International Journal of Mathematical, Engineering and Management Sciences
Vol. 10, No. 4, 1013-1031, 2025
https://doi.org/10.33889/IJMEMS.2025.10.4.049

Ram Arti Publishers

Multivariate Statistical Analysis of Water Quality of Major Rivers in Pune: A Case Study

Pooja Singh
Symbiosis Centre for Waste Resource Management,
Symbiosis International (Deemed University), Pune, Maharashtra, India.
E-mail: pooja.singh@scwrm.siu.edu.in

Javid Gani Dar
Symbiosis Institute of Technology,
Symbiosis International (Deemed University), Pune, Maharashtra, India.
E-mail: Javid.dar@sitpune.edu.in

Manikprabhu Dhanorkar
Symbiosis Centre for Waste Resource Management,
Symbiosis International (Deemed University), Pune, Maharashtra, India.
E-mail: head@scwrm.siu.edu.in

Arundhati Warke
Symbiosis Institute of Technology,
Symbiosis International (Deemed University), Pune, Maharashtra, India.
Corresponding author: krinda2003@gmail.com

(Received on January 9, 2024; Revised on July 15, 2024 & October 22, 2024 & January 7, 2025 & February 27, 2025; Accepted on March 3, 2025)

Abstract
Accurate statistical analysis of the complex and extensive water quality data is crucial to assess the severity of pollution and determine the contribution of each parameter to the overall water quality of aquatic bodies at various locations. This paper describes precise multivariate statistical analysis of the water quality parameters, assessed for rivers Mula and Mutha traversing Pune city, Maharashtra state, India. It examines long-term trends in water quality, and variations before and after the Coronavirus-19 (COVID-19) pandemic. Data from the Maharashtra Stata Pollution Control Board was analysed for eight locations and revealed that Khadakwasla dam had the cleanest water, with a mean highest Dissolved oxygen (DO) of 6.2 mg/l in 2022. At the same time, the Deccan Bridge in the heart of the city was the most polluted stretch of the river, with biochemical oxygen demand (BOD) of 12.23 mg/l and mean faecal coliform load of 25.67 MPN/100 ml in 2022. Principal component 1 (PC1) in 2018 (before COVID-19) and both PC1 and PC2 in 2022 (after COVID-19) covered comparatively maximum variance. High correlations were observed between DO and pH and between chemical oxygen demand (COD) and BOD. Seasonal assessment of water quality at multiple locations can reveal a clear picture of the state of Pune's rivers. This work is important to provide baseline data on the significance of statistical water quality monitoring, and aid in devising suitable interventions to maintain water quality and public health system in the rapidly expanding Pune city.

Keywords- Correlation analysis, Multivariate analysis, Principal component analysis, River pollution, Water quality assessment.

1. Introduction

Several developing, as well as developed countries, across the globe, are battling with severe water crises in recent times. This crisis is further exacerbated by the increasing demand for freshwater as well as the spatial and temporal variations in the hydrological cycles due to increasing climate change (Putro et al., 2016). Water quality deterioration due to increasing water pollution further deepens the global concern about water availability and its quality. Surface run-offs and point discharge sources, both are responsible for this deterioration and add to the water woes of a large number of the population. Contaminated water bodies pose a health risk to all the population dependent on the same, as well as detrimental to the native flora and fauna. It is hence imperative to prevent and control water bodies' pollution, and at the same time continuously monitor and maintain their health. There are both spatial and temporal variations in the water quality of various freshwater reserves, including rivers. Hence systematic water quality monitoring and accurate evaluation is the need of the hour (Yasin et al., 2020). Extensive water quality monitoring programs generate huge amounts of data that need skilled interpretation. In the absence of any scientific statistical analysis and interpretation, the huge and complex data matrix generated upon water analysis is much more complicated to analyse and infer appropriate conclusions. All changes in water quality in water systems should be reflected in the qualitative variables that are assessed at monitoring stations.

Water Quality Indices (WQI) have been one of the most widely used and appropriate ways to express water quality, where a complex set of data is transformed into a single value that is indicative of water quality and spatiotemporal variations (Ji et al., 2016). Analysis using WQI can reveal both seasonal and interannual changes, as was observed during studies on water quality in Lake Poyang, China (Wu et al., 2017), Kolong River, Assam, India (Bora and Goswami, 2017), Tigris River, Iraq (Ewaid et al., 2020), various rivers in Turkey (Aydin et al., 2021; Ustaoglu et al., 2020) and many others. Various indices have been developed and WQI has been modified over the years to make it more flexible and universally accepted. An integrated water quality index (IWQI) was developed to test the suitability of drinking water in an industrial area of Solapur city, Maharashtra state, India (Mukate et al., 2019). A refined WQI model was applied for the assessment of the water quality of Cork Harbour, Ireland. The model used a machine learning algorithm, XGBoost, to rank and select water quality indicators for inclusion based on relative importance to overall water quality status. A seasonal variation in water quality was accurately identified with low ambiguity. However, the study was only of one-year data, and the application of the model to study variation over multiple years was not tested (Uddin et al., 2022).

Apart from WQI, data-driven Multivariate Analysis (MVA) has emerged as a widely used tool, mainly with Principal Components Analysis (PCA) and Cluster Analysis, that aids in achieving a better understanding of the spatial and temporal dynamics of water quality and is useful for water quality analysis, monitoring, and assessment (de Andrade Costa et al., 2020). Since MVA takes into consideration more than one factor of independent variables, the variability of dependent variables is deeply assessed, and hence the conclusion drawn is more authentic, accurate, and realistic. The temporal patterns, seasonal variations, and long-term trends in water quality parameters can be deciphered more accurately by MVA, and hence increasing number of statistical studies are using MVA for data analysis and interpretation (Fatima et al., 2022; Malsy et al., 2017; Njugana et al., 2020; Patil et al., 2020). Multivariate techniques such as PCA, discriminant analysis (DA), and multiple linear regression analysis (MLRA) were done on the Tigris River, Baghdad, by using twenty-five water quality parameters. PCA helped identify pollution source inputs, DA reduced the data and identified significant parameters while MLRA identified the most significant variable that affected the critical parameter (Abed et al., 2019). In another study on one of the largest hydrographic basins in the European Union, WQI, PCA, and response surface method (RSM) were used to analyse 18 physiochemical parameters of the Danube River water. Indices were correlated and the interdependencies obtained between the selected parameters were then used to determine the potential sources of pollution (Iticescu et al., 2019). The two-way multivariate analysis method, called HJ-Biplot, was used to find the variable relationships of the parameters studied from the Gamboa and Paraiso regions, located in the Gatun Reservoir, part of the Panama Canal Watershed. The analysis aided in the proper interpretation and quantification of the hydrological balance through various determining components such as rain, evaporation, discharge, runoff, and/or leakage (Carrasco et al., 2019).

Pune region in the Maharashtra state of western India is an economic and academic hub and one of the fastest developing and expanding cities in Asia. Hundreds of large, medium, and small industries, including automotive, chemical, forging, food, electronic & information technology, have their centres in this region, making Pune a hotspot for a large number of migrant populations from across India (Gohain et al., 2021). Pune is traversed by two major rivers, Mula and Mutha. These rivers serve as a major source of water for the population of the city for various purposes, and hence the quality of surface waters has a direct impact on human health. Extensive urbanization, large-scale migration, and uncontrolled release of sewage and other wastewater at multiple points along the course of the rivers have resulted in extreme degradation of river water quality (Dhananjay et al., 2021). Regular scientific monitoring and defined interventions for improving the health of rivers are lacking. Although there are few reports on the study of river water quality of the rivers flowing through Pune, there is no work reported to date on the extensive statistical analysis of water quality over multiple months and for more than one year (Jain et al., 2022; Rane et al., 2020). Although water quality assessment is regularly being done for the rivers in the Pune region by the Maharashtra State Pollution Control Board (MPCB), no clear and accurate analysis and study on the shift in water quality over the years has been undertaken. Hence exact identification of point sources of pollution, most susceptible river stretches, and prediction for relevant interventions for the same is lacking. This work focuses on the accurate multivariate statistical analysis of the water quality parameters, assessed for the selected water bodies in and around Pune district, Maharashtra state.

