Egypt J. Aquat. Biol. & Fish., Vol. 15, No. 3: 151-163 (2011) ISSN 1110 – 6131

Identification of pollution sources and homogenous regions in lake Nasser water: A multivariate statistical analysis

Mohsen M. Yousry
Water Research Center, Nile Research Institute, El-Qanater El-Khairiya, Egypt

ABSTRACT

Surface water resources have played an important function throughout the history of human civilization. Lake Nasser water is a major source used for drinking, irrigation and domestic purposes in Egypt.

The purpose of the present study is to present a strategy that identifies the pollution sources and homogenous regions along Lake Nasser. The data collected from 14 sites through four years monitoring program were standardized and subjected to principal components analysis (PCA)/Factor analysis (FA) extraction to define the parameters responsible for the main variability in water quality variance for Nasser Lake water. The data also, subjected to Cluster Analysis (CA) to define the homogeneity region along the Nasser Lake. The PCA produced four significant main components explain more than 86% of the variance, that represent 45.528%, 17.527%, 13.893% and 9.539%, respectively of the total variance of water quality in Nasser Lake. Factor analysis shows that Nasser Lake water quality data consists of four components the geological effects, phytoplankton and DO budget, silicate weathering and non-point sources of pollution. Cluster analysis suggests that Lake monitoring stations can be consolidated, and the Lake classified as three main zones. Cluster Analysis is useful in design a future sampling monitoring strategy in an optimal manner. This study showed the importance of the multivariate statistical techniques for getting better information about the water quality and design of monitoring network for effective management of water resources.

Keywords: Principal component analysis, Factor analysis, Cluster analysis, Lake Nasser.

INTRODUCTION

Water is essential to sustain life. At a very basic level, a minimum amount of water is required for consumption on a daily basis for survival and therefore access to some form of water is essential for life. The water demand is continuously increasing mainly due to population growth and raising needs in agriculture, industrial uses and domestic services. Integrated management has a strong impact on long-term protection and sustainability (Singh *et al.,* 2005). Water quality performs important role in health of human, animals and plants. Improving access to safe drinking water can result in tangible benefits to health (WHO 2004). The quality of surface water within a region is governed by both natural processes and anthropogenic effects (Nouri *et al.,* 2008).

The High Dam Lake is one of the largest man-made lakes in Africa. It is bounded by latitude 24°N in Egypt and 21°S in Sudan. It includes Lake Nasser which extends for about 330 Km in Egypt and Lake Nubia 160 Km in Sudan. Nasser Lake water is a major source used for drinking, irrigation and domestic purposes in Egypt.

The application of different multivariate statistical techniques, such as principal component analysis (PCA), factor analysis (FA) and cluster analysis (CA) helps in the interpretation of complex data matrices to better understand the water quality and ecological status of the studied systems. This better understanding allows the identification of possible factors sources that influence water systems and offers a valuable tool for reliable management of water resources, as well as rapid solution to pollution problems (Lee *et al.,* 2001; Reghunath *et al.,* 2002; Wunderlin *et al.,* 2001). Also in recent years, the PCA and FA methods have been exerted for a variety of environmental applications, containing evaluation of ground water monitoring wells and hydrographs, examination of spatial and temporal patterns of surface water quality, identification of chemical species related to hydrological conditions and assessment of environmental quality indicators (Ouyang *et al.,* 2006; Voutsa *et al.,* 2001). Multivariate statistical methods were proved as one of useful tools to extract the meaningful information from data set, Simeonov *et al.* (2003) and Astel *et al.* (2006) applied CA to declinate the monitoring sites.

In the present study, large data sets, obtained during a four year monitoring program, were subjected to PCA, FA and CA. Its main objectives are as follows:

1) To identify the most important variables describing the quality of water at Nasser Lake and define the pollution sources, and
2) To identify the homogenous regions and find out the similarity groups between the sampling stations.

MATERIALS AND METHODS

Study area and sampling

In total, 14 sites were located along the Lake, from Dabarosa at the Sudanese-Egyptian border to El Daka at the southern end of the Lake (Figure 1). Water samples were collected for 4 years along Nasser Lake (2004, 2006, 2007 and 2008).

Monitored parameters and analytical methods

The data for 14 water quality monitoring sites, consisting of 21 water quality parameters monitored for four years (2004, 2006, 2007 and 2008) during winter, were obtained. The sampling, preservation, transportation, and analysis were performed according to standard methods (APHA, 1989). Table (1) summarizes the water quality parameters, units and methods of analysis.

