Statistical techniques as investigatory tools for the Evaluation of Water Quality of Freshwater Ecosystems

Anish Dua[1] , Prabha[2,*]

[1]Aquatic Biology Laboratory, Department of Zoology Guru Nanak Dev University, Amritsar-143005, Punjab, India
[2*]Assistant professor, PG Department of Zoology Kanya Maha Vidyalaya, Jalandhar (144001), Punjab, India

*Corresponding Author; Prabha
* (e-mail: prabhasyall@gmail.com)

Abstract

Statistics plays crucial role in every field of mankind. Basic statistical techniques help researchers to design experiments, verify conclusions and systematic interpretations of results in various branches of natural as well as social sciences. Aquatic ecosystems being indispensable part of nature are important for the planet earth and human life. Amongst all the aquatic ecosystems, freshwater ecosystems play vital role in the survival of mankind as well as flora and fauna. Rivers are the most engrossing and multiplex freshwater ecosystem on the earth. Due to increased anthropogenic and natural activities these freshwater resources are getting exploited and becoming vulnerable to pollution. Evaluation of these systems becomes crucial as their hydrology is greatly affected by anthropogenic resources and climatic conditions. Being a developing country, Indian freshwater ecosystems are also getting affected by manmade and climatic changes. In present study River Beas (Punjab region) is taken into consideration being one of the major rivers flowing in India. With increasing anthropogenic activities and climatic changes in recent years, river water quality has become an issue of concern and there is need to incorporate multivariate statistical techniques for analysis of quality of the water. Statistical analysis provides crucial insight of many physico-chemical parameters. For this study The Kolmogorov-Smirnov (K-S), principal component analysis/ factor analysis (PCA/FA) and two way ANOVA were applied that reduced huge data in small valuable data set. Two way analysis of variance (ANOVA) followed by Tukey's HSD was applied to mean values to evaluate spatio temporal variations in selected water quality parameters. Principal component analysis specified the responsible variables for variations in water quality and included soluble salts, nutrients, organic pollutants and some heavy metals from both natural and anthropogenic sources.

Keywords: Multivariate statistical techniques, Principal component analysis, Pollution, River Beas, Two way ANOVA

1. Introduction

Multivariate statistical techniques provide reliable estimations and hence are used in most of water quality studies (Bu *et al*. 2014; Kaur *et al*; 2014; Isiyaka *et al*; Kaur and Dua 2016). Multivariate techniques like principal component analysis and factor analysis reduce the large data sets of variables into smaller set of variables without losing information of original set of variables (Dunteman 1989 and Hamzah *et al*. 2016). Analysis of Variance (ANOVA) analyse total variation in data that is possibly attributed to various sources or causes of variations (Gupta 2007) and therefore two way ANOVA was used to analyse possible variations among all the sources of variance.

Among all the freshwater ecosystems systems, rivers are the most engrossing and multiplex ecosystem on the earth. These ecosystems are considerable freshwater resources for human consumption, agricultural needs and industrial as well as recreational purposes (Wang *et al*. 2017; Razmkhaha *et al*. 2010). Due to increased anthropogenic and natural activities these freshwater resources are getting exploited and becoming vulnerable to pollution. Evaluation of these systems becomes crucial as their hydrology is greatly affected by anthropogenic resources and climatic conditions. River water quality is greatly affected by seasonal variations in precipitation, surface runoff and groundwater flow (Varol *et al*. 2012). This leads to necessity of regular river monitoring programmes for the reliable information of the water quality (Singh *et al*. 2017). A number of studies have been carried out to assess spatial and temporal variations in the water quality of rivers worldwide (Ogwueleka 2014; Septong and Chuersuwan 2015).

Being a developing country, Indian freshwater ecosystems are also getting affected by manmade and climatic changes. In present study River Beas (Punjab region) is taken into consideration being one of the major rivers flowing in India.