The study aims to undertake extensive water quality analysis using MVA and assess the shift in water quality and changes in significant determinants, especially before and after the COVID-19 pandemic. Monthly data from different sections (8 in number) for 6 parameters across the river path has been considered to ascertain the most sensitive and polluted sections. Correlation analysis (CA) used in the study evaluates interrelationships between different parameters that further aid in scientifically identifying and evaluating the primary and critical determinants of water quality. This helps to discover the primary sources of pollution and specifically predict river stretches that need critical attention and monitoring. This plays a crucial role in establishing more stringent monitoring protocols and ensuring the quality and safety of water resources. The analysis will be helpful to environmental regulators and policymakers to accurately identify points of intervention and optimize the treatment processes to minimize and prevent pollution of prominent Pune rivers in the future. There are currently no available reports on extensive statistical analysis of water quality parameters of rivers in Pune, that can enable accurate predictions for interventions. The correlation between water quality parameters and geographical locations of Pune has not been statistically established earlier.

2. Methodology

2.1 Study Area

Pune lies on the western side of the Deccan plateau in India on the leeward side of the Sahyadri mountain range (**Figure 1**).

It lies between 18 degrees 32" North latitude and 73 degrees 51" East longitude and is located 560 m above sea level covering an area of around 700 square kilometres. The region is home to many rivers with three rivers Mula, Mutha, and Pawana mainly traversing across the twin cities of Pune and Pimpri-Chinchwad. These rivers originate in the Sahyadri range and traverse across the Pune city. This study is limited to the Mula and Mutha rivers that travel across Pune city. Mutha River originates in the Western Ghats around 45 km west of Pune, from the confluence of two smaller rivers, Ambi and Moshi, and travels east through Pune. Varsgaon dam on Moshi and Panshet dam on Ambi River feed the waters of the Mutha River. Mutha River enters Khadakwasla dam, which also receives water from Temghar dam, before traveling further towards Pune. Mula River originates from the Mulshi dam and travels East meeting the Mutha River inside

Pune city. Mula-Mutha River thus formed, travels further East before meeting the Bheema River. Bheema River further joins the Krishna River to drain into the Bay of Bengal (**Figure 2**).

Figure 1. Geographical location of Pune.
(http://www.maphill.com/india/maharashtra/pune/location-maps/political-map/).

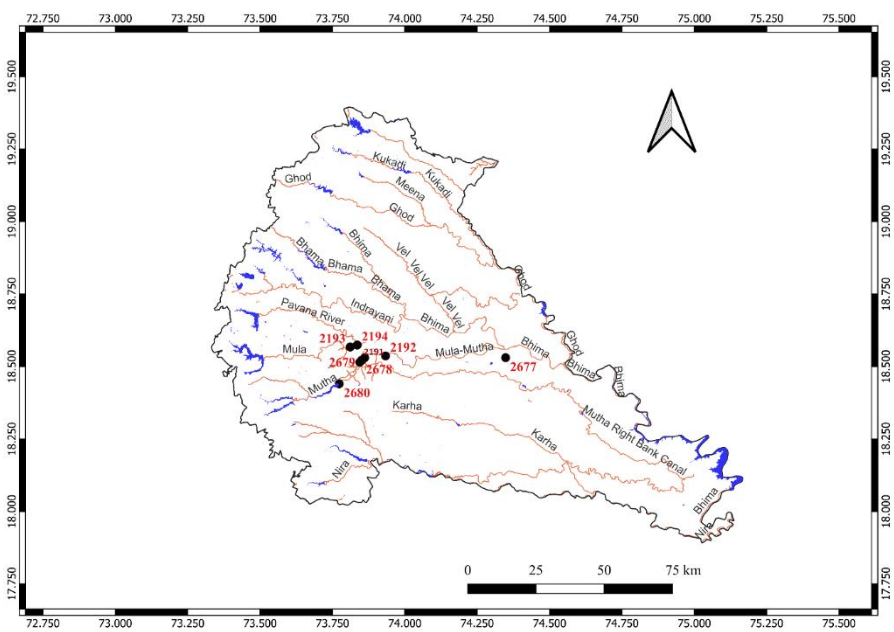

Figure 2. Geological map of the study area with the rivers under study and distribution of sampling locations.

Within the Pune Municipal Corporation limits, Mula River covers 22.2 km, 10.4 km is covered by Mutha River and 11.8 km is traversed by Mula Mutha River. These rivers cover a total area of 44 km and impact a total population of around 5 million people (https://www.pmc.gov.in/en/riverfront). As per the last census conducted in 2011, the population of Pune city was 3,124,458 which is estimated to have expanded to over 5 million as per the calculated growth rate (Dhananjay et al., 2021; Population Census, 2011).

2.2 Data Source and Data Set used in the Study

Mula and Mutha rivers were targeted for this study and the data used was obtained from the regular monitoring program of the Maharashtra Pollution Control Board (MPCB), Government of Maharashtra, India (https://www.mpcb.gov.in/water-quality/pune/17). Eight different sampling points were identified along the three rivers, as per the sites specified by MPCB as shown in **Figure 2** and **Table 1**.

Table 1. Locations along the rivers selected for the current study.

Location on map	Sample code	River	Location
2193	S1	Mula River	Aundh bridge
2194	S2	Mula River	Harrison bridge
2680	S3	Mutha River	Khadakwasla dam
2679	S4	Mutha River	Deccan bridge
2678	S5	Mutha River	Veer Sawarkar bhawan
2191	S6	Mutha River	Sangam Bridge
2192	S7	Mula-Mutha River	Mundwa Bridge
2677	S8	Mula-Mutha River	Theur Gaon

Total 6 parameters were monitored regularly at these points along the path of the river through the city by the local authority, Pune Municipal Corporation, and hence these parameters were selected for the current study, including pH, dissolved oxygen (DO) (mg/l), biological oxygen demand (BOD) (mg/ml), chemical oxygen demand (COD) (mg/ml), nitrate (mg/ml), and total faecal coliforms (MPN /100 ml). The water quality Index (WQI) as per MPCB data was also studied and water quality classification was done using the criteria specified in **Table 2** (Adelagun et al., 2021; US-EPA, 2023).

Table 2. Water quality index classification according to MPCB (https://www.mpcb.gov.in/water-quality/pune/2017).

WQI range	Quality classification	Pollution status
63-100	Good to excellent	Non polluted
50-63	Medium to good	Non polluted
38-50	Bad	Polluted
38 and less	Bad to very bad	Heavily polluted

2.3 Multivariate Analysis

Comparison of the physicochemical parameters between the different sampling locations across Pune city was done by using multivariate statistical techniques. Correlation Analysis (CA) and Principal Component Analysis (PCA) were performed using R software. The interrelationships among different parameters were evaluated using Karl Pearson’s Correlation Analysis coefficient. The strength of the reputed linear association between variables is indicated by the correlation coefficient that varies between -1 to +1, with zero indicating no linear relationship. A positive correlation indicates a direct relationship between the parameters. A negative correlation, on the other hand, indicates that the parameters are inversely related (Mukaka et al., 2012). **Figure 3** describes the overall methodology adopted in the current study.

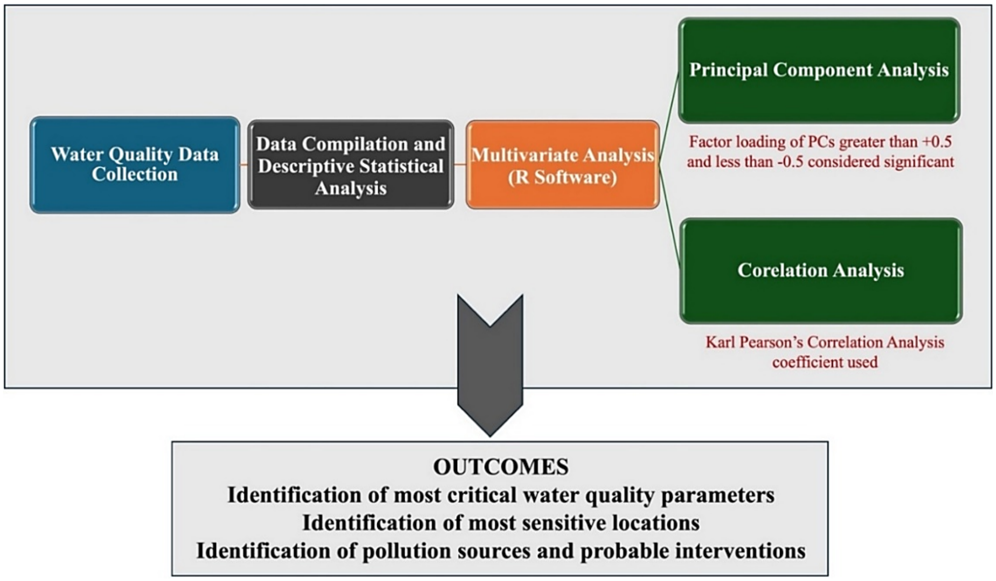

Figure 3. Work plan followed for the statistical analysis of river water quality.

