RASĀYAN *J. Chem.*
180-188| Special Issue | 2022
ISSN: 0974-1496 | e-ISSN: 0976-0083 | CODEN: RJCABP
http://www.rasayanjournal.com
http://www.rasayanjournal.co.in

GROUNDWATER QUALITY EVALUATION USING STATISTICAL APPROACH AND WATER QUALITY INDEX IN AURANGABAD, BIHAR

Aftab Alam✉ and Anshuman Singh
Department of Civil Engineering, National Institute of Technology Patna, (Bihar) 800005, India
✉Corresponding Author: aftaba.phd19.ce@nitp.ac.in

ABSTRACT

Groundwater is the primary source of drinking water in south Bihar, and its quality can potentially affect human health. To accurately anticipate and evaluate water quality, 112 samples of groundwater were collected from Aurangabad, Bihar. Samples were taken from hand pumps and bore/tube wells. Physiochemical quality parameters, namely pH, total dissolved solids, alkalinity, hardness, magnesium, calcium, sulphate, nitrate, chloride, dissolved oxygen, and electrical conductivity were measured and compared to the Bureau of Indian Standards (BIS10500:2012). The physicochemical properties of samples were evaluated to determine their water quality index (WQI). Various water quality indicators were examined and correlated with other variables. All statistical analysis was done using RStudio4.1.1 software. On water quality data, the PCA method was employed, and four components with a total variance of 75.66% were extracted. WQI of more than 98% of the samples, either of excellent or good quality. Overall, the groundwater is drinkable. Comprehensive regulation and coordinated groundwater management are needed to protect groundwater quality and to achieve one of the Sustainable Development Goals.

Keywords: Groundwater, Water Quality Index(WQI), Correlation, Principal Component Analysis, Aurangabad.

RASĀYAN *J. Chem.*, Special Issue, 2022

This manuscript is focusing **SDG-6**: Clean Water and Sanitation

INTRODUCTION

The most precious natural resource is water, which all living things, including plants, animals, and people need to survive.[1-2] Water is the most important component for aquaculture, tourism, small businesses, industry, and agriculture.[3] Water supply constraints are one of the biggest challenges facing the water industry globally.[4] Groundwater is required to supply demand in many locations with limited surface water supplies and insufficient precipitation, usually without treatment.[5] Groundwater is among the world's most significant natural water resources, with widespread usage in dry and semi-arid areas for drinking, agriculture, and industry.[6] Agricultural development, fertilizer overuse, extensive evaporation, and little rainfall have all contributed to the growth of groundwater pollution in many geologic terrains.[7] PCA was used to reduce the dataset's dimensionality, and the clusters identified the water samples as coming from surface water or groundwater.[8] Many researchers deal with water quality using Principal component analysis and Cluster analysis.[9] Water quality characteristics were used to classify the water networks.[10] The relationship between the study's physical, chemical, and microbiological components was investigated using Euclidean distance and PCA.[11] It's worth emphasizing that the water is of high quality and satisfies all requirements. So, understanding the factors affecting groundwater quality is important due to the scarcity of fresh groundwater.

Study Area

Bihar is located in the center of India's Gangetic Plain. Figure-1 depicts the study area of the district, which is located in the South Bihar Plains (SBP) and is a portion of the Ganga Basin's peripheral alluvial plains, with a geographical area of 3389 Km^2 between longitudes of 84^o 00' - 84^o 45' E and latitudes of 24^o 30' - 25^o 15' N. The district is bordered on the west by the Sone River, and on the south by the Chhota Nagpur Granitic Gneissic Complex (CGGC) of Jharkhand state, which is part of peninsular India. On the north and east, the district is bordered by the Arwal and Gaya districts.

People in the Aurangabad district rely on groundwater for their basic needs especially drinking, except in a few urban areas that obtain their water from neighbouring rivers; for example, because there is no possible

http://doi.org/10.31788/RJC.2022.1558191

aquifer underneath Aurangabad, a portion of the town's drinking needs are fulfilled by the Batane River. Previously, people relied on dug wells/dug-cum-bore wells to get groundwater.

