



Surface Water Quality in Punjab, India: Tracking Human and Farm Animal-Specific Adenoviral Contamination and Correlation with Microbiological and Physiochemical Parameters

Asima Zehra · Simranpreet Kaur · Randhir Singh · J. P. S. Gill

Received: 22 May 2020 / Accepted: 2 October 2020 / Published online: 19 October 2020
© Springer Nature Switzerland AG 2020

Abstract The capability of faecal indicator bacteria (FIB) to be environmental in origin and the deficient specificity of markers can mean uncertain outcomes when just a single marker is distinguished. This can give a bogus thought regarding water quality. A water quality analysis and the correlation between FIB, multiple physiochemical parameters and a viral marker in recreational waters were studied. For water concentration, carefully validated virus filtration techniques were used coupled with PCR-based assays for virus quantification. Around 32.0% of analysed surface water samples were positive for human adenovirus (HAdV, qPCR), 16.0% for bovine adenovirus (BAdV, nested PCR) and 11.0% for porcine adenovirus (PAdV, qPCR). For all the types of

the adenoviruses, detection frequencies were not significantly different except for HAdV that showed higher frequency in water samples from tehsil Ludhiana. Only a sporadic significant correlation was observed with the few components; there was no constant, stable correlation between the microbiological and physiochemical parameters. Although both canals and river showed a stable and significant positive correlation between bacteriological parameters (coliform and enterococci) and between biological oxygen demand (BOD) and chemical oxygen demand (COD), regression analysis revealed no significant correlations between concentration of enterococci count to a concentration of AdVs, unlike coliform that showed sporadic but significant correlation with PAdV and no correlation with HAdVs/BAdVs. On principal component analysis, the first factor had strong positive loadings on BOD, COD, total dissolved solids (TDS), turbidity and PAdV and is largely a pollution loading factor which is also supported by their mutual cluster in cluster analysis. Factor 2 had strong loading on enterococci/coliform count due to the natural quality of water and domestic waste and factor 3 had loading on HAdV. This factor shows the impact of human activities overwhelmed in the catchment area of the canals and river. Hence, a significant level of certainty on the contamination source can be acquired if more than one marker is distinguished. Our study thus gives a case of how quantifying both physiochemical, FIB and adenoviral markers in surface water can improve water quality evaluation and help tailor CPCB -India programs for impaired water bodies in the future.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11270-020-04892-5>) contains supplementary material, which is available to authorized users.

A. Zehra (✉) · S. Kaur · R. Singh
School of Public Health and Zoonoses, Guru Angad Dev
Veterinary and Animal Sciences University (GADVASU),
Ludhiana, Punjab 141004, India
e-mail: assimazehra@gmail.com

S. Kaur
e-mail: simranthind18@gmail.com

R. Singh
e-mail: sainirandhir74@gmail.com

J. P. S. Gill
GADVASU, Ludhiana, Punjab 141004, India
e-mail: gilljps@gmail.com

Keywords Microbial source tracking · Coliform · Adenoviruses · Physiochemical parameters · Cluster analysis · Principal component analysis

1 Introduction

Numerous epidemiological studies of waterborne illness in countries like India indicate that the common aetiological agents are bacteria, viruses and parasitic protozoa (Hamner et al. 2007; Patel et al. 2002) and the FIB (coliform and enterococci) presence is an important parameter for determining water pollution levels. But the ubiquitous nature of FIB especially coliform, the majority being environmental contaminant, had raised doubt about their authenticity as an indicator of the pathogen (CPCB 2002). Also, poor relationships have been accounted for between waterborne human viruses or protozoa and FIBs (CPCB 2002; Schriewer et al. 2015). This can give a bogus thought regarding water quality. Hence recently, CPCB used BOD, or the dissolved oxygen required for water bodies to bolster eco-systems, as a reliable parameter for estimation of river quality (CPCB 2011; CPCB 2015). However, they provide judgement in terms of individual parameters and do not provide a complete picture of the scenario (Ali et al. 2014; Rosemond et al. 2009).

Furthermore, many epidemiology studies have failed to find a correlation between human health outcomes and FIB levels, particularly when the pollution is not from a known point source such as a wastewater treatment plant (WWTP) (Dwight et al. 2004; Colford et al. 2007). Also, the determination of the FIBs does not provide information about the pollution source. Nonetheless, water regulatory agencies have yet to come to terms with the inherent problems resulting from reliance on a single parameter as currently determined. Hence, several water quality indices (WQI) have been developed globally to monitor freshwater quality for direct human consumption and other uses (Sun et al. 2016). As there is a variety of chemical, physical and biological water quality parameters, several researchers have proposed other parameters or markers that can be incorporated in the WQI so that not only the level of pollution can be indicated but also the source of pollution can be identified (Zeinalzadeh and Rezaei 2017). Fortunately, new index organisms for some pathogens look promising like *C. perfringens*, phages and viruses like adenoviruses. As these index organisms are relatively untested worldwide especially

in India, extensive trials are necessary before their general acceptance in microbial risk assessment. Also, most of the microbial source tracking (MST)-based studies focused on faecal contamination from human sources and their backtracking. But as the understanding of zoonoses potential expands, it will be worth to involve water contamination from animal sources and its backtracking.

In India, a lot of the studies have been carried out using multiple parameters for constructing water quality index in assessing the water quality (Bora and Goswami 2016; Ponsadailakshmi et al. 2018), and in Punjab, most of the extensive studies were on heavy metal contamination (Setia et al. 2020). But none of the studies could give the conclusive results related to which parameters must be used in the water quality index for assessing policies related to water quality. And the determination of the FIB does not provide information about the pollution source. Hence, the need to develop effective methodologies for evaluation of groundwater and surface water resources for sustainable development and safety of human health arises. In the current study, water pollution levels were therefore monitored in the midstream and at both river-sides as well as in the river Sutlej's most important tributaries. To track the origin of faecal pollution, host-specific genetic faecal markers for different host groups were determined simultaneously at all the sampling sites along the river and the tributaries.

Lastly, the simultaneous detection of human-specific markers (like *B. thetaiotaomicron*, *M. smithii*, HPyVs and HAdVs etc.) with *E. coli* and enterococcus supports the likelihood of recent faecal contamination, since the human-specific markers are unable to regrow in natural surface waters (Liang et al. 2015; Das Kangabam et al. 2017). Multiple-linear-regression results further confirm that the inclusion of HAdVs, together with traditional indicators, would better predict the occurrence of pathogens. Further study is needed to determine the applicability of such models to different geographical locations and environmental conditions (Liang et al. 2015; Farkas et al. 2018).