3. Results and Discussion

A significant amount of industrial and domestic waste from various residential and commercial establishments flows into the Mula and Mutha rivers across their progression through Pune city. This substantially undermines the water quality of these major rivers, which are a primary source of water for the city's population serving various domestic, agricultural, and industrial needs (Rane et al., 2020).

3.1 Water Quality Status (Descriptive Statistics)

Statistical analysis using mean, quartiles, variance, and standard deviation of water quality data obtained from the MPCB data source for two years was done across 8 different locations to deduce any changes in the water pollution levels post-COVID pandemic and to identify critical locations for pollution control and remediation interventions (**Table 3**).

pH of two river waters at all locations in 2018 and 2022 was fairly constant and within the recommended levels of 6.5-8.5. This was similar to an earlier study on the Mula-Mutha rivers of Pune wherein no seasonal impact on the pH of river water was observed (Rane et al., 2020). pH of surface water is a chemically and biologically important parameter for healthy aquatic life and other dependent human and animal populations. It plays a significant role in the ability of aquatic organisms to regulate basic metabolic processes, including respiratory processes, thereby playing a significant role in their life cycle. DO of the water body plays a significant role in aesthetics and ecosystem functioning and is an indicator of ecological health. Generally, DO above 3 mg/l is considered to be optimum for good aquatic life while DO less than 1 mg/l is considered as a hypoxic condition and detrimental to aquatic plant and animal life (US-EPA, 2000). According to WHO specifications, water with a minimum of 7 mg/l DO is classified as potable water (Rane et al., 2020). DO across all the locations under study was found to be low as per permissible limits and found to vary between 1 and 6.8 mg/l across 2018 with the highest average value (6.508 mg/l) observed at S3 and lowest (3.2 mg/l) at S4. In the year 2022, DO varies from 3.8 to 6.8 mg/l across all the locations. DO as low as 1 mg/l was seen downstream of the Mula-Mutha River indicating an influx of heavy organic load in the water.

Table 3. Analysis of data collected for 8 locations across the three rivers for the years 2018 and 2022 (compared to the standard specified by WHO (2017), US-EPA (2018), and CPCB (2017) for freshwater intended for urban use).

Location		Parameters											
		pH		DO (mg/l)		BOD (mg/l)		COD (mg/l)		Nitrate (mg/l)		Faecal coliform (MPN/100ml)	
		2018	2022	2018	2022	2018	2022	2018	2022	2018	2022	2018	2022
S1	Min	6.9	7.2	1.7	4.2	5.8	4.4	24	19	0.05	0.9	125	9
	Max	8.6	8.5	5.6	5.7	8.66	15	52	48.4	7.2	9.8	900	35
	Mean	7.87	7.64	3.88	4.83	14.5	8.65	32.67	29.67	2.74	3.13	290.8	13.83
	Q1	7.63	7.4	3.23	4.58	6.4	7.35	28	23	1.38	1.27	198.8	10
	Q3	8.13	7.7	5.03	4.95	9.58	9.35	37	33.5	3.95	4.22	256.2	12
	Var	0.15	0.14	1.85	0.22	7.19	8.30	75.15	92.67	4.88	8.37	39753.79	62.87
	SD	0.39	0.38	1.36	0.46	2.68	2.88	8.66	9.63	2.20	2.89	199.38	7.92
S2	Min	7	7.1	1.4	4.1	7.5	6.8	24	17.8	0.05	0.77	200	11
	Max	8.4	8.4	4.8	5.1	15.5	16.9	48	56.1	7.4	9.85	1800	45
	Mean	7.798	7.54	3.3	4.61	11.42	9.74	39.67	34.83	2.578	3.24	383.3	21.58
	Q1	7.5	7.3	2.28	4.43	9.875	8.175	32	27.82	1.58	1.073	243.8	14.75
	Q3	8.125	7.63	4.45	4.83	13.5	10.75	48	41	3.33	4.88	275	22.5
	Var	0.18	0.15	1.68	0.09	5.76	8.23	68.24	117.37	3.85	7.77	200378.8	103.17
	SD	0.43	0.39	1.29	0.31	2.40	2.86	8.26	10.83	1.96	2.78	447.63	10.15
S3	Min	7.2	7.6	6.1	5.8	2	2.2	8	6.4	0.07	0.3	2	4
	Max	8.5	8.4	6.8	6.8	3.4	3.2	20	12	0.8	0.79	14	12
	Mean	7.834	7.91	6.51	6.24	2.44	2.39	10.67	8.6	0.26	0.34	7.5	6.83
	Q1	7.65	7.7	6.35	6.08	2.2	2.2	8	8	0.1	0.3	6	6
	Q3	8.03	8.03	6.73	6.45	2.53	2.45	12	8.33	0.33	0.30	9.25	7.25
	Var	0.13	0.06	0.06	0.09	0.12	0.09	15.51	2.70	0.05	0.01	11.18	4.33
	SD	0.36	0.25	0.25	0.31	0.35	0.30	3.93	1.64	0.23	0.14	3.34	2.08
S4	Min	7	7.2	1.8	4	8	8.5	24	31.8	0	0.61	195	12
	Max	8.28	8.2	4.2	4.8	25.5	17.9	108	63.8	9	6.55	1600	50
	Mean	7.598	7.52	3.2	4.37	14.96	12.23	54	47.06	2	1.94	380.8	25.67
	Q1	7.35	7.28	2.93	4.1	10.62	9.5	32	43	0.5	0.96	250	14.75
	Q3	7.775	7.75	3.5	4.55	19.25	13.75	65	52	2.4	2.03	293.8	31.25
	Var	0.15	0.13	0.58	0.08	35.83	8.35	653.09	95.57	6.62	3.05	149612.9	134.42
	SD	0.39	0.36	0.76	0.29	5.98	2.88	25.55	9.77	2.57	1.74	386.79	11.59
S5	Min	7	7.1	1.4	4	6.8	7.4	24	27.7	0.267	0.45	250	10
	Max	8.36	8.3	4.4	4.433	23	18.4	100	67.9	8.8	5.09	1600	45
	Mean	7.63	7.575	3.208	4.9	13.86	12.02	48.67	45.77	2.22	1.594	387.5	20.25
	Q1	7.35	7.275	2.85	4.275	11.25	9.5	36	42.6	0.65	1.058	250	13.75
	Q3	7.825	7.925	3.65	4.6	16.62	13.12	57	49.6	2.7	1.643	275	25
	Var	0.15	0.16	0.68	0.07	19.82	8.88	406.78	104.50	5.94	1.35	148125	100.38
	SD	0.39	0.40	0.82	0.27	4.65	2.98	20.16	10.22	2.43	1.16	368.48	10.01
S6	Min1	7	7.2	1.7	3.8	8.6	5.6	24	22.1	0.02	0.53	170	14
	Max	8.42	8.4	5	5.4	23.5	16.4	108	56.1	7.6	6.13	900	45
	Mean	7.74	7.43	3.33	4.49	13.6	11.517	50	42.76	2.3605	1.914	269.6	25.58
	Q1	7.68	7.275	2.75	4.275	9.875	9.375	32	39.15	0.8015	1.015	192.5	16.5
	Q3	7.93	7.43	3.98	4.63	17.25	13.75	61	46.83	3.03	1.84	250	31.25
	Var	0.16	0.11	1.05	0.15	24.82	10.81	580.36	104.02	4.28	2.50	40588.45	103.71
	SD	0.40	0.33	1.02	0.39	4.98	3.28	24.09	10.19	2.06	1.58	201.46	10.18
S7	Min	7.2	7.3	1	4.1	8.5	4.6	32	15.8	0.04	0.91	140	14
	Max	8.52	8.4	5.2	5.5	16.5	12.4	56	48	5.2	8.7	1600	35
	Mean	7.835	7.52	3.39	4.63	12.08	9.23	42	34.98	2.41	2.893	332.1	21.5
	Q1	7.58	7.38	2.08	4.38	10	7.8	36	27	1.62	1.12	200	15
	Q3	8.1	7.53	4.5	4.73	13.75	11.43	45	44	3.75	4.94	250	26.25
	Var	0.14	0.08	2.30	0.14	6.35	5.80	50.90	115.92	2.49	6.65	162115.7	49.54
	SD	0.38	0.29	1.51	0.37	2.52	2.40	7.13	10.76	1.58	2.58	402.63	7.03
S8	Min	7.2	7.2	3.2	4.6	3.2	5.2	16	16	0.09	0.99	40	8
	Max	8.6	8.3	5.8	5.4	8.6	8.6	32	32	8.8	8.9	250	25
	Mean	7.912	7.59	4.86	5.09	6.08	7.003	25.33	24.27	2.99	4.12	123.33	15.5
	Q1	7.575	7.3	4.8	4.9	4.875	6.25	23	20.4	1.125	1.755	68.75	10
	Q3	8.335	7.8	5.23	5.33	7.6	7.8	29	28	4.45	5.98	170	21.25
	Var	0.24	0.10	0.56	0.07	3.29	1.3	27.15	25.34	7.73	7.72	3946.97	44.27
	SD	0.49	0.32	0.75	0.27	1.81	1.14	5.21	5.03	2.78	2.77	62.8	6.65
Standard (WHO)		6.5-8.5		7		3		250		45		0	
Standard (US-EPA)		6.5-8.5		-		10		-		10		-	
Standard (CPCB)		6.5-8.5		-		3		250		45		500	