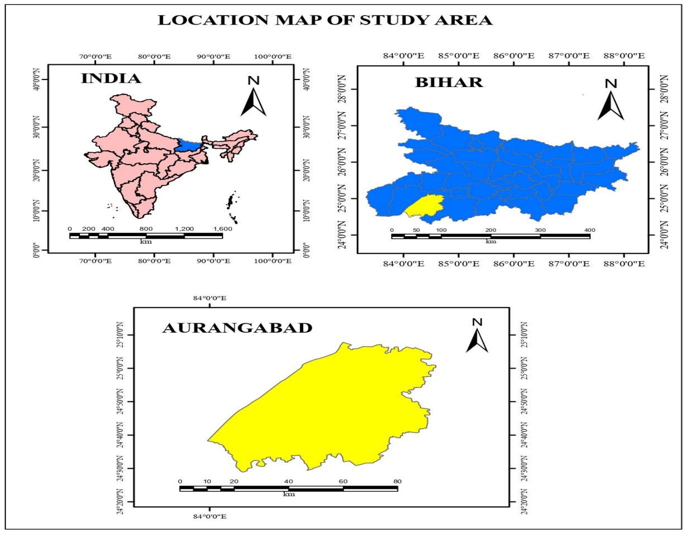

Fig.-1: Location Map of Study Area

In the state of Bihar's Minor Irrigation Census of 1993-94, 9056, groundwater abstraction structures were recorded in the district. In the research area each grid selected one exemplary sample based on the tenant population, totaling 112 drinking water samples. Each sample location's (Fig.-2) lat. and long. were calculated using the Global Positioning System (GPS). During the collection of samples, each source was left running for 2–4 minutes as a flushing time. Water samples were collected in 1-liter PVC bottles, which were treated with 5 M nitric acid and then cleaned with distilled water.[6,12] Groundwater samples were collected before the monsoon season in 2017.

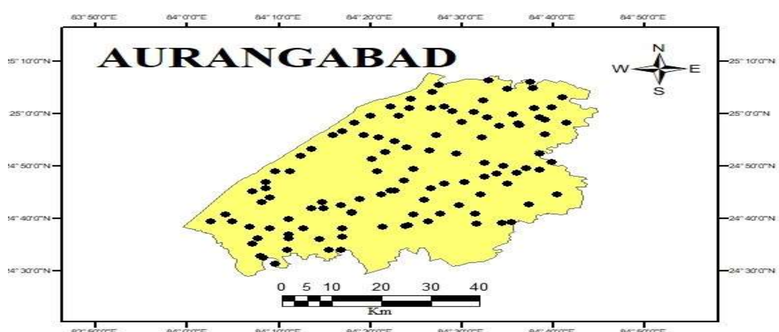

Fig.-2: Sample Location Map of Study Area

EXPERIMENTAL

Material and Methods

A 0.45μm membrane filter (syringe filter) was used to filter the collected samples. The Thermo Scientific Orion VERSA STAR Multi-Parameter Kit was used to assess in-situ factors, including pH, TDS, and electrical conductivity (EC), just after the sample was taken.[13] After that, the samples were carried back to the laboratory for further analysis. A portion of the samples was acidified after filtering to prevent wall deposition. It was necessary to keep unacidified samples at 4 °C in the dark until the analysis was complete to prevent structural disintegration.[14-16] Colorimetric method was used to assess nitrate in a water sample.[17] The spectrophotometer apparatus (Thermo Scientific Evoluion 201) is used for chloride, sulphate, and phosphate. They were measured using $AgNO_3$ titration, turbidimetric, and colorimetric methods, respectively.[6] The calcium and magnesium concentrations were determined using the conventional EDTA titration technique.[7] All the processes used in the investigation were conventional methods for examining water and wastewater.[15] In hydrological research, multivariate statistical analysis is commonly employed to investigate multi-dimensional issues. Multivariate statistical analysis methods have been adopted to monitor groundwater quality.[19-20] All statistical analysis is done using RStudio Software 4.1.1. and mapping is done using Arc- GIS 10.3 software.