Thus, the current study incorporated to monitor faecal coliforms, enterococci, HAdVs, BAdVs and PAdVs and other physiochemical properties in surface water bodies. The objectives for this study were (1) quantifying FIB and adenoviral markers in surface water, (2) determining the physiochemical values of the surface water and (3) testing for correlations among physiochemical, FIB and adenoviral markers using the statistical models.

2 Material and Methods

2.1 Surface Water Samples from Canals and River Sutlej

The river Sutlej flows through the Ludhiana district of Punjab, and Sindhwan and Neelon-Ropar canals are tributaries that join river Sutlej. Samples were collected from these canals and Sutlej River (Fig. 1) from February 2017 to January 2018. Around 52 surface water samples (17 water samples from each Sidhwan canal and Neelon-Ropar canal and 18 water samples from river Sutlej) were collected in sterile polypropylene bottles, properly labelled and immediately placed in an insulated box containing ice packs. A 0.2 ml of 3% sodium thiosulfate ($\text{Na}_2\text{S}_2\text{O}_3$) was added to 200-ml water samples kept for bacteriological analysis, to neutralize any chlorine present. The insulated boxes carrying samples were transported to the School of Public Health and Zoonoses lab and stored at a low temperature of 4 °C. The pH and turbidity were tested immediately after sampling as they may change during storage and transport.

2.2 Physiochemical Parameters

The pH, temperature, EC (electrical conductivity), TDS, turbidity and dissolved oxygen (DO), BOD and COD were estimated in this study (Supplementary material).

2.3 Enumeration of Culturable Indicator Organisms

E. coli and *Enterococcus* spp. in surface water samples were enumerated using the Environmental Protection Agency's standard membrane filtration methods (EPA 1997; U.S. EPA 2002). Minimum analytical QC requirement for the analysis of samples using Method 1603 was followed (Supplementary material).

2.4 Virus Concentration Methods and Nucleic Acid Extraction

The following two virus filtration methods were optimized and used in this study:

i Protocol I: direct nucleic acid extraction from negatively charged membranes (Zehra 2018) and were applied to the highly turbid samples (Supplementary material).

j Protocol II: adsorption/elution-based protocol with negatively charged membranes (Zehra 2018) and were applied to the river and canal water samples with the low turbidity (Supplementary material).

2.5 Molecular Assays

Touch-down PCR, semi-nested PCR and the qPCR were first optimized/validated using the known faecal samples and PCR inhibition assay (Sketa-22 Assay) was studied before applying these methods to the environmental samples (Zehra 2018). The details of these assays are given in Supplementary material.

3 Statistical Analysis

The statistical analyses were done using IBM Statistical Package for Statistical Sciences (SPSS V24.0) software, and before, any statistical analysis normality of data collected was tested. A one-way analysis of variance (ANOVA) was performed to evaluate a significant difference in physiochemical and bacteriological parameters between water samples collected from different tehsils of Ludhiana. If a significant difference was acquired ($p \leq .05$), the post hoc analysis utilizing Tukey's test was performed.

When assessing the virus concentration, the average number of virus particles per sample was \log_{10} transformed and plotted using a 95% confidence interval (CI). Since it is obscure whether an estimation of zero-copy number/sample was because of the lack of AdV DNA, inhibition or a sample DNA below the limit of detection, all zeros were removed before \log_{10} transformation and then put back in zeros. The relationship between various water quality parameters was determined by calculating Pearson correlation coefficients and was accounted for as " r ". Differences were considered significant when $p \leq .05$ and two-sided tests were performed for all analysis. Observations of markers were converted to binary data and binary logistic regression model (SPSS V24.0) was used to assess relationships between physiochemical parameters, FIB concentration and the presence or absence of the viral marker. The strength of relation was reported as Nagelkerke's R square; the stronger association had the value closer to 1.0. Relationships were considered significant when the p value for the model Chi-square was $\leq .05$ and the CI

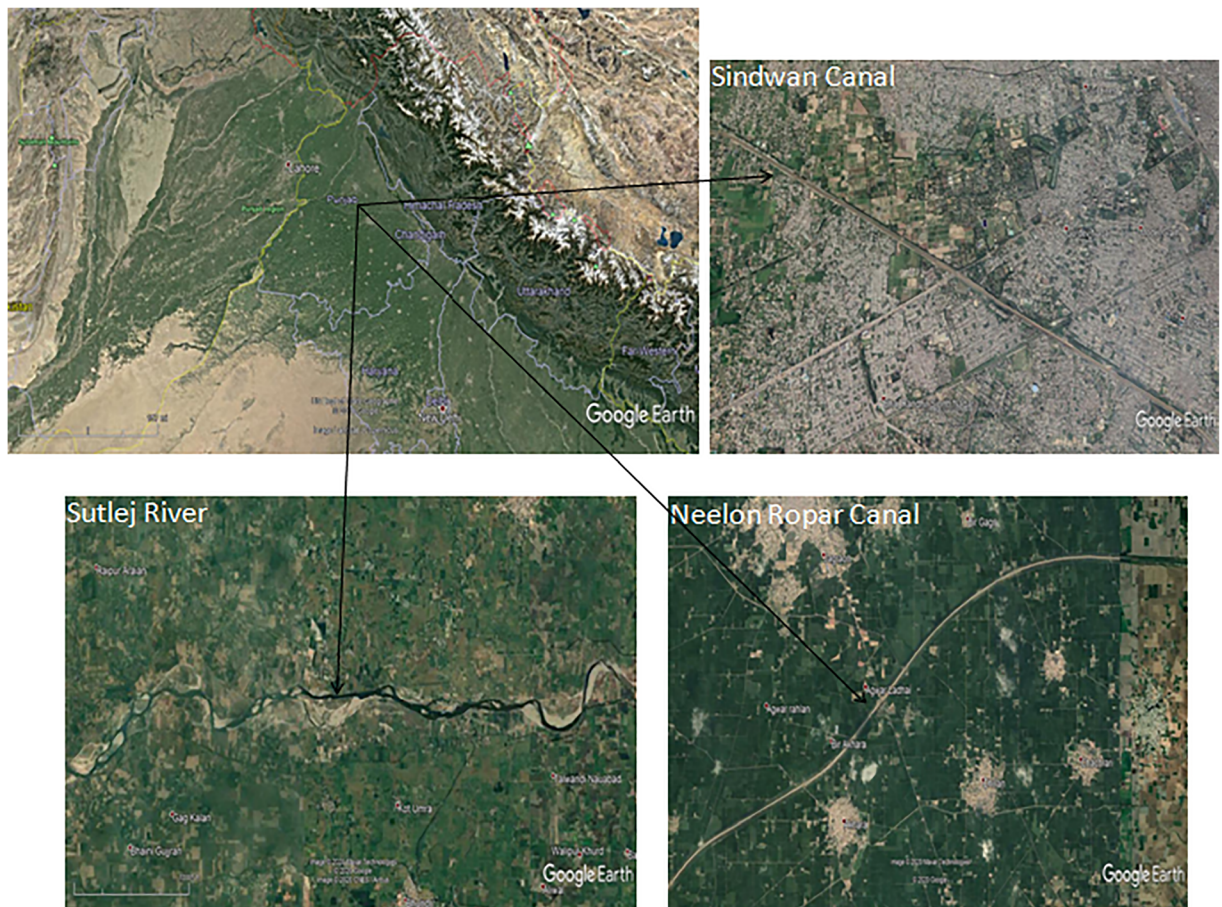


Fig. 1 These maps of the Punjab India, pulled from the Google Earth, show the canals and the river Sutlej of Punjab. Sampling was done after every 3 to 5 km along the stretch of the river and canals

