International Journal of Recent Research and Review, Vol. XVIII, Issue 1, March 2025
ISSN 2277 – 8322

Correlation and Regression Analysis of Physico-Chemical Parameters in Groundwater: A Python Approach

Sumit Kasara
M.Tech Scholar
Department of Computer Science and Engineering
Rajasthan Institute of Engineering & Technology, Jaipur

Ms. Pragya Bharti
Assistant Professor
Department of Computer Science and Engineering
Rajasthan Institute of Engineering & Technology, Jaipur

Abstract: Groundwater is a crucial natural resource that serves as the primary source of drinking water for millions across India. However, its quality is increasingly affected by various natural and anthropogenic factors, including industrialization, agricultural runoff, urbanization, and climate change. Assessing groundwater quality is essential to ensure public health and sustainable water management.There is a relationship between variables which shows that one variable actually causes changes in another variable. In this paper, a statistical regression analysis method on ground water of four states from each zone was carried out. This technique was based on the study and calculating the correlation coefficients between various physicochemical parameters of drinking water and ranking of selected districts is done from each state.

The findings of the proposed work offer valuable insights into regional groundwater quality trends and can aid policymakers in implementing effective water management strategies. By identifying critical influencing parameters, this research contributes to improving water resource planning, ensuring sustainable usage, and mitigating potential health risks associated with groundwater contamination.

Keywords: Groundwater Quality Assessment, Physico-Chemical Parameters, Correlation and Regression Analysis, Statistical Modeling in Water Quality, Python-Based Data Analysis.

INTRODUCTION

1.Importance of Groundwater in India
Groundwater is one of the most vital natural resources in India, supporting drinking water supply, irrigation, and industrial activities. India is the largest user of groundwater in the

world, accounting for nearly 25% of the global extraction. Approximately 85% of rural and 50% of urban populations rely on groundwater for drinking, while nearly 60% of irrigation needs are met through it. The increasing dependence on groundwater highlights the urgent need to assess its quality to ensure sustainability and public health.[1]

2. Groundwater Quality: A Key Concern Groundwater quality varies significantly across India due to natural geological formations and anthropogenic activities. The presence of chemical, physical, and biological contaminants can severely impact its usability. Contaminants such as fluoride, arsenic, nitrates, heavy metals, and microbial pathogens have been reported in various regions, leading to severe health risks and environmental challenges.[2]

3. Parameters Affecting Groundwater Quality Groundwater quality is determined by several parameters, which can be classified into three categories:

Types of Bioremediation:

A. Microbial Bioremediation: Uses bacteria, fungi, and algae to break down contaminants. Example: Arsenic-oxidizing bacteria that convert toxic arsenic into less harmful forms.

B. Phytoremediation (Plant-Based Bioremediation): Uses plants to absorb and remove contaminants. Example: Water hyacinth and vetiver grass absorb heavy metals.

C. Bioaugmentation: Introduces specific microbial strains to accelerate contaminant degradation. Used in industrial pollution sites.

D. Natural Attenuation: Relies on natural microbial degradation over time. Effective for petroleum and organic waste contamination.

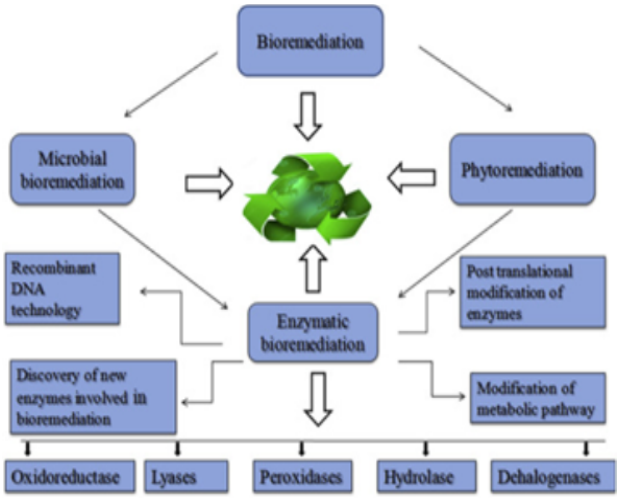

Fig.1.Types of Bioremediation

OBJECTIVES AND SIGNIFICANCE OF THE STUDY

1. To study Physico-chemical parameters of groundwater in Tamil Nadu, Rajasthan, Punjab, Odisha i.e. in different zone of India.
2. To study pre-monsoon and post-monsoon variation in Physico-chemical parameter in Tamil Nadu and Shahjahanpur.
3. To study the BOD(Biochemical oxygen demand) of Rajasthan in 2015 to 2016.