Water Quality Index (WQI)

The WQI Method, put out by Horton in 1965, is used in this article to construct the Water Quality Index utilizing twelve factors.[21] These variables include pH, Chlorides, Electrical conductivity, Total Alkalinity, dissolved oxygen (DO), Fluoride, Total dissolved solids (TDS), Calcium, Magnesium, Total hardness, nitrate, and sulphate. The Horton recommends using this approach to determine the water quality index. In the following procedures, the standard value of parameters is considered with the help of World Health Organization (WHO) recommendations and Indian Standards.[14,22] Here, the Water Quality Index Based on the weighted arithmetic index shown below, the calculation was made.[23]

The calculation for water quality rating:

$$Q_n = 100 * [\frac{(Vn - Vi)}{(Vs - Vi)}]$$

Where,

Q_n = Value of quality rating or Sub-index

Vn = Actual concentration of the water sample,

Vs = Standard value of n^{th} parameter,

Vi = Ideal value of pure water sample.

Calculation of Unit Weight (Wn):

Unit weight for the respective parameter had a value that was inversely related to the Sn standard value that was suggested. $Wn = \frac{K}{Sn}$

Where, W_n = unit weight of each parameter used in water quality evaluation.

S_n = No. of standard value for each parameter,

K = Constant of proportionality,

$K = \frac{1}{\frac{1}{s1}+\frac{1}{s2}+\frac{1}{s3}.......+\frac{1}{sn}} = \frac{1}{\sum\frac{1}{Sn}}$

To calculate the WQI of a water sample, each water quality parameter's subindex is added up.

Water Quality Index = $\frac{\sum Q_n \mathrm{Wn}}{\sum \mathrm{Wn}}$

Where, $\sum Wn = 1$

RESULTS AND DISCUSSION

Hydrochemistry

Groundwater quality is essential to evaluate since it is the most important aspect in its suitability for drinking, residential, agricultural, and industrial uses. Physiochemical parameters, as well as statistical measurements such as minimum and maximum concentrations, mean concentration standard deviation, skewness, and kurtosis, are included in the Table-1. This research generally involves a considerable number of physiochemical variables and observations. The physicochemical parameters assessed in groundwater samples were combined for principal component analysis (PCA) in the current investigation. The goal is to use principal component analysis to reduce the dimensionality of data obtained throughout water quality monitoring, as well as to uncover correlations between the physio-chemical under consideration. The findings help the drinking water supply system & management and investment operations examine the data more effectively. Managers may use this strategy to acquire a global view of water quality in study areas, identify the components that fluctuate most, and benefit from more monitoring.[10]

Descriptive Statistics and Correlation Matrix

The statistical methods discussed here are empirical yet offer a way to find correlations between variables related to water quality, which are frequently the consequence of functional relationships. The strongly associated factors are found using statistical tests because the strength of these associations varies between streams. The correlation coefficient matrix for groundwater samples is shown in Fig.-3, while descriptive data is shown in Table-1. Hydrochemistry processes are clarified by the correlation matrix because it illustrates the separate roles of each parameter. Pearson's correlation matrix considers r values of + 1 and -1 as high correlation values which indicate complete correlation, i.e., functional dependency, between the two variables. A relationship between two variables is not significant if the numbers are close to zero. An r value of greater than or equal to 0.7 indicates a high correlation, whereas an r value of between 0.4 and 0.7 indicates a moderate correlation. An analysis using principal component analysis is used to determine if there is a correlation between the empirically observed parameters and the loadings of the factors.[24]

Table1: Descriptive Statistics of Water Quality Parameters in Aurangabad

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	SD Dev.	Skewness	Kurtosis
pH	6.43	7.013	7.11	7.083	7.192	7.52	0.213	-0.72	4.027
TDS	101.4	260.7	305.7	366.6	412.1	1470	201.15	2.61	12.096
EC	380	531.2	623.7	743.5	845.9	1200	415.95	2.5	11.70
DO	0.008	1.748	2.09	2.178	2.413	4.99	0.647	1.34	8.38
F	0.063	0.327	0.45	0.518	0.605	2.1	0.325	2.1	8.88
Cl	16.81	28.82	40.83	57.88	63.05	422.71	54	3.6	20.7
NO_3^-	1.11	6.728	12.215	17.741	25.637	52.22	14.85	0.997	2.75
SO_4^{2-}	0.12	5.57	13.44	24.89	26.82	220.63	34.88	3.245	15.94
T. Alk	13.8	31.9	36.3	38.72	43.77	88.7	12.45	0.982	4.78
TH	134	230	269.5	303.9	341	801	125.93	1.976	7.28
Ca^{2+}	46	138	170	183.1	207	640	83.44	2.645	12.58
Mg^{2+}	9	69	99	120.8	152	414	78.5	1.5	5.2

pH is in numerical value, EC is in µS/cm, all others are in mg/L.