for the odds ratio did not overlap 1.0. Only correlation in which p values were $< .05$ was noted. A non-significant correlation was denoted by NS. Fischer's exact test was used to evaluate significant differences in the frequency of observation of binary marker data. An alpha of 0.05 was used as the cut-off for significance.

In the multiple regression, coliform/enterococci counts were the dependent variable and physiochemical parameters were the independent variables. The same analysis was repeated using adenoviruses as the dependent variable (McCulloch 2015). Likewise, in a multinomial logistic regression using canals as the “dependent variable”, all other quantitative water quality parameters as “covariate” and qualitative water quality parameters as factors.

Cluster analysis comprises series of multivariate methods which are used to find clusters within the data (Du et al. 2017). The predefined algorithms were used to calculate similarity at which observations are clustered

and used to build a dendrogram (Kumar et al. 2018). Cluster analysis of the considerable number of variable in surface water was made based on the estimation of Pearson's correlation coefficient and centroid method. Likewise, principal component analysis for 8–10 water quality parameters was performed to determine the principal components which explains the maximum variance of the data. In factor analysis, new factors that cause variation in physiochemical parameters and microbiological concentration were extracted by varimax/direct oblimin rotation of the PCA.

4 Results and Discussion

The changeability of a load of genetic marker and prevalence in populations from other geographical areas proposes that the utilization of MST markers developed in a geographical area requires a priori characterization

of the assay performance at each watershed of interest before being actualized (Yahya et al. 2017). Subsequently, in the first instance, the chosen MST markers were tested in faecal samples from known sources. The virus concentration methods were optimized/validated and PCR inhibition assay was studied before applying these methods to the environmental samples (Zehra 2018). All the PCR markers demonstrated high sensitivity and specificity, albeit none of them accomplished 100% for the two parameters. Albeit a portion of the MST markers was identified in hosts other than the expected ones, their load in the target group was in every case significantly higher than in the non-target hosts, showing their reasonableness to recognize between sources of pollution.

The river Sutlej flows through the Ludhiana district of Punjab, and Sidhwan and Neelon-Ropar canals are tributaries that join Sutlej river. Viral load and other nine water quality parameters including BOD, COD, DO, TDS, turbidity, pH, temperature, coliform and enterococci count were studied at various sampling sites throughout 63 km between Samrala and Jagroan tehsil of district Ludhiana.

4.1 Detection of Adenoviruses in Surface Water Samples

Human enteric viruses are common surface water contaminants throughout the world. In this study, the data also showed the occurrence of HAdVs and bovine/porcine AdVs in samples of surface water. From 52 water samples, 32% analysed samples were positive for HAdV (qPCR), followed by 16% samples positive for BAdV (nested PCR) and 11% positive for PAdV (qPCR). In this study, the concentration of HAdV in water samples ranged from 10^1 to 10^2 copy no./L. These outcomes are like those found in surface waters (Hundesha et al. 2006; Wyn-Jones et al. 2011). However, viral loads higher in 2 logs have been detailed in other studies. Albinana-Gimenez et al. (2009) (10^1 – 10^4 gc/L), Choi and Jiang (2005) (10^2 – 10^4 gc/L) and Haramoto et al. (2010) (10^3 – 10^5 gc/L), pointed out that the HAdV viral loads present in water are exceptionally subject to the geographic level, temporal variation, environmental damage and the methods utilized for detection. Also, the concentration of PAdVs in the present study was around 10 copy no./L. This study is only among the few studies like Bortagaray et al. (2019) that reported PAdV in the surface water.

The frequency of adenovirus detection in each tehsil was analysed and summarized in Fig. 2. For all type of adenoviruses, detection frequencies were not significantly different except for HAdV that showed higher frequency in surface water samples from tehsil Ludhiana (Binary logistic—HAdVqPCR was taken as dependent variable and tehsil was taken as a categorical covariate) (Fig. 2).

There were points in tehsil Ludhiana showing maximum HAdV viral load (around 10^2 copy no./L). This may be because of higher population density in tehsil Ludhiana compared with other tehsils of district Ludhiana.

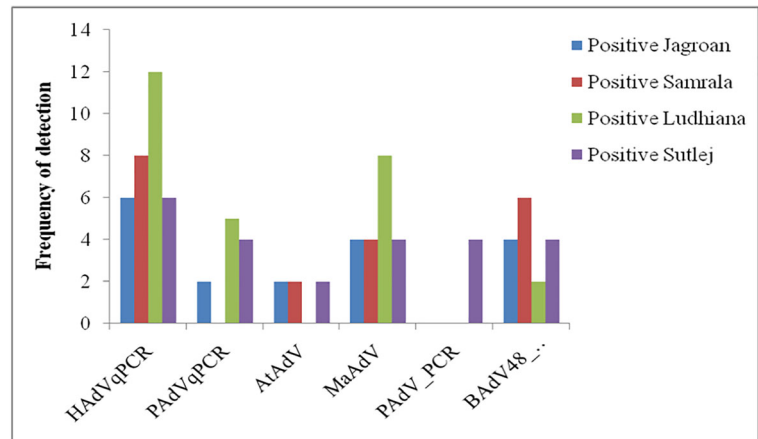
4.2 Analysis of Physiochemical and Bacteriological Parameters of River and Canals

There was a significant difference in physiochemical and bacteriological parameters between water samples collected from different tehsils of Ludhiana ($p < .05$, one-way ANOVA) (Table 1). On post hoc analysis using Tukey's test, a significant difference in enterococci count, coliform count, turbidity and TDS was observed ($p < .05$). The average concentration of both enterococci and coliform at Samrala was significantly larger than those at Jagroan and river Sutlej ($p < .05$, ANOVA). Although the average concentration of enterococci and coliforms among Jagroan, Ludhiana and river Sutlej were not significantly different (Table 1), there was a significant difference between turbidity and TDS of water samples from river Sutlej with any other water sample ($p < .05$).