Data Processing

Most multivariate statistical methods require variables to confirm to the normal distribution, thus, the normality of the distribution of each variable was checked by applying the Kolmogorov-Smirnov test and by analyzing kurtosis and skewness, it could be found that the original data were far from normal distribution. Since most of the values of kurtosis or skewness were greater or less than zero, the original data were transformed in the form of natural logarithm (Ln value). After transformation, most variables were normally distributed (at 5% level of significance). TDS was chosen as the variable responsible for salinity along with TSS which is some times attributed to the organic load.

Statistical Methods

All mathematical and statistical calculations were implemented using Statistica 8, Minitab 12 and SPSS 8 for windows.

Principal Component Analysis and Factor Aanalysis

PCA is designed to transmute the original variables into new, uncorrelated variables (axes), called the principle components, which are linear combinations of the original variables (Shrestha and Kazama, 2007). PCA provides information on the most meaningful parameters which describe the whole data set interpretation, data reduction and to summarize the statistical correlation among constituent in the water with minimum loss of original information (Helena *et al.,* 2000). FA follows PCA which is a linear combination of observable water quality variables, whereas FA can include unobservable, hypothetical, latent variables (Helena *et al.,* 2000; Vega *et al.,* 1998). PCA of the normalized variables was executed to extract significant components (PCs) and to further reduce the contribution of variables with minor significance; these PCs were subjected to varimax rotation generating factors (Singh *et al.,* 2004, 2005). The main applications of FA are to reduce the number of variables and to discover structure in the relationships between variables that is to classify variables.

Cluster Analysis

CA is a group of multivariate techniques that divides a large group of cases into smaller groups or cluster of relatively similar cases that are dissimilar to other groups. Hierarchical CA, the most common approach, starts with each case in a separate cluster and joins the clusters together step by step until only one cluster remains (Lattin *et al.,* 2003; McKenna, 2003). The Euclidean distance usually gives the similarity between two samples and a distance can be represented by the difference between analytical values of the samples (Otto, 1998). In this study, hierarchical agglomerative CA was performed on the normalized data set by means of the Ward's method with Euclidean distance as a measure of similarity.

RESULTS AND DISCUSION

Descriptive statistics including the mean, median, standard deviation, minimum, maximum, kurtosis and skewness for Nasser Lake (2004, 2006, 2007, and 2008) are presented in Table (2). It can be seen that, the mean values of each variable was found within the permissible limits of Law 48/1982. Its worth to mention that, the increase of pH values in some sites (9.04) may be related to high temperature, photosynthesis and growth of aquatic plants (El-Wakeel and Wahby 1970). The results of kurtosis and skewness showed that, the original data were far from normal distribution. Since most of the values of kurtosis or skewness were greater or less than zero. Normality tests are used to determine whether a random variable is normally distributed, or not. Many data analysis methods (ANOVA, Regression and Multivariate statistical analysis) depend on the assumption that data were sampled from a normal distribution. The results of the application of Anderson-Darling test (Anderson and Darling 1952), Kolmogrov-Smirnov test with Lilliefors correction (Lilliefors 1967) and Shapiro Wilk test (Shapiro and Wilk 1965) showed that most of the water quality variables along the Nasser Lake do not pass the normality. Consequently, nature logarithmic (Ln) transformation helped normalize the distribution of these values.

Results of principal components analysis and factor analysis:

Principal component analysis/factor analysis was executed on 12 variables for the 14 different sampling stations in four years; this is identified important water quality parameters. The Scree plot (Fig. 2) was used to identify the number of principal components to be retained to comprehend the underlying data structure (Vega *et al.,* 1998). An eigenvalue gives a measure of the significance of the factor: the factors with the highest eigenvalues are the most significant. Eigenvalues of 1 or greater are considered significant (Shrestha and Kazama, 2007). The interpretation of un-rotated principal components is relatively more difficult to explain physically. Therefore, the first four components were rotated using Varimax normalized rotation in order to make interpretation easier. In the present study PCA by the unrotated component matrix and PCA by varimax rotation were both performed. Corresponding, variable loadings (before and after varimax rotation) and explained variance are presented in Tables from 3 to 5 and strong loading values have been highlighted (greater than 0.7). From Table (3): it is observed that, the first two varifactors were dominant and together accounted for 63.06% of the total variance. Whereas, the remaining two varifactors were secondary and accounted for 13.89% and 9.54% of the variance respectively. Table (6) shows the correlation components matrix (component score covariance matrix) of varimax rotated four PC; it note that there are no correlation between components, this is, each component represent a discrete unit from others.