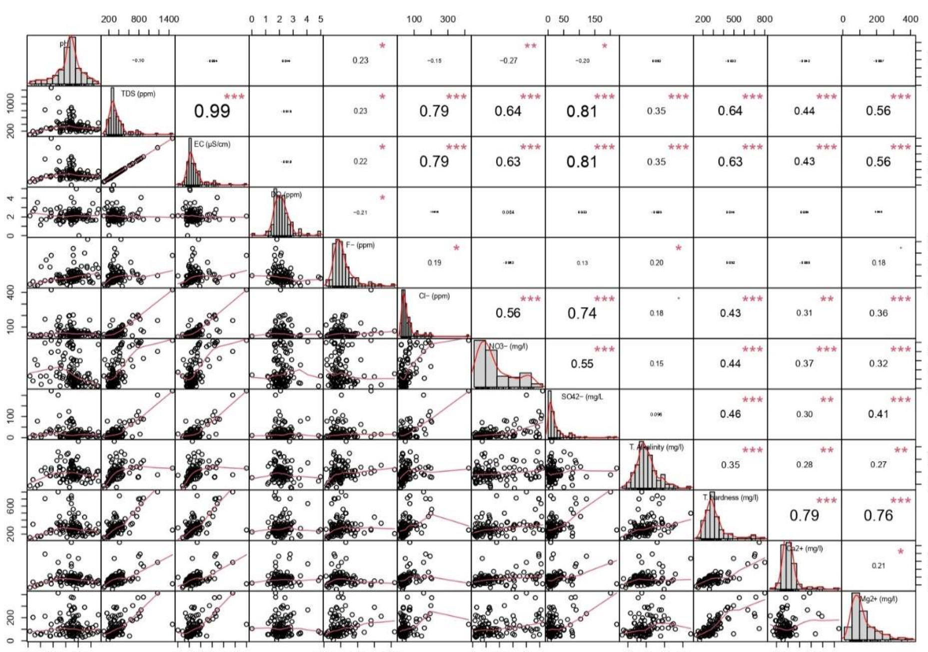

Fig.-3: Spearman's Correlation Coefficient Between Groundwater Measured Parameters. ~~that Have Been Measured~~

Spearman's correlation is shown in Fig.-3 for groundwater samples with varying water quality parameters. pH, DO, Fluoride, and alkalinity histograms and normally distributed curves indicate that they were well distributed. On the other hand, the TDS and EC indicate an anomaly in the dataset. Furthermore, other parameters, such as chloride and sulphate, were extremely skewed, with a lognormal distribution. Fluoride has a weak connection with other water quality indicators. The correlation of chloride with nitrate(NO_3^-) and sulphate(SO_4^{2-}) shows pollution caused by an anthropogenic interventions such as fertilizer application, wastewater percolation, or nutrient-rich irrigation water.[9] Furthermore, there were no significant one-to-one correlations between sulphate and calcium, and magnesium ($p < 0.05$, $r = 0.30, 0.41$). A study of the

studied region indicates either the gradual disintegration of calcium-bearing rocks such as calcite and dolomite or the simultaneous dissolution of magnesium-containing silicates and carbonates.[25]
GIS techniques are utilized to outline the geographical interpolation of groundwater contaminants to draw a WQI map (Fig.-5) that shows spatial variation of different water classes using the Inverse Distance Weighted (IDW) approach.