In the present study, high numbers of coliform and enterococci were observed in the water of the canal over all locales which could be almost certain because of the presence of faecal contamination from sewage leakage, animal sources and environmental sources. The coliform count was generally 10-fold higher than enterococci along most of the sites.

The BOD and COD values represented in Table 1 showed higher standard deviation, which means that samples from river Sutlej were showing high variation in physiochemical parameters, which could be because of entry of highly contaminated water from Budha nallah, a highly polluted seasonal water stream that drains into river Sutlej. There was significant variation between water quality parameter from upstream and downstream of confluence point of Budha nallah with river Sutlej (Table 2, Fig. 3).

Fig. 2 Frequency of detection of adenoviruses (HAdV/BAdV/PAdV) in each tehsil of Ludhiana



4.3 Correlation Between Physiochemical, Bacteriological and Viral Markers in Surface Water Samples

To the best of our knowledge, this study is among the first in India to analyse the extent and recurrence of observation of culturable FIB, multiple physicochemical parameters and viral markers in recreational waters. Among the most significant findings were the relationships of human-associated MST markers with one another and with adenovirus detection. Data were from all the locales of the canal and river Sutlej excluding values of samples collected downstream Sutlej. These values were outlier therefore not statistically analysed with other values from water samples (Fig. 3). The average \log_{10} -transformed concentrations of bacterial and viral markers at each site are summarized in Fig. 4 and the relationship between the physiochemical, bacteriological and viral indicators is presented in Table 3.

A correlation between analytes at each site was now and again yet not generally noted. Generally, only a sporadic significant correlation was observed with the few components; there was no constant, stable correlation between the microbiological and physiochemical parameters (Table 3). Their correlation was positive between the coliform and enterococci and between the COD and BOD (Table 3). In river water samples, qPCR detection of HAdVs and PAdVs was weakly to strongly correlated with the log concentrations of coliforms (HAdV: Nagelkerke's $R^2 = 0.446$, odds ratio = 6.829, $p = .019$ and with PAdV: Nagelkerke's $R^2 = 0.688$, odds ratio = 19.764, $p = .05$). PAdVs detection rate also

showed strong correlation with TDS (Nagelkerke's $R^2 = 0.686$, odds ratio = 0.886, $p = .044$). Nested PCR detection of BAdVs was not correlated with any water quality parameter except COD (Nagelkerke's $R^2 = 0.632$, odds ratio = 1.487, $p = .021$). This correlation could be significant or nonsignificant at different sites or for different samples. No significant relationships were observed at sites from Sidhwan canal and Neelon-Ropar canal (multinomial logistic regression using canal as the "dependent variable", all other quantitative water quality parameters as "covariates" and qualitative water quality parameters as "factors"). But when considering coliform, BOD, DO and pH together for classifying samples as per CPCB (2008) then odds of finding the class D (Propagation of Wildlife and Fisheries, Irrigation, Industrial Cooling, Controlled Waste disposal) samples in Neelon-Ropar canal was 13.93 and 15.83 times higher than samples from Sidhwan canal and Sutlej river, respectively ($p = .001$). The lack of correlation between the physiochemical, bacteriological and virological parameters had already been well documented by other studies. Lee et al. (2013) have demonstrated that the presence of total coliform and faecal coliform were not related to the presence of enteric viruses in a study conducted with surface water in South Korea over 4 years. Likewise, Vecchia et al. (2015) observed correlation only with a few components while evaluating 78 physiochemical parameters (physiochemical parameters, metal components, pesticides and organic compounds) and its correlation with microbiological parameters. In their study, an inverse correlation was observed between DO and canine AdV while a positive correlation was observed

Table 1 Summary of physicochemical and bacteriological parameters of water samples

	Log_entero (CFU/100 ml)	Log_coli (CFU/100 ml)	COD (mg/l)	BOD (mg/l)	pH	Turbidity (NTU)	TDS (mg/l)
Jagroan	3.16 ± 0.45	4.64 ± 0.94	18.63 ± 4.55	3.55 ± 0.87	7.33 ± 0.33	3.29 ± 0.71	152.16 ± 18.17
Samrala	4.35 ± 0.97	5.88 ± 0.61	26.02 ± 8.05	4.96 ± 1.53	7.33 ± 0.36	2.82 ± 0.95	158.94 ± 10.28
Ludhiana	3.56 ± 0.72	5.38 ± 0.63	17.94 ± 5.50	3.65 ± 1.11	7.23 ± 0.34	8.68 ± 1.45	171.68 ± 13.89
Sutlej	3.49 ± 1.02	4.68 ± 1.34	93.37 ± 151.09	9.45 ± 11.25	7.24 ± 0.29	25.36 ± 17.17	277.59 ± 136.92

between DO and *E. coli* count. Unlike the study reported by Waso et al. (2018), a positive correlation was observed between the coliform and human adenovirus.

In the present study, both the canals and river Sutlej showed a stable and significant positive correlation between bacteriological parameters (coliform and enterococci) and between BOD and COD. However, regression analysis revealed no significant correlations between concentrations of enterococci count with the concentration of AdVs, unlike coliform that showed correlation with PAdV and no correlation with HAdVs/BAAdVs, similar to the studies that reported such type of correlations with HAdVs (Lee et al. 2013; Vecchia et al. 2015).

In nutshell, the concentration of *E. coli* and *Enterococcus* spp. concentration varied along the line of Sidhwan and Neelon-Ropar canal but was detected at all the points evaluated. There was no correlation between the HAdV and bacterial indicators when it comes to point-source contamination.

4.4 Cluster Analysis and Principal Component Analysis

One important aspect of the current study was to understand the source allocation of the water quality

parameters in the study area using cluster analysis (CA) and principal component analysis (PCA).

In the dendrogram (Fig. 5), cluster analysis separates the different water quality parameters into four major clusters (C1 to C4). PAdV forms a cluster C1 with BOD and COD that in turn forms a sub-cluster with turbidity and TDS (C2). Bacterial indicators and HAdV forms a separate cluster as C3 and C4, respectively, thereby represents no similarity with the other clusters.

Principal component analysis of the data ends up being a successful tool for data reduction as the initial three principal components of all the water quality parameters clarified 91% variance. Factor analysis outlined three factors of basic water quality (Kumar et al. 2016; Kumar et al. 2018; Tripathi and Singal 2019). Factor 1 comprised parameters like BOD, COD, TDS, turbidity and PAdV. Factor 2 was a bacteriological water quality determinant and explained maximum variance in coliform and enterococci. Factor 3 comprised HAdV.

