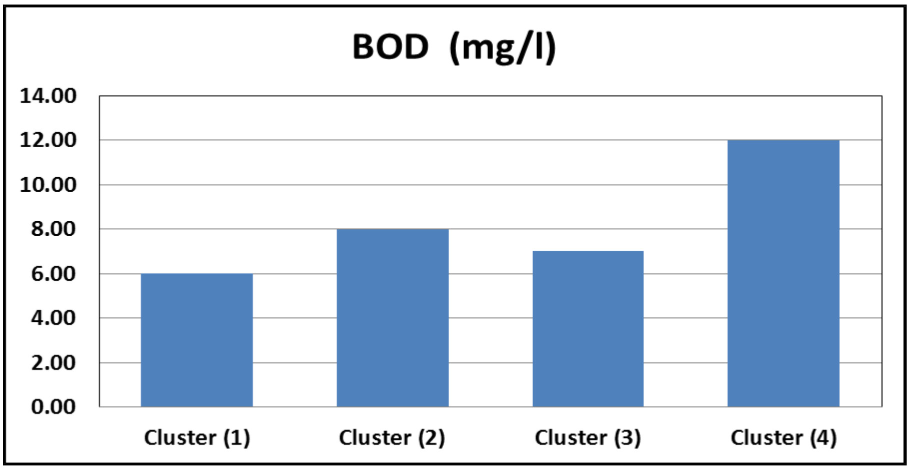


Fig. 3: the change in BOD (mg/l) between clusters

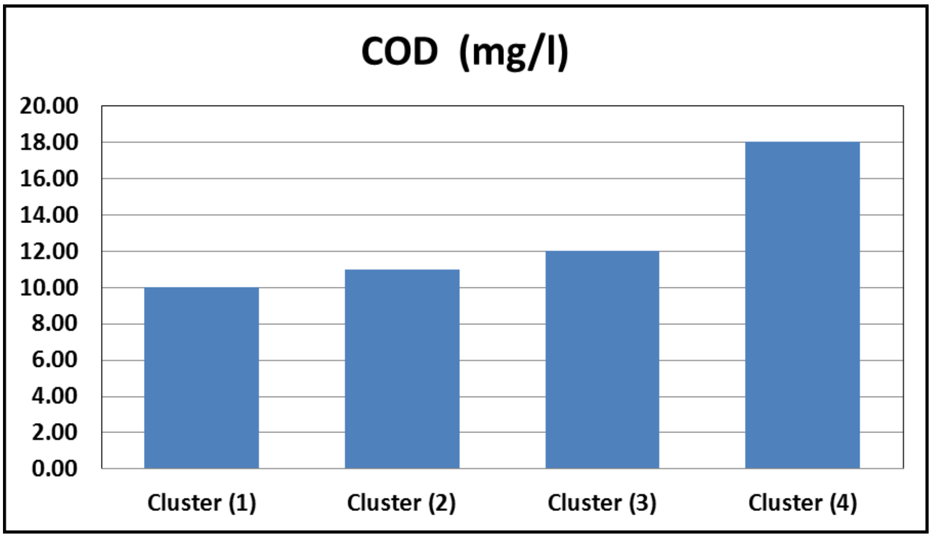


Fig. 4: the change in COD (mg/l) between clusters

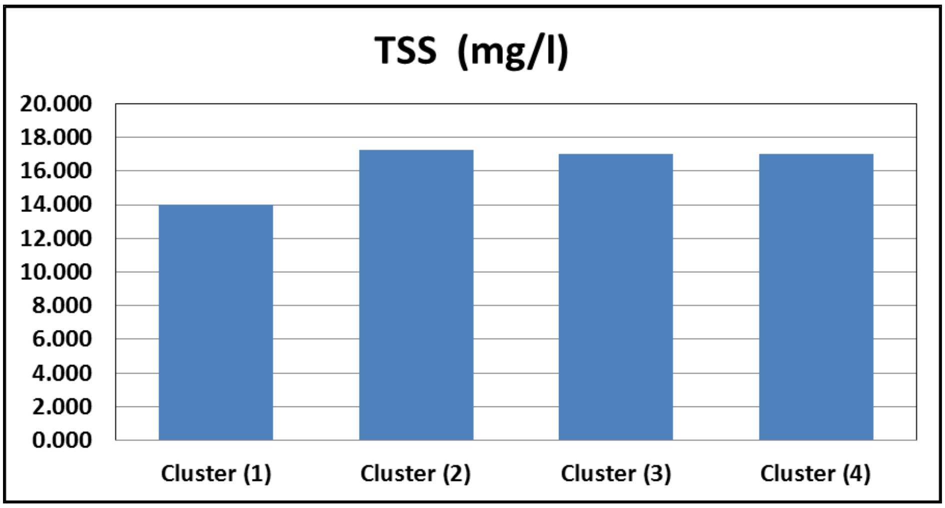


Fig. 5: the change in TSS (mg/l) between clusters