The first varifactor, with an eigenvalue of 5.463, explained 45.53% of the total variance it is clearly dominated by the pH, Ca, TDS, SD (positive loading) and TSS (negative loading) as shown in Table 5 and Figure 3. This varifactor reflect the geological effects of Nasser Lake and the mineral composition of Lake water (natural source). The high positive loading between SD and pH and negative loading with TSS indicate the effect of penetration of sunlight into the water column on the phytoplankton growth. TDS and Ca loading on varifactor 1 can be explained by the dissolution of soils and minerals in sediments such as calcite $CaCO_3$ (mineral weathering) Wayland *et al.,* 2003.

The second varifactor, with an eignvalue of 2.1 and accounting for 17.53% of the total variance was also significant. The variables Chla and DO exhibited high positive loading.

This varifactor seems to represent the biological processes like phytoplankton and DO budget (oxygen production by photosynthesis). Thus, this varifactor seems to indicate the connectivity between phytoplankton or algal biomass (represent by Chla) and the oxygen production (DO). It's worth mentioning that, the variation of dissolved oxygen values can be related to the water temperature and photosynthesis activity (Willoughby, 1976).

The third and the fourth varifactors explain relatively lesser variance, as compared to the first two. They seem represent the nutrients, as they include SiO_2 and NO_3. These two factors reflect silicate weathering and the non-point sources of pollution along Nasser Lake such as re-suspension of bed sediment or agricultural runoff.

The data reduction from FA in this case study was not as great as we had expected. However, the FA/PCA can serve as an important means to identify the pollution sources.

In summary the four extracted components representing four different sources are:

- Geological effects of Lake
- Productivity of Lake
- Silicate weathering
- Non-point source of pollution

Cluster Analysis (CA)

Factor analysis was used to combine water quality parameters into homogenous groups; it is also possible to evaluate whether water quality samples at various location can be combined into homogenous regions so the number of sampling sites can be reduced. Cluster analysis was applied to find out the similarity between the sampling stations (homogenous regions). CA rendered a dendrogram as shown in Figures 4 to 6, grouping all of fourteen sampling sites into three zones. The first zone (cluster1) included sites from El Daka to Madeek Amka corresponds to riverine zone. The second zone (cluster 2) included Second Cataract corresponds to transition zone, while the third zone (cluster 3) represents area from Halfa to Dabrousa corresponds to lacustrine zone. The results of CA reveal that, the three zones have different characteristics in terms of selected parameters and indicate that

each region received pollution from different sources (natural or anthropogenic).
It is evident that the CA is useful tool in affording reliable classification of water quality in the whole region and will make possible to design a future monitoring network, hence, the number of sampling sites and respective cost in the future monitoring plans can be lessen taking into consideration all parameters.

Cluster of variables

The goal of cluster analysis of variables is to detect the hierarchy of interrelations among a set of variables of a data matrix. Figure 7 shows the dendrogram for the 12 studied parameters of all samples data set. Along Nasser Lake different clusters were extracted as follows:

Cluster 1: pH, TDS, Ca, SD and K
Cluster 2: Chla, DO, OP, TP and TSS
Cluster 3: SiO_2
Cluster 4: NO_3

This data analysis gives an idea of how the single water quality parameters should be compared and related to one another. For instance, within all a group of all samples, there was a stronger correlation between the groups of parameters. The results of cluster analysis of variables confirmed the classification obtained by PCA.

CONCLUSION

The multivariate statistical techniques, namely, cluster analysis and factor analysis are important analytical techniques for the processing of water quality parameters and powerful tools for identification of sources of pollution as well as for identification of homogenous regions. PCA and FA assisted to extract and recognize the factors responsible for water quality variations. PCA and FA show that water quality data consists of four components (Geological effects of Lake, productivity of Lake, silicate weathering and non-point source of pollution), which can be distinguished using pH, Ca, TDS, SD, TSS, Chla, DO, SiO_2 and NO_3 respectively. The first two factors were dominant and together explaining about 63.06% of the data contains the key variables of Lake water. Cluster analysis grouped 14 sampling sites into three clusters (riverine, transition and lacustrine) of similar water quality features. Cluster Analysis is useful in design a future sampling monitoring strategy in an optimal manner hence, the number of sampling sites and respective cost can be diminished in future monitoring plans. The results of cluster analysis of variables confirmed the classification obtained by PCA.