2. Material and methods

2.1 Study area

The river originates from south of the Rohtang Pass located in Kullu and it flows through the height of 13,326 feet. Then it follows the path through district Mandi and further goes to district of Kangra at Sanghol that is located at a height of 1920 feet. Once reaching near Hoshiarpur, the river changes its course to a sharp northern turn and passes through the Kangra district. Later, it takes another sharp curve at foot of Shivalik hills and turns the path to southern direction while

extricating Gurdaspur and Hoshiarpur districts. As it reaches the Jalandhar district, it separates the districts of Amritsar and Kapurthala. It takes the course through the southern west route of district Kapurthala in the state of Punjab traversing total path of 290 miles and ends before merging into Sutlej River. Beas is a crucial river as it is auctioned annually by Punjab Fisheries department in the month of September. Out of six sampling sites, site I was in Himachal Pradesh and sites II-VI were in Punjab region. Geographical coordinates of the selected sampling sites are given in Table 1.

Table 1: Geographical coordinates of sampling sites

Site	Latitude	Longitude	Altitude (msl)
SITE I (Pong Dam)	32°02′22 N	76°50′37 E	1869
SITE II (Naushehra Pattan)	31°59′15 N	75°33′12 E	237.0
SITE III (Beas)	31°59′15 N	75°33′12 E	239.0
SITE IV (Goindwal Sahib)	31°02′22 N	75°15′37 E	226.5
SITE V (Confluence)	31°08′30.15 N	75°03′44.18 E	210.0
SITE VI (Harike Pattan)	31°08′30.15 N	75°03′44.18 E	210.0

2.2 Analytical procedure

A data set of 16 water quality variables viz. Dissolved oxygen (DO), Electrical conductivity (EC), pH, Turbidity (TD), Water temperature (WT), Alkalinity (AK), Total hardness (TH), Magnesium hardness (Mg), Total dissolved solids (TDS), Chemical oxygen demand (COD), Five days Biological oxygen demand (BOD_5), Total nitrogen (∑N), Cadmium (Cd), Lead (Pb), Nickel (Ni) and Chromate (Cr) collected from six sites for two years, in four seasons (spring, summer, rains and winter) was monitored. Sampling was done during early hours of morning for each season and samples were collected in pre-treated and clean plastic sampling bottles for physico-chemical analysis and were preserved according to standards given in (VishnuRadhan *et al*. 2017) manual. WT, pH, DO, EC and TDS were measured on site with the help of portable water analysis kit (WTW Multi 340i/SET). Five days Biological oxygen demand (BOD_5) was calculated using Oxitop measuring system for five days at 20°C in a thermostat (TS 606-G/2-i). Clarity of the water (TD) was measured using sechhi disk as well as NTU turbiditimeter at the sampling site. AK, TH and Mg were analysed in laboratory following method given by (Trivedi and Goel 1986). Other variables were analyzed using Merck cell test kits on UV/VIS spectrophotometer (Spectroquant® Pharo 300).

2.3 Water quality analysis and Data treatment

The Kolmogorov-Smirnov (K-S) test was applied on raw data to test the goodness-of-fit of the data to normal distribution. According to K-S test, observed water quality variables showed normal distribution. Spatial and temporal observations were interpreted by using principal component analysis/factor analysis (PCA/FA). Further spatio-temporal observations were compared using two way ANOVA to observe variations due to sampling sites and seasons (Mustapha 2008). Statistical computations were done using Past (version 3), Microsoft office Excel 2010, SPSS (Version 16), MINITAB (Version 18) and XLSTAT software.

2.4 Data structure treatment

Kaiser-Meyer-Olkin (KMO) and Barlett's test were applied to examine data suitability for PCA/FA. KMO measures adequacy of the samples that ascertain proportion of common variance. KMO value close to 1 indicates usefulness of PCA/FA (Wang *et al*. 2012). For present study KMO=0.50. Further Bartlett's test indicates correlation among variables. Significant correlation among variables was observed (P = 0.00).