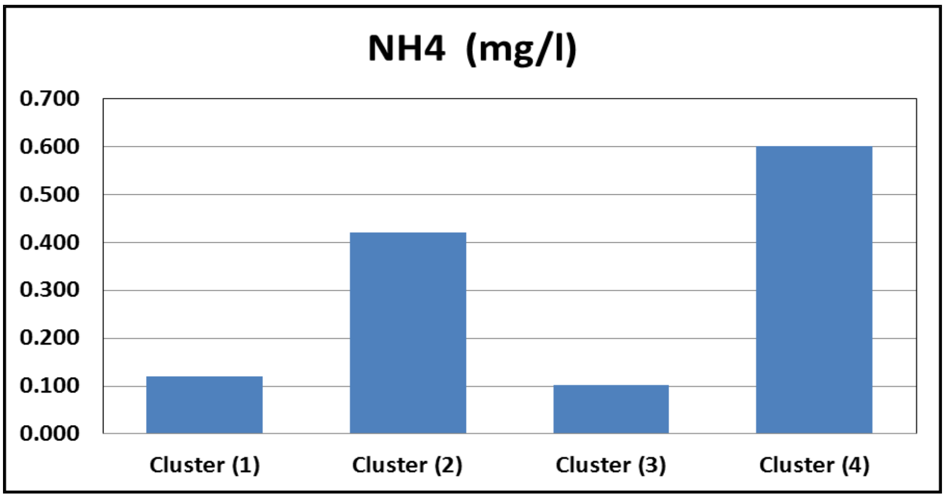


Fig. 6: the change in NH_4 (mg/l) between clusters

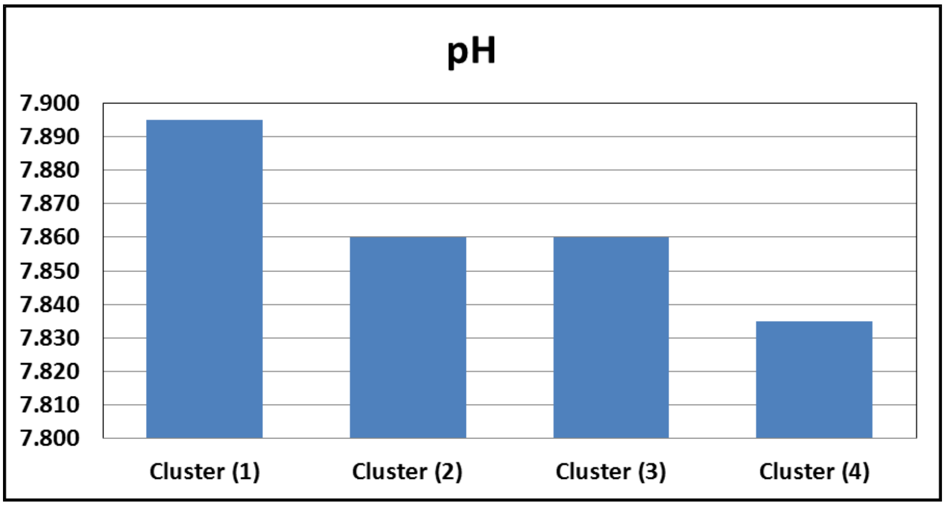


Fig. 7: the change in pH between clusters

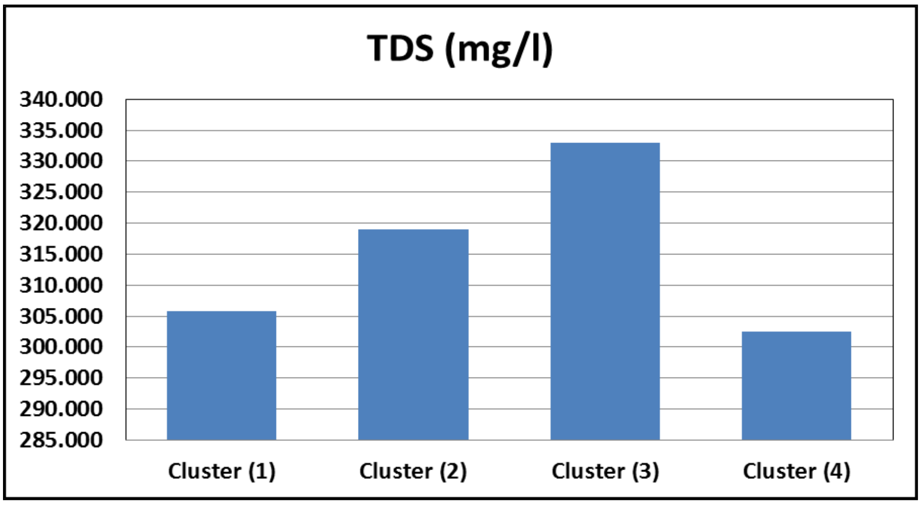


Fig. 8: the change in TDS (mg/l) between clusters