REFERENCES

Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. Annals of Mathematical Statistics 23: 193–212. APHA (1989). Standards methods for the examination of water and wastewater. 17th Ed., Washington: American Public Health Association.

Astel, A.; Biziuk, M.; Przjazny, A. and Namiesnik, J. (2006). Chemometrics in monitoring spatial and temporal variations in drinking water quality. Water Res., 8:1706-1716.

El-Wakeel, S. K. and S. D. Wahby (1970). Hydrography and chemistry of Lake Mansalah, Egypt, Arch. Hydrobiol., (67): 173 – 200.

Helena, B.; R. Pardo; M. Vega; E. Barrado; J. M. Ferna´ndez and L. Ferna´ndez (2000). Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water Res., 34 (3): 807-816.

Lattin, J.; D. Carroll and P. Green (2003). Analyzing multivariate data. New York: Duxbury.

Lee, J. Y.; J. Y. Cheon; K. K. Lee; S. Y. Lee and M. H. Lee (2001). Statistical evaluation of geochemical parameter distribution in a ground water system contaminated with petroleum hydrocarbons. J. Environ. Qual., 30: 1548-1563.

Lilliefors, H. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown, J. of the American Statistical Association, (62): 399-402.

McKenna, J. (2003). An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. Environmental Modelling and Software, 18: 205-220.

Nouri, J.; A. R. Karbassi and S. Mirkia (2008). Environmental management of coastal regions in the Caspian Sea. Int. J. Environ. Sci. Tech., 5 (1): 43-52.

Otto, M. (1998). Multivariate methods. In: Kellner, R., Mermet, J. M., Otto, M., Widmer, H. M. (Eds.), Analytical Chemistry. WileyeVCH, Weinheim, Germany, 916.

Ouyang, Y.; P. Nkedi-Kizza; Q. T. Wu; D. Shinde and C. H. Huang (2006). Assessment of seasonal variations in surface water quality. Water Res., 40 (20): 3800-3810.

Reghunath, R.; T. R. S. Murthy and B. R. Raghavan (2002). The utility of multivariate statistical techniques in hydrogeochemical studies: An example from Karnataka, India. Water Res., 36 (10): 2437-2442.

Shapiro, S. S. and M. B. Wilk (1965). An analysis of variance test for normality (complete samples), *Biometrika*, 1965; 52: 591-611.

Shrestha, S. and F. Kazama (2007). Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environ. Model. Software, 22 (4): 464-475.

Simeonov, V.; J. A. Stratis; C. Samara; G. Zachariadis; D. Voutsa; A. Anthemidis; M. Sofoniou and T. Kouimtzis (2003). Assessment of the surface water quality in Northern Greece. Water Res., 37 (17): 4119-4124.

Singh, K. P.; A. Malik; D. Mohan and S. Sinha (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India): A case study. Water Res., 38 (18): 3980-3992.

Singh, K. P.; A. Malik and S. Sinha (2005). Water quality assessment and apportionment of pollution sources of Gomti River (India) using multivariate statistical techniques: A case study. Analytica. Chimica. Acta., 538 (1-2): 355-374.

Vega, M.; R. Pardo; E. Barrado and L. Deban (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Res., 32 (12): 3581-3592.

Voutsa, D.; E. Manoli; C. Samara; M. Sofoniou and I. Stratis (2001). A study of surface water quality in Macedonia, Greece: Speciation of nitrogen and phosphorus. Water Air Soil Pollut., 129 (1-4): 13-32.

Wayland, K.; D. Long; D. Hyndman; B. Pijanowski; S. Woodhams and K. Haak (2003). Identify relationships between base flow geochemistry and land use with synoptic sampling and R-Mode factor analysis. J Environ Qual., 32: 18—190.

Willoughby, L. G. (1976). Fresh water biology, fresh water biology association, Ambleside, Cumbria.

WHO (2004). Guidelines for drinking water quality, Vol 1. World Health Organization, Geneva.

Wunderlin, D. A.; M. P. Diaz; M. V. Ame; S. F. Pesce; A. C. Hued and M. A. Bistoni (2001). Pattern recognition techniques for the evaluation of patial and temporal variations in water quality. A case study: Suquia river basin (Cordoba, Argentina). Water Res., 35 (12): 2881-2894.