The significance of this study is to help policy makers to spread awareness about water quality

degradation and the parameters affecting it and to rank selected districts form each state taken from the four zones of the country. The significance of other objectives is to determine the pre-monsoon and post monsoon variation in physico-chemical parameters.

STUDY DESIGN

The secondary data about the of the physico-chemical parameters present in water in the state of Rajasthan, Punjab, Odisha and Tamil Nadu is taken which includes parameters such as pH, Electrical Conductivity (EC), Turbidity ,Total Dissolved Solids (TDS), Total Hardness (TH), Calcium (Ca2+), Magnesium (Mg2+), Chloride (Cl-), Sulphate (SO4), Total Alkalinity (TA) and Dissolved Oxygen (DO) etc.

Correlation, regression and Difference of mean tests are used in this study.

The degree to which the relative motions of two variables are connected is determined by a statistical measure known as the correlation coefficient. The values fall between -1.0 to 1.0. If the calculated result was greater than 1.0 or less than -1.0, the correlation measurement was inaccurate. A correlation of -1.0 indicates a perfect negative correlation, while a correlation of 1.0 indicates a perfect positive correlation.

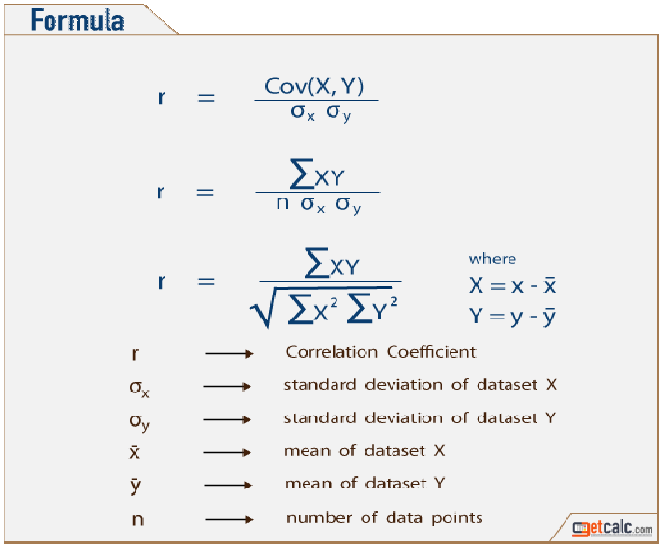

Fig.2.1.MATHEMATICAL MODEL

CORRELATION MATRIX OF DIFFERENT PARAMETERS OF WATER QUALITY IN TAMIL NADU.

	pH	Turbidity	TDS	TH	E.C	Mg	Ca
pH	1.000000	0.200385	0.466960	0.389609	0.353874	0.130193	0.429990
Turbidity	0.200385	1.000000	0.383990	0.278050	0.520994	0.435719	0.302374
TDS	0.466960	0.383990	1.000000	0.908101	0.941791	0.797905	0.888443
TH	0.389609	0.278050	0.908101	1.000000	0.774930	0.844163	0.952276
E.C	0.353874	0.520994	0.941791	0.774930	1.000000	0.793677	0.782660
Mg	0.130193	0.435719	0.797905	0.844163	0.793677	1.000000	0.737576
Ca	0.429990	0.302374	0.888443	0.952276	0.782660	0.737576	1.000000

Heat Map corresponding to above matrix -

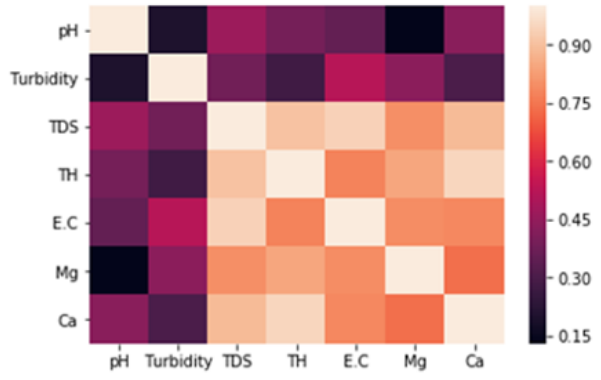

Fig.2. Mltiple linear regression between 'tds' & ca, mg ,th, e.c, ph

TDS = a +b(Ca) +c(Mg) + d(TH)+e(E.C) +f(pH)

TDS = -94.36 -0.25(Ca) - 0.24(Mg) +1.10(TH) + 0.38(E.C) + 15.85(pH)

Here b,c,d,e,f are the regression coefficient respectively and a is the intercept.