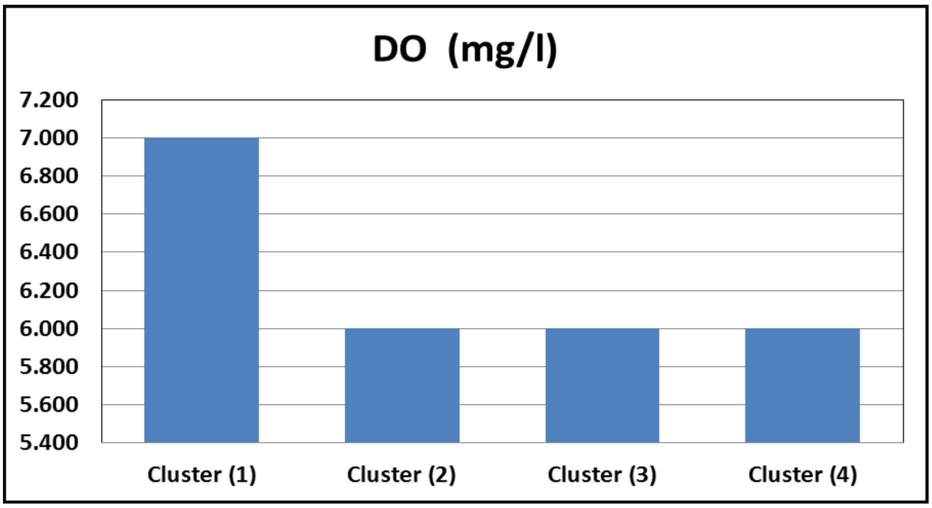


Fig. 9: the change in DO (mg/l) between clusters

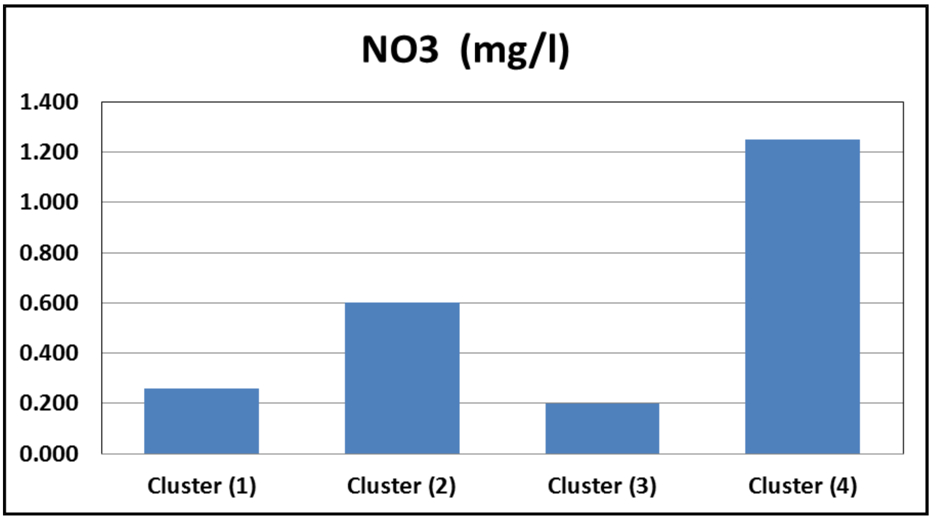


Fig. 10: the change in NO_3 (mg/l) between clusters

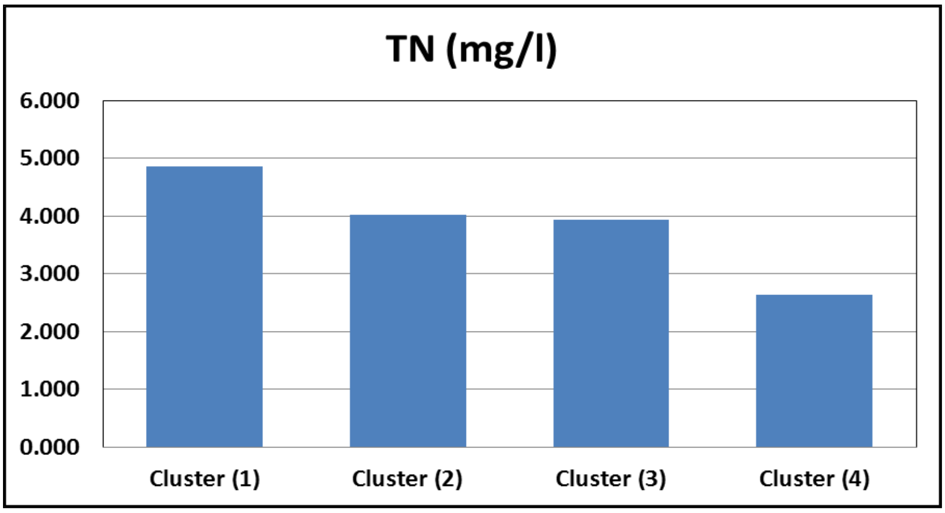


Fig. 11: the change in TN (mg/l) between clusters

Table 4: Water quality standards for irrigation canals in Egypt

Parameters in	Water quality standard Law 48, 1982 and its amendment 2013 (Article 49)
BOD (mg/l)	< 6
COD (mg/l)	< 10
TSS (mg/l)	< 10
NO3 (mg/l)	< 2
NH4 (mg/l)	< 0.5
TN (mg/l)	< 3.5
TDS (mg/l)	< 500
DO (mg/l)	> 6
pH	> 6 - < 8.5