At first, the factorability of the 10 water quality parameters was examined using varimax and oblimin with Kaiser-Meyer-Olkin normalization as rotation methods. An oblimin rotation gave the best-characterized factor structure. A few very much perceived rules for the factorability of a correlation were

Table 2 Variation in mean values of various physicochemical parameters at six selected sites (S1–S6)

Parameters	S1–4 (upstream Budha nallah) mean ± SD	S5–6 (downstream Budha nallah) mean ± SD
pH	7.34 ± 0.28	7.13 ± 0.30
Turbidity (NTU)	13.89 ± 2.55	48.83 ± 13.17
DO (mg/l)	7.71 ± 0.48	2.23 ± 1.16
TDS (mg/l)	193.94 ± 27.49	475.34 ± 75.64
BOD ₅ (mg/l)	3.00 ± 0.54	25.29 ± 8.89
COD (mg/l)	8.71 ± 0.39	305.74 ± 121.39
Enterococci count (log ₁₀ , CFU/100 ml)	3.13 ± 0.73	4.72 ± 0.09
Coliform count (log ₁₀ , CFU/100 ml)	3.95 ± 0.43	6.66 ± 0.15

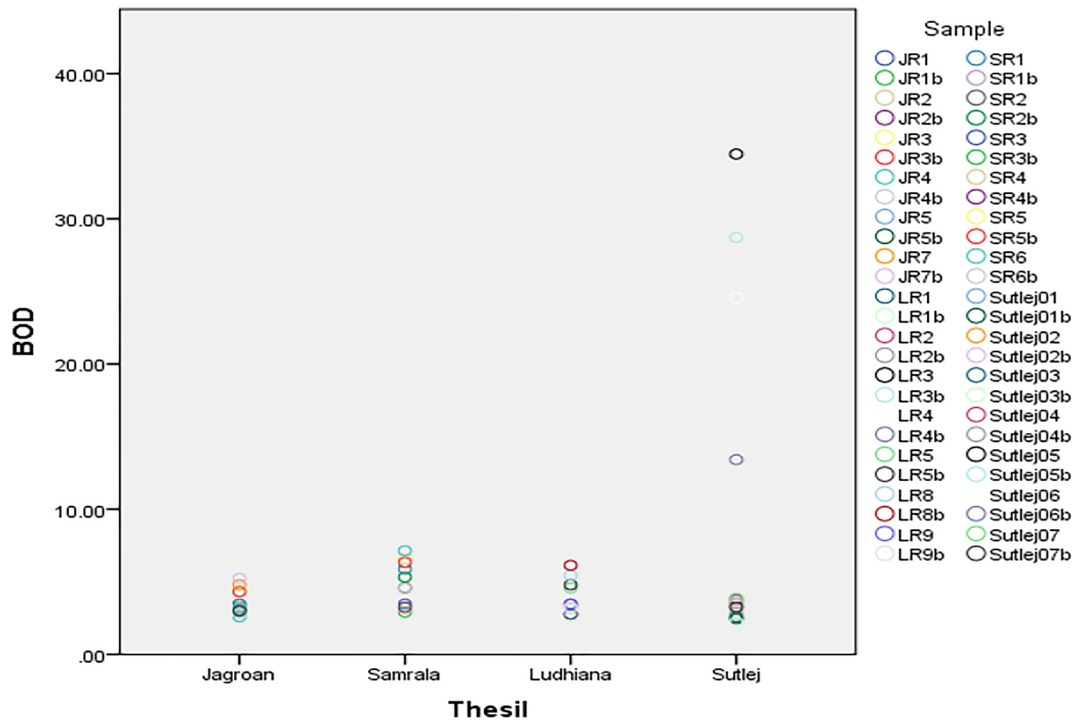


Fig. 3 Scatter plot showing outliers representing samples from river Sutej after joining of Budha nallah

utilized. Firstly, it was observed that 7 of the 10 parameters correlated at least 0.3 with at least one other item, suggesting reasonable factorability. Secondly, the Kaiser-Meyer-Olkin measure of sampling adequacy was 0.775, above the commonly recommended value of 0.6 and Bartlett’s test of sphericity was significant (χ^2

(45) = 520.502, $p < .05$). Additionally, the anti-image correlation matrix’s diagonals were all over 0.5 and the commonalities were all above 0.3 except for HAdV. To increase the reliability of PCA, the two parameters (DO and pH) that were not correlated with any other parameter were removed for second PCA analysis.

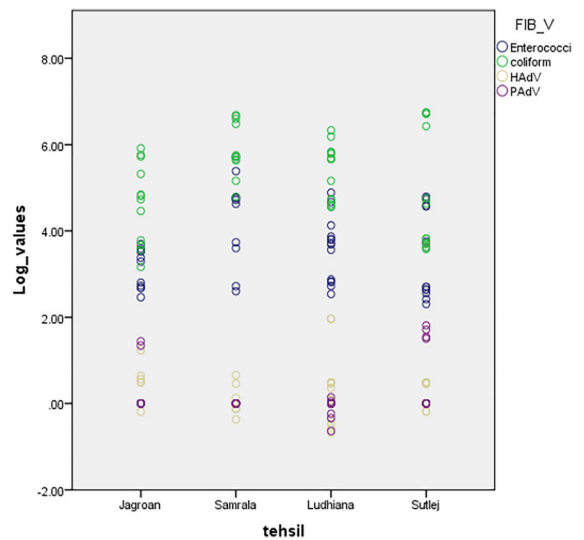
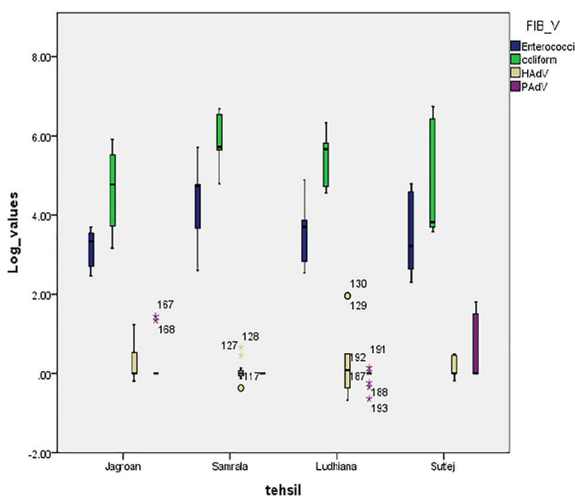


Fig. 4 Log₁₀-transformed concentrations of enterococci, coliform and adenoviruses at different tehsils. Error bars represent standard deviations. Indicator bacteria are reported as log₁₀ CFU/100 ml and AdVs are reported as log₁₀ copy number/100 ml

Table 3 Relationship between indicators and markers of surface water samples from canals and river Sutlej