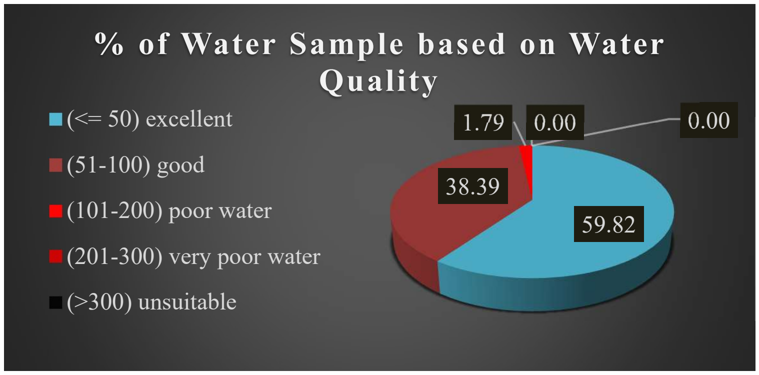

Fig.-4: Class of Water Based on WQI

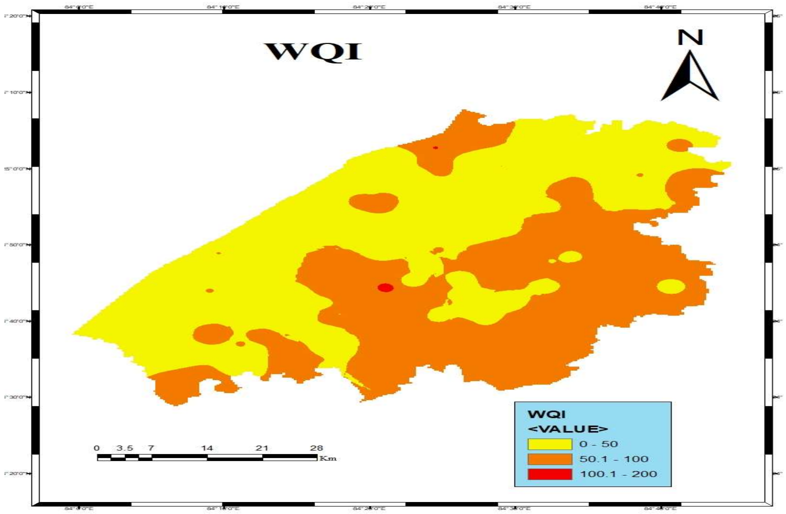

Fig.-5: Map Displays the Geographic Fluctuation of Different Water Classes using WQI Values

Results of Multivariate Analysis and Water Quality Parameter Distribution

The results of descriptive data on the groundwater quality in the Aurangabad Districts are described as follows. The pH of the water ranged from 6.43 to 7.52, with an average of 7.083, indicating that the groundwater is somewhat alkaline. Fluoride has a maximum concentration of 2.1 mg/l and a minimum concentration of 0.06 mg/l. The WHO recommends a fluoride content in drinking water between 0.5 and 1.5 mg/l. Fluorosis affects more than 200 million individuals worldwide due to increasing fluoride levels in drinking water. The EC ranged from 380 to 1200µs/cm. Higher EC indicates water salinity and substantially impacts groundwater ion exchange and dissolution. In groundwater, total dissolved solids varied from 101.14 to 1470 mg/l. TDS levels beyond a certain threshold can induce gastrointestinal discomfort, cardiac disease, and kidney stones in humans. In groundwater, Ca^{2+} concentrations ranged from 46 to 640 mg/l with an average of 183.1 mg/l. It is possible to develop kidney stones, strokes, and hypertension if there is not enough calcium in water consumption. A sulphate concentration of 220.63 mg/l is the highest in groundwater, while a sulphate concentration of 0.12 mg/l is the lowest. In the study area, magnesium concentrations ranged from 9 to 414 mg/l in the groundwater.

PCA Scree Plot

On the other hand, a scree plot (Fig.-6) is a diagnostic tool for determining if principal component analysis (PCA) is appropriate for data. PC1 captures the most incredible variety, PC2 the second most, and so on. Principal components are generated in the sequence of increasing coverage. Here PC1 captures more than

5 Eigenvalues and PC2, PC3 and PC4 capture greater than 1. A PCA has as many principal components as qualities, and each provides some information about the data.