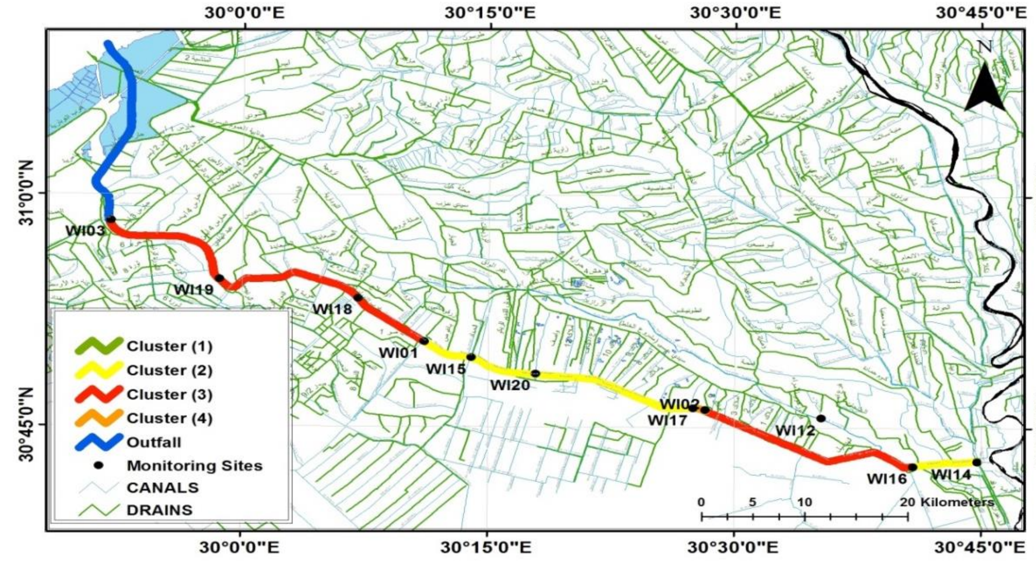


Fig. 12: Water quality clustering of El-Nubaria canal

According to their level of pollution, water quality of El-Nubaria canal was statistically classified into 4 clusters. Each of them had a major dependency on specific water quality indicators which have a considerable physical effect on the canal water quality. Generally, as illustrated in table (1) and figure (12), the water qualities of the monitoring sites (WI12; WI02; WI19, WI18; WI03) follow cluster (clusters 3). On the other side, the water qualities of the sites (WI16; WI20; WI01; WI15) follow cluster (clusters 2). The water quality of site (WI14) described only with cluster (1) where the intake of the El-Nubaria canal from El-Rayah El-Beheiry. The water quality of site (WI17) described only with cluster (4) before the intake of El-Bustan canal.

Figure (13) presents the clustering results in Dendrogram chart obtained for the monitoring sites in the El-Nubaria canal based upon data collected during the period from August 2007 to July 2016. The results showed that the monitoring sites in the El-Nubaria canal can be classified into four clusters.

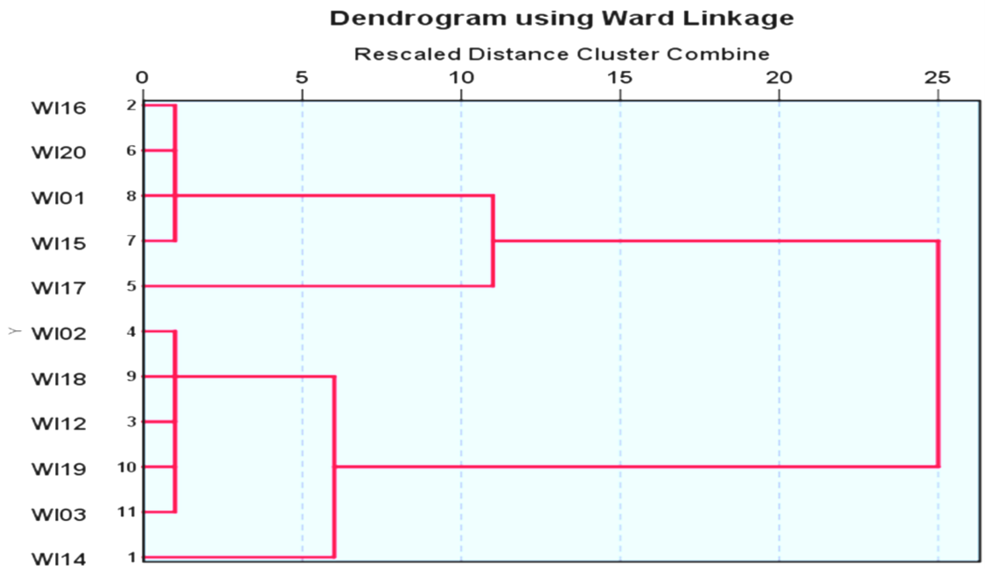


Fig. 13: Dendrogram obtained for the monitoring locations in El-Nubaria canal system.

The highest median values of WQ parameters BOD, COD, NH_4 and NO_3 was in cluster 4 as presented in (table 3) and figures (3, 4, 6, 10 and 12) in this reach, it can be concluded the effect of El-Bustan drainage P.S. effluent to El-Nubaria.

5. Conclusions

The GIS is a very comprehensive tool may provide the decision makers with a simple and reliable overview of the status of surface water quality. Therefore, researchers could rely on this tool in the design of framework for cost-effective pollution control measures. As a result, GIS extensively enable the decision support systems designed to manage the national water quality.

This research work may enable the decision makers to control the pollution of El-Nubaria canal where the clustering process gives an effective overview about the sites where intensified monitoring activities are required. Consequently, the results obtained make a good contribution in the assessment and redesign of the water quality monitoring strategy.

Generally, El-Nubaria canal systems can be classified statistically into the following clusters:

- Cluster 1: low biological pollution, low salt constituents, and moderate nutrients.
- Cluster 2: moderate biological pollution, moderate salt and high nutrients.
- Cluster 3: low biological pollution, high salt constituents, moderate nutrients and high iron concentrations.
- Cluster 4: high biological pollution, low salt constituents and high nutrients.