Correlation between vectors of values		Log_entero	Log_coli	COD	BOD	DO	pH	Turbidity	TDS	HAdVqPCR	PAdVqPCR	BAdV48
Log_entero (<i>r</i>)	1.00											
Log_coli (<i>r</i>)	<i>0.776</i>	1.00										
COD (<i>r</i>)	0.399	0.557	1.00									
BOD (<i>r</i>)	0.370	0.479	<i>0.929</i>	1.00								
DO (<i>r</i>)	0.067	-0.062	0.265	0.532	1.00							
pH (<i>r</i>)	0.002	-0.046	-0.110	0.074	0.068	1.00						
Turbidity (<i>r</i>)	-0.248	-0.418	-0.780	-0.384	0.360	-0.051	1.00					
TDS (<i>r</i>)	-0.098	-0.148	-0.398	-0.169	0.323	0.164	0.566	1.00				
HAdVqPCR (<i>R</i> ²)	NS	0.446	NS	NS	NS	NS	NS	NS	1.00			
PAdVqPCR (<i>R</i> ²)	NS	<i>0.688</i>	NS	NS	NS	NS	NS	<i>0.688</i>	NS	1.00		
BAdV48 (<i>R</i> ²)	NS	NS	<i>0.632</i>	NS	NS	NS	NS	NS	NS	NS	1.00	

Italicized figures represents strong positive correlation between the different parameters

Now, factorability of the eight water quality parameters was examined. It was observed that seven out of eight parameters correlated and the Kaiser-Meyer-Olkin measure of sampling adequacy was 0.806 and Bartlett’s test of sphericity adequacy was significant ($\chi^2 (28) = 484.476, p < .05$). Additionally, the anti-image correlation matrix’s diagonals were all over 0.5 and communalities were all above 0.3, further affirming that every parameter imparted some common variance with other

parameters. Given these overall indicators, factor analysis was regarded to be reasonable with each of the eight parameters. A tabular and graphical representation for factor loading of principal components (eigenvector > 1) alongside their cumulative variance (%) has been given in Table 4 and Fig. 6.

Extraction method: principal component analysis

Factor 1 accounted for 59.70% of the total variance and had strong positive loadings on BOD, COD, TDS,

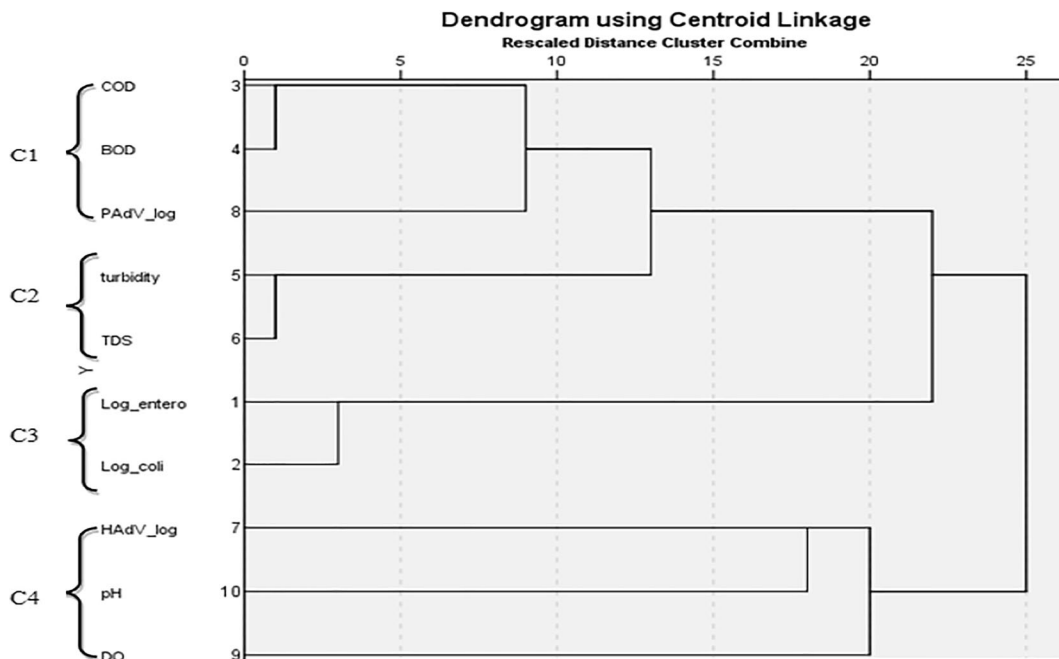


Fig. 5 Dendrogram representing similar groups clustered together. C1 to C4 represent the four clusters

Table 4 PCA of different water quality parameters of water samples from Sutlej and its tributaries (component matrix)

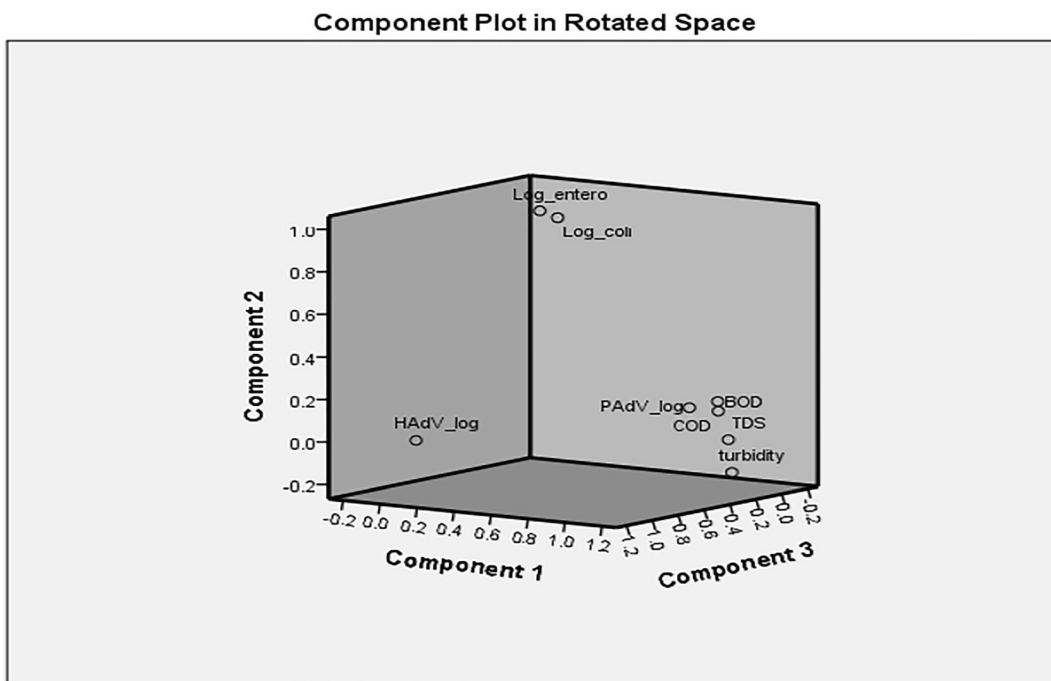
	Component			Communalities
	1	2	3	
BOD	0.969			0.947
COD	0.968			0.953
TDS	0.944			0.951
Turbidity	0.875	-0.375		0.909
PAdV_log	0.825			0.688
Log_entero	0.487	0.795		0.897
Log_coli	0.568	0.755		0.918
HAdV_log		0.302	0.952	1.00

turbidity and PAdV which are also supported by their mutual cluster in CA. Since all these parameters are pollution-related from industrial pollutant and farm wastewater, the first factor is largely a pollution loading factor leading to eutrophication. Factor 2 accounted for 19.05% of the total variance and had strong loadings on enterococci and coliform count due to the natural quality of water and domestic waste. Factor 3 represented 12.06% of the total variance and had strong positive loadings on HAdV same as indicated in CA. This factor

shows the impact of human activities overwhelmed in the catchment area of the canals and river.