Principal Component Analysis

To better recognize the variation in the data set of connected variables, the principal component analysis (PCA) was given as an important method.[26] When using PCA, correlations are extracted and the amount of data is reduced to components that each account for a percentage of the variation between physiochemical parameters. The multivariate covariance model minimizes the dimensionality of the original, uncorrelated variables by creating a dataset of correlated variables based on weighted linear combinations of the original variables.[27-28]

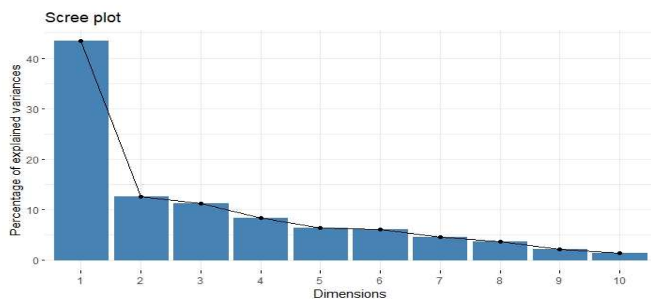

Fig.-6: Scree Plot Eigen Value Graph

According to this information, the first four components collectively account for 75.66% of the total variation in the data set. Of the overall variation, the first main component accounts for 43.45% and the second 12.62%. The extracted factor results suggest (Table-2) that TDS, EC, chloride, nitrate, sulphate, and hardness have high positive factor loadings in factor 1. Weathering and leaching of host rocks may be regarded as the source of Factor 1. Factor 2 has a significant positive loading for pH, fluoride, and Alkalinity and accounts for 12.62 % of the overall variation. Therefore, it may be claimed that Factor 2 reflects the impact of human activity. Chloride (Cl^-), and sulphate (SO_4^{2-}) have a positive loading with factor 3. In the third principal component, 11.21% of the total variance is contributed by factor 3. DO and pH have a high negative correlation in factor 4. Only Alkalinity and calcium hardness have a positive relation with factor 4. And fourth principal component contribution is 8.36% of the total variance. Analyzing the principal components, a two-dimensional graphic known as a biplot (or PCA biplot) illustrates the link between PC1 and PC2(Fig.-7).

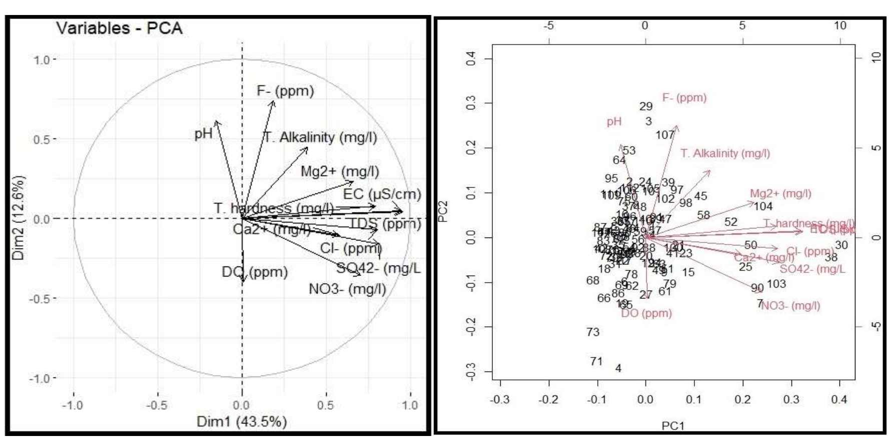

Fig.-7: An Explanation of the Relationship Between Physiochemical Parameters using Biplots of Principal Components

Table-2: Component Matrix

	PC1	PC2	PC3	PC4
pH	-0.066820538	0.49788341	-0.200498	-0.52918218
TDS	0.416879292	0.03333313	0.1376837	-0.06679305
EC	0.414184758	0.03713208	0.1403091	-0.07184524

