Research Journal of Recent Sciences ISSN 2277-2502
Vol. **11(1),** 13-19, January **(2022)** *Res. J. Recent Sci.*

Evaluation of spatio-temporal variation in water quality parameters of river ecosystem in tropical climate using multivariate statistical techniques

Suyog Gupta* and Sunil Kumar Gupta
Department of Environmental Science and Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad, Jharkhand, India
suyog.iitdh@gmail.com

Available online at: www.isca.in, www.isca.me
Received 11th September 2021, revised 13th December 2021, accepted 30th December 2021

Abstract

River water quality is an essential issue in several nations due to its utility for drinking, wildlife and industries. Therefore, it is a great need to analyze the river water quality parameters (RWQPs) to describe the spatio-temporal variability of the river ecosystem. Hence, the river water quality status was investigated using multivariate statistical techniques, i.e., analysis of variance (ANOVA), correlation analysis (CA), and principal component analysis (PCA), at different locations on the river Damodar. In this study, river water samples were gathered temporally from fifteen sampling locations to analyze the twelve RWQPs: biological oxygen demand (BOD_5), chlorides (Cl^-), electrical conductivity (EC), fluorides (F^-), iron (Fe), lead (Pb), dissolved oxygen (DO),nitrates (NO_3^-), potential of hydrogen (pH), sulphates (SO_4^{2-}), total coliform (TC), and total dissolved solids (TDS). The present study revealed the significant ($p<0.05$) spatio-temporal variations in RWQPs and distinguished the source of that variation in river ecosystems. The water quality of the Damodar river exposed an increasing trend in pre-monsoon compared to monsoon and post-monsoon periods on account of the dilution effect in the monsoon and its enduring impact in the post-monsoon period.

Keywords: Analysis of variance, Correlation analysis, Principal component analysis, River water quality, spatio-temporal variation.

Introduction

River water quality (RWQ) is an essential issue in several nations because of its utility for drinking, wildlife, irrigation and industries. Fluctuations in RWQ, primarily due to variation in river water quality parameters (RWQPs)[1,2]. Rivers express a highly heterogeneous spatial nature, and numerous researchers have recognized this heterogeneity due to the physio-chemical dynamics of the rivers. In a river basin, an assortment of factors like hydrogeology, meteorology, land use and anthropogenic activities govern the pollution loads. Therefore, nowadays, RWQ assessment has become a prerequisite for river health monitoring and management[1-5]. It needs the intensified efforts to evaluate RWQ due to its spatio-temporal variability[1,4]. It helps make and manage the river water policies and allocate the optimal uses of the river water1. As no single parameter can demonstrate the RWQ significantly, it should be estimated by analyzing a series of physio-chemical parameters of river water[1,5].

Numerous research studies determined and concluded that the multivariate statistical techniques better understand possible variations in RWQ[1,5-11]. In several studies, PCA was used for evaluating the quality of water to identify the water quality (WQ) deteriorating factors[8,11-16]. It was applied to estimate the correlation between elements and decrease the number of WQPs for facilitating RWQ assessment[6,14]. Pujar et al. used the ANOVA on the river Krishna sampling data and found this statistical technique very effective in carrying out WQ analysis[17]. Miyittah et al. exhibited that the ANOVA technique was able for displaying the variation in WQPs among sampling stations[18]. Kothari et al. used the CA to identify the highly correlated and interrelated WQPs which might influence the WQ of the area[19]. Gupta and Gupta also used the CA to determine the interrelationship between the RWQPs of the river Damodar[1].

It is hardly possible to suggest the best management option for the river ecosystem without RWQ monitoring[1,6]. Therefore, it is necessary to monitor and analyze the spatio-temporal variability in WQPs to describe the river ecosystem. Therefore, this study proposed a critical concept of RWQ analysis. The study aimed to identify inter-relationship, source apportionment and spatio-temporal variations in RWQPs employing multivariate statistical techniques.