In accordance with Egyptian quality standards (Law 48 for 1982 and its amendment 2013 - article 49), the median values for WQ parameters (BOD, COD, DO, TDS, NH4, NO3, TN, TSS and pH) at for the four clusters at the selected monitoring sites through the stream of El- Nubaria canal were in the permissible law range (MAB, 1983).

References

Anand Kumar, 2004. Water quality mapping using GIS. Development Alternatives Group Newsletter 14 (3). http://www.devalt.org/pub.htm Kishangarh Vasant Kunj, New Delhi.

Camp, J.S., LeBoeuf, E.J., Abkowitz, M.D., 2010. Application of an enhanced spill management information system to inland waterways. J. Hazard. Mater. 175, 583–592.

Dabberdt, W.F., Frederick, G.L., Hardesty, R.M., Lee, W.C., Underwood, K., 2004. Advances in meteorological instrumentation for air quality and emergency response. Meteorol. Atmos. Phys. 87, 57–88.

Dobbie, M., Harch, B., Packett, R., 2003. Statistical Analysis of The Water Quality of the Fitzroy River Estuary (FE3): Recommendations for Improving Current Monitoring Practices. Technical Report 10. Cooperative Research Center for Coastal Zone Estuary andWaterway Management. Available from:.http://www.coastal.crc.org.au/pdf/TechnicalReports/10-fitzroy_water_qual.pdf.

Donia N. S. and Farag H. A. 2010, "A wast load model analysis for El-Noubariy canal drinking water abstraction, Egypt", Fourteenth International Water Technology Conference, IWTC 14, Cairo, Egypt., p-675

El-Gammal H. and Ali H. 2008. Commissioning of Abandoned Drainage Water Reuse Systems in Egypt: A Case Study of Upgrading the Umoum Project, Nile Delta, Proc.

3rd Int. Conf. on Water Resources and Arid Environments and the 1st Arab Water Forum.
Güler, C., Thyne, G.D., McCray, J.E., Turner, K.A., 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeology Journal 10, 455e474.
MAB, 1983. Law 48/1982 Regarding The Protection of the River Nile and Waterways from Pollution. MAB National Committee, Egypt.
NWRP-2005; “National Water Resources Plan, Facing the Challenge”, Ministry of Water Resources and Irrigation Integrated Water Resources Management Plan for 2017.
Odom K.R., 2003. Assessment and redesign of the synoptic water quality monitoring network in the Great Smoky Mountains National Park. The Dissertation Presented for the Doctor of Philosophy Degree, The University of Tennessee, Knoxville.
Ryberg, K.R., 2006. Cluster Analysis of Water-Quality Data for Lake Sakakawea, Audubon Lake, and McClusky Canal, 1990. 2003-Scientific Investigations Report 2006-5202.
Santos-Roman, M.D., Warner, S.G., Scatena, F., 2003. Multivariate analysis of water quality and physical characteristics of selected watershed in Puerto Rico. Journal of the American Water Resources Association, 829e839.
Samuels, W.B., Amstutz, D.E., Bahadur, R., Pickus, J.M., 2006. RiverSpill: a national application for drinking water protection. J. Hydraulic Eng. 132, 393–403.
Theodoridis, S., Koutroumbas, K., 1999. Pattern Recognition. Academic Press, New York.
Veerabhadram, K., 2003. Mapping ofWater Quality Index (WQI) Using Geographical Information System (GIS) as Decision Supporting System Tool. Department of Environmental Studies, College of Engineering GITAM, Visakhapatnam-530045, India. Available from: http://www.gisdevelopment.net/application/environment/water
Warda, K.T., Johnson, G.R., 2007. Geospatial methods provide timely and comprehensive urban forest information. Urban Forest. Urban Greening 6, 15–22.
World Bank, 2005. “Country Environmental Analysis (1992e2002)” Arab Republic of Egypt, Water and Environment Department, The Middle East and North Africa Region. World Bank.
Zeng, X., Rasmussen, T.C., 2005. Multivariate statistical characterization of water quality in Lake Lanier, Georgia, USA. Journal of Environmental Quality 34,1980e1991.