Journal of Applied Science, Innovation & Technology (JASIT)

Journal homepage: https://prakritimitrango.com/documents-reports-and-publications/journal/

Research Article

Multivariate statistical techniques for the evaluation of groundwater quality of Mathura (India)

Vinod Kumar Kushwah[1*], Neha Singh[2], Kunwar Raghvendra Singh[1*]
[1]Civil Engineering Department, GLA University Mathura-281406, India
[2] Zoology Department, A. N. College, Patna-800013, India
ARTICLE INFOR: *Received: 04 June 2023; Revised: 17 June 2023; Accepted: 17 June 2023*
CORRESPONDING AUTHORS: E-mail: *vinnikushwah@gmail.com (V.K. Kushwah); kunwar.raghvendra@gla.ac.in (K.R. Singh)*

Abstract
The preservation and protection of ground water quality has received worldwide attention due to its adulteration by natural or human activities. In addition to being harmful to aquatic ecosystems, untreated municipal and industrial waste discharges, agricultural runoffs, leachates, and other activities have introduced a variety of trace elements. As a result, numerous programs for monitoring and evaluating water quality have been emerged worldwide to provide accurate information on activities that contribute to the resource degradation. However, utilizing the vast amounts of data, generated by monitoring programs to obtain valuable information about water quality has presented a global challenge. Multivariate statistical techniques (MSTs) can be utilized for treating huge and complex water quality. In the current study, cluster analysis (CA) and principal component analysis (PCA) were used for ground water quality assessment. CA was used to assess the similarities between the sites according to the similarities in their characteristics. Analysis showed that three principal components (PCs) had eigenvalues greater than one. These three PCs explain approximately 92 percent of the variance in the data. Present work will be very useful for policy makers in making policies for control of water pollution and preservation of water resources.
Keywords: Groundwater; Mathura; Cluster analysis; Principal Component Analysis

1. Introduction

Groundwater located in aquifers below the ground is one of the country's most important natural resources. About 30 percent of the world's readily available fresh water is groundwater. Groundwater is often cheaper, more convenient and less polluted than surface water (Carrard et al., 2019). Therefore, it is usually used in public water supply. India is the enormous consumer of groundwater in the world. Today, groundwater is the only resource of water for most Indians and provides most of the water for agriculture and domestic use. While groundwater fuelled the Green Revolution that made India a food-secure nation, the widespread exploitation of this precious resource has led to its alarming decline (Singh, 2000). As climate change makes rainfall patterns even more unpredictable, groundwater will become even more important. Already almost 63 % of the regions of the India are at risk of falling underground levels. In many cases, this water is contaminated. Most groundwater is free of pathogenic organisms and does not require treatment for domestic or industrial use. Short-term droughts do not significantly affect groundwater storage and are available in many areas without reliable surface water storage. Since, last few decades groundwater resources are threatened by pollution from agrochemicals, landfills lechate and septic tanks, and other sources (Stefanakis et al., 2017). Such contamination can make groundwater unusable and is costly and not easy to clean up.

Water quality is one of the greatest challenges of 21st century societies, adversely affecting human health, food production, ecosystem functions and economic growth (Kanwar, 2009). Poor national cultivation, irrigation practices, waste practices, and increasing chemical concentrations are all factors affecting groundwater quality, leading to contamination as a chemical hazard (Stefanakis et al. 2017). Deterioration of water quality directly leads to environmental, social and economic problems. Water pollution threatens our health and the environment. The main task of water treatment centres is to collect and treat water, identify potential contaminants, protect human health, and provide and identify resources needed to improve water quality. The type and extent of water treatment depend on the level of pollution or contamination which can be determined by the regular water quality monitoring. The results of monitoring helped to identify spatial-temporal variation in water quality. These regular monitoring is an important aspect of identifying existing or potential future problems (Singh et al., 2004). Information gathered from water quality monitoring is very useful in planning pollution prevention and control strategies. Today, governments, communities and businesses must meet multiple water quality goals. One of the major problems associated with monitoring is the managing of large and vast data sets (Dixon and Chiswell, 1996). Application of various MSTs offers a preferable assessment of data (Singh et al., 2004; Fan et al., 2010; Singh et al., 2017; Bouguerne et al., 2017; Singh et al., 2020). In the current research, cluster analysis (CA) & principal component analysis (PCA) have been used to determine the Mathura groundwater quality and determine the current water quality condition. As a result, strategies to lessen the increased strain on these groundwater resources will be easier to develop for policymakers.