Methodology

Study area: The Damodar river basin (DRB) area extends from the circumscribes parts of Jharkhand (Ramgarh, Bokaro, Dhanbad) and West Bengal (Asansol, Durgapur, Burdwan and Howrah). It originates from the Chotanagpur plateau of Jharkhand and ends at the river Hooghly after flowing about 541 KM in the eastern part of India. The geology of DRB is mainly characterized by rocks consisting of granites, sandstones of the Gondwana age and the recent alluvial[20]. It typically experiences a dry, hot summer (~40°C) and cold winter (~2°C). The mean annual rainfall of about 1350mm occurs due to the south-western monsoon. Most of the rain (>80%) happens throughout the monsoon season. The direct and indirect discharges from steel plants, thermal power plants, coke oven plants and numerous coal washeries are the primary pollution sources into the river course[1,20]. The comprehensive details of the study area are showed in Figure-1[1, 20].

Sampling and analytical procedures: To accurately reflect the WQ of the Damodar river, fifteen representative water sampling locations were selected. The water samples were gathered in one-litre pre-washed bottles during pre-monsoon (PreM), monsoon (Mon) and post-monsoon (PostM) periods along the river in 2019. The bottles were rinsed three times with river water before filling them and then labelled accordingly. Samples were taken to the laboratory in an ice-box for physio-chemical analysis. River water samples were analyzed for the potential of hydrogen (pH), biochemical oxygen demand (BOD_5), total dissolved solids (TDS), chlorides (Cl^-), sulphates (SO_4^{2-}), nitrates (NO_3^-), electrical conductivity (EC), total coliforms (TC), dissolved oxygen (DO), fluorides (F^-), iron (Fe), lead (Pb). DO, pH, and EC were measured in-situ, whereas further RWQPs were measured in the laboratory according to American Public Health Association guidelines[1,21].

Multivariate statistical techniques: The multivariate statistical techniques, i.e., analysis of variance (ANOVA), principal component analysis (PCA), and correlation analysis (CA) were smeared in this study to determine autocorrelation of RWQPs and spatio-temporal distribution of RWQ[8,14,17,22,23]. One-way ANOVA was performed for estimating the temporal and spatial variability of RWQ. PCA was applied for the identification of the sources or responsible factors for RWQ fluctuations. This derived information on the most significant RWQPs due to spatial and seasonal disparities. Moreover, the normality of RWQ data was smeared using the Shapiro-Wilk test. The non-normal distribution of RWQ data was transformed using the min-max normalization technique. Pearson's correlation was castoff to check the interrelationship among RWQPs. These statistical analyses were conducted using SPSS 25.0 software.

Results and Discussion

Spatial variation in RWQPs: Figure-2 displayed the spatial dispersal of RWQPs, which showed the location wise mean variation of different RWQPs. The outcomes mirrored the alkaline character of the river. The augmentation in pH and NO_3^- was because of industrial effluent and agricultural runoff. A high NO_3^- concentration in river water stimulated the rapid growth of the algae, which deteriorated RWQ. The level of DO was adequate aside from a couple of areas for different physiological conditions and high aeration. TDS and EC were increased in the destinations close to the massive industries. The sites showed a high alliance of BOD_5 and Cl^-, where the stream got the metropolitan wastes. The higher BOD_5 concentration was the indication of the microorganisms because of effluent release to the river. The level of TC was high at a large portion of the sampling locations. The primary benefactors of SO_4^{2-} were mine squanders, mechanical releases, and overflow from rural terrains[1,6,7,20].