BOD is one of the most prominent criteria for assessing water quality. It provides information on the biodegradable fraction of the organic matter present in the water. Thus, high organic load indicates a high BOD and a lower DO due to reduced oxygen availability (de Andrade Costa et al., 2020). The lowest BOD at S3 was 2 and 2.2 mg/l in 2018 and 2022 respectively while S4 had the highest BOD 25.5 mg/l in 2018 and S5 had the highest BOD 18.4 mg/l in 2022. Average BOD was found to decrease at 7 out of 8 locations in 2022, as compared to BOD in 2018. High BOD and low DO of water at location S4 (at the Deccan bridge location) indicated low water quality in this stretch of the river as compared to location S3 (at the source of the Mutha River; Khadakwasla dam). This is because S3 is at the source of the river with less urban interventions and S4 is a location in the city area and exposed to unsolicited garbage dumping. As water travels through Pune city, several domestic and industrial discharges and other activities in and around the river cause an increase in the organic matter load, leading to low DO and high BOD at S4. An increase in organic load in water favours microbial growth leading to an increase in BOD and a decrease in DO. Low DO can also be a result of excessive water hyacinth growth in this stretch of the river which tends to deplete the oxygen from water. Low DO coupled with high organic load in water leads to increased acidification and a lowering of the pH of river water. Higher BOD levels for the same river water have been reported by other research groups working on the analysis of river water quality (Abhyankar et al., 2020). In another work, 173 mg/l BOD was earlier reported from a river stretch within the Pune city region upon water quality analysis conducted in the year 2019-2020 (Jadhav and Jadhav, 2020). This data is markedly higher than that reported by the Governmental analysis, and hence river water quality needs to be critically and extensively analyzed by independent and neutral analytical entities, so that suitable site-specific interventions can be designed and deployed. COD in the current analysis for 2018 was found to be in accordance with the reported BOD levels and was highest at S4 and S6, the locations where the higher values of BOD were reported. Similar to the trend observed for BOD, COD in 2022 (6.4-67.9 mg/l) for most of the locations were found to be reduced as compared to 2018 (8-108 mg/l). The lowest COD was reported at S3.

From the data set, the lowest DO was seen in the winter months of December and the lowest BOD and COD were observed in August at most of the locations, which coincides with the rainy season in the area. In a study carried out in the cold months of January and February 2022, COD in the range of 11-208 mg/l was reported from various locations along the Mula River, with higher COD observed at locations where river meets the Mutha River in the heart of the city (Ahire et al., 2023). This range is higher than the data available at the MPCB portal of the same river for 2022. However, in this work, the locations selected for the study are different from the locations selected by MPCB for water quality data collection. Earlier reports from over 15 years ago, show higher pollution levels in the Mula River (S1-S2) in comparison to Mutha and Mula-Mutha rivers (S3-S8). COD as high as 348.7 mg/ml was reported from a location along this river during the summer months (Fadtare and Mane, 2007). Higher BOD and COD have also been reported by an earlier study in 2019-2020, where the BOD and COD ranged between 2.6-142 mg/l and 5.6-240 mg/l respectively in Mula-Mutha River (Rane et al., 2020). Reduction in BOD and COD in recent years, as retrieved from MPCB portal, indicates a probable restriction in the inflow of untreated sewage in river waters as it traverses the city.

Apart from sewage discharge, nitrate in water is generally considered as a result of both agricultural and urban uses. Analysis of the reported nitrate levels at the selected locations revealed a shift in nitrate levels from toward the source to sink of the river flow and a seasonal variation as well. March saw the lowest levels of nitrate in river water in 2018 while in 2022 nitrate levels were low all year round with the lowest being observed pre- and post-monsoon months. Nitrate levels in Mula River (locations S1 and S2) were marginally higher in 2022 as compared to 2018, but Mutha River (locations S3 and S4) saw a reduction in nitrate levels, which was similar to the trend observed for BOD and COD. In a previous study in 2017, the nitrate content in the Mula and Mutha Rivers was reported to be higher at the source and lower as the river flows through the city, rising again at the depositional zone of Mula-Mutha River (Dhananjay et al., 2021). The values for nitrate in the current study are higher than that reported in the study conducted in 2017 indicating a rise in nitrate load in the river water. Nitrates in water could be linked to the inflow of wastewater from industrial zones or runoffs from heavily fertilized fields that are located along the banks of the river. More detailed analysis at multiple locations and source-to-sink studies will reveal exact point sources for this influx.

A study on faecal coliform content reveals the microbial load in water and reflects on the disposal of human and animal Faecal matter in river water. As per the MPCB reports, total coliforms decreased from 2018 to 2022 at all the locations under this study (**Table 3**). In 2018, the level of faecal coliform was as high as 1800 MPN/100 ml in Mula River which made the water unfit not just for drinking but also for bathing and other outdoor activities. A similar high microbial load was reported in another source-to-sink study from 2018 to 2019 in Mula-Mutha Rivers (Rane et al., 2020). However, the levels were drastically reduced in 2022 and the number of faecal coliforms was found to lie within the permissible limit specified for water to be used for drinking purposes with simple treatment processes (50 MPN/100ml). Based on the parameters of the year 2018, the month of September, which is the peak of the rainy season, was seen to be the most polluted month with an elevated faecal coliform load as high as 1600 MPN/100 ml. The pre-monsoon months gave the lowest pollution indication, while microbial load post-monsoon was found to be higher as compared to values reported before the monsoon season. As with other parameters, location S3, which lies before the populated areas of the city, had the lowest microbial load as compared to other locations. High microbial load in water can be a potential health risk not just for waterborne diseases but also for the development and spread of antimicrobial resistance genes. Hence locations with highest faecal load, indicating domestic / hospital waste discharge, need to be especially mapped for inflow of sewage or hospital wastewater.

However, a reduction in microbial load in 2022 indicates an improvement in the sanitary conditions along the river and control in point discharges. WQI data available at the MPCB portal indicated S3, with the highest WQI index, to be the cleanest and least polluted stretch of the river in 2018 and 2022. WQI was in the range of 80-90 indicating 'good' water quality (Adelagun et al., 2021). WQI criteria as specified by US-EPA also indicate water with a WQI score from 80-100 as good quality water (US-EPA, 2023). WQI at other locations varied from 30-50 in 2018 to 50-70 in 2022 indicating 'poor' and 'medium' water quality respectively. The overall analysis of the data at all 8 locations for 2 years reveals S4, S5, and S6 along Mutha River to be the most dynamic locations while S3 exhibited an almost monotonic behavior over time, and not much seasonal variation can be observed. Seasonality is in line with the rainfall regime of the region, with a dry period from November to May and a wet period from July to October. Due to the complexity of the parameters and their interdependency on each other, the need for a methodological, reliable, and replicative statistical framework, capable of assessing the health of water bodies in simple numeric terms, is widely accepted and needs to be adopted in the regular water quality monitoring regime.

3.2 Principal Component Analysis

Assessment of water quality using multivariate analysis and pollution indicators is a critical sustainability aspect (Ustaoglu and Tepe, 2019). This will aid in identifying variables and ensuring the safe and healthy availability of fresh water for human and animal needs. The term "Environmetrics" has been used by many researchers to characterize and document freshwater and sediment quality using multivariate analytical tools (Matta et al., 2022). PCA is a multivariate technique of covariance modeling that reduces the dimensionality of originally correlated variables with the lowest possible loss of information. A new set of variables containing orthogonal uncorrelated variables is formed from the data set of correlated variables;

those are linear combinations of original correlated variables known as principal components. Thus, PCA technique reduces the data, helps find common patterns within a large dataset, and also identifies the sources, or factors, which are accountable for variations in the water quality parameters. Accurate identification of effective pollutant factors in the river water quality parameters gives PCA an advantage over other methods (Zeinalzadeh and Rezaei, 2017). PCA is considered a mathematical model, and a statistical technique, used for dimensionality reduction that transforms a dataset into a set of orthogonal (uncorrelated) components. In rapidly changing environments, other mathematical models may become outdated quickly. Analytical studies that focus on current observations are more relevant and accurate.