- RANKING

I have rank different district of Tamil Nadu based on Ph, TDS, E.C, TH, Turbidity.

District		
D5	(-0.7313497547181644+0j)	1
D1	(-0.41335721197579134+0j)	2
D3	(-0.4133572119757913+0j)	3
D2	(-0.3179925427423734+0j)	4
D4	(-0.14926456260319138+0j)	5

CORRELATION MATRIX OF DIFFERENT PARAMETERS OF WATER QUALITY IN RAJASTHAN:

	PH	EC	Cl	TDS	Ca	Mg	F
PH	1.000000	-0.105595	-0.090251	-0.105557	-0.178583	0.134235	-0.000124
EC	-0.105595	1.000000	0.919141	1.000000	0.620306	0.612181	0.404169
Cl	-0.090251	0.919141	1.000000	0.919161	0.604545	0.640071	0.364317
TDS	-0.105557	1.000000	0.919161	1.000000	0.620305	0.612203	0.404174
Ca	-0.178583	0.620306	0.604545	0.620305	1.000000	0.469875	0.301257
Mg	0.134235	0.612181	0.640071	0.612203	0.469875	1.000000	0.055547
F	-0.000124	0.404169	0.364317	0.404174	0.301257	0.055547	1.000000

Heat Map corresponding to above matrix –

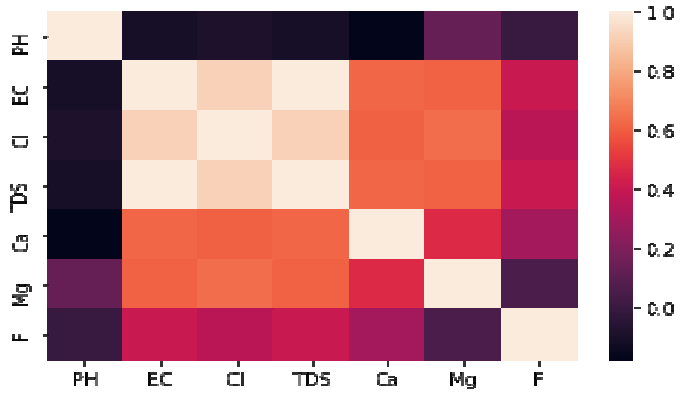

Fig.3. Multiple linear regression between 'tds' and 'ca' , 'mg' , 'f' , 'cl' , 'ec'

TDS = a + b(Ca) + c(Mg) + d(F) +e(Cl) +f(EC)

TDS = 0.2 -0.00023(Ca) + 0.00051(Mg) + 0.017(F) + 0.0003(Cl) + 0.6499(EC)

Here, b ,c ,d , e,f are the regression coefficients respectively and a is the intercept

- RANKING

  Ranking of 5 major districts of Rajasthan is based on the following parameters PH ,F ,TDS ,EC ,Cl , on the following states .

  1. AJMER
  2. ALWAR
  3. JAIPUR
  4. DAUSA
  5. JHUNJHUNU

  A matrix was created based on the above parameters if the PH lies between 6-8.5 then the district is given a point '1' ,if does not lie in that range then it will be awarded '0' , If Cl is less than 500 mg/l then it is given '1' otherwise '0' , if TDS is more than 900 mg/l it is given '1' otherwise '0' , if Fluorine is more than 1.5mg/l District is awarded '0' otherwise '1'. Eigen vector and values were found for this matrix , district were ranked according to this matrix.

	PH	F	TDS	EC	Cl
AJMER	1	0	0	0	0
ALWAR	1	0	0	0	1
JAIPUR	1	0	0	0	1
DAUSA	1	1	0	0	1
JHUNJHUNU	0	0	1	1	1

eigen vector of each district

District	
AJMER	0j
ALWAR	(-0.3378422856337929+0j)
JAIPUR	(-0.33784228563379287+0j)
DAUSA	(-0.4951319347363809+0j)
JHUNJHUNU	(-0.725651119533125+0j)