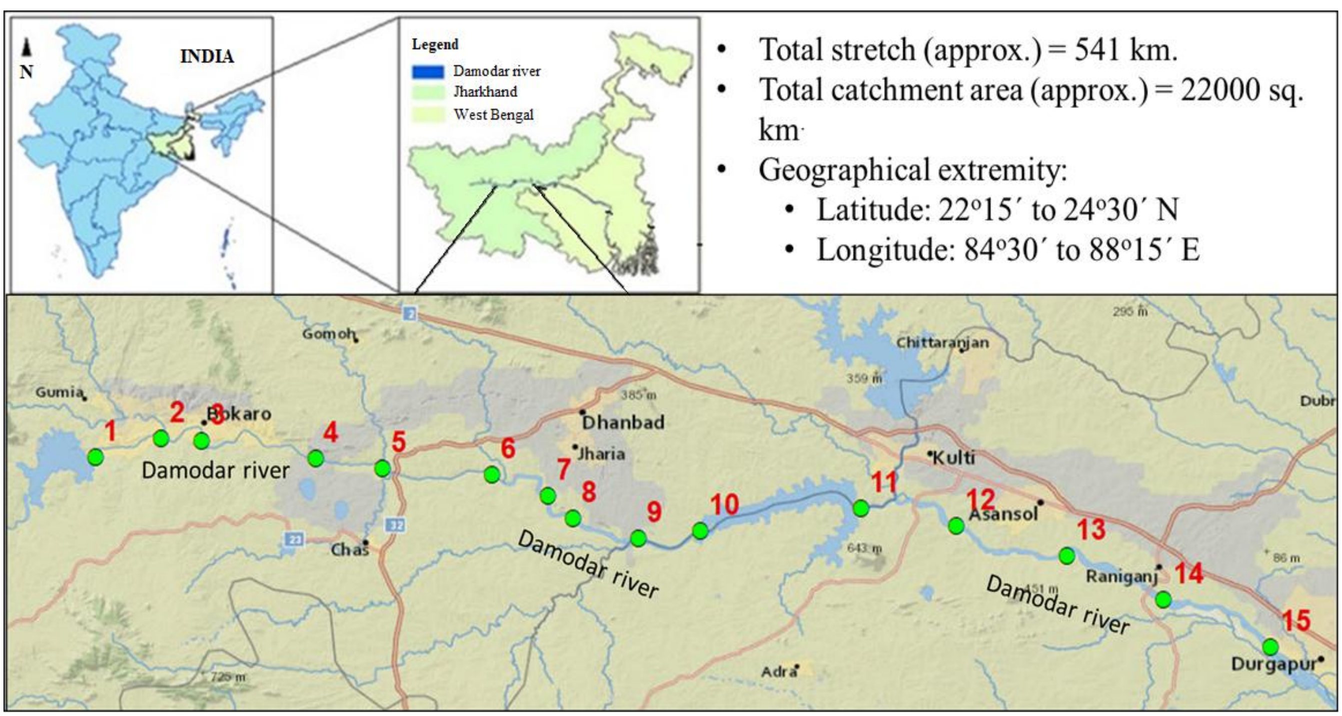
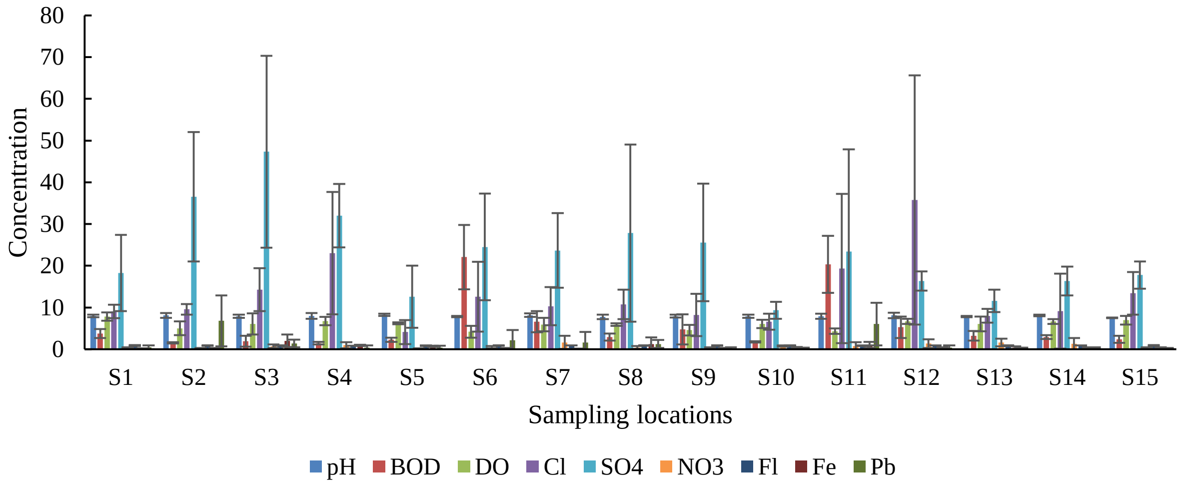


Figure-1: Study area map and sampling locations.