Using the river water quality data procured from the MPCB website, PCA was carried out to deduce the most significant factor that showed similarity between the different sampling sites in 2015 as indicated in **Table 4.**

Table 4. PCA at different locations for the year 2018.

PCA	S1	S2	S3	S4	S5	S6	S7	S8
PC1	0.998	0.9996	0.5898	0.9959	0.9857	0.9857	0.9997	0.9924
PC2	0.0018	0.0004	0.3976	0.004	0.015	0.014	0.0003	0.0058

Similar analysis was done for the river water quality data for the year 2022 as indicated in **Table 5**.

Table 5. PCA at different locations for the year 2022.

PCA	S1	S2	S3	S4	S5	S6	S7	S8
PC1	0.631	0.544	0.699	0.729	0.625	0.739	0.729	0.582
PC2	0.316	0.417	0.272	0.242	0.294	0.195	0.229	0.314
PC3	0.042	0.028	0.0179	0.017	0.0577	0.0336	0.0333	0.099
PC4	0.009	0.011	0.006	0.011	0.0222	0.029	0.006	0.0063
PC5	0.0008	0.0005	0.004	0.0039	0.001	0.003	0.0005	0.0009
PC6	0.00006	0.0001	0.0003	0.0001	0	0	0.0003	0.0007

The influence of parameters can be described by six PCs obtained from the data of 2018 and 2022 for eight locations (S1-S8), as shown in **Figure 4** and **Figure 5.**

In the year 2018 (**Figure 4**), PC 1 was the most prominent and covered more than 99 % of the variation at all the locations except S3, where PC1 (58.98%) and PC2 (39.76%) accounted for around 97% of the total variations. Analysis of water in the year 2022 revealed PC1 to be most prominent followed by PC2 (**Figure 5**). Only at location S8, out of the total variation, 89.5% was explained by PC1 and PC2, and 9.9% variance was explained by PC3. It indicates that only (PC1 and PC2) accounted for over 95% of the total variation and were the most significant factors.

In the current study, factor loading of PCs greater than + 0.5 and less than **-** 0.5 were considered significant and considered to have a major contribution to the associated factor (de Andrade Costa et al., 2020). This is in line with a previous study on Ganga River water, wherein PCs above 0.5 were considered to have a significant contribution to the associated factor (Tyagi et al., 2020). For the current study, the loadings that compose the first two components (PC1 and PC2) showing maximum cumulative variations are compiled in **Table 6** for the year 2022.

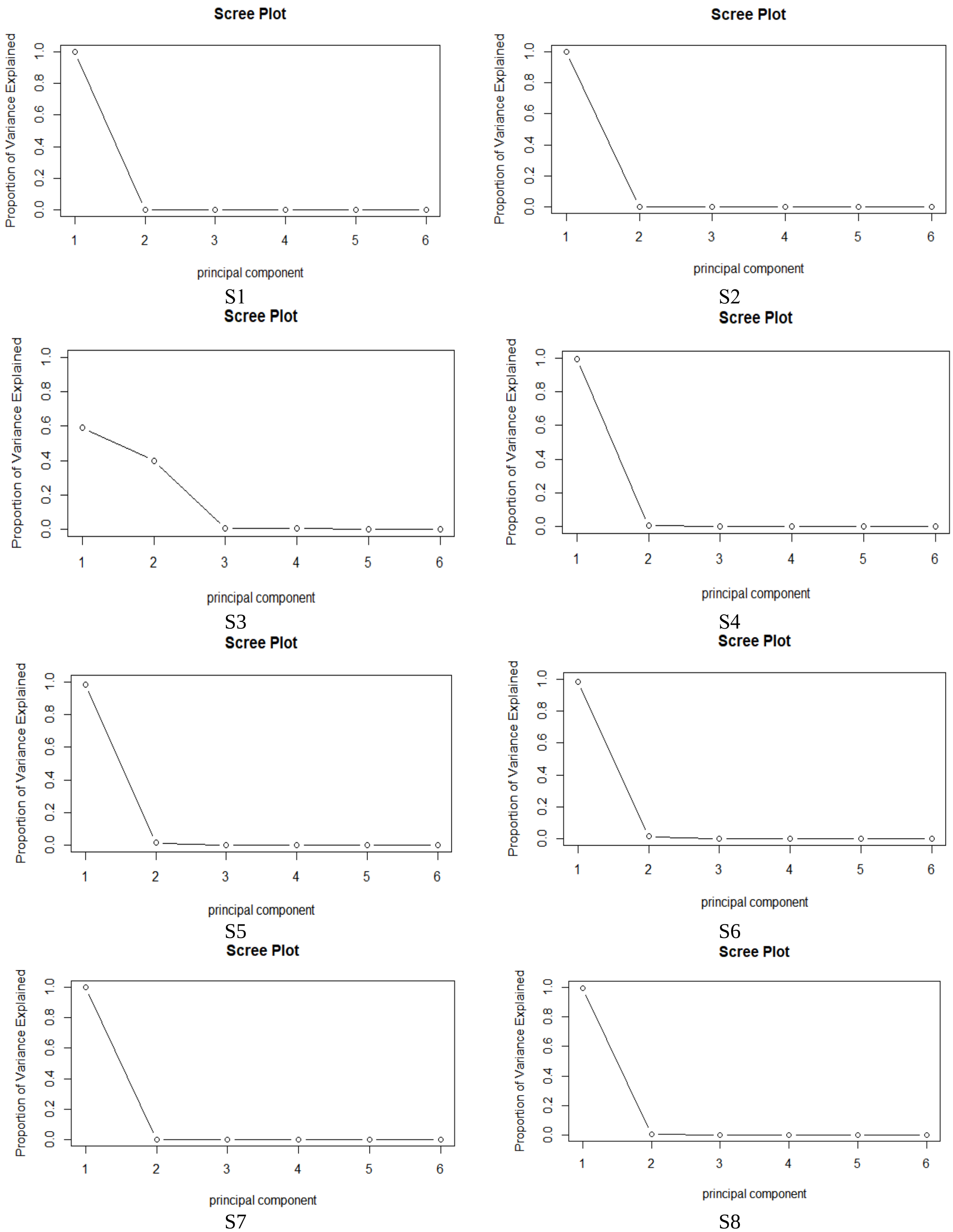

Figure 4. Scree plot for eight locations (S1-S8) showing proportion of variation for 2018.

