Correlation Analysis and Comprehensive Health Evaluation for River Water Quality Management

Navneet Kumar[1] and Gaurav Kumar Rastogi[2]
[1]Associate Professor, Department of Chemistry, Faculty of Engineering, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh 244001, India.

Email: navkchem@gmail.com, drnavneet.engineering@tmu.ac.in

[2] Professor, Department of Chemistry, R.D. Engineering College, Ghaziabad, Uttar Pradesh, India.

Email: drgauravrdec@gmail.com, **Corresponding Author**

Abstract
Water exists in various forms in nature, and monitoring its quality is crucial to determine if it meets the requirements for its intended purposes. An experiment was conducted to assess the concentration levels of different water quality parameters. This involved a statistical regression analysis of thirteen data points from sites along the Gagan River in Moradabad. The aim was to develop a management strategy for water quality by studying correlations between various water quality indices. The physico-chemical properties analyzed in the study revealed that the water in this region is contaminated. A comparison of the measured values with the World Health Organization (W.H.O.) drinking water standards confirmed that the water quality does not meet acceptable safety levels for human consumption. Therefore, the results highlight the need for effective water quality management in the area to protect public health and the environment. The results of the regression analysis indicate a substantial correlation between the conductivity of river water and each of the fourteen water quality measures that were examined. It might be argued that regulating the water's conductivity is a very effective way to check the quality of Gagan river water in Moradabad. The current study could be viewed as a first step in managing water quality. Strong environmental policy implementation and widespread public awareness campaigns are two possible ways to improve the health of rivers. Modeling water quality is used to assess the health of the Gagan rivers. This study uses biological and physiochemical characteristics to assess the health of rivers. Public health, environmental preservation, and river water resource management are all examined in this study.

Keywords: Water quality parameters, correlation, regression models and quality

Introduction

For their survival, numerous ancient societies and human settlements depended on the freshwater supplied by rivers. Approximately 9000 km3/year of world water are used by human activity. This limited water source has become poisoned due to overuse. Monitoring river water quality is so essential. To effectively divide water resources and put water protection measures into action, surface water quality monitoring is necessary (Da Luz et al. 2020, Khatri et al. 2021). Accurately determining the water quality of a river and locating possible pollution sources are essential for both sustainable water management and pollution control. Rivers strengthen an economy. The amount of river water is infinite. In addition, they offer sand mining, transportation, garbage disposal, irrigation, entertainment, sports, and power. Given that rivers are a significant source of freshwater, their water quality is significant to the ecosystem (Ighalo et al. 2020, Kumar 2022). Rivers have been severely contaminated for a long time due to improper trash disposal. Authorities are able to identify the source and location of pollution, transmit an alarm, and continuously monitor of thanks to the real-time water quality monitoring system. After the industrial revolution, rivers were no longer able to filter waste. The quality and use of river water are influenced by both natural and anthropogenic processes, such as erosion, precipitation, and crustal weathering. Internal river channels influence the absorption and transportation of pollutants from many sources, therefore their use is also dependent on them. River pollution is influenced not only by the connections but also by groundwater flow, inflow and outflow pumps, interflow, seasonal precipitation, and surface variation (Kharake 2021).

Activities both natural and man-made will have an impact on the quantity and quality of water, which will change the ecosystem and result in hydrological extremes. Urbanization, industry, and population growth raise the need for freshwater. The management of freshwater resources is getting harder as the world's population increases. Water quality modeling is used to evaluate the general condition of India's rivers (Khatri et al. 2020). This study evaluates the health of rivers by combining biological, chemical, and physiological aspects. The northern region of India's public health, environmental preservation, and water resource management are the main topics of this study (Mokarram et al. 2020).

It is common knowledge that having access to clean water is crucial for maintaining good health. Every human person on the planet has the basic need for an adequate quantity of clean, fresh water, but millions of people around the world are said to be without it. The majority of water-related illnesses and health issues are brought on by insufficient and careless water resource management (Schot et al. 2022). It is only possible to ensure equitable access, sustainability, and safe water for all. Using correlation studies and statistical regression analysis of eight data points from the Gagan River water in Moradabad, the current study aims to recommend some practical solutions for the management of river water quality. (Khan, 2005, Madhuri, 2004, Ramamurthy, 2007).

Statistical regression analysis is a useful tool for achieving the correlation between different elements. Correlation analysis measures how closely certain independent and dependent variables are related. The analysis provides a mechanism for forecasting or prediction by seeking to ascertain the nature of the relationship between the variables. (Mulla, 2007, Draper, 1966, Snedecor, 1967, Kumar, 2005).