District	eigen vector of each district	rank
JHUNJHUNU	(-0.725651119533125+0j)	1
DAUSA	(-0.4951319347363809+0j)	2
ALWAR	(-0.3378422856337929+0j)	3
JAIPUR	(-0.33784228563379287+0j)	4
AJMER	0j	5

RESULT AND DISCUSSION

OBJECTIVE 1.1:

The coefficient of correlation r values between EC and TH, pH Ca, Mg, TDS are ,0.35,0.521,0.774,0.94,0.79 respectively. These values indicate that there is strong positive relation with EC, TH, HCO3, Na, Ca and Mg. Ca has strong correlations of 0.84with TDS and has linear correlation of 0.13,0.43,0.79,0.73 with pH , E.C, Mg. TDS has strong correlations of 0.94, 0.90 ,0.88 with EC, TH, and Ca and linear correlation of 0.79, 0.38 with, Turbidity, Mg .

TH has very strong correlations of 0.93 and 0.92 with Ca and Mg and linear correlation of 0.60, 0.77, 0.64 for TDS, E.C. But TH has very low correlations of 0.38, 0.27 with pH , Turbidity.

When analyzing data on groundwater quality and connecting it to particular hydrogeological processes, correlation analysis is incredibly

helpful 14, 15. Electrical conductivity and total dissolved solids can be utilized to distinguish one another. Total hardness has a positive correlation with calcium, magnesium, and chloride ions, and electrical conductivity is proportionate to the total dissolved solids. Changes in land use, mining, and inappropriate effluent discharge in the research area may be the cause of the high association between these characteristics.

OBJECTIVE 1.2:

The water quality analysis for different groundwater has been carried out for PH, TDS, Ca ,Mg ,Cl , F and EC . It has been found that amount of fluorine present in groundwater of various districts is very high, beyond the permissible range i.e. beyond 1.5 mg/l. In many districts amount of salt concentration(mg/l) is beyond the permissible limits and for Rajasthan the correlation coefficient between TDS and EC is 1.

TDS has strong correlation with electrical conductivity, Ca, Mg, Cl with correlation coefficients 1, 0.61, 0.62, 0.9. It has moderate correlation with F with correlation coefficient being 0.4.

PH has weak correlation with EC, Mg, Ca, Cl, TDS which shows PH is independent of all these parameters.

EC has strong correlation with Mg, Ca ,Cl , TDS which is evident as EC is nothing but the amount of salt present in water .

F has low correlation with the parameters. F is independent of all the parameters.

The naturally occurring increased quantity of F-bearing minerals in the host rocks and sediments is the reason for the abnormally high F concentration found in the groundwater of all 32 districts of Rajasthan. The fluorotic minerals that account for F in the range of 180 to 3100 parts per million (average) are found in the important rocks, which include granites, gneisses, mica, schists, limestone, sandstone, phosphorite, shales, clays, acid igneous rocks, basalts, and alluvium.

The 5 districts of Rajasthan have been ranked in the following order from the best to poor groundwater quality – Jhunjhunu ,Dausa ,Alwar , Ajmer and Jaipur .

OBJECTIVE 1.3:

The water quality analysis for different groundwater has been carried out for PH , EC, Ca ,Mg , Cl, F and some other parameters . In many districts amount of salt concentration(mg/l) is beyond the permissible limits .

Cl has strong correlation with EC ,Ca ,Mg , Na and weak with F with correlation coefficient being -0.4 .

PH has weak correlation with EC ,Mg ,Ca ,Cl , TDS which shows PH is independent of all these parameters.

EC has strong correlation with Mg ,Ca ,Cl , SO4, Na and weak with F which is evident as EC is nothing but the amount of salt present in water .

F has weak correlation with the parameters .Therefore F is independent of all the parameters.

Punjab's ground water quality varies greatly from one area to another. From the north to the south/southwest, the quality of the ground water shifts from good to bad. Different quantities of soluble salts are present in the ground water in Punjab's South Western districts of Mansa, Bathinda, Muktsar, Ferozepur, and Faridkot. Using this water for irrigation has a negative impact on agricultural output. Three categories—good, marginal, and poor—have been established for the state's groundwater. The fluctuation in Electrical Conductivity (EC) and Residual Sodium Carbonate (RSC) values has served as the foundation for this. RSC must be less than 2.5 me/L and EC must be less than 2.00 dS/m for ground water to be considered high quality.

The 5 districts of Punjab have been ranked in the following order from the best to poor groundwater quality – Ludhiana ,Jalandhar ,Amritsar , Sangur and Faridkot.