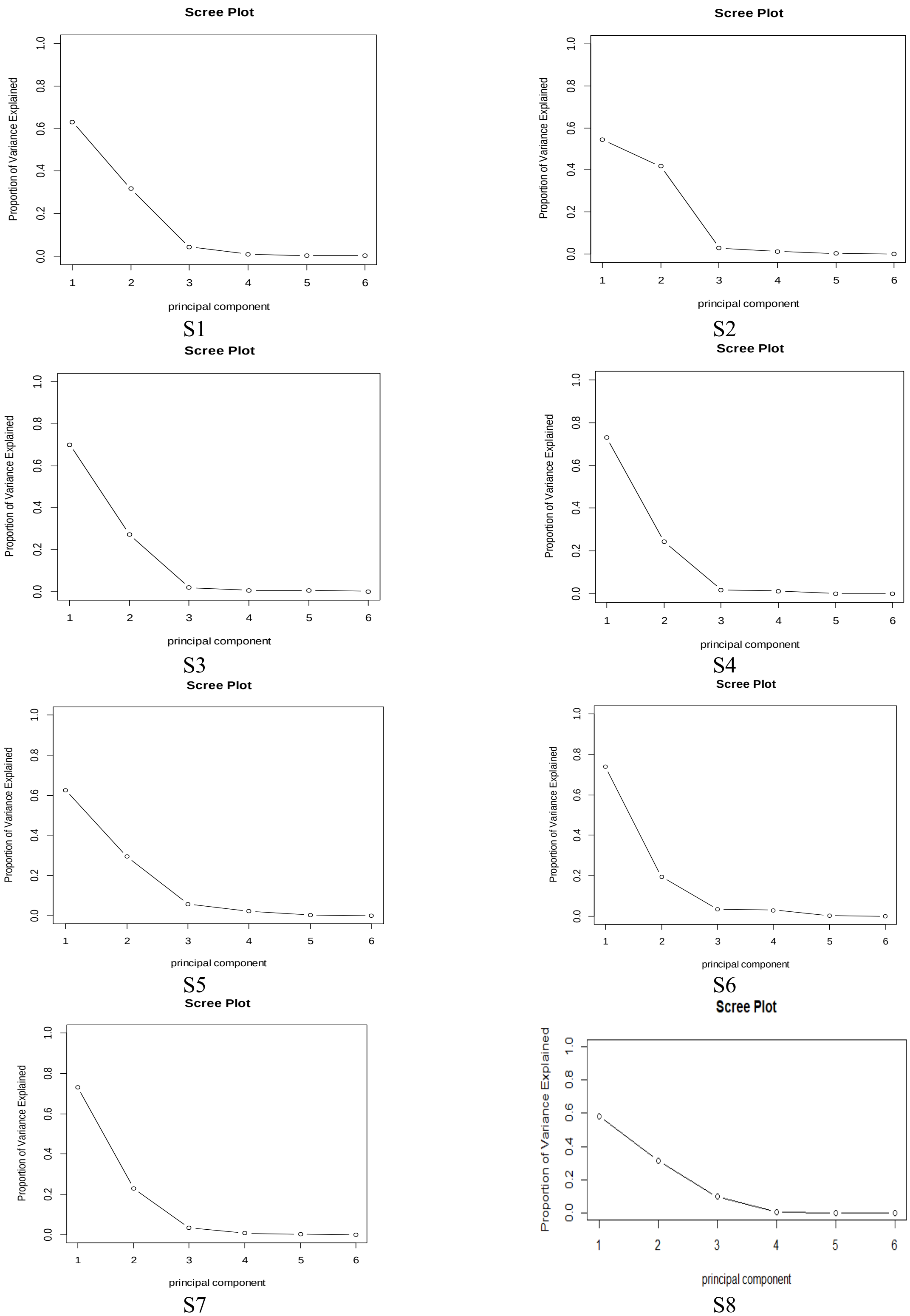

Figure 5. Scree plot for eight locations (S1-S8) showing the proportion of variation for 2022.

Table 6. PCA loading of different parameters for the year 2022.

PCA	pH	DO	BOD	COD	Nitrate	Faecal coliforms
PC1						
S1	0.00078	0.241	-0.238	**-0.879**	0.106	-0.397
S2	0.0029	0.0174	-0.204	**-0.8901**	0.0553	0.403
S3	0.0048	-0.032	-0.0468	**-0.492**	-0.0188	**0.868**
S4	0.01	0.013	-0.085	**-0.586**	0.0065	**-0.805**
S5	0.015	-0.017	0.24	**0.884**	-0.009	-0.4
S6	0.01	0.0222	-0.177	**-0.708**	0.003	**-0.682**
S7	0.0007	0.016	-0.19	**-0.926**	0.011	0.326
S8	0.002	0.01	-0.042	-0.26	-0.059	**-0.962**
PC2						
S1	-0.0144	0.0016	-0.109	-0.389	-0.047	**0.913**
S2	0.0069	0.0156	-0.07	-0.391	0.092	**-0.913**
S3	-0.039	0.0562	-0.119	**-0.858**	-0.0582	**-0.491**
S4	0.0006	0.018	-0.245	**-0.771**	0.009	**0.588**
S5	0.022	0.01	-0.032	-0.405	0.026	**-0.913**
S6	0.009	-0.017	0.201	**0.651**	-0.07	**-0.728**
S7	-0.015	-0.008	-0.01	0.329	-0.145	**0.933**
S8	-0.006	0.014	-0.16	**-0.947**	-0.074	0.269

Maximum loading for PC1 in the year 2022 was on COD and faecal coliforms. At locations S1, S2, S3, S4, S6 and S7, PC1 loaded negatively with COD, while at location S5 PC1 loaded positively with COD. PC1 had significant positive loading on faecal coliforms at location S3 and negative loading on faecal coliform at locations S4, S6, and S8. PC2 had significant positive loading on COD at S6 and on faecal coliform at S1, S4, and S7. Negative loading was observed on COD at S3, S4, and S8 and on coliforms at S2, S3, S5, and S6. However, in 2018 maximum positive loading of faecal coliform at locations S1, S5, and S6, and negative loading of faecal coliform at S2, S4, and S8 was observed. At location S3 negative loading on COD by PC1 and on faecal coliform by PC2 have also been observed (data not shown).

A positive loading indicates that a variable contributes, to some degree, to the principal component, and a negative loading indicates its absence to some degree to the principal component. The larger loading is in relative magnitude, the more important is its presence, or absence, to the principal component. Previous research on multivariate analysis has demonstrated the significance and impact of specific factors influencing river water quality. In a study on the water quality of Terme River, Turkey, four Principal Components were found to be significant. These components highlighted the pivotal role of point and diffuse sources, such as basin rock types, soil erosion, domestic wastewater discharge, and agricultural use of inorganic fertilizers (Ustaoglu et al., 2021). In the current study, COD and faecal coliforms emerge as the prominent parameters at maximum locations studied, indicating their dominance in contributing to quality of river water in and around Pune. Considering the prominence of faecal coliforms in water quality, domestic discharge in river water can be considered to be the main point of concern while monitoring the water quality. Discharge from domestic sources and agricultural runoffs, hence, need to be targeted and specifically identified.

3.3 Correlation Analysis

Correlation Analysis was done using Karl Pearson's Correlation Analysis coefficient. Correlation analysis between the physicochemical water quality parameters exhibited varying interdependencies from strong to weak relationships. **Figure 6** gives the heat maps observed for all locations giving the relationship and interdependency between various parameters.

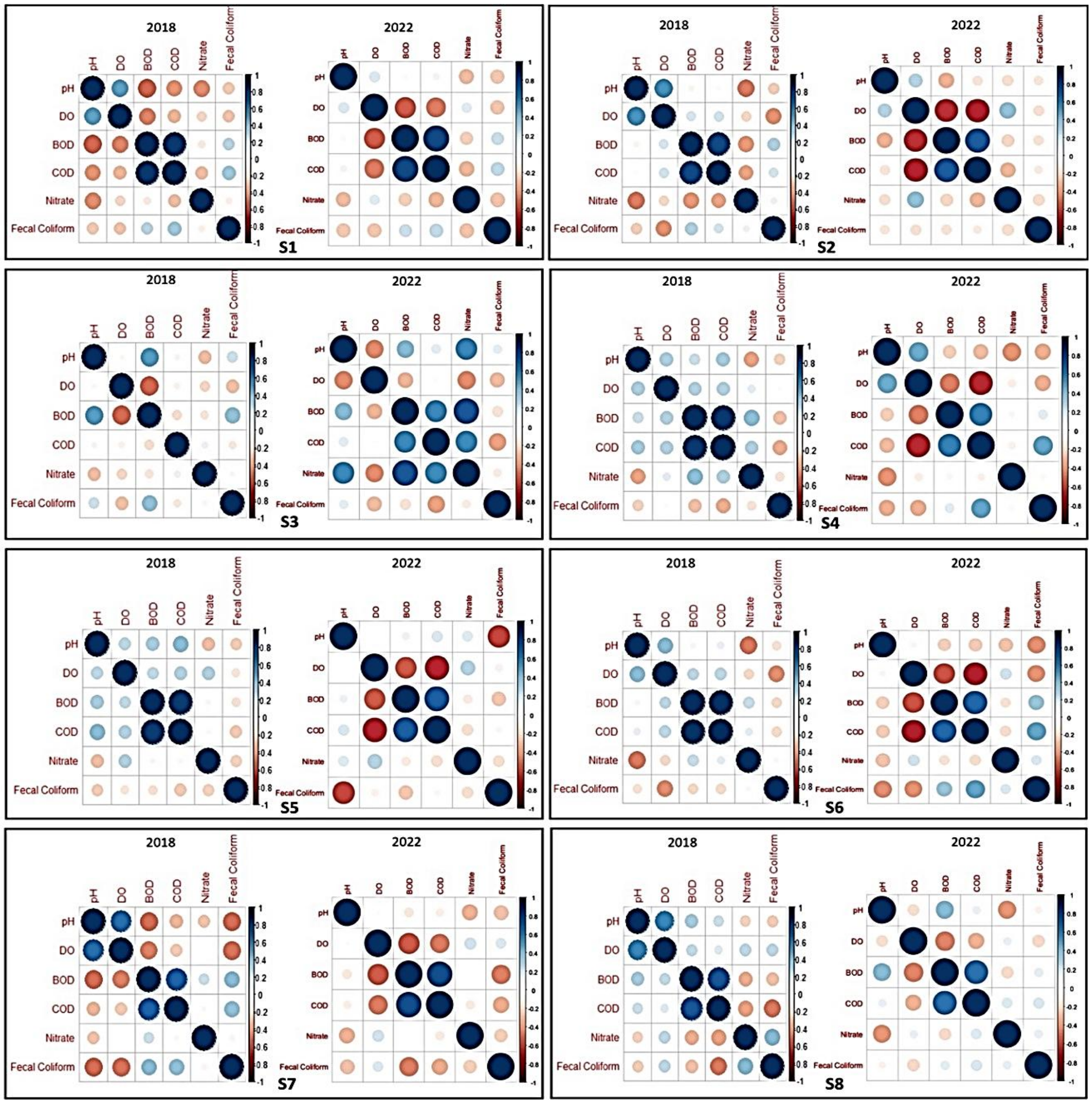

Figure 6. Heat map of two years for each component at all the locations (S1-S8) studied for the rivers traversing Pune city.