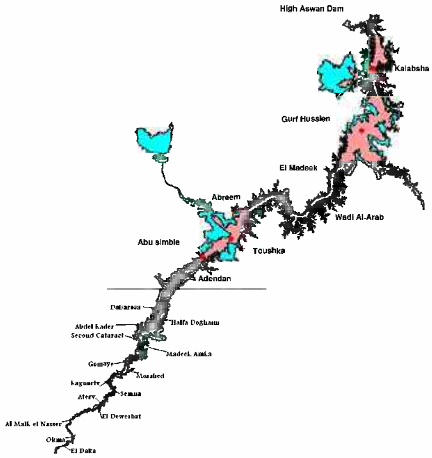

Fig. 1: Sampling sites along Nasser Lake

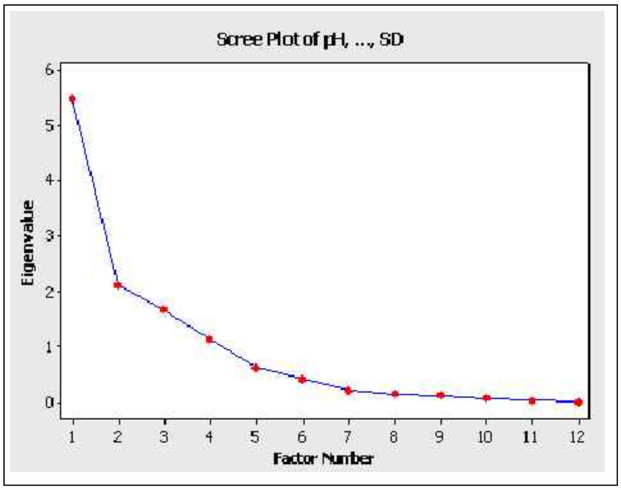

Fig. 2: Scree plot graph for components with its eigenvalues

Table 1: Water quality parameters, units and methods of analysis

parameters	Units	Analytical Methods
Water Temperature (WT)	°C	Thermometer
Hydrogen ion concentration (pH)	unit	pH meter (WTW, pH 197)
Dissolve Oxygen (DO)	mgl^{-1}	DO - meter (WTW, Oxi 197)
Electrical Conductivity (EC)	μmohs/cm	Conductinity meter (WTW, LF 197)
Turbidity	NTU	Turbidimeter (HACH 2100)
Total Suspended Solids (TSS)	mgl^{-1}	Gravimetric
Total Dissolved Solids (TDS)	mgl^{-1}	Gravimetric
Nitrate (NO_3)	mgl^{-1}	Ion selective electrode (Orion ionanalyzer EA 940)
Phosphate	mgl^{-1}	Molybdate method (spectrophotometric DR 2400 HACH)
Silicate (SiO_2)	mgl^{-1}	Spectrophotometric (DR 2400 HACH)
Transparency	cm	Scchi disc
Bicarbonate (HCO_3)	mgl^{-1}	Titrimetric
Calcium, Magnesium and Total Hardness	mgl^{-1}	Titrimetric (by EDTA)
Sodium and Potassium	mgl^{-1}	Flame Photometer (Jenway)
Chloride (Cl)	mgl^{-1}	Argentometric method
Sulfate (SO_4)	mgl^{-1}	Turbidimeteric method
Chlorophyll a	μgl^{-1}	Spectrophotometric method

Table 2: Descriptive statistics of water quality for Nasser Lake.