OBJECTIVE 1.4:

The water quality analysis for different groundwater has been carried out and 'pH', 'Turbidiy_NTU', 'DO_mg/L', 'BOD_mg/L', 'NO3-_mg/L', 'Cl-_mg/L', 'Ca2+_mg/L', 'Mg2+_mg/L', 'Na+_mg/L', 'Fe_mg/L', 'Cr_mg/L.

It has been found that amount of chloride present in groundwater of various stations is very high, beyond the permissible range set by WHO. There is strong relation between Fe and Cr of 0.90 and moderate relation between turbidity and Cr, TDS and Cr, BOD and Fe, pH and TDS, Na and Cl .

pH has weak correlation with all the factors and moderate with DO. TDS and Cr have good correlation with all other factors.

DO has negative correlation with every parameter except pH.

Presence of high Cr concentration in groundwaters of all stations is due to chromite mineral found there in large amount. Many other minerals are also present there as Odisha has maximum no of mines in India. This thereason of salts found in water are in a bit high concentration.

The 5 stations of Odisha have been ranked on the following criteria and among them best was Jalapa Nadi at Handibhanga

OBJECTIVE 2.1:

In case of pH effect of pre monsoon on water quality is less than effect of post monsoon on water quality

Alternate hypothesis is accepted

In case of TDS effect of pre monsoon on water quality is less than effect of post monsoon on water quality

Alternate hypothesis is accepted

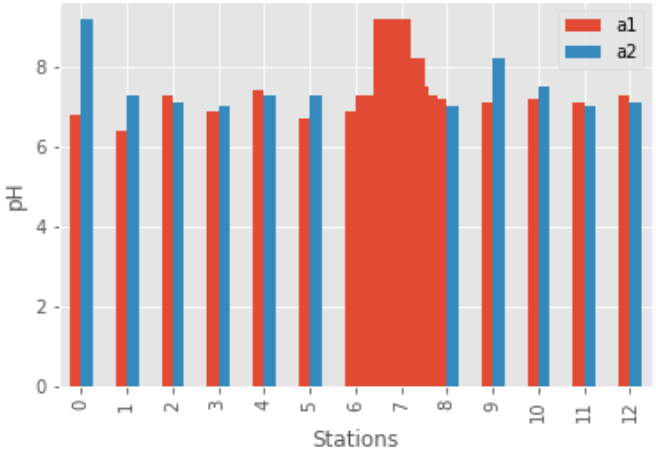

Fig.4. pH effect of onmonsoon water quality

OBJECTIVE 2.2:

In case of pH and Cl the null hypothesis is rejected and hence there is a significant difference in these parameters pre-monsoon and post-monsoon whereas in case of TDS the null hypothesis is accepted and hence there is no significant difference in it.

The study reveals that the quality of groundwater was far better during the post-monsoon season, as compared to that of pre-monsoon season. Main reason behind this is during monsoon surface water filtrates through various layers of the earth and recharge the Ground water. This filtration makes its quality better.