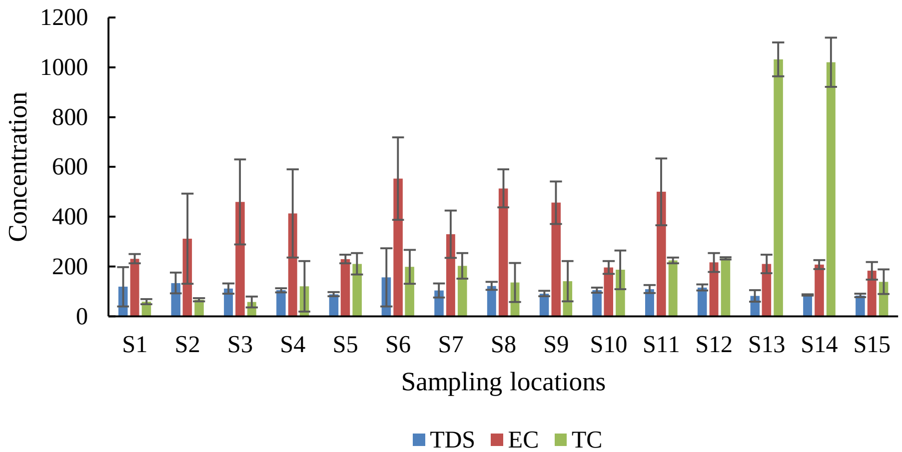


Figure-2: Spatial distribution of RWQPs.

Temporal variation in RWQPs: A one-way ANOVA test was employed to monitor the temporal variances between RWQPs in this study. The outcomes exhibited a significant ($p<0.05$) variation in RWQPs among the three periods: PreM, Mon, and PostM. As shown in Table-1, most parameters were significantly different from PreM to Mon and PostM periods. It indicated that the RWQ was highly affected by the rapid flow of the river during the Mon which altered the concentration of the RWQPs. The concentrations of BOD_5 and DO were quite different between the PreM and PostM periods due to pH, temperature, presence of microorganisms and the organic and inorganic substances dissolved in the river water. There was no significant difference in pH of studied river water as the variation in pH mainly depends upon the geological formation of the area, i.e., type of parent rock present in a specific location and sometimes the dissolved solids present in river water.

Moreover, concentrations of TDS and EC were also significantly different and higher in PreM than in Mon and PostM periods. It might be because of the weathering of parent rocks, climatic conditions, and the extent of anthropogenic activities that increase the chances of adding organic or inorganic substances. Similarly, NO_3^- concentration was significantly different between PreM and PostM periods due to runoff dynamics, seepage from fertilized agricultural or industrial wastewater, refused dumps, urban drainage, and sewage disposal systems. The Cl^- concentrations were also found significantly different between PreM and PostM periods due to the presence of chloride-containing parent rock in the geological formation of the study area, agricultural runoff, and effluent wastewater from different industries or wastewater treatment plants. The SO_4^{2-} concentrations were significantly different between PreM and PostM periods due to municipal or industrial discharges, parent rock containing gypsum and atmospheric deposition. The TC counts were quite different between PreM and PostM periods due to the organic load in the studied river water. The heavy metals such as Fe and Pb concentrations were significantly different between PreM and PostM periods due to the effluents of steel plants with thermal and hydroelectric power stations near the river bank[1,20].