The interpretation of the corelation coefficient was done as indicated in **Table 7.**

Table 7. Interpretation of correlation coefficient.

Correlation coefficient (r)	-1	-1< r < 0	0	0 < r < 1	1
Interpretation	Perfect negative	Negative	No correlation	Positive	Perfect positive

The correlation coefficient obtained from the study of parameters exhibited a strong positive correlation between BOD and COD for all locations in both the years, except S3 in 2018. DO was found to be always positively correlated to pH in 2018 except at S3. However, in 2022, a positive correlation in DO and pH can be seen only at S1, S2, and S4. The relationship between pH and DO in water is crucial in understanding aquatic ecosystems. Changes in pH influence the availability of oxygen in water. Low pH level reduces the solubility of oxygen directly affecting aquatic life. In more acidic conditions, the presence of hydrogen ions can react with oxygen molecules, thereby reducing the availability of oxygen in water. The decomposition of excessive organic material by microorganisms consumes oxygen, further leading to a decrease in dissolved oxygen levels in water. At some locations, faecal coliform was positively correlated to BOD indicating the increase in water contamination and inflow of domestic waste in river water. A similar trend has been observed in earlier studies on the Ganga River as well, wherein a strong correlation among the parameters indicated the interdependency of the variables to maintain the quality of the river water (Tyagi et al., 2020). This interdependency may be caused by various geological characteristics of the river basin as well as changes in the water level in the river, directly and/or indirectly affecting the natural structure and quality of the water (Tas et al., 2019). In the current analysis, presence of nitrate in water affected the other parameters (except faecal coliforms) at S3, as evident from analysis for the year 2022, indicating increase in agricultural and animal waste at this location. Hence, in the current study, agricultural runoffs and domestic anthropogenic activities around the rivers in the Pune region seem to be the two significant areas to be targeted for controlling river water quality. This data obtained, hence, is useful to conclude the most significant locations and most significant probable sources of waste that need to be monitored at these locations.

4. Conclusion

This study aimed to evaluate the contribution of water quality parameters, assessed by the Government monitoring station, to temporal variations in surface water quality of rivers in Pune. Further aim was to identify the types and sources of contamination affecting the water quality. The discharge of chemical products from residential and commercial establishments, treatment plant effluents, and dumping of domestic garbage, animal manures, and road dust waste in the river contribute to the organic and inorganic load in water, making the Mula-Mutha River as one of the most polluted rivers stretches in the country. High BOD and low DO at the locations in the city area indicate low water quality as compared to other locations. The influence of parameters is described by six Principal components (PC) for eight locations. In 2018, PC 1 was the most prominent, showing approximately 99% variations at all locations except at S3, where PC1 and PC2 accounted for around 97% of the total variations. Wherein, analysis of water in 2022 revealed that PC1 is most prominent followed by PC2. PC1 and PC2 are the most significant factors contributing 95% of the total variation.

A strong positive correlation has been observed between BOD and COD in both years at all locations. DO and pH showed a positive correlation in 2018 almost at all the locations as compared to only three locations in 2022. A positive correlation between Faecal Coliform and BOD shows an increase in domestic waste discharge as prominent source of contamination of water by faecal organisms. Effect of nitrate on other water quality parameters at Khadakwasla dam area is a point of concern. Increased use of nitrate fertilizers by adjacent farms and rise in the number of tourist residential establishments in and around the dam area can be probable reasons for this variation, and need to be further investigated.

Despite providing an accurate statistical analysis of the water quality of the river, the study faced major limitations due to a lack of extensive water quality data. Critical water quality parameters, including hardness, total dissolved solids, phosphates, and chlorides were not available for analysis. Extensive data collection through actual sample analysis is recommended to overcome these limitations over time. Our study shows inconsistencies in the scattered limited data from independent studies, and the data available on the MPCB portal in the public domain. The locations selected in the Government monitoring program appear inadequate for accurately assessing the health of rivers in Pune, a rapidly expanding city with increasing amounts of waste generated per capita. The actual state of domestic and industrial wastewater discharged in river water is also not documented. The factors responsible for the deterioration of water quality need to be considered for future planning and management of rivers across the Pune region. Since the three rivers traverse the expanse of the Pune region, the sampling locations along the river flow need to be revamped and increased. Regular and extensive water monitoring at new locations needs to be undertaken to investigate and ascertain exact pollution sources to ensure high public health standards. This monitoring should include locations adjacent and downstream from operational sewage treatment plants along the river bank. In addition, it is essential to include other challenging-to-assess locations in the regular analysis regime for more comprehensive monitoring of river water quality from source to sink.

Although there are certain limitations on the data available for this study, PCA and Correlation analysis, and seasonal assessment of water quality at multiple locations, extending beyond those specified by the Government, can reveal a more accurate picture of the state of Pune rivers, and aid in devising suitable interventions for maintenance of water quality and public health.

Abbreviations

BOD: Biochemical oxygen demand.
COD: Chemical oxygen demand.
DO: Dissolved oxygen.
MPCB: Maharashtra Pollution Control Board.
MVA: Multivariate Analysis.
PCA: Principal Component Analysis.

Conflict of Interest

There is no conflict of interest to declare for this publication.

AI Disclosure

The author(s) declare that no assistance is taken from generative AI to write this article.

Acknowledgments

The authors are thankful to Symbiosis International (Deemed University), Pune, India for providing the necessary facilities for carrying out the research. The authors would like to thank the editor and anonymous reviewers for their comments that helped improve the quality of this work.

References

Abhyankar, P.S., Ambade, S.V., & Patwardhan, R.B. (2020). Water pollution analysis of Mutha river across Pune city using BOD as primary parameter. *Pollution Research*, *39*(November Suppl.), S70-S74.

Adelagun, R.O.A., Etim, E.E., & Godwin, O.E. (2021). Application of water quality index for the assessment of water from different sources in Nigeria. *Promising Techniques for Wastewater Treatment and Water Quality Assessment*, *267*, 25. https://doi.org/10.5772/intechopen.98696.

Ahire, S.G., Patil, A.V., Jadhav, A.B., Sawale, D.D., Phalke, D.H., Bote, P.D. & Majik, S.T. (2023). Water pollution intensity analysis of Mula River in Pune using BOD, COD, and heavy metal concentrations as primary parameters. *The Pharma Innovation Journal*, *12*(1), 1527-1530.

Abed, S.A., Ewaid, S.H., & Al-Ansari, N. (2019). Evaluation of water quality in the Tigris River within Baghdad, Iraq using multivariate statistical techniques. *Journal of Physics: Conference Series*, *1294*, 072025. https://doi.org/10.1088/1742-6596/1294/7/072025.

Aydin, H., Ustaoğlu, F., Tepe, Y., & Soylu, E.N. (2021). Assessment of water quality of streams in northeast Turkey by water quality index and multiple statistical methods. *Environmental Forensics*, *22*(1-2), 270-287. https://doi.org/10.1080/15275922.2020.1836074.

Bora, M., & Goswami, D.C. (2017). Water quality assessment in terms of water quality index (WQI): a case study of the Kolong River, Assam, India. *Applied Water Science*, *7*, 3125-3135. https://doi.org/10.1007/s13201-016-0451-y.

Carrasco, G., Molina, J.L., Patino-Alonso, M.C., Castillo, M.D.C., Vicente-Galindo, M.P., & Galindo-Villardón, M.P. (2019). Water quality evaluation through a multivariate statistical HJ-Biplot approach. *Journal of Hydrology*, *577*, 123993. https://doi.org/10.1016/j.jhydrol.2019.123993.