Parameters	Mean	Median	Minimum	Maximum	Range	Variance	Std.Dev.	Skewness	Kurtosis
pH (unit)	8.1745	8.1900	7.7700	9.0400	1.2700	0.096	0.30954	0.820044	0.48842
DO (mg/l)	8.8168	8.7100	7.5400	9.9200	2.3800	0.188	0.43314	0.595573	1.39678
WT (°C)	18.5161	17.6000	15.5000	24.1000	8.6000	4.107	2.02653	0.667148	-0.40648
E.C (μmohs/cm)	200.8321	186.5000	172.0000	247.0000	75.0000	765.185	27.66198	0.686521	-1.27171
SD (cm)	47.1429	35.0000	15.0000	180.0000	165.0000	1121.688	33.49162	2.248218	5.09266
Turbidity (ntu)	64.9213	68.0000	6.1000	150.0000	143.9000	1758.926	41.93956	0.208854	-1.32332
NO_3 (mg/l)	0.2938	0.2300	0.1800	0.6600	0.4800	0.015	0.12349	1.499604	1.25695
OP (mg/l)	0.1254	0.1300	0.0500	0.1800	0.1300	0.001	0.03063	-0.593136	0.20654
TP (mg/l)	0.3054	0.3100	0.1500	0.5400	0.3900	0.012	0.10812	0.474905	-0.47650
SIO2 (mg/l)	11.2893	10.5500	6.1000	17.7000	11.6000	11.082	3.32903	0.186405	-1.05442
HCO_3 (mg/l)	91.4464	90.0000	78.0000	108.0000	30.0000	66.470	8.15290	0.620496	-0.39094
SO_4 (mg/l)	5.1618	4.4400	2.0000	13.0000	11.0000	6.960	2.63822	1.251219	1.45359
Cl (mg/l)	4.9866	4.4300	3.5000	12.0000	8.5000	2.913	1.70664	2.683320	7.62937
TDS (mg/l)	128.9554	119.8000	111.8000	158.0000	46.2000	300.097	17.32331	0.704410	-1.26781
Ca (mg/l)	18.2314	17.8400	13.6000	25.6500	12.0500	7.838	2.79958	0.526517	-0.10388
Mg (mg/l)	6.9452	6.8100	4.3700	10.0800	5.7100	3.225	1.79583	0.318696	-1.15500
T.H (mg/l)	74.4464	70.0000	54.0000	96.0000	42.0000	163.088	12.77059	0.139172	-1.33380
Na (mg/l)	13.9893	14.5000	10.0000	18.0000	8.0000	5.070	2.25160	-0.203937	-1.16337
K (mg/l)	3.8429	3.8000	3.0000	6.0000	3.0000	0.369	0.60749	1.209262	2.22114
TSS (mg/l)	69.0179	74.0000	8.0000	140.0000	132.0000	1772.018	42.09534	0.008874	-1.55784
Chla (μg/l)	16.2250	14.7000	3.4000	35.8000	32.4000	76.043	8.72024	0.530314	-0.76410

Table 3: Explains initial components

Components	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	5.463	45.528	45.528
2	2.103	17.527	63.055
3	1.667	13.893	76.948
4	1.145	9.539	86.488
5	.623	5.193	91.680
6	.415	3.457	95.137
7	.198	1.653	96.790
8	.145	1.207	97.996
9	.112	.936	98.933
10	7.561E-02	.630	99.563
11	3.962E-02	.330	99.893
12	1.284E-02	.107	100.000

Table 4: Unrotated component matrix

Variable	F1	F2	F3	F4
pH	0.847	0.327	0.319	0.017
NO_3	0.177	-0.081	-0.872	0.150
TP	-0.837	0.340	-0.103	-0.211
SiO_2	0.018	-0.722	-0.311	-0.552
Ca	0.817	0.030	-0.404	-0.246
K	0.590	-0.175	0.507	-0.416
Chla	-0.178	0.903	0.061	-0.111
TDS	0.912	-0.174	0.255	-0.091
OP	-0.698	-0.047	0.214	-0.507
TSS	-0.921	-0.303	0.096	0.086
DO	-0.291	0.526	-0.279	-0.486
SD	0.846	0.323	-0.274	-0.135

Table 5: Varimax rotated component matrix

Variable	F1	F2	F3	F4
pH	0.835	-0.155	0.317	0.323
NO_3	0.229	-0.002	-0.219	-0.849
TP	-0.648	0.661	0.096	-0.064
SiO_2	-0.021	-0.055	-0.958	-0.054
Ca	0.874	-0.003	-0.272	-0.233
K	0.524	-0.168	-0.257	0.663
Chla	0.073	0.708	0.594	0.070
TDS	0.799	-0.422	-0.128	0.321
OP	-0.600	0.475	-0.279	0.359
TSS	-0.971	0.033	-0.112	0.013
DO	-0.014	0.817	-0.011	-0.093
SD	0.938	0.078	0.045	-0.162