3. Results and discussion

3.1 Descriptive statistics

Mean, maximum (max.) & minimum (min.) values and coefficient of variation (CV %) for various water quality variables are given in Table 2. Increased concentration of DO at higher altitude as compare to decreased concentration at lower altitude was also reported by Das and Acharya 2003; Hart and Zabbey 2005. Maximum value for EC recorded was 388.08 µS/cm (spring/site V) and min. value was 153.72µS/cm (summer/site IV). This trend is also supported by Mazurek *et al*. (2012). According to BIS (1991) and EPAR (2012) observed pH range was within permissible limits. Increased turbidity was seen at site V during winter season (101.23 NTU) and water of site I was fond to be the most clear during summer (0.56 NTU). Maximum temperature recorded was 31.08°C (summer/site V) and min. 15.05 °C (winter I). This trend of temperature variation is possibly due to illumination (Manjare *et al*. 2010). Similar trend was identified by (Shinde *et al*. 2011).Value of alkalinity ranged from 47.50 mg/L (rains/site I) to 197.50 mg/L (rains/VI). An increase in alkalinity during rainy season was also noticed by Simpi *et al*. (2011). Total hardness ranged between 59.33 mg/L (summer/site VI) to 140.50 mg/L (spring/site V). Higher conc. of total dissolved solids possibly could be one of the reasons for an increase in conc. of TH at site V. Values of Mg, TDS, COD and BOD ranged between 3.67 mg/L (summer/site III) to 24.31 mg/L (rains/site V),41.99 mg/L (rains/site I) to 246.27mg/L (spring/site V), 32.92 mg/L (summer/site IV) to 159.58 mg/L (summer/site II) and 0.00 mg/L (summer/site II-VI) to 19.80 mg/L (spring/site III) respectively and were within maximum permissible limits as prescribed by BIS (1999) and EPAR (2012). For ∑N

maximum value recorded was 11.50 mg/L (spring and winter/ site V) and min. value 3.33 mg/L (summer/site V). For Cd, Pb, Ni and Cr observed values between 0.02mg/L (spring/site IV) to 0.70 mg/L (winter/site V), 0.15mg/L (spring/site I) to 1.09 mg/L(spring/site V), 0.13mg/L (rains/site IV) to 0.62 mg/L (summer/site V) and 0.03mg/L (spring/site II) to 0.46 mg/L (rains/site I & VI) respectively. Observed values of Cd, Pb and Cr were slightly higher than prescribed limits by EPAR (2012).

Table 1 Spatial and temporal descriptive statistics for physico-chemical variables

Variable	(Mean±S.E.)	Maximum	Minimum	CV (%)	MPL(BIS/EPAR)
DO	8.58±0.28	10.67	6.62	10.18	09
EC	222.58±13.10	388.08	153.72	18.55	-
Ph	7.36±0.08	8.16	6.75	3.49	8.50
TD	23.05±4.61	101.23	0.56	63.23	300
WT	23.30±0.93	31.08	15.05	12.54	-
AK	110.63±5.28	197.50	47.50	15.09	600
TH	92.72±4.44	140.50	59.33	15.15	600
Mg	9.24±1.26	24.31	3.67	43.29	30
TDS	111.46±9.82	246.27	41.99	27.86	2000
COD	74.46±4.70	159.58	32.92	19.98	250
BOD	3.72±0.67	19.80	0.00	56.57	30
N	6.57±0.34	11.50	3.33	16.37	-
Cd	0.36±0.04	**0.70**	0.02	37.30	0.01
Pb	0.48±0.0.05	**1.09**	0.15	31.23	0.05
Ni	0.27±0.02	0.62	0.13	26.06	3.0
Cr	0.12±0.01	**0.46**	0.03	45.07	0.1

Note: BIS (Bureau of Indian Standards); EPAR (Environmental (Protection) Amendment Rules); MPL-Maximum Permissible Limit; CV-Coefficient of variation. All parameters were measured in mg/L except WT-°C; pH – pH; EC - μS/cm.