2. Materials and methods

Water samples have been collected from various locations in Mathura. Mathura is a city in the northern Indian state of Uttar

Pradesh (as shown in Fig. 1). It lies approximately north of Agra and southeast of Delhi. It is the administrative centre of Mathura district in the state of Uttar Pradesh. The total area of Mathura is approximately 3340 sq km and its elevation is 174 m. Population density is nearly 763 per sq. km. It is part of the northern plains of India and on the west bank of the Yamuna River. The climate of Mathura is typical of North Indian nature. In summer, the temperature is reached up to 45°C. During the monsoon, Mathura receives a generous rainfall (20 mm on average from June to August). Winters (from December to February), on the other hand, remain beautiful and pleasant, although it can feel cold at times. During this period, the lowest temperature often falls below 4-5°C, while the maximum mostly remains at 12-14°C.

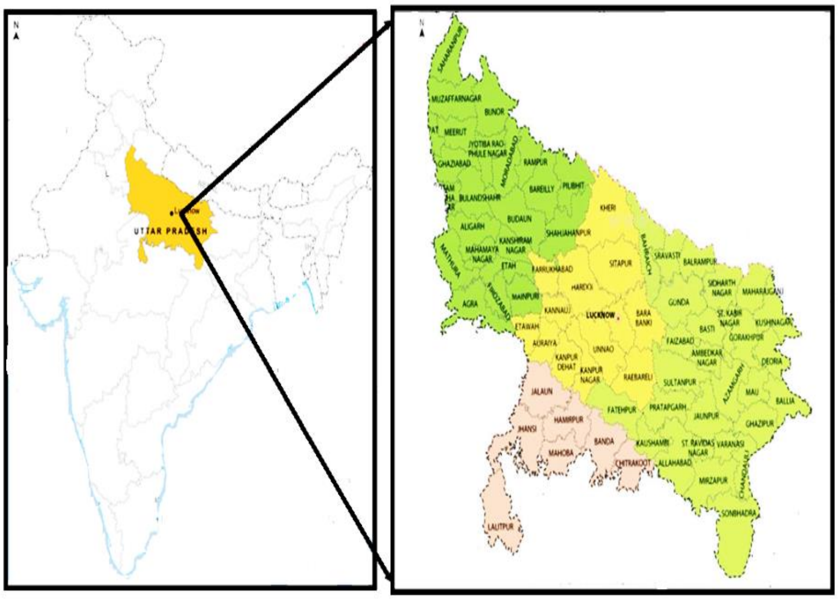

Fig. 1 Study Area

Samples have been collected in clean one-liter bottles that were free of dirt and dust. Before collection, bottles had been washed with the water to be collected. Protected from accidental contamination, all bottles arrived at the sampling site clean and intact. Only identical bottles have been used for the observed parameters. Plastic gloves were used to keep the hands from getting into the bloodstream. The date and site number were printed on labels for the sample bottles.

CA is a statistical approach that describes the grouping of data with similar characteristics without specific results. A typical cluster analysis results in placing data points into groups based on similarity - items within a group are similar to each other, while different groups are different (Singh et al., 2020). Many methods are available for grouping (agglomerative, distributive, non-hierarchical, etc.). CA is unique as a statistical method because it is performed without an assumed basis of principle or fact. PCA is a popular technique for studying large datasets, enhancing data understanding while maintaining the most information. PCA is a tool for mitigating the number of dimensions in huge data sets.

3. Results and discussion

In this study, the groundwater quality of Mathura region has been studied. Water samples has been collected from 7 different sites and analyzed in the laboratory. Descriptive statistics have given in Table 1. It was observed that the turbidity of the samples was well below the permissible limit but the total dissolved solid (TDS) concentration was very high. Maximum TDS was observed as 2800 mg/L and average TDS was 1943 mg/L. Average concentration of chloride, fluoride, iron, nitrate and total hardness (TH) were also above permissible limit (IS 10500:2012). Hence, water is not suitable for drinking purposes.