Approximately 38 lac people live in Moradabad, a B class city in western Uttar Pradesh. Positioned on the banks of the Ram Ganga River, Moradabad rises to an elevation of around 670 feet above sea level. From the Himalaya in the north to the Chambal River in the south, it is stretched out. 28°20', 29°15', 79°E, and 78°4' are its coordinates. District Budaun is located north of district Moradabad, district Bijnor and Nainital are located north of district Moradabad, while district Rampur is located east of the Ganga River. Moradabad's modernization and population have increased significantly during the last several decades. The main industries include steel, brass, and various allied ancillaries such as paper mills, sugar mills, crushers, dye factories, and so on.

The Gagan River rises in a pond in Nazibabad's Harganpur, in the Bijnor district. It travels around 150 kilometers, passing through Moradabad and J.P. Nagar. It merges with the Ram Ganga river in Seekanderpur-Patti, southeast of Moradabad. along Moradabad City, two small rivers—Bann and Karula-I—that originate in the Bijnor district are also combining along the right bank of the Gagan River. On both sides of the Gagan River are several heavily populated communities. For thousands of individuals, river water is essential to their everyday lives. It appears that the water in the Gagan River is extremely contaminated and unsafe for use by people or animals.

The objective of the study to determine whether or not the rivers' water quality satisfies the requirements and is thus fit for human consumption. To determine the factors causing variations in the rivers' water quality and a comprehensive evaluation of Gagan River's health.

Material and Methods

Following conventional methods and procedures of sampling and estimation, fifteen distinct physico-chemical parameters of water quality, including conductivity, were measured at thirteen separate Gagan river water locations in Moradabad (APHA 2012, Merck, 1974). For the quantitative study, all compounds of Anal R grade were utilized. The Century CP901 pH meter, RI 215 R conductivity meter, and Hach spectrophotometer 2010 were utilized, in that order, to measure the following parameters: pH, conductivity, sulphate, and phosphate. Table 1 provides a brief overview of the sampling locations used for the quantitative estimation of water quality parameters.

Table 1: Brief description of sampling sites

S.No.	No. and name of site	Noticed activities
1	I, U/S River at Sirsa Manihar	Nil
2	II, Bann river at Sirsa Manihar	Occasional bathing and fishing
3	III,D/S River at Sirsa Manihar	Nil
4	IV, River at Taiya-Moda	Nil

5	V, River at Chaudharpur	Very occasional funeral activities
6	VI, River at Malgadda	Receives agricultural run off
7	VII, U/S River at Mbd-Dlh bridge	Sand digging, cattle bathing and laundering of clothes
8	VIII, Mixed discharge at Mbd-Dlh bridge	Nil
9	IX, D/S River water at Moradabad-Delhi bridge	Sand digging, cattle bathing
10	X, U/S River water at Moradabad-Sambhal bridge	Sand digging, cattle bathing and laundering of clothes
11	XI, Mixed discharge at Moradabad-Sambhal bridge	Nil
12	XII, D/S River water at Moradabad-Sambhal bridge	Sand digging, cattle bathing
13	XIII, River water at Seekanderpur-Patti	Receives agricultural run off

The Karl Pearson's correlation coefficient, or r, is used to find the link between two parameters, x and y. It can be calculated as follows:

$$r = \frac{n \sum x y - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (1)$$

here, n = number of data points ; x = values of x–variable ;y = values of y–variable

To evaluate the straight–line by linear regression, following equation of straight line can be used –

$$Y = a x + b \quad (2)$$

here, y = dependent variable ; x = independent variable ; a = slope of line; b = intercept on y–axis

$$a = \frac{n \sum x y - \sum x \sum y}{n\sum x^2 - (\sum x)^2} \quad (3)$$

$$\text{and } b = \bar{y} - a\bar{x} \quad (4)$$

here, $\bar{x}$ = mean of all values of x ; $\bar{y}$ = mean of all values of y

Using the computer program SPSS, a regression analysis was performed to examine the relationship between certain water quality measures.

Results and Discussion

The Gagan River's fifteen physico-chemical water quality characteristics are statistically represented in Table 2 (W.H.O.1984), along with the accompanying W.H.O. standards. In the research area, there is an urgent need for water quality management, as demonstrated by the comparison of estimated values of several parameters with World Health Organization requirements, which revealed contaminated river water for all parameters studied.