Finally, in evaluating the surface water of two canals and river Sutlej, there was a small association between viruses and few other variables, but this was not consistent for the present study to establish a reliable relation. In general, the correlations showed that linearity and nonlinear correlations were found only at point source contaminations. The findings of the presented study were evaluated in polluted areas and not so polluted areas, judging by the bacteriological indicators. This sort of clustering can likewise be seen in other studies like Kirschner et al. (2017). Their study detailed microbial faecal contamination of the river as an independent component and did not cluster with any other estimated environmental parameters within the environmental data matrix. This study shows that the application of host-associated genetic microbial source tracking markers working together with the customary idea of microbial faecal contamination observing based on FIB essentially improves the information on the extent and origin of microbial faecal contamination patterns in large rivers. It establishes an incredible tool to guide target-oriented water quality management in large river basins.

In conclusion, the capability of coliform to be environmental in origin and the deficient specificity of

**Fig. 6** Component plot in the rotated space

markers can mean uncertain outcomes when just a single marker is distinguished. In any case, each marker's predictive probability of success is increased when more than one marker is identified at the same site. For all markers utilized in this study, an epidemiological study evaluating the human health risks associated with the marker would all the more accurately characterize the usefulness of each assay. Hence, a significant level of certainty on the contamination source can be acquired if more than one marker is distinguished. Predictive statistical models represent a novel technique for advising the management organizations regarding likely factors impacting faecal contamination intensity and how to alleviate future faecal contamination occasions. Our study thus gives a case of how quantifying physiochemical, FIB and adenoviral markers in surface water can improve water quality evaluation and help tailor Central Pollution Control Board (CPCB) programs for impaired water bodies in the future.

Acknowledgements We thank the Farmers Commission Punjab and GADVASU Ludhiana (India) for supporting Dr. Simranpreet Kaur's visit to CSIRO Land and Water Flagship, Brisbane, Australia. We also thank Dr. Sachin Kumar (IIT Guwahati) for providing Human Adenovirus standard for this study.

Authors' Contribution Zehra A.—lab work, data analysis and manuscript writing; Kaur S.—guidance and manuscript editing; Kaur S., Singh R. and Gill JPS.— technical program of work.

Funding This work was financially supported by School of Public Health and Zoonoses, Guru Angad Dev Veterinary Animal Sciences University (GADVASU), Punjab, India.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethics Approval Not applicable.

References

- Albinana-Gimenez, N., Miagostovich, M. P., Calqua, B., Huguet, J. M., Matia, L., & Girones, R. (2009). Analysis of adenoviruses and polyomaviruses quantified by qPCR as indicators of water quality in source and drinking-water treatment plants. *Water Research*, 43, 2011–2019.
- Ali, S., Hussain, S., Abbas, Q., Hussain, M., Ali, W., & Akber, M. et al. (2014). Water quality assessment of Gilgit river, using fecal and total coliform as indicators. *Journal Of Biodiversity And Environmental Sciences*, 5(4), 343–347.
- Bora, M., & Goswami, D. (2016). Water quality assessment in terms of water quality index (WQI): A case study of the Kolong River, Assam, India. *Applied Water Science*, 7(6), 3125–3135. <https://doi.org/10.1007/s13201-016-0451-y>.
- Bortagaray, V., Lizasoain, A., Piccini, C., Gillman, L., Berois, M., Pou, S., et al. (2019). Microbial source tracking analysis using viral indicators in Santa Lucia and Uruguay Rivers, Uruguay. *Food And Environmental Virology*, 11(3), 259–267. <https://doi.org/10.1007/s12560-019-09384-2>.
- Choi, S., & Jiang, S. C. (2005). Real-time PCR quantification of human adenoviruses in urban rivers indicates genome prevalence but low infectivity. *Applied and Environmental Microbiology*, 71(11), 7426–7433.
- Central Pollution Control Board (CPCB). (2002). Water quality in India – status and trend (1990 - 2001) (pp. 1–20) Delhi: Ministry of Environment & forests, Govt. of India.
- Central Pollution Control Board (CPCB). (2008). Guidelines for water quality management. Parivesh Bhawan, East Arjun Nagar, Delhi.
- Central Pollution Control Board (CPCB). (2011). Status of water quality in India—2011. New Delhi: *Central Pollution Control Board* (Ministry of Environment and Forestry) <http://www.cpcb.nic.in>
- Central Pollution Control Board (CPCB). (2015). River stretches for restoration of water quality. New Delhi: *Central Pollution Control Board* (Ministry of Environment, Forest and Climate change) report.
- Colford, J.M., Wade, T.J., Schiff, K.C., Wright, C.C., Griffith, J.F., Sandhu, S.K. et al. (2007). Water quality indicators and the risk of illness at beaches with nonpoint sources of fecal contamination. *Epidemiology*, 18(1), 27–35.
- Das Kangabam, R., Bhoominathan, S., Kanagaraj, S., & Govindaraju, M. (2017). Development of a water quality index (WQI) for the Loktak Lake in India. *Applied Water Science*, 7(6), 2907–2918.
- Du, X., Shao, F., Wu, S., Zhang, H., & Xu, S. (2017). Water quality assessment with hierarchical cluster analysis based on Mahalanobis distance. *Environmental Monitoring and Assessment*, 189(7), 335.
- Dwight, R.H., Baker, D.B., Semenza, J.C., & Olson, B.H. (2004). Health effects associated with recreational coastal water use: Urban versus Rural California. *American Journal of Public Health*, 94, 565–567.
- U.S. Environmental Protection Agency (EPA) (1997) Method 1600: membrane filter test method for enterococci in water. EPA/821/R-97/004. US Environmental Protection Agency, Washington, DC.
- Farkas, K., Cooper, D., McDonald, J., Malham, S., de Rougemont, A., & Jones, D. (2018). Seasonal and spatial dynamics of enteric viruses in wastewater and in riverine and estuarine receiving waters. *Science of the Total Environment*, 634, 1174–1183. <https://doi.org/10.1016/j.scitotenv.2018.04.038>.
- Hamner, S., Tripathi, A., Mishra, R., Bouskill, N., Broadaway, S., & Pyle Band Ford, T. (2007). The role of water use patterns and sewage pollution in incidence of water-borne/enteric diseases along the Ganges river in Varanasi, India.