Table 1 Descriptive Statistics for different water quality parameters

	Turbidity	TDS	pH	Chloride	Fluoride	Iron	Nitrate	TH
Minimum	1.00	800.00	7.00	226.00	1.00	0.10	15.00	520.00
Maximum	2.90	2800.00	8.20	1220.00	1.60	0.80	50.00	3280.00
Mean	2.13	1942.86	7.54	711.29	1.21	0.46	25.71	1080.00
[1]StdError	0.29	254.35	0.15	120.54	0.09	0.09	4.42	369.71
Median	2.50	2000.00	7.50	660.00	1.20	0.40	25.00	800.00
[2]StdDev	0.77	672.95	0.40	318.92	0.25	0.25	11.70	978.16
Kurtosis	-1.51	0.22	0.21	0.40	-1.06	-1.24	3.92	6.64
Skewness	-0.67	-0.51	0.57	0.22	0.78	0.06	1.73	2.55

1. Standard Error, 2. Standard Deviation

Table 2 is displaying correlation coefficients between different variables. The correlation between TDS and chloride was high. Similarly, correlation between pH and Fluoride, Iron, Nitrate was very high. It indicates that TDS was mainly influenced by chloride and pH is greatly influenced by concentration of fluoride, iron and nitrate ion.

Table 2 Correlation matrix between different water quality parameters

	Turbidity	TDS	pH	Chloride	Fluoride	Iron	Nitrate	TH
Turbidity	1.00							
TDS	0.68	1.00						
pH	-0.15	-0.03	1.00					
Chloride	0.62	**0.74**	0.31	1.00				
Fluoride	0.21	0.51	**0.74**	0.46	1.00			
Iron	-0.43	-0.57	**0.76**	-0.11	0.33	1.00		
Nitrate	0.01	0.13	**0.80**	0.62	0.51	0.50	1.00	
TH	0.22	0.53	0.03	0.09	0.55	-0.18	-0.36	1.00

CA was used to see the similitude among the sampling locations. The CA resulted in a dendrogram, which makes it easier to classify the clusters, provides visual information, and makes it easy to identify with the variation in water quality (Kim et al., 2007). Cluster analysis was grouped the sampling locations into 3 cluster based on observed parameters. 1st group [fourth sampling site (SS4) and fifth sampling site (SS5)] represents the low polluted sampling sites, cluster 2 (SS2, SS3, SS6 & SS7) represents the medium polluted sites and cluster 3 (SS1) represents the highly pollutes sampling sits. SS represents sampling site.

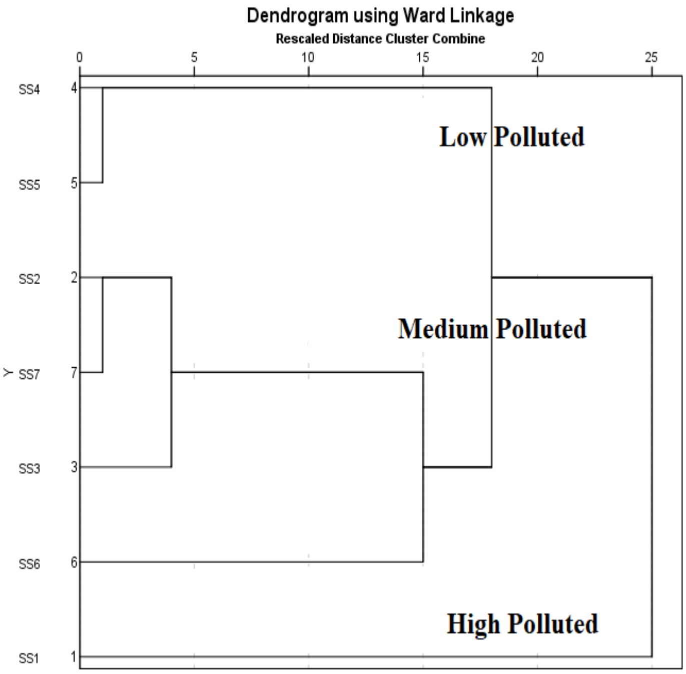

Fig. 2 Cluster analysis

In PCA, the scree graph has been used to decide the number of factors to retain. In scree plot, eigenvalues of factors arranged in descending order of magnitude from the left to the right side. The goal is to determine the factors that can be taken by locating the factors' point of inflection. Therefore, the significant number of factors that must be extracted for your factor analysis is the number of factors prior to the curve flattening. PCA gives information on the most important variables for accurately describing the entire data (Iscen et al. 2008).