Table 2: Data on the quality of water Physico-chemical characteristics and their WHO standards

S.No.	Parameters	Units	Data points	Min	Max	Sum	Mean	Standard Deviation	WHO Standards
1	Conductivity	µS/cm	13	0.398	0.924	8.243	0.63	0.21	0.3
2	pH	-	13	6.81	8.1	97.07	7.46	0.46	7.0-8.5
3	Chloride	ppm	13	13	90	539	41.46	26.50	200
4	Total Hardness	ppm	13	150	230	2364	181.84	27.60	100
5	Magnesium	ppm	13	11	90	671	51.61	25.57	30
6	Dissolved Oxygen	ppm	13	0.1	5.3	36.4	2.8	1.99	5
7	Biological Oxygen Demand	ppm	13	13.5	85	611.5	47.03	31.07	6
8	Chemical Oxygen Demand	ppm	13	60	270	2062	158.61	75.68	10
9	Total Solids	ppm	13	338	924	7902	607.84	205.41	500
10	Total Dissolved Solids	ppm	13	260	670	5284	406.46	124.33	500
11	Total Suspended Solids	ppm	13	78	384	3064	235.69	121.10	-
12	Free CO2	ppm	13	61.6	158.6	1526	117.38	36.12	10

13	Alkalinity	ppm	13	200	416	3616	278.15	88.71	-
14	Phosphate	ppm	13	0.75	7.8	55.5	4.26	2.90	100
15	Sulphate	ppm	13	2	18	106	8.15	6.12	-

Statistical regression analysis of the water quality indicators produced the following regression equations, with conductivity serving as the dependent variable for each of the thirteen river water data points in Moradabad, India.

Regression Equations	r, correlation coefficient	F, variance ratio	S, standard error of estimate	Equations number
Conductivity=-0.453xpH+4.032	r = 0.943	F = 41.676	S = 0.04	5
Conductivity=+0.009xChloride+0.270	r = 0.972	F = 76.495	S = 0.03	6
Conductivity=+0.005xTotal Hardness-0.570	r = 0.992	F = 428	S = 0.01	7
Conductivity=+0.004xMagnesium+0.289	r = 0.891	F = 16.793	S = 0.07	8
Conductivity=-0.058xDissolvedOxygen+0.722	r = 0.939	F = 25.777	S = 0.03	9
Conductivity = +0.005xBiological Oxygen Demand+0.354	r = 0.964	F = 62.610	S = 0.03	10
Conductivity=+0.002xChemicalOxygenDemand+0.295	r = 0.932	F = 23.732	S = 0.04	11
Conductivity=+0.001xTotalSolids-0.087	r = 0.968	F = 68.814	S = 0.04	12
Conductivity=+0.002xTotalDissolvedSolids-0.233	r = 0.942	F = 43.412	S = 0.04	13
Conductivity=+0.002xTotalSuspendedSolids+0.137	r = 0.947	F = 37.880	S = 0.04	14
Conductivity=+0.002xCO_2+0.002	r = 0.926	F = 21.896	S = 0.04	15
Conductivity=+0.001xAlkalinity+0.021	r = 0.968	F = 68.285	S = 0.04	16
Conductivity=+0.043xPhosphate+0.349	r = 0.924	F = 31.913	S = 0.04	17
Conductivity=+0.041xSulphate+0.271	r = 0.827	F = 13.33	S = 0.07	18

Mean values were taken into consideration as characteristic values to verify the variations in order to evaluate the influence of fluctuations on river water quality. The association between variables is ascertained by correlation analysis. It is a helpful tool for examining correlations between various variables and for determining the degree and direction of the interactions between the variables.

A critical and logical analysis of the aforementioned regression equations reveals the following information regarding correlation studies among different physico-chemical parameters when conductivity is taken into account as a dependent variable. Conductivity exhibits a significant correlation with all fourteen water quality parameters: pH, phosphate, sulphate, alkalinity, total hardness, magnesium, alkalinity, total suspended solids concentration of water, biological oxygen demand, dissolved oxygen, chloride, free CO2, and chemical oxygen demand. The regression coefficient, or r, must be equal to or greater than 0.90 or near 0.90, i.e., there is more than 90% association in the data. The degree of association or correlation between two variables, one of which is considered the dependent variable, is measured by this correlation coefficient. Regression variables are more relevant and have a better fit when the regression coefficient value is higher. Significant correlation also requires low standard error of estimate (S) and high variance ratio (F), both of which have high values.