Correlation analysis: In this study, Pearson's CA was conducted to determine the inter-relationship among RWQPs to recognize their possible inducing ability towards the spatio-temporal variation in RWQ[1,22,23]. The Pearson's CA matrix of RWQPs wasdescribed in Figure-3 to measure the association between the two continuous RWQPs. As shown in Figure-3, DO has a moderate negative relationship with BOD_5 and Pb in the PreM period. Similarly, SO_4^{2-} has a moderate negative relationship with NO_3^-. Additionally, BOD_5 has a moderate positive relationship with TDS and EC. Similarly, SO_4^{2-} also has a moderate positive relationship with Fe, while it has a strong positive relationship with EC. Analogous relationship between NO_3^- and TC was also occurred. In Mon, a moderate positive relationship has arisen between DO and F^-, Cl^- and NO_3^-, SO_4^{2-} and EC, and EC and Fe. A similar relationship occurred between BOD_5, EC and Pb, DO and F^-, TDS and Fe, EC and Pb, as well as Fe and Pb in PostM. And a moderate negative relationship has occurred between DO and EC, TDS and TC, F^- and EC. Pb has a robust positive relationship with BOD_5, EC and Fe. A strong negative correlation existed between BOD and F^-. The moderate or strong correlation between the physicochemical parameters showed that these parameters influenced the RWQ. A positive correlation among BOD_5, EC and TDS designated the occurrence of a high concentration of organic and inorganic substances, which might be ascribed to domestic, agricultural and industrial pollution sources.

A negative correlation between BOD_5 and DO signify that the increment in BOD_5 level exhausted the DO level in the river due to untreated municipal wastewater and domestic sewage. The correlation between NO_3^- and TC indicated the occurrence of domestic causes of pollution. This singularity showed that RWQ was emphatically influenced via industrial as well as domestic pollution sources located at DRB[1,20]. Due to the irregular and unsteady flow dynamics in river water during the Mon period, the correlation observed among RWQPs was insufficient to draw valid conclusions regarding their interrelationships with RWQ[1].

Table-1: Results of ANOVA.

RWQPs	Units	Periods		
		PreM	Mon	PostM
pH	-	8.071 ± 0.312^{b}	7.934 ± 0.409^{a}	7.835 ± 0.468^{a}
BOD_5	mg/L	6.860 ± 8.004^{b}	4.882 ± 6.036^{a}	4.887 ± 6.656^{a}
DO	mg/L	5.168 ± 1.265^{b}	5.822 ± 1.168^{b}	6.844 ± 1.450^{a}
TDS	mg/L	133.267 ± 55.609^{a}	91.067 ± 15.471^{b}	99.733 ± 17.982^{b}
Cl^-	mg/L	19.204 ± 17.494^{a}	7.662 ± 3.494^{b}	$11.940 \pm 6.441^{a, b}$
SO_4^{2-}	mg/L	30.799 ± 18.567^{a}	16.469 ± 7.396^{b}	$21.354 \pm 12.444^{a, b}$
NO_3^-	mg/L	0.649 ± 0.881^{a}	0.895 ± 0.700^{a}	0.753 ± 0.793^{a}
EC	µS/cm	415.600 ± 183.180^{a}	281.333 ± 114.009^{b}	$306.333 \pm 143.433^{a, b}$
TC	MPN/100ml	296.667 ± 332.109^{a}	283.733 ± 320.596^{a}	226.133 ± 292.221^{a}
F^-	mg/L	0.866 ± 0.035^{a}	0.887 ± 0.051^{a}	0.896 ± 0.066^{a}
Fe	mg/L	0.659 ± 0.882^{a}	0.371 ± 0.250^{b}	0.651 ± 0.866^{a}
Pb	mg/L	8.466 ± 25.414^{b}	0.740 ± 1.920^{b}	1.123 ± 1.978^{a}
Different alphabets (a, b) represent significant difference among the RWQPs in three periods at $p < 0.05$				