CPCB (2017). Central pollution control board. water quality data. https://cpcb.nic.in/water-quality-criteria-2/. Assessed October 10, 2023.

de Andrade Costa, D., Soares de Azevedo, J.P., Dos Santos, M.A., & dos Santos Facchetti Vinhaes Assumpção, R. (2020). Water quality assessment based on multivariate statistics and water quality index of a strategic river in the Brazilian Atlantic Forest. *Scientific Reports*, *10*(1), 22038. https://doi.org/10.1038/s41598-020-78563-0.

Dhananjay, P.S., Tejal, V., Vilas, M.A., & Ashwini, S. (2021). Investigation of inorganic nutrients available in Mula and Mutha river from origin to Pune City, India. *International Research Journal of Environmental Sciences*, *10*(1), 48-54.

Ewaid, S.H., Abed, S.A., Al-Ansari, N., & Salih, R.M. (2020). Development and evaluation of a water quality index for the Iraqi rivers. *Hydrology*, *7*(3), 67. https://doi.org/10.3390/hydrology7030067.

Fadtare, V.V., & Mane, T.T. (2007). Studies on water pollution of Mula, Mutha, and Pawana rivers in the summer season in the Pune city region. *Nature Environment and Pollution Technology*, *6*(3), 499-506.

Fatima, S.U., Khan, M.A., Alamgir, A., Mahmood, N., & Sulman, N. (2022). Multivariate and spatial methods-based water quality assessment of Chu Tran Valley, Gilgit Baltistan. *Applied Water Science*, *12*(6), 129. https://doi.org/10.1007/s13201-022-01661-z.

Gohain, K.J., Mohammad, P., & Goswami, A. (2021). Assessing the impact of land use land cover changes on land surface temperature over Pune city, India. *Quaternary International*, *575-576*, 259-269. https://doi.org/10.1016/j.quaint.2020.04.052.

Iticescu, C., Georgescu, L.P., Murariu, G., Topa, C., Timofti, M., Pintilie, V., & Arseni, M. (2019). Lower Danube water quality was quantified through WQI and multivariate analysis. *Water*, *11*(6), 1305. https://doi.org/10.3390/w11061305.

Jadhav, S.D., & Jadhav, M.S. (2020). Study of water quality parameters of Mula-Mutha river at Pune, Maharastra (India). *International Journal of Lakes and Rivers*, *13*(1), 95-103.

Jain, N., Yevatikar, R., & Raxamwar, T.S. (2022). Comparative study of physico-chemical parameters and water quality index of river. *Materials Today: Proceedings*, *60*(2), 859-867. https://doi.org/10.1016/j.matpr.2021.09.508.

Ji, X., Dahlgren, R.A. & Zhang, M. (2016). Comparison of seven water quality assessment methods for the characterization and management of highly impaired river systems. *Environment Monitoring and Assessment*, *188*(1), 15. https://doi.org/10.1007/s10661-015-5016-2.

Malsy, M., Flörke, M., & Borchardt, D. (2017). What drives the water quality changes in the Selenga Basin: climate change or socio-economic development? *Regional Environmental Change*, *17*(7), 1977-1989. https://doi.org/10.1007/s10113-016-1005-4.

Matta, G., Kumar, A., Nayak, A., Kumar, P. (2022). Appraisal of spatial-temporal variation and pollution source estimation of Ganga River system through pollution indices and environmetrics in upper Ganga basin. *Applied Water Science*, *12*(3), 33. https://doi.org/10.1007/s13201-021-01552-9.

Mukaka, M.M. (2012). A guide to the appropriate use of correlation coefficient in medical research. *Malawi Medical Journal, 24*(3), 69-71.

Mukate, S., Wagh, V., Panaskar, D., Jacobs, J.A., & Sawant, A. (2019). Development of a new integrated water quality index (IWQI) model to evaluate the drinking suitability of water. *Ecological Indicators*, *101*, 348-354. https://doi.org/10.1016/j.ecolind.2019.01.034.

Njuguna, S.M., Onyango, J.A., Githaiga, K.B., Gituru, R.W., & Yan, X. (2020). Application of multivariate statistical analysis and water quality index in health risk assessment by domestic use of river water. Case study of Tana River in Kenya. *Process Safety and Environmental Protection*, *133*, 149-158. https://doi.org/10.1016/j.psep.2019.11.006.

Patil, V.B.B., Pinto, S.M., Govindaraju, T., Hebbalu, V.S., Bhat, V., & Kannanur, L.N. (2020). Multivariate statistics and water quality index (WQI) approach for geochemical assessment of groundwater quality-a case study of Kanavi Halla Sub-Basin, Belagavi, India. *Environmental Geochemistry and Health*, *42*(9), 2667-2684. https://doi.org/10.1007/s10653-019-00500-6.

Population Census (2011) https://www.census2011.co.in/census/city/375-pune.html#:~:text=Pune%20city%20is%20governed%20by,population%20is%20estimated%20at%206%2C975%2C000%20. Assessed July 5, 2023.

Putro, B., Kjeldsen, T.R., Hutchins, M.G. & Miller, J. (2016). An empirical investigation of climate and land-use effects on water quantity and quality in two urbanizing catchments in the southern United Kingdom. *Science of the Total Environment, 548-549*, 164-172. https://doi.org/10.1016/j.scitotenv.2015.12.132.

Rane, A.R., Supekar, A.H., Karmalkar, N.R., & Raman, S. (2020). Physico-chemical and biological parameters investigation of Mula-Mutha river in India: from source to sink. *Pollution Research, 39*(3), 798-810.

Taş, B., Tepe, Y., Ustaoğlu, F., & Alptekin, S. (2019). Benthic algal diversity and water quality evaluation by the biological approach of Turnasuyu Creek, NE Turkey. *Desalination and Water Treatment*, *155*, 404-417. https://doi.org/10.5004/dwt.2019.24225.

Tyagi, S., Dubey, R.C., Bhutiani, R., & Ahamad, F. (2020). Multivariate statistical analysis of river Ganga water at Rishikesh and Haridwar, India. *Analytical Chemistry Letters*, *10*(2), 195-213. https://doi.org/10.1080/22297928.2020.1756405.

Uddin, M.G., Nash, S., Rahman, A., & Olbert, A.I. (2022). A comprehensive method for improvement of water quality index (WQI) models for coastal water quality assessment. *Water Research*, *219*, 118532. https://doi.org/10.1016/j.watres.2022.118532.

US-EPA (2000). United States environment protection agency. ambient aquatic life water quality criteria for dissolved oxygen (saltwater). Cape Cod to Cape Hatteras. US Environment Protection Agency, Washington D.C.

US-EPA (2018). *Edition of the drinking water standards and health advisories tables*. US Environment Protection Agency, Washington D.C.

US-EPA (2023). https://www.epa.gov/salish-sea/freshwater-quality. Assessed October 6, 2023.

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. https://doi.org/10.1016/j.iswcr.2018.09.001.

Ustaoğlu, F., Taş, B., Tepe, Y., & Topaldemir, H. (2021). Comprehensive assessment of water quality and associated health risk by using physicochemical quality indices and multivariate analysis in Terme River, Turkey. *Environmental Science and Pollution Research*, *28*(44), 62736-62754. https://doi.org/10.1007/s11356-021-15135-3.

Ustaoğlu, F., Tepe, Y., & Taş, B. (2020). Assessment of stream quality and health risk in a subtropical Turkey river system: A combined approach using statistical analysis and water quality index. *Ecological Indicators*, *113*, 105815. https://doi.org/10.1016/j.ecolind.2019.105815.

WHO (2017). *Guidelines for drinking-water quality, 4th ed. incorporating the 1st addendum (chapters).* World Health Organization, Geneva, Switzerland.

Wu, Z., Zhang, D., Cai, Y., Wang, X., Zhang, L., & Chen, Y. (2017). Water quality assessment based on the water quality index method in lake Poyang: the largest freshwater lake in China. *Scientific Reports*, *7*(1), 17999. https://doi.org/10.1038/s41598-017-18285-y.

Yasin, S.N.T.M., Yunus, M.F.M., & Wahab, N.B.A. (2020). The development of a water quality monitoring system using the Internet of Things. *Journal of Educational and Learning Studies, 3*(1), 14-20. https://doi.org/10.32698/0852.

Zeinalzadeh, K., & Rezaei, E. (2017). Determining spatial and temporal changes of surface water quality using principal component analysis. *Journal of Hydrology: Regional Studies, 13*, 1-10. https://doi.org/10.1016/j.ejrh.2017.07.002.