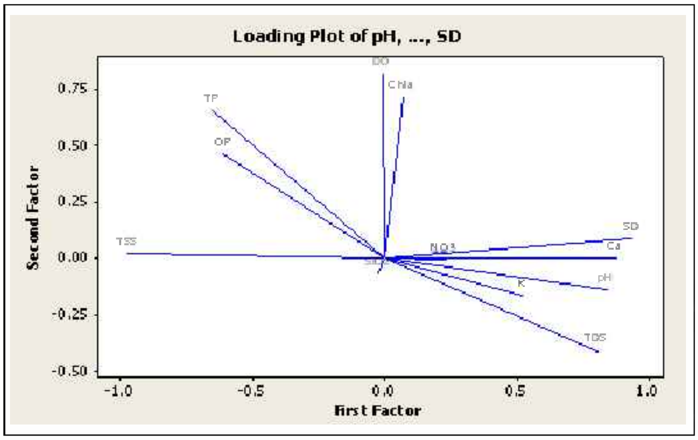

Fig. 3: Varimax normalized rotated factors along Nasser Lake

Table 6: Components score covariance matrix

Component	1	2	3	4
1	1.000	.000	.000	.000
2	.000	1.000	.000	.000
3	.000	.000	1.000	.000
4	.000	.000	.000	1.000

Table 7: The most significant parameters that contribute to water quality variations

Factors	Parameters with strong positive loadings	Parameters with strong negative loadings
F1	pH, Ca, TDS, SD	TSS
F2	Chla, DO	
F3		SiO2
F4		NO3