PCA was applied to the normalized data sets in this study, in order to extract the valuable factor, eigenvalue-one criterion (Kaiser, 1960) was employed.

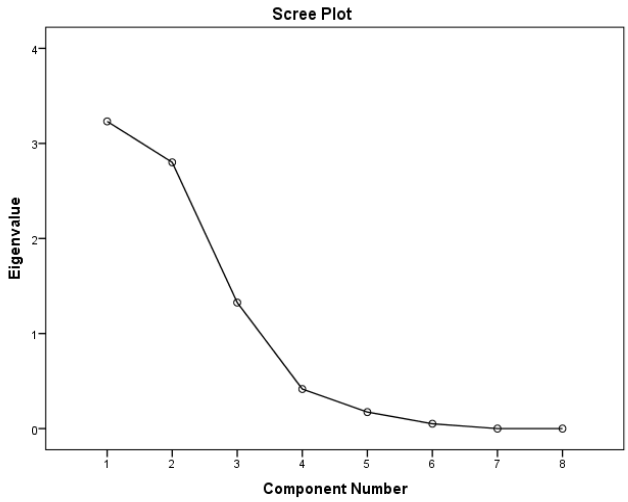

Fig. 3 Scree plot

The significance of the component is determined by its Eigenvalue. The most significant components have the most Eigen values. To make the results easier to understand and to maximize the variance of each factor, the varimax rotation method was used. Three principal components had eigen values larger than one. About 92% of the data sets' total variance can be attributed to these three main components. About 38% of the variance was accounted for by the first component. About 34 % percent of the total variance came from the second component. About 21 % of the variance was explained by the third component.

The rotated component matrix has been given in Table 4. The first component has extreme positive loading on pH, iron, nitrates and fluoride. This component represents the mineral content of groundwater. The second part has extreme positive loading on TDS, turbidity and chloride. This component may be related to processes known as "soil leaching." The third component has a strong positive loading on total hardness.

The components of a space-rotated graph (Fig. 4) enable a clear presentation in which the data's position in relation to the axes and their relationship to one another reveal environmental data's similarities.

Table 3 Rotation of sums of squared loadings

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	Percentage of Variance	Cumulative Percentage	Total	Percentage of Variance	Cumulative Percentage	Total	Percentage of Variance	Cumulative Percentage
1	3.232	40.400	40.400	3.232	40.400	40.400	3.012	37.647	37.647
2	2.801	35.014	75.415	2.801	35.014	75.415	2.712	33.894	71.540
3	1.326	16.575	91.990	1.326	16.575	91.990	1.636	20.449	91.990
4	.416	5.202	97.192						
5	.174	2.172	99.364						
6	.051	.636	100.000						
7	2E-16	2.7E-15	100.000						
8	7E-17	9.1E-16	100.000						

Table 4 Rotated component matrix

	Component		
	1	2	3
Tur	-0.151	0.834	0.096
TDS	-0.038	0.881	0.419
pH	0.976	-0.044	0.126
Chloride	0.393	0.877	-0.049
Fluoride	0.699	0.296	0.616
Iron	0.784	-0.517	-0.059
Nitrate	0.879	0.302	-0.326
TH	-0.085	0.144	0.972