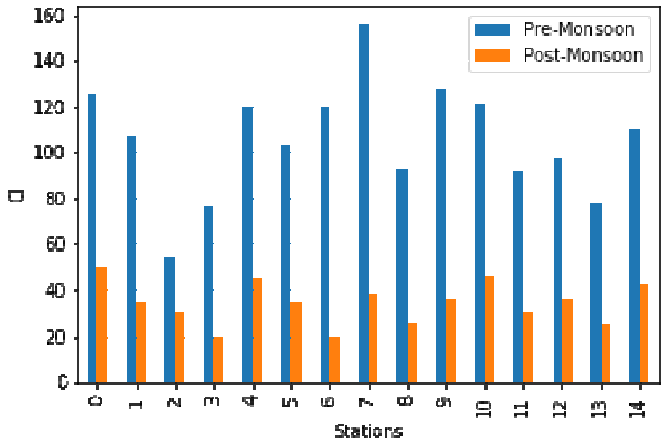

Fig.5. Graph between pH and post monsoon

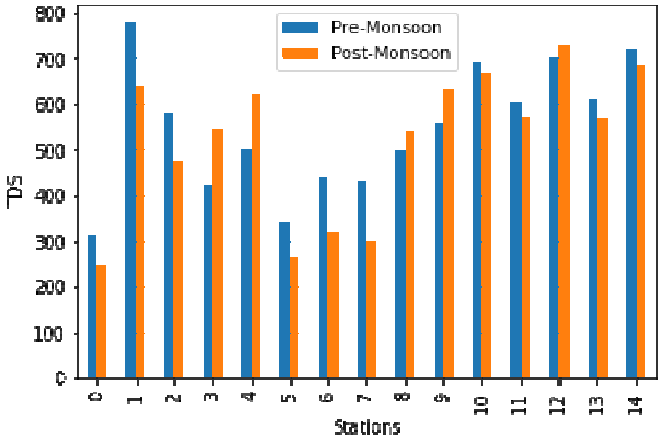

Fig.6. Graph between pH and pre and post monsoon

REFERENCES

1. In"About the Sustainable Development Goals - United Nations Sustainable Development." United Nations. https://www.un.org/sustainabledevelopment/sustainable-development-goals/.

2. Soni Chaubey and Mohan Kumar Patil, 2015. “Applications of Artificial Neural Network in Assessing Water Quality: A Case Study”, International Journal of Current Research Vol 7.

3. Dr. Mrugesh M Trivedi, M. (2016). Comparative Study of Pre and Post Monsoon Water Quality in and Around Gandhidham, Kachchh, Gujarat, India, for the Year 2015. *International Journal Of Science And Research (IJSR)*, *5*(6), 319-326. doi: 10.21275/v5i6.nov164049

4. Mazhar Nazeeb S.M Khan and Ravi Kumar., An Interpretation of Groundwater Quality using Correlation and Linear Regression analysis from Tiruchengode taluk, Namakkal di strict, Tamilnadu, Journal of Chemical and Pharmaceutical Research., (2012).

5. Altman S. J, Parizek, R. R., Dilution of nonpoint source nitrate in ground water, Jour. Environ. Quality (24) (1995),707-717.

6. Subramani T, Elango L and Damodarasamy S. R., Groundwater quality and its suitability for drinking and agricultural use in Chithar River Basin, Tamil Nadu, India, Environ Geol., 47(2005), 1099–1110.

7. N. Manivasagam, Physico-chemical examination of water, sewage and industrial effluents, Pragati Prakashan, Meerut. (1984).

8. A. H. Mahvi, J. Nouri, A. A. Babiei, and R. Nabizadeh, Int. J. Environ. Sci. Tech., 2 (1), 41 (2005).

9. Antony Ravindran A., Azimuthal Square Array Resistivity Method and Ground water Exploration in Sanganoor, Coimbatore District, Tamilnadu, India, Research Journal of Recent Sciences, 1(4), 41-45 (2012).

10. Ramaraju Anirudh and Giridhar M.V.S.S., “Evaluation of surface water quality in and around GHMC” Proceedings of National Conference on Climate Change and Sustainable Water Resources Management, (CSWM2015) organized by department of civil engineering, NIT Warangal on 3–5 September 2015.

11. Sharda S. “Status of drinking water quality in Ludhiana district, Punjab”

12. Burrough PA, Goodchild MF, McDonnell RA, Switzer P, Worboys M (1998) Principles of geographic information systems. Oxford University Press, Oxford

13. Falivene O, Cabrera L, Tolosana-Delgado R, Saez A (2010) Interpolation

algorithm ranking using cross-validation and the role of smoothing effect: a coal zone example. ComputGeosci 36(4):512–519

14. Gotway CA, Ferguson RB, Hergert GW, Peterson TA (1996) Comparison of kriging and inverse-distance methods for mapping soil parameters. Soil Sci Soc Am J 60(4):1237–1247
15. Seaborn.pydata.org. (2019). Multiple linear regression — seaborn 0.9.0 documentation. https://seaborn.pydata.org/examples) /multiple_regression.html
16. Sahu, R., Patnaik, D., & Behera, P. (2023). Impact of coal mining on groundwater quality in Odisha. *Environmental Science and Pollution Research, 30*(1), 45-56.
17. Sharma, N., Yadav, A., & Bansal, P. (2021). Groundwater contamination in the Upper-Doab region: A hydrogeochemical approach. *Journal of Environmental Management, 282*, 112-125.
18. Kumar, S., & Srinivas, R. (2021). Heavy metal contamination in groundwater: A study from peri-urban Bengaluru. *Environmental Pollution, 276*, 116-124.
19. Mishra, A., & Tiwari, P. (2019). Hydrogeochemistry and groundwater quality in Madhya Pradesh. *Groundwater for Sustainable Development, 8*, 123-134.