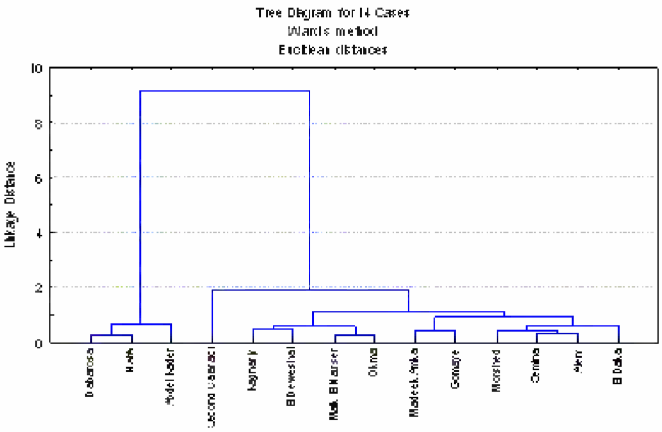

Fig. 4: Dendrogram depicted clustering of monitoring sites for Chla

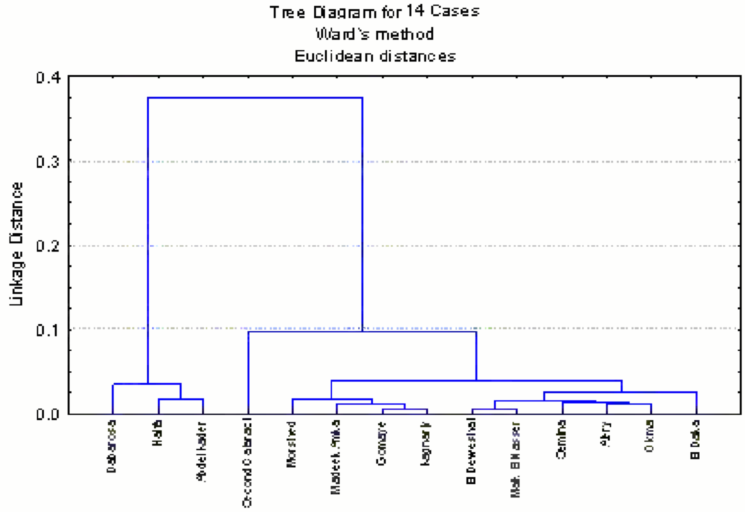

Fig. 5: Dendrogram depicted clustering of monitoring sites for pH

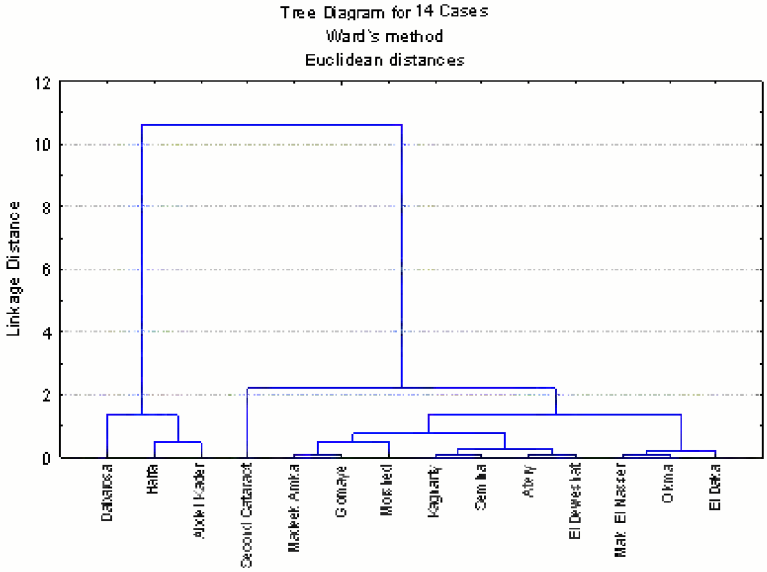

Fig. 6: Dendrogram depicted clustering of monitoring sites for TSS.

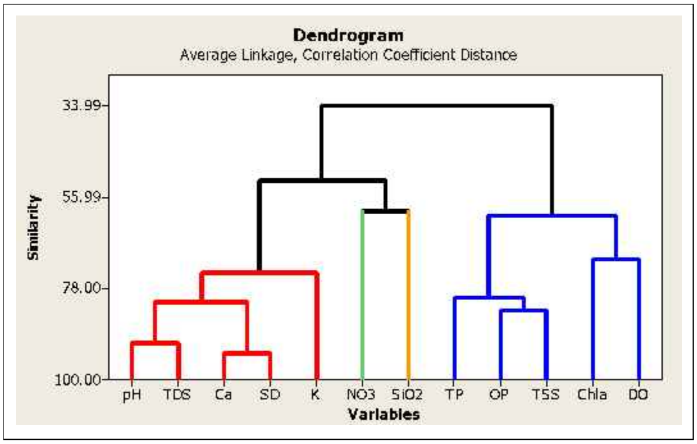

Fig. 7: Dendrogram depicted clustering of monitoring variables

التعرف علي مصادر التلوث والمناطق المتجانسة في مياه بحيرة ناصر: التحليل الإحصائي متعدد التغير

محسن محمود يسري
باحث بالمركز القومي لبحوث المياه - معهد بحوث النيل - القناطر الخيرية- مصر.

مصادر المياه السطحية تلعب دور مهم في تاريخ الشعوب المتحضرة. تستخدم مياه بحيرة ناصر كمصدر رئيسي مهم للشرب، الري والأغراض المنزلية في مصر.

الغرض الرئيسي من هذا البحث هو تقديم إستراتيجية للتعرف علي مصادر التلوث والمناطق المتجانسة علي طول بحيرة ناصر. النتائج التي تم جمعها من ١٤ موقع خلال ٤ سنوات لبرنامج الرصد تم معايرتها وتعرضها لتحليل المركبات الأساسية/التحليل العاملي للتعرف علي المتغيرات المسؤولة عن التغير الرئيسي في نوعية مياه بحيرة ناصر. تعرضت النتائج أيضا لتحليل المجموعات لتحديد المناطق المتجانسة علي طول بحيرة ناصر.

أنتج تحليل المركبات الرئيسية أربعة مكونات رئيسية ذات معني تفسر أكثر من ٨٦% من التغاير الكلي في نوعية مياه بحيرة ناصر والتي تمثل ٤٥.٥٢٨%، ١٧.٥٢٧%، ١٣.٨٩٣%، ٩.٥٣٩% علي الترتيب.

أظهرت نتائج التحليل العاملي أن نتائج نوعية مياه بحيرة ناصر تتكون من أربعة مكونات هي التأثير الجيولوجي، الهائمات النباتية والأكسجين الذائب، السليكات المحللة بالعوامل الجوية ومصادر التلوث الغير محددة. أقترح تحليل المجموعات أن محطات الرصد بالبحيرة يمكن إختزالها وتقسيمها لثلاث مناطق رئيسية. يعتبر تحليل المجموعات مهم فى التصميم المستقبلى لأستراتيجية شبكة الرصد بطريقة مثالية.

هذه الدراسة أظهرت أهمية تقنية التحليل الإحصائي متعدد التغير للحصول علي معلومات أفضل عن نوعية المياه وتصميم شبكات الرصد للإدارة الفعالة للموارد المائيه.

الكلمات الدالة: تحليل المركبات الرئيسية، التحليل العاملي، تحليل المجموعات، بحيرة ناصر.