3.1.1 Principal component analysis (PCA)

PCA/FA was applied on the mean values of data sets of 16 variables. Useful information and insignificant values i.e. eigen values less than 1 were eliminated through varimax rotation (Shreshtha and Kazama 2007; Yang *et al*. 2010; Nazar and Khan 2012; Wang *et al*. 2013; Bu *et al*. 2014; Kaur and Dua 2016). Components for PCA were retained as criteria given by Wang *et al*. 2013. Values obtained from PCA/FA were classified as strong (>0.75), moderate (0.75-0.50) and weak (0.50-0.30) on the basis of classification given by Yang *et al*. 2010. Number of principal components (PCs) for physicochemical parameters during spatial and temporal sampling was identified with the help of scree plots (Figure 1a and 1b). Spatial and temporal loadings of water quality variables are given in tables 3 and table 4.

Table 3 Spatial loadings of 16 variables for six sampling sites on four significant principal components and varifactors

Variable	*Component Matrix*				*Varimax Rotated Component*			
	PC1	PC2	PC3	PC4	VF1	VF2	VF3	VF4
DO	**-0.901**	0.208	0.287	-0.137	**-0.851**	-0.112	0.092	-0.459
EC	**0.756**	-0.462	-0.432	-0.162	**0.982**	-0.096	-0.127	0.093
pH	**-0.795**	0.422	0.261	0.133	**-0.930**	0.052	-0.023	-0.170
TD	**0.763**	-0.144	*0.586*	-0.036	0.474	0.167	*0.725*	0.412
WT	**0.829**	*0.511*	-0.213	-0.007	*0.519*	**0.764**	-0.118	0.355
AK	*0.693*	*0.549*	-0.249	-0.158	0.455	**0.786**	-0.141	0.156
TH	**0.912**	-0.242	0.088	-0.222	**0.858**	0.177	0.364	0.218
Mg	**0.883**	0.307	0.087	0.002	*0.541*	*0.602*	0.201	0.431
TDS	**0.943**	-0.232	-0.033	-0.215	**0.917**	0.195	0.255	0.217
COD	-0.098	-0.218	**0.962**	-0.096	-0.283	-0.201	**0.932**	0.038
BOD	-0.130	**-0.856**	-0.064	0.409	0.143	**-0.916**	-0.058	0.241
N	**0.877**	0.237	0.020	0.377	0.472	0.444	0.043	*0.739*
Cd	0.452	-0.155	**0.823**	-0.292	0.233	0.107	**0.957**	0.093
Pb	*0.646*	-0.129	0.332	*0.675*	0.242	-0.041	0.261	**0.933**
Ni	**0.955**	0.090	0.028	0.048	*0.692*	0.423	0.184	0.481
Cr	0.084	**0.889**	0.299	0.176	-0.453	**0.785**	0.093	0.294
Eigen value	8.601	2.899	2.611	1.078	6.244	3.570	2.723	2.652
Variance (%)	53.76	18.12	16.32	06.74	39.02	22.32	17.02	16.57
Cumulative (%)	53.76	71.88	88.20	94.30	39.03	61.34	78.36	94.93

Note: Strong loadings are denoted by bold format and moderate loadings are denoted by italicised format.