DO	0.003150122	-0.32444684	-0.2900473	-0.74300457
F^-	0.082140048	0.59972686	0.3038091	-0.11479528
Cl^-	0.351483805	-0.05928145	0.298821	-0.11398243
$N0_3^-$	0.30936037	-0.29342009	0.0764187	0.01997956
$S0_4^{2-}$	0.356014144	-0.13345173	0.2860046	-0.12820936
T. Alkalinity	0.170495486	0.3615428	-0.2530373	0.26554757
T. hardness	0.34775052	0.05975155	-0.4535984	0.09684845
Ca^{2+}	0.253156244	-0.08662497	-0.5151719	0.18355245
Mg^{2+}	0.288726088	0.18835562	-0.1805986	-0.04115776

Table-3: Total Variance in Data Set

Factors	eigenvalues	% of variances	cumulative % of variances
comp1	5.214860972	43.45717476570	43.45717
comp2	1.514531189	12.62109324038	56.07827
comp3	1.345730356	11.21441963144	67.29269
comp4	1.004121907	8.36768255851	75.66037
comp5	0.769162423	6.40968686086	82.07006
comp6	0.727834186	6.06528488423	88.13534
comp7	0.553284388	4.61070323286	92.74605
comp8	0.442049912	3.68374926297	96.42979
comp9	0.250868569	2.09057140473	98.52037
comp10	0.16826338	1.40219483474	99.92256
comp11	0.009286642	0.07738868705	99.99995
comp12	6.07638E-06	0.00005063652	100.00000

CONCLUSION

Descriptive statics were produced on the entire groundwater quality data to obtain a quick overview of the data and explain the variation. Spearman's correlation matrix is constructed and illustrated to conclude the data better using the scatter matrix graphic. The correlation matrix is beneficial because it explains the connection between variables and the importance of each chemical parameter in various impact factors. Because most of what we know about water quality comes from numerical datasets, statistical approaches are key in water quality study. Various statistical approaches for analyzing water quality data have been introduced in this paper. Correlation matrix, scatter plot, and Principal Component Analysis are just a few examples. This study accurately predicts the Water Quality Index (WQI) for a groundwater sample, which is important for analyzing and reducing environmental consequences and, ultimately, ensuring good health. Using the PCA approach, four components with a combined variance of 75.66% are extracted from the water quality measurements. There are samples with a WQI of either good or very good in more than 98% of the samples. This means that the water is generally suitable for drinking. Anthropogenic activity is a significant contributor to groundwater pollution. Local communities' awareness of proper groundwater usage and preventing untreated domestic sewage from entering groundwater might be important for protecting water resources from future pollution. As a result, this study established that multivariate statistical methods (MSTs) and the Water Quality Index (WQI) are appropriate ways of assessing groundwater quality. This research also reveals that groundwater is suitable for irrigating and cultivating salt-tolerant crops. The study's conclusions will help decision-makers shortly as they manage groundwater resources and create sustainable designs.

ACKNOWLEDGMENTS

The Board of Research and Nuclear Sciences, under the Department of Atomic Energy in India, provided funding for this project's data collection of water quality parameters (BRNS Project Ref. No.: 36(4)/14/10/2014-BRNS). The authors are profoundly grateful to the reviewers and the editor for the careful examination of the draft of the manuscript and their many valuable comments and suggestions to help improve the manuscript.

CONFLICT OF INTERESTS

The authors declare that they have no conflict of interest.

AUTHOR CONTRIBUTIONS

All the authors contributed equally to the preparation of this manuscript. All authors read and approved the final manuscript The research profile of the authors can be verified from their ORCID ids, given below:

Aftab Alam http://orchid.org/0000-0002-7783-6186
Anshuman Singh http://orchid.org/0000-0003-2705-6453