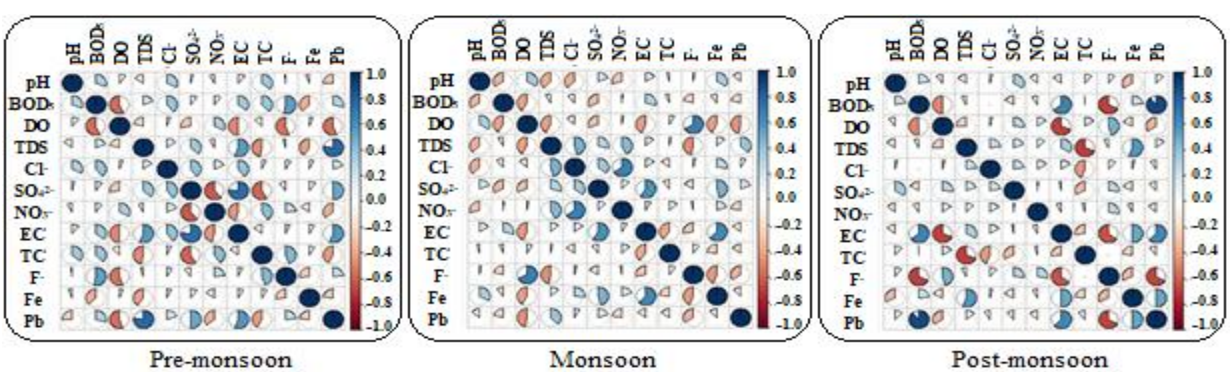


Figure-3: Correlation between RWQPs.

Principal component analysis: In the present study, PCA with the association of Kaiser normalization and varimax rotation was performed by extracting eigenvalues (EV), eigenvectors, percent of variance (%Var) and cumulative percentage (Cum %) for obtaining an accurate estimation of influential parameters responsible for the temporal variation of RWQPs. The scree plots helped to select the principal components (PCs) and comprehend the elementary data structure[10]. The factors with an eigenvalue greater than one were selected as the PCs. Four PCs were extracted with 76.7% and 74.7% of the total data variance for PreM and PostM periods, respectively. Five PCs were extracted with 83.5% of the total data variance for the Mon period (Table-2). According to PCA results, the RWQPs, i.e.,BOD5, DO, NO3-, TC, SO42-, Fe, pH, and Cl in PreM, EC, Fe, F-, Pb, Cl-, NO3-, BOD5, and TC in Mon and BOD5, EC, F-, Pb, TDS, TC, pH, and NO3-in PostM were interpreted as the most influential highly loaded parameters in respective PCs. The different loadings of RWQPs in the different periods may be due to the atmospheric conditions, heavy rainfall, and unsteady river flow, which governed the drastic changes in the river's water quality[1,20,24,25].

Conclusion

This study concluded that the multivariate statistical techniques, i.e., CA, ANOVA and PCA, were the most helpful for determining and describing the inter-relationship and spatio-temporal variation in WQ parameters of the river ecosystem. The spatial distribution of RWQPs displayed the alkaline nature of the river. Moreover, BOD5, Fe and TC were the major contaminants present in the river water. This study also revealed the significant temporal variations in RWQPs and distinguished the source of that variation in river ecosystems. It determined that the Damodar river water quality was ardently influenced by domestic and industrial sources mushroomed at the bank of the river. Overall, this approach can be recommended as an effective tool for sustainable river health monitoring and management.

Acknowledgement

The authors concede the Indian Institute of Technology (Indian School of Mines), Dhanbad, India, to support this study's completion.

Table-2: Results of PCA.