Table 4 Temporal loadings of 16 variables for four seasons on three significant principal components and varifactors

Variable	Component Matrix			Varimax Rotated Component		
	PC1	PC2	PC3	VF1	VF2	VF3
DO	0.289	0.105	**-0.952**	0.025	0.367	**0.930**
EC	**0.939**	-0.024	-0.343	**0.762**	0.486	0.428
pH	0.443	**0.769**	0.461	*0.743*	*-0.670*	-0.004
TD	0.425	**-0.847**	-0.320	0.077	**0.997**	-0.031
WT	-0.492	*0.738*	0.462	-0.127	**-0.980**	-0.153
AK	**0.897**	-0.351	0.269	**0.810**	*0.524*	-0.265
TH	**0.882**	-0.092	0.461	**0.919**	0.226	-0.323
Mg	-0.227	*-0.541*	**0.810**	-0.124	0.082	**-0.989**
TDS	**0.991**	0.054	0.123	**0.963**	0.265	0.057
COD	**0.833**	0.400	-0.383	**0.766**	0.099	*0.635*
BOD	**0.931**	0.314	0.185	**0.994**	-0.002	0.109
N	**0.875**	-0.358	-0.326	0.620	*0.742*	0.255
Cd	**-0.821**	*-0.571*	-0.032	**-0.917**	0.205	-0.342
Pb	**0.949**	-0.078	0.305	**0.940**	0.295	-0.169
Ni	-0.477	*0.753*	-0.454	-0.368	*-0.649*	*0.667*
Cr	**-0.964**	-0.256	-0.071	**-0.977**	-0.101	-0.189
Total	9.257	3.642	3.101	8.370	4.296	3.334
Variance (%)	57.86	22.76	19.38	52.31	26.86	20.84
Cumulative (%)	57.86	80.62	100.0	52.31	79.16	100.0

Note: Strong loadings are denoted by bold format and moderate loadings are denoted by italicised format.

3.1.2 Spatial PCA Analysis

Four components were retained for analysis as rest of components had eigen value less than 0. Out of 16 variables 10 variables contributed to the water quality of river Beas viz. DO, EC, pH, TD, WT, TH, Mg, TDS, N and Ni. These variables have strong loadings, Ni with the highest loading (>0.90). PC 1 explained 53.76% (eigenvalue-8.601) of total variance contributed by minerals & nutrients, organic and agricultural pollution. PC 2 and PC 3 explained 18.12% (eigen value-2.899) 16.32% of the variance (eigen value-2.611) resp. contributed by natural, agricultural and organic pollution. PC 4 explained 6.674% of the variance (eigen value-1.078) with weak loadings (>0.50)) except Pb with moderate loadings. For obtaining more significant results varimax rotation was done that resulted in four varifactors (VF). VF1 39.02% (Eigen value 6.244) explained less variance as compare to PC1. This result is supported by Septong and Chuersuwan (2015) who worked on spatial variation in water quality of the Lower Lam Takhong River. VF1 showed positive loadings on EC, TH and TDS. DO and pH had negative loadings. These loadings justify that variation in river water quality was explained by nutrients, agricultural runoff and domestic wastewater. Our results are in accordance with work done on water quality of Songhua River (Wang *et al.* 2012). VF2 explained 22.32 % variance (eigen value-3.570) with strong positive loadings on WT, AK, Cr and negative loadings on BOD. VF3 explained 17.02% (eigen value-2.723) of the variance with strong COD and Cd loadings contributed by natural and organic pollution. VF4 explained 16.57% variance (eigen value-2.652). Biplot of spatial PCA loadings indicated that water quality of site V was different from other five sites and variance was contributed mainly by TD, TDS, COD, AK, Mg, TH, WT, BOD and EC (Figure 1c).