- International Journal Of Environmental Health Research*, 16(2), 113–132.
- Haramoto, E., Kitajima, M., Katayama, H., & Ohgaki, S. (2010). Real-time PCR detection of adenoviruses, polyomaviruses and torque Teno viruses in river water in Japan. *Water Research*, 44(6), 1747–1752.
- Hundesda, A., Maluquer de Motes, C., Bofill-Mas, S., Albinana-Gimenez, N., & Girones, R. (2006). Identification of human and animal adenoviruses and polyomaviruses for determination of sources of faecal contamination in the environment. *Applied and Environmental Microbiology*, 72(12), 7886–7893.
- Kirschner, A., Reischer, G., Jakwerth, S., Savio, D., Ixenmaier, S., Toth, E., et al. (2017). Multiparametric monitoring of microbial faecal pollution reveals the dominance of human contamination along the whole Danube River. *Water Research*, 124, 543–555. <https://doi.org/10.1016/j.watres.2017.07.052>.
- Kumar, V., Sharma, A., Chawla, A., Bhardwaj, R., & Thukral, A. (2016). Water quality assessment of river Beas, India, using multivariate and remote sensing techniques. *Environmental Monitoring and Assessment*, 188(3), 137.
- Kumar, M., Ramanathan, A., Tripathi, R., Farswan, S., Kumar, D., & Bhattacharya, P. (2018). A study of trace element contamination using multivariate statistical techniques and health risk assessment in groundwater of Chhaprola Industrial Area, Gautam Buddha Nagar, Uttar Pradesh, India. *Chemosphere*, 166, 135–145.
- Lee, G. C., Jheong, W. H., Kim, M. J., Choi, D. H., & Baik, K. H. (2013). A 5-year survey (2007–2011) of enteric viruses in Korean aquatic environments and the use of coliforms as viral indicators. *Microbiology and Immunology*, 57, 46–53.
- Liang, L., Goh, S.G., Vergara, G.G.R.V., Fang, H.M., Rezaeinejad, S., Chang, S.Y., Bayen, S., Lee, W.A., Sobsey, M.D., Rose, J.B., & Gina, K.Y.H. (2015). Alternative fecal indicators and their empirical relationships with enteric viruses, Salmonella enterica and Pseudomonas aeruginosa in surface waters of a tropical urban catchment. *Applied and Environmental Microbiology*, 81(3), 850–860.
- McCulloch, K. J. (2015). Thesis: Analysis of the relationship between water quality parameters and stream sediment with faecal bacteria in hidden creeks, Rock Hill, SC. paper 14, Winthrop University. http://digitalcommons.winthrop.edu/graduate_theses.
- Patel, K., Shukla, A., Tripathi, A., & Hoffmann, P. (2002). Heavy metal concentrations of precipitation in East Madhya Pradesh of India. *Water, Air and Soil Pollution*, 130(1/4), 463–468.
- Ponsadailakshmi, S., Sankari, S., Prasanna, S., & Madhurambal, G. (2018). Evaluation of water quality suitability for drinking using drinking water quality index in Nagapattinam district, Tamil Nadu in Southern India. *Groundwater for Sustainable Development*, 6, 43–49. <https://doi.org/10.1016/j.gsd.2017.10.005>.
- Rosemond, S., Duro, D., & Dubé, M. (2009). Comparative analysis of regional water quality in Canada using the Water Quality Index. *Environmental Monitoring And Assessment*, 156(1–4), 223–240.
- Schriewer, A., Wuertz, S., Misra, P., Clasen, T., Panigrahi, P., Jenkins, M., & Odagiri, M. (2015). Human and animal fecal contamination of community water sources, stored drinking water and hands in rural India measured with validated microbial source tracking assays. *The American Journal of Tropical Medicine and Hygiene*, 93(3), 509–516.
- Setia, R., Dhaliwal, S., Kumar, V., Singh, R., Kukal, S., & Pateriya, B. (2020). Impact assessment of metal contamination in surface water of Sutlej River (India) on human health risks. *Environmental Pollution*, 265, 114907. <https://doi.org/10.1016/j.envpol.2020.114907>.
- Sun, W., Xia, C., Xu, M., Guo, J., & Sun, G. (2016). Application of modified water quality indices as indicators to assess the spatial and temporal trends of water quality in the Dongjiang River. *Ecological Indicators*, 66, 306–312.
- Tripathi, M., & Singal, S. (2019). Use of principal component analysis for parameter selection for development of a novel water quality index: a case study of river Ganga India. *Ecological Indicators*, 96, 430–436. <https://doi.org/10.1016/j.ecolind.2018.09.025>.
- U.S. Environmental Protection Agency (2002) Method 1603: *Escherichia coli* in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar (modified mTEC). EPA/821/R-02/023. US Environmental Protection Agency, Washington, DC.
- Vecchia, D. A., Rigotto, C., Staggemeier, R., Solimon, C. M., Gil de Souza, F., Henzel, A., & Santos, L. E. (2015). Surface water quality in the Sinos river basin, in Southern Brazil: tracking microbiological contamination and correlation with physiochemical parameters. *Environmental Science and Pollution Research*, 22(13), 9899–9811.
- Waso, M., Khan, S., & Khan, W. (2018). Microbial source tracking markers associated with domestic rainwater harvesting systems: correlation to indicator organisms. *Environmental Research*, 161, 446–455.
- Wyn-Jones, A. P., Carducci, A., Cook, N., D'Agostino, M., Divizia, M., Fleischer, J., Gantzer, C., Gawler, A., Girones, R., Holler, C., de Roda Husman, A. M., Kay, D., Kozyra, I., Lopez-Pila, J., Muscillo, M., Nascimento, M. S., Papageorgiou, G., Rutjes, S., Sellwood, J., Szewzyk, R., & Wyer, M. (2011). Surveillance of adenoviruses and noroviruses in European recreational waters. *Water Research*, 45(3), 1025–1038.
- Yahya, M., Blanch, A. R., Meijer, W. G., Antoniou, K., Hmaied, F., & Ballesté, E. (2017). Comparison of the performance of different microbial source tracking markers among European and North African Regions. *Journal of Environmental Quality*, 46(4), 760.
- Zehra, A. (2018). Molecular assays targeting species-specific adenoviruses in different aquatic environments and their application for library-independent microbial source tracking (PhD Dissertation). Guru Angad Dev Veterinary and Animal Sciences University (GADVASU), Ludhiana-India.
- 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.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.