One of the most delicate ecosystems on the planet is that of rivers and streams. Rivers contain intrinsic values that are not dependent on human wishes; these values include environmental, spiritual, artistic, social, cultural, and economic ones. These river values are determined by their health. River ecosystems are a crucial component of natural ecosystems (Qu et al. 2020). They establish a connection between terrestrial and aquatic ecosystems through the exchange of materials, energy, and information. Ecologically significant river environments support and foster a wide variety of flora and fauna, as well as providing a significant cultural and scientific legacy. Healthy rivers provide water for drinking, agriculture, and industry, serve as water purification systems, and are crucial to the biogeochemical cycle (Liu et al. 2021). River health and ecosystems are increasingly being studied in depth for river bank sustainability. The health of the river depends on its ability to maintain vital functions such sediment movement, as well as its ability to recover from harm and support the local biota, including human populations. In general, a river that is in good health may preserve its natural purity (Varol 2020). A variety of environmental benefits provided by healthy rivers meet the needs and goals of society. The monitoring and visualization of water quality through cloud-based data sharing platforms can lead to enhanced environmental management at watershed sizes. Global issues including urbanization, economic growth, and climatic variability are causing environmental difficulties that are rekindling worries about the aquatic ecosystem, especially in the intricate nature-human interaction. Global issues including urbanization, economic growth, and climate variability are to blame for these issues (Kundu et al. 2009). Gaining a deeper understanding of the environmental consequences of various industries is vital to mitigate the associated ecological harms (Singh 2022).

The idea of formally initiating a river health evaluation may have originated in the USA and some other developed European countries. The idea of river health was first established legally by the US Clean Water Act (CWA) in 1972. Most countries in the globe today plan to maintain primary natural resources like rivers through management strategies.

In many countries, managing rivers is contingent upon the completion of river health assessments. The river's condition is crucial. Healthy rivers provide opportunities for energy production, transportation, entertainment, and flood buffers in addition to providing water, fish, and other consumable goods for industry, agriculture, and drinking water. Rivers lose their capacity to provide these priceless goods and services when they get ill. An in-depth assessment of the current state of ecological rivers is necessary to preserve and improve the health of the river. Yon et al 2022 and others suggest that this process should ideally include monitoring and assessment.

- The identification of rivers that pose little risk to health.
- Determine plausible causes for health issues, like sources of pollutants.
- To provide funding for river restoration top priority and oversee good management practices, especially in the most underserved catchments.
- To assess the effectiveness of management initiatives, which may be crucial if a substantial amount of public funds is allocated to improving the health of rivers.
- To make river health reports possible and raise public awareness of the waterway's current state within the government and the community at large.
- To encompass every vital element of the riverine ecosystem.

Conclusions

Rivers are essential to the wellbeing of human cultures and ecosystems. The rivers that traverse the Indian peninsula are impacted by pollution, deforestation, and overuse. The complete health of these rivers must be understood and analyzed for effective management and conservation. This research looks at the socioeconomic implications and uses water quality models to assess the condition of India's rivers. As on the correlation analysis, one can conclude that there is varying degrees of correlation among the readings, offering a fairly accurate assessment of the water quality.

From the above explanation, it is evident that conductivity is an essential physico-chemical characteristic of water quality. A strong and significant relationship exists between conductivity and each of the fourteen studied parameters across all thirteen data points. These parameters include total dissolved solids, total solids, total suspended solids, pH, phosphate, alkalinity, biological oxygen demand, dissolved oxygen, chloride, total hardness, free CO2, and chemical oxygen demand. Additionally, a moderate correlation is observed between conductivity and two other parameters magnesium and sulphate in the water samples examined.

Utilizing this methodology will significantly streamline the process of monitoring water pollution efficiently and cost-effectively, which is crucial for any pollution study aimed at proposing practical solutions for water quality management. The framework can also be adapted to explain the roles of other parameters. This study confidently suggests that water conductivity can be effectively managed to assess the quality of the Gagan River in Moradabad. Furthermore, this approach can be applied to water quality management in other locations as well. In the context of river water, this research represents an important initial step toward achieving equity and sustainability in water resource management.

Implementing targeted preventative measures can significantly enhance water quality. Monitoring river systems is essential to ensure that human activities do not compromise water purity. The use of pesticides that could harm aquatic life along riverbanks is discouraged to foster aquatic growth and maintain clean water. Establishing wastewater treatment facilities is a vital step in reducing the volume of wastewater released into these waterways. Additionally, a robust model for estimating water conductivity based on chemical data has been developed through stepwise regression analysis. This model identifies how various chemical parameters uniquely influence conductivity. Such reliable predictions can aid environmentalists and water resource managers in making informed decisions. Water quality modeling serves as a valuable tool for evaluating the overall health of rivers in India, enhancing our understanding of the impact of human actions and guiding effective management strategies. This research not only contributes to public health and environmental preservation but also promotes the sustainable management of water resources by aligning scientific findings with societal needs.