3.1.3 Temporal PCA Analysis

For temporal PCA three components were retained and explained 100% of variance. PC1 explained 57.86% (eigen value-9.257) of the variance and had strong positive loadings on EC, AK, TH, TDS, COD, BOD, N & Pb and negative loadings on Cd and Cr. Similar findings were made by Kaur and Dua (2016) while studying water quality of Tung Dhab Drain. PC2 had positive loadings on pH and negative loading on TD. PC3 explained 19.38% variance (eigen value-9.257) with positive loadings on DO and Mg. Negative TD loadings and positive DO loadings were also observed by Singh et al. (2004). VF1 explained 52.31% variance (eigen value-8.370) with loadings on EC, AK, TH, TDS, COD, BOD, Cd, Pb and Cr, contributed by natural, agricultural, organic and industrial pollution. VF2 explained 26.86% of the variance (eigen value-8.370) and had strong loadings on TD and WT (>0.90). VF3 explained 20.84% of the variance (eigen value-3.334) with negative DO loadings and positive Mg loadings. Biplot (figure 1d) for temporal loadings revealed that amongst all the seasons water quality of summer and rains was different from winter and spring seasons and variation was contributed by N, WT, DO, BOD, COD, TD, AK, TH, TDS and Mg.
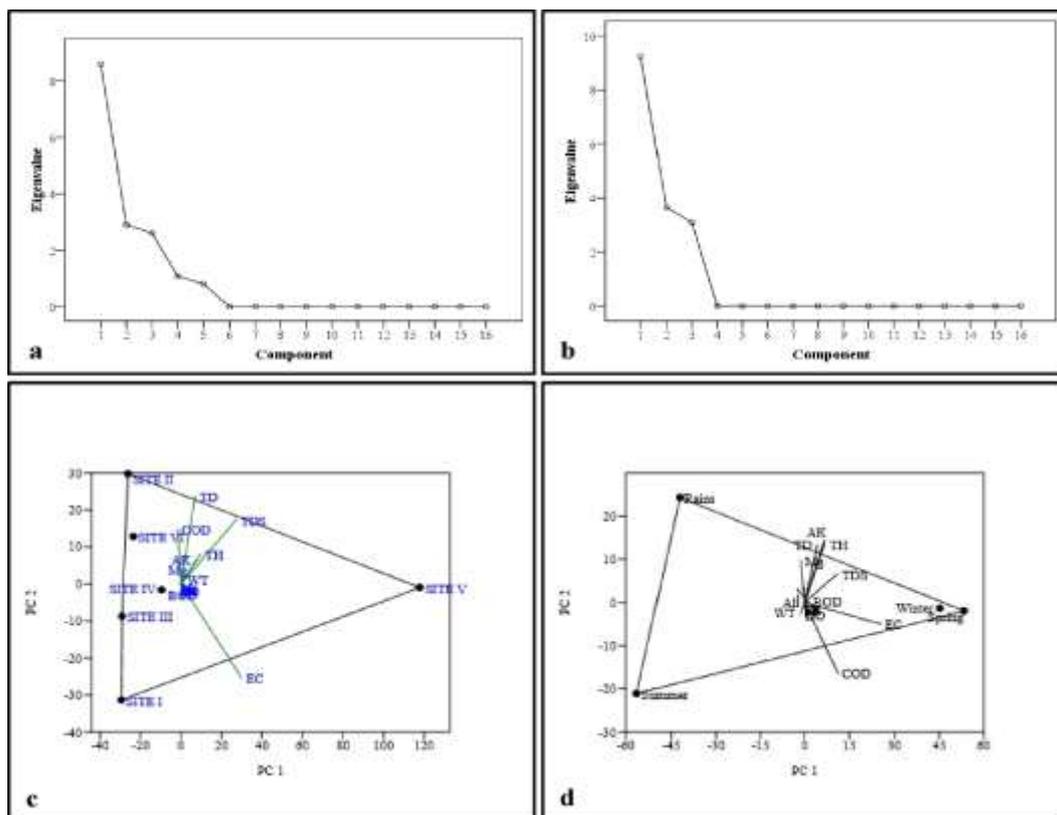
Figure 1 a-scree plot for spatial loadings; b-scree plot for temporal loadings; c-biplot for spatial PCA (PC1+PC2); d- biplot for temporal PCA (PC1+PC2).