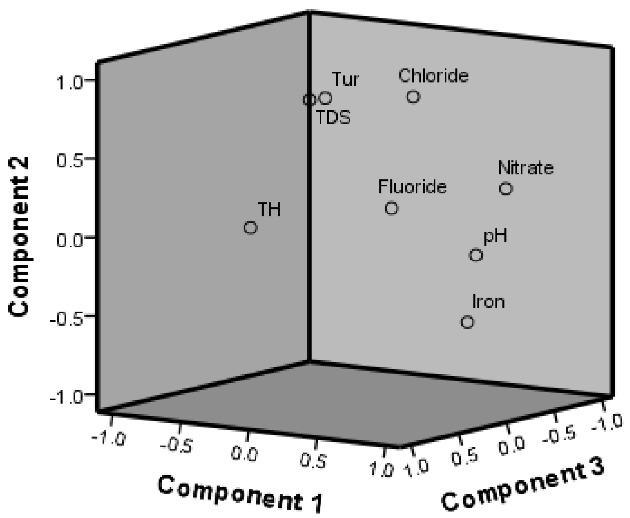

Fig. 4 Component plot in rotated space

4. Conclusion

Multivariate statistical techniques have been used in the present study to observe important aspects of water quality monitoring. Evaluation of groundwater quality was accomplished with the assistance of testing and examination. Study showed that water is not fit for drinking purpose. Treatment will be required for supply of this water to users. In order to determine the latent pollution sources of the various sampling locations; the dataset was then subjected to a variety of Multivariate statistical techniques with principal component analysis and cluster analysis. As a result, this study demonstrates how these techniques can help to effectively manage ground water quality by identifying factors that cause variations in water quality and analysing and interpreting complex data sets.

Authors Contribution

Vinod Kumar Kushwah: Survey of study area, collection and analysis of samples in laboratory**; Neha Singh:** Drafted the manuscript; **Kunwar Raghvendra Singh:** Performed statistical analysis and critical revision of the article.

Conflict of interest

There is no conflict of interest.

Acknowledgement

I would like to convey my profound gratitude to Prof. Sudhir Kumar Goyal, HOD of Department of Civil Engineering, for his continuous support and guidance.

References

Bouguerne, A., Boudoukha, A., Benkhaled, A. Mebarkia, A.H., 2017. Assessment of surface water quality of Ain Zada dam (Algeria) using multivariate statistical techniques. Int. J. River Basin Manag. 15(2), 133-143.

Carrard, N., Foster, T., Willetts, J., 2019. Groundwater as a source of drinking water in Southeast Asia and the Pacific: A multi-country review of current reliance and resource concerns. Water 11(8), 1605.

Dixon, W., Chiswell, B. 1996. Review of aquatic monitoring program design. Water Res. 30(9), 1935-1948.

Fan, X., Cui, B., Zhao, H., Zhang, Z., Zhang, H., 2010. Assessment of river water quality in Pearl River Delta using multivariate statistical techniques. Procedia Environ. Sci. 2, 1220-1234.

Iscen, C.F., Emiroglu, Ö., Ilhan, S., Arslan, N., Yilmaz, V., Ahiska, S., 2008. Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey. Environ. Monit. Assess. 144(1-3), 269-276.

Kaiser, H.F. 1960. The application of electronic computers to factor analysis. Educ. Psychol. Meas. 20(1), 141-151.

Kim, J.O., Mueller, C.W., 1987. Introduction to factor analysis, what it is and how to do it, Quantitative Applications in the Social Science Series, Sage University Press, Newbury park.

Kanwar, R., 2009. Sustainable water systems for agriculture and 21st century challenges. J. Crop Improv. 24(1), 41-59.

Singh, K.P., Malik, A., Mohan, D., 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 Res. 38(18), 3980-3992.

Singh, K.R., Bharti, N., Kalamdhad, A.S., Kumar, B., 2017. Surface water quality assessment of Amingaon (Assam, India) using multivariate statistical techniques. Water Pract. Technol. 12(4), 997-1008.

Singh, K.R., Goswami, A.P., Kalamdhad, A.S., Kumar, B., 2020. Assessment of surface water quality of Pagladia, Beki and Kolong rivers (Assam, India) using multivariate statistical techniques. Int. J. River Basin Manag. 18(4), 511-520.

Singh, R.B., 2000. Environmental consequences of agricultural development: a case study from the Green Revolution state of Haryana. Agric. Ecosyst. Environ. 82 (1-3), 97-103.

Stefanakis, A.I., Zouzias, D., Marsellos, A., 2017. Groundwater pollution: human and natural sources and risks. Environ. Sci. Eng. 4, 82–102.

Cite this article:

Kushwah, V.K., Singh, N., Singh, K.R., 2023. Multivariate statistical techniques for the evaluation of Groundwater quality of Mathura (India). J. Appl. Sci. Innov. Technol. 2 (1), 47-51.