REFERENCES

1. K. Boudeffa, F. Fekrache, and N. Bouchareb, *Rasayan Journal of Chemistry,* **13(1)**, 168(2020), https://doi.org/10.31788/RJC.2020.1315255
2. P. Vijayalakshmi and M. Abraham, *Rasayan Journal of Chemistry*, **11(2)**, 728(2018), https://doi.org/10.7324/RJC.2018.1122051
3. M. Upadhyay and M. K. Jindal, *Rasayan Journal of Chemistry,* **13(3)**, 1777(2020), https://doi.org/10.31788/RJC.2020.1335853
4. D. J. Reymond and K. Sudalaimuthu, *Rasayan Journal of Chemistry* **14(3)**, 1910(2021), https://doi.org/10.31788/RJC.2021.1436358
5. A. A. Mohammadi, *International Journal of Environmental Analytical Chemistry*, **1**, 11(2020), https://doi.org/10.1080/03067319.2020.1807533
6. D. Kumar, A. Singh, R. K. Jha, B. B. Sahoo, and S. K. Sahoo, *Arabian Journal of Geosciences*, **12(11)**, 2(2019), https://doi.org/10.1007/s12517-019-4519-5
7. N. Adimalla and H. Qian, *Ecotoxicology and Environmental Safety*, **176(126)**, 153(2019), https://doi.org/10.1016/j.ecoenv.2019.03.066
8. V. B. B Patil, S. M. Pinto, T. Govindaraju, V. S. Hebbalu, V. Bhat, and L. N. Kannanur, *Environmental Geochemistry, and Health*, **42(9)**, 2667(2020), https://doi.org/10.1007/s10653-019-00500-6
9. H. S. Kim and S. R. Park, *Water (Switzerland)*, **8**, 8(2016), https://doi.org/10.3390/w8080345
10. M. Maiolo and D. Pantusa, *Water (Switzerland),* **13**, 1766(2021), https://doi.org/10.3390/w13131766
11. P. Prusty and S. H. Farooq, *Environmental Processes*, **7(3)**, 825(2020), https://doi.org/10.1007/s40710-020-00453-4
12. IS:3025, Methods of sampling and test (physical and chemical) for water and wastewater, Part 39 (1991).
13. D. Kumar, A. Singh, R. K. Jha, S. K. Sahoo, and V. Jha, *Exposure and Health*, **11(2)**, 139(2019), https://doi.org/ 10.1007/s12403-018-00293-6
14. WHO Guidelines for Drinking-water Quality, fourth ed. World Health Organization, WHO (2011)
15. APHA, *Standard Methods for the Examination of Water and Wastewater,* 21st Edition, American Public Health Association, Washington, DC (2005).
16. D. Kumar, A. Singh, R. K. Jha, B. B. Sahoo, S. K. Sahoo, and V. Jha, *Environmental Earth Sciences*, **77(19)**, 702(2018), https://doi.org/10.1007/s12665-018-7889-1
17. U. Transformed and R. Results, Evolution 201 and 220 Innovation to Keep You Ahead Build Your Analyzer With CUE Software.
18. T. Nineteenth and E. Editions, Standard methods for the examination of water and waste water, **37(20)**, 2792(2000), https://doi.org/10.5860/choice.37-2792
19. S. Li, S. Gu, W. Liu, H. Han, and Q. Zhang, *Catena*, **75(2)**, 216(2008),https://doi.org/ 10.1016/j.catena.2008.06.005
20. M. Varol, *Environmental Pollution*, **266**, 115417(2020), https://doi.org/10.1016/j.envpol.2020.115417
21. S. Kumar and R. Singh, *Materials Today,* **45**, 5695(2021), https://doi.org/10.1016/j.matpr.2021.02.503
22. BIS, Indian Standard Drinking Water Specification (Second Revision), IS 10500, 1(2012),
23. A. N. Gupta, D. Kumar, and A. Singh, *Journal of the Geological Society of India*, **97(9)**, 1063(2021),

https://doi.org/10.1007/s12594-021-1821-0
24. C. K. Singh, S. Shashtri, and S. Mukherjee, *Environmental Earth Sciences*, **62(7)**, 1405(2011), https://doi.org/10.1007/s12665-010-0625-0
25. M. Kumar, A. Ramanathan, M. S. Rao, and B. Kumar, *Environmental Geology*, **50(7)**, 1025(2006), https://doi.org/10.1007/s00254-006-0275-4
26. C. A. B. Garcia, *Modern Environmental Science, and Engineering*, **3(10)**, 690(2017).
27. M. Tripathi and S. K. Singal, *Ecological Indicators*, **96**, 430(2019), https://doi.org/10.1016/j.ecolind.2018.09.025
28. M. Singh, P. Sahu, K. Tapadia, and D. Jhariya, *Water Supply*, **22(4)**, 1(2022), https://doi.org/10.2166/ws.2022.024

[RJC-8191/2022]