References

1. APHA 2012. Standard methods for examination of water and wastewater, American Public Health Association, Washington, D.C. 22th Edn.
2. Da Luz, Nelson, Kumpel, Emily, 2020. Evaluating the impact of sampling design on drinking water quality monitoring program outcomes. Water Res. 185.
3. Draper, N.R., and Smith, H., 1966. Applied Regression Analysis. Wiley, New – York.
4. Ighalo, Joshua O., Adeniyi, Adewale George, 2020. A comprehensive review of water quality monitoring and assessment in Nigeria. Chemosphere 260.
5. Kharake, A.C., Raut, V.S., 2021. An assessment of water quality index of Godavari river water in Nashik city, Maharashtra. Appl. Water. Sci. 11, 101.
6. Khatri, N., Tyagi, S., Rawtani, D., et al., 2020. Assessment of river water quality through application of indices: a case study River Sabarmati, Gujarat, India. Sustain. Water Resour. Manag. 6, 101.
7. Khatri, Nitasha, Tyagi, Sanjiv, Rawtani, Deepak, Tharmavaram, Maithri, 2020. Assessment of river water quality through application of indices: a case study River Sabarmati, Gujarat, India. Sustain Water Resour. Manag.
8. Khatri, N., Raval, K., Jha, A.K., 2021. Integrated water quality monitoring of Mahi river using benthic macroinvertebrates and comparison of its biodiversity among various stretches. Appl. Water. Sci. 11, 143.
9. Kumar, M., 2022. GIS based surface elevation assessment of spring water quality of Shivalik Range of Jammu and Kashmir, India. Ground. Sustain. Dev. 19, 100851.
10. Kumar, J., Jana, A.K., Bansal, A. and Garg, R., 2005. Development of correlation Between BOD and COD for refinery waste. Indian J. Env. Prot., 25(5), 405-409.
11. Kundu, S., Thakur, D., Gill, G.S., Tuli, N., Dadwal, V., 2009. Assessment of groundwater quality adjoining the Ghaggar River using statistical analysis. Bull. Ind. Geol. Assoc. 42 (1–2), 73–85.

12. Liu, Wen, Engel, Bernard A., Chen, Weiping, Wei, Wei, Wang, Yu, Feng, Qi, 2021. Quantifying the contributions of structural factors on runoff water quality from green roofs and optimizing assembled combinations using Taguchi method. J. Hydrol. (Amst.) 593.
13. Madhuri, U. Srinivas, T. and Sirresha, K., 2004. A study on ground water quality in commercial area of Visakhapatnam.. Poll. Res., 23(3), 565-568.
14. Merck, E., 1974. The Testing of Water, Federal Republic of Germany, Darmstadt.
15. Mulla, J.G. Farooqui, M.and Zaheer, A., 2007. A correlation and regression equations among water quality parameters. Int. J. Chem. Sci., 5(2), 943-952.
16. Mokarram, Marzieh, Saber, Ali, Sheykhi, Vahideh, 2020. Effects of heavy metal contamination on river water quality due to release of industrial effluents. J. Clean. Prod. 277.
17. Qu, Xiao, Chen, Yushun, Liu, Han, Xia, Wentong, Lu, Ying, Gang, Daniel-Dianchen, Lin, Lian-Shin, 2020. A holistic assessment of water quality condition and spatiotemporal patterns in impounded lakes along the eastern route of China's South-to-North water diversion project. Water Res. 185.
18. Schot, P., Beard, J., Hissink, R., Silberbauer, M., Griffioen, J., 2022. Adapting classical water quality diagrams for ecohydrological and policy applications. J. Hydrol. X 17, 100137.
19. Singh, S., Rai, S., Singh, P., Mishra, V.K., 2022. Real-time water quality monitoring of River Ganga (India) using internet of things. Ecol. Inform. 71, 101770.
20. Snedecor, G.W.and Cohran, W.O., 1967. Statistical Method, The Lowa State University Press, Ames, 6th Edn.
21. Varol, Memet, 2020. Use of water quality index and multivariate statistical methods for the evaluation of water quality of a stream affected by multiple stressors: a case study. Environ. Pollut. 266 (Part 3).
22. Verma, M., Loganathan, V.A., Bhatt, V.K., 2022. Development of entropy and deviation based water quality index: case of river Ganga, India. Ecol. Indic. 143, 109319.
23. Yin, Dameng, Wang, Le, Zhu, Zhenduo, Clark, Susan Spierre, Cao, Ying, Besek, Jordan, Dai, Ning, 2021. Water quality related to Conservation Reserve Program (CRP) and cropland areas: evidence from multi-temporal remote sensing. Int. J. Appl. Earth Observ. Geoinform. 96.
24. W.H.O., 1984. WHO Guidelines for Drinking Water, Geneva, Switzerland. Vol.1.