3.1.2 Two way ANOVA

Two way ANOVA was applied on raw data obtained from six sampling sites during four seasons for two years. Values obtained were further subjected to Tukey's HSD test for analysis of grouping among the sampling sites and seasons (VishnuRadhan et al. 2017). Results yielded from Two Way ANOVA are shown in Table 5. R-square represents percentage variation in the response variable, explained by its relationship with other variables. Spatial and temporal mean value comparisons and groupings are given in Table 5. Interactions plots for various sampling sites and seasons are given in Figure 2 and Figure 3. For DO significant variation between seasons ($F_{3, 5}$ = 88.24, $p < 0.05$) and sites ($F_{3, 5}$ = 15.15, $p < 0.05$) were observed. R^2 statistic suggested that predictor variables (seasons and sites) explained 75.13% of variation is in response variable (DO). Interaction plot resulted in roughly parallel lines that indicated absence of interaction between seasons and sites. EC showed significant variation for all sources of variations i.e. seasons ($F_{3, 5}$ = 32.93, $p < 0.05$), sites ($F_{3, 5}$ = 18.41, $p < 0.05$) and interaction (seasons*sites) ($F_{3, 5}$ = 6.01, $p < 0.05$) (R^2= 70%). pH revealed significant variation for all sources of variations i.e. seasons ($F_{3, 5}$ = 18.55, $p < 0.05$), sites ($F_{3, 5}$ = 3.90, $p < 0.05$) and interaction (seasons*sites) ($F_{3, 5}$ = 2.31, $p < 0.05$) (R^2= 47.79%). Significant variation for all sources of variations i.e. seasons ($F_{3, 5}$ = 22.73, $p < 0.05$), sites ($F_{3, 5}$ = 14.92, $p < 0.05$) and interaction (seasons*sites) ($F_{3, 5}$ = 7.46, $p <0.05$) were observed for TD (R^2= 67.98%). All sources of variations were found to be significant for temperature i.e. seasons ($F_{3, 5}$ = 155.18, $p < 0.05$), sites ($F_{3, 5}$ = 17.50, $p < 0.05$) and interaction (seasons*sites) ($F_{3, 5}$ = 5.23, $p < 0.05$) (86.85%). AK showed significant variations for seasons ($F_{3, 5}$ = 5.49, $p < 0.05$), sites ($F_{3, 5}$ = 6.07, $p < 0.05$) and interaction (seasons*sites) ($F_{3, 5}$ = 3.88, $p < 0.05$) (R^2= 46.67%). For TH, significant variations were observed for all the sources of variance i.e. seasons ($F_{3, 5}$ = 45.78, $p < 0.05$), within sites ($F_{3, 5}$ = 23.68, $p < 0.05$) and interaction ($F_{3, 5}$ = 3.95, $p < 0.05$) (R^2= 72.41%). For Mg, seasons ($F_{3, 5}$ = 86.05, $p < 0.05$) and sites ($F_{3, 5}$ = 5.55, $p < 0.05$) showed significant variations (R^2= 72.08%). For TDS significant variations in all sources i.e. seasons ($F_{3, 5}$ = 28.12, $p < 0.05$), sites ($F_{3, 5}$ = 59.70, $p < 0.05$) and interaction (seasons*sites) ($F_{3, 5}$ = 10.82, $p < 0.05$) (R^2= 81.96%) were observed. COD showed significant variations for all the sources of variance i.e. seasons ($F_{3, 5}$ = 112.03, $p < 0.05$), within sites ($F_{3, 5}$ = 13.86, $p < 0.05$) and interaction ($F_{3, 5}$ = 28.24, $p < 0.05$) (R^2= 87.35%). BOD indicated significant variations for all the sources of variance i.e. seasons ($F_{3, 5}$ = 75.56, $p < 0.05$), within sites ($F_{3, 5}$ = 13.18, $p < 0.05$) and interaction ($F_{3, 5}$ = 26.82, $p < 0.05$) (R^2= 85.87%). For $\sum N$, significant variations for all the sources of variance i.e. seasons ($F_{3, 5}$ = 22.94, $p < 0.05$), within sites ($F_{3, 5}$ = 15.84, $p < 0.05$) and interaction ($F_{3, 5}$ = 9.39, $p < 0.05$) (R^2= 70.65%). Cd had significant variations in all the sources of variance i.e. seasons ($F_{3, 5}$ = 1076.61, $p < 0.05$), sites ($F_{3, 5}$ = 69.70, $p < 0.05$) and interaction ($F_{3, 5}$ = 49.59, $p < 0.05$) (R^2= 97.30%). Significant variations for all the sources of variance were observed for Pb i.e. seasons ($F_{3, 5}$ = 5.51, $p < 0.05$), sites ($F_{3, 5}$ = 6.12, $p < 0.05$) and interaction ($F_{3, 5}$ = 1.48, $p < 0.05$) (R^2= 97.30%). For Ni significant variations in all the sources of variance i.e. seasons ($F_{3, 5}$ = 12.36, $p < 0.05$), sites ($F_{3, 5}$ = 14.09, $p < 0.05$) and interaction ($F_{3, 5}$ = 1.48, $p < 0.05$) (R^2= 68.00%). Cr showed significant variations in all the sources of variance i.e. seasons ($F_{3, 5}$ = 7.48, $p < 0.05$), sites ($F_{3, 5}$ = 2.76, $p < 0.05$) and interaction ($F_{3, 5}$ = 3.93, $p < 0.05$) (R^2= 44.22%).

Table 5 Spatial and temporal analysis of variance of physicochemical variables (two way ANOVA).

| Parameter	Source of variation	F-value	P-value	R^2 (%)
	Between seasons	88.24*	0.000	75.13
DO	Within sites	15.15*	0.000	
	Interaction	1.47ns	0.129	
	Between seasons	32.93*	0.00	70.08
EC	Within sites	18.41*	0.00	
	Interaction	6.01*	0.00	
	Between seasons	18.55*	0.000	47.79
pH	Within sites	3.90*	0.003	
	Interaction	2.31*	0.006	
	Between seasons	22.73*	0.000	67.98
TD	Within sites	14.92*	0.000	
	Interaction	7.46*	0.000	
	Between seasons	155.18*	0.000	86.85
WT	Within sites	17.50*	0.000	
	Interaction	5.23*	0.000	
	Between seasons	5.49*	0.001	46.67
AK	Within sites	6.07*	0.000	
	Interaction	3.88*	0.000	
	Between seasons	45.78*	0.000	72.41
TH	Within sites	23.68*	0.000	
	Interaction	3.95*	0.000	
	Between seasons	86.05*	0.000	72.08
Mg	Within sites	5.55*	0.000	
	Interaction	1.60ns	0.085	
	Between seasons	28.12*	0.000	81.96
TDS	Within sites	59.70*	0.000	
	Interaction	10.82*	0.000	
	Between seasons	112.03*	0.000	87.35
COD	Within sites	13.86*	0.000	
	Interaction	28.24*	0.000	
	Between seasons	75.56*	0.000	85.27
BOD	Within sites	13.18*	0.000	
	Interaction	26.82*	0.000	
	Between seasons	22.94*	0.000	70.65
N	Within sites	15.84*	0.000	
	Interaction	9.39*	0.000	
	Between seasons	1076.6*	0.000	97.30
Cd	Within sites	69.7*	0.000	
	Interaction	49.6*	0.000	
	Between seasons	5.51*	0.000	51.96
Pb	Within sites	6.12*	0.000	
	Interaction	1.48*	0.002	
	Between seasons	12.36*	0.000	68.16
Ni	Within sites	14.09*	0.000	
	Interaction	9.96*	0.000	
	Between seasons	7.48*	0.000	44.22
Cr	Within sites	2.76*	0.000	
	Interaction	3.93*	0.000	

Note: * = significant at 0.05%, ns = non-significant