RWQPs	PreM				Mon					PostM			
	PC 1	PC 2	PC 3	PC 4	PC 1	PC 2	PC 3	PC 4	PC 5	PC 1	PC 2	PC 3	PC 4
pH	0.05	0.01	-0.18	0.78	0.49	-0.29	-0.56	0.44	0.11	0.13	-0.14	0.85	-0.09
BOD_5	0.85	0.06	-0.07	0.24	0.01	0.10	-0.05	-0.95	-0.07	0.93	-0.14	0.15	0.07
DO	-0.89	-0.16	-0.12	0.16	-0.31	-0.82	-0.07	0.34	0.08	-0.68	-0.12	0.03	0.05
TDS	0.66	-0.36	-0.17	-0.17	0.04	0.63	0.58	0.13	0.09	0.11	0.88	-0.21	0.00
Cl^-	-0.04	-0.02	0.13	0.76	0.13	0.04	0.89	0.14	-0.21	0.00	0.49	0.25	0.51
SO_4^{2-}	0.27	-0.37	0.82	0.08	0.62	0.24	0.23	0.44	-0.18	-0.24	0.46	0.65	-0.04
NO_3^-	-0.25	0.89	-0.13	-0.07	0.15	-0.12	0.86	-0.07	0.25	-0.14	0.00	-0.15	0.88
EC	0.67	-0.28	0.57	0.12	0.82	0.10	0.07	-0.25	-0.32	0.83	0.32	-0.14	0.04
TC	-0.15	0.93	-0.15	0.07	-0.21	-0.01	0.03	0.05	0.90	-0.08	-0.86	-0.08	-0.12
F^-	0.48	0.67	-0.21	-0.04	-0.37	-0.77	0.07	0.18	-0.15	-0.80	0.05	0.19	0.22
Fe	-0.18	-0.04	0.90	-0.13	0.89	0.14	0.10	0.06	0.00	0.39	0.47	-0.54	-0.19
Pb	0.71	-0.22	0.11	0.03	-0.28	0.77	-0.06	0.37	-0.30	0.87	0.06	0.00	-0.10
EV	3.32	2.54	2.01	1.34	2.48	2.44	2.25	1.68	1.19	3.76	2.40	1.64	1.18
% Var	27.70	21.18	16.72	11.14	20.64	20.31	18.78	13.96	9.88	31.35	19.98	13.63	9.82
Cum %	27.70	48.88	65.60	76.74	20.64	40.95	59.73	73.69	83.57	31.35	51.33	64.96	74.78

References

1. Gupta, S., & Gupta, S. K. (2021). Development and evaluation of an innovative Enhanced River Pollution Index model for holistic monitoring and management of river water quality. *Environmental Science and Pollution Research*, 28(21), 27033-27046.

2. Teixeira de Souza, A., Carneiro, L. A. T., da Silva Junior, O. P., de Carvalho, S. L., & Américo-Pinheiro, J. H. P. (2021). Assessment of water quality using principal component analysis: a case study of the Marrecas stream basin in Brazil. *Environmental technology*, 42(27), 4286-4295.

3. Chen, S. K., Jang, C. S., & Chou, C. Y. (2019). Assessment of spatiotemporal variations in river water quality for sustainable environmental and recreational management in the highly urbanized Danshui River basin. *Environmental monitoring and assessment*, 191(2), 100.

4. Das, S., & Sarkar, R. (2021). Monitoring and evaluating the spatiotemporal variations of the water quality of a stretch of the Bhagirathi-Hugli River, West Bengal, India, using geospatial technology and integrated statistical methods. *Environmental Science and Pollution Research*, 28(13), 15853-15869.

5. Fathi, E., Zamani-Ahmadmahmoodi, R., and Zare-Bidaki, R. (2018). Water quality evaluation using water quality index and multivariate methods, Beheshtabad River, Iran. *Applied Water Science*, 8(7), 1-6.

6. Dimri, D., Daverey, A., Kumar, A., & Sharma, A. (2021). Monitoring water quality of River Ganga using multivariate techniques and WQI (Water Quality Index) in Western Himalayan region of Uttarakhand, India. *Environmental Nanotechnology, Monitoring & Management*, 15, 100375.

7. Mitra, S., Ghosh, S., Satpathy, K. K., Bhattacharya, B. D., Sarkar, S. K., Mishra, P., & Raja, P. (2018). Water quality assessment of the ecologically stressed Hooghly River Estuary, India: a multivariate approach. *Marine pollution bulletin*, 126, 592-599.

8. Jahin, H. S., Abuzaid, A. S., & Abdellatif, A. D. (2020). Using multivariate analysis to develop irrigation water quality index for surface water in Kafr El-Sheikh Governorate, Egypt. *Environmental Technology & Innovation*, 17, 100532.

9. Ling, T. Y., Soo, C. L., Liew, J. J., Nyanti, L., Sim, S. F., & Grinang, J. (2017). Application of multivariate statistical analysis in evaluation of surface river water quality of a tropical river. *Journal of Chemistry*.

10. Pramanik, A. K., Majumdar, D., & Chatterjee, A. (2020). Factors affecting lean, wet-season water quality of Tilaiya reservoir in Koderma District, India during 2013–2017. *Water Science*, 34(1), 85-97.

11. Sun, X., Zhang, H., Zhong, M., Wang, Z., Liang, X., Huang, T., & Huang, H. (2019). Analyses on the temporal and spatial characteristics of water quality in a seagoing river using multivariate statistical techniques: A case study in the Duliujian River, China. *International journal of environmental research and public health*, 16(6), 1020.

12. Sahoo, M. M., Patra, K. C., & Khatua, K. K. (2015). Inference of water quality index using ANFIA and PCA. *Aquatic Procedia*, 4, 1099-1106.

13. Maji, K. J., & Chaudhary, R. (2019). Principal component analysis for water quality assessment of the Ganga river in Uttar Pradesh, India. *Water Resources*, 46(5), 789-806.

14. Ustaoğlu, F., & Tepe, Y. (2019). Water quality and sediment contamination assessment of Pazarsuyu Stream, Turkey using multivariate statistical methods and pollution indicators. *International Soil and Water Conservation Research*, 7(1), 47-56.

15. Varol, M. (2020). Spatio-temporal changes in surface water quality and sediment phosphorus content of a large reservoir in Turkey. *Environmental Pollution*, 259, 113860.

16. Rajab, J. M., MatJafri, M. Z., & Lim, H. S. (2013). Combining multiple regression and principal component analysis for accurate predictions for column ozone in Peninsular Malaysia. *Atmospheric Environment*, 71, 36-43.

17. Pujar, P. M., Kenchannavar, H. H., Kulkarni, R. M., & Kulkarni, U. P. (2020). Real-time water quality monitoring through Internet of Things and ANOVA-based analysis: a case study on river Krishna. *Applied Water Science*, 10(1), 1-16.

18. Miyittah, M. K., Tulashie, S. K., Tsyawo, F. W., Sarfo, J. K., & Darko, A. A. (2020). Assessment of surface water quality status of the Aby Lagoon System in the Western Region of Ghana. *Heliyon*, 6(7), e04466.

19. Kothari, V., Vij, S., Sharma, S., & Gupta, N. (2021). Correlation of various water quality parameters and water quality index of districts of Uttarakhand. *Environmental and Sustainability Indicators*, *9*, 100093.

20. Verma, R. K., Murthy, S., Tiwary, R. K., & Verma, S. (2019). Development of simplified WQIs for assessment of spatial and temporal variations of surface water quality in upper Damodar river basin, eastern India. *Applied Water Science*, 9(1), 21.

21. American Public Health Association, American Water Works Association, Water Pollution Control Federation, & Water Environment Federation (2017). *Standard methods for the examination of water and wastewater*, 23. American Public Health Association.

22. Singh, G., Patel, N., Jindal, T., Srivastava, P., & Bhowmik, A. (2020). Assessment of spatial and temporal variations in water quality by the application of multivariate statistical methods in the Kali River, Uttar Pradesh, India. *Environmental Monitoring and Assessment*, 192, 1-26.

23. Haldar, K., Kujawa-Roeleveld, K., Dey, P., Bosu, S., Datta, D. K., & Rijnaarts, H. H. (2020). Spatio-temporal variations in chemical-physical water quality parameters influencing water reuse for irrigated agriculture in tropical urbanized deltas. *Science of the Total Environment*, *708*, 134559.

24. Singh, R. K., Chaturvedi, A., & Kumari, K. (2019). Water-quality assessment of Damodar River and its tributaries and sub tributaries in Dhanbad Coal mining areas of India based on WQI. *Sustainable Water Resources Management*, 5(2), 381-386.

25. Maurya, P., & Kumari, R. (2021). Spatiotemporal variation of the nutrients and heavy metals in mangroves using multivariate statistical analysis, Gulf of Kachchh (India). *Environmental Research*, *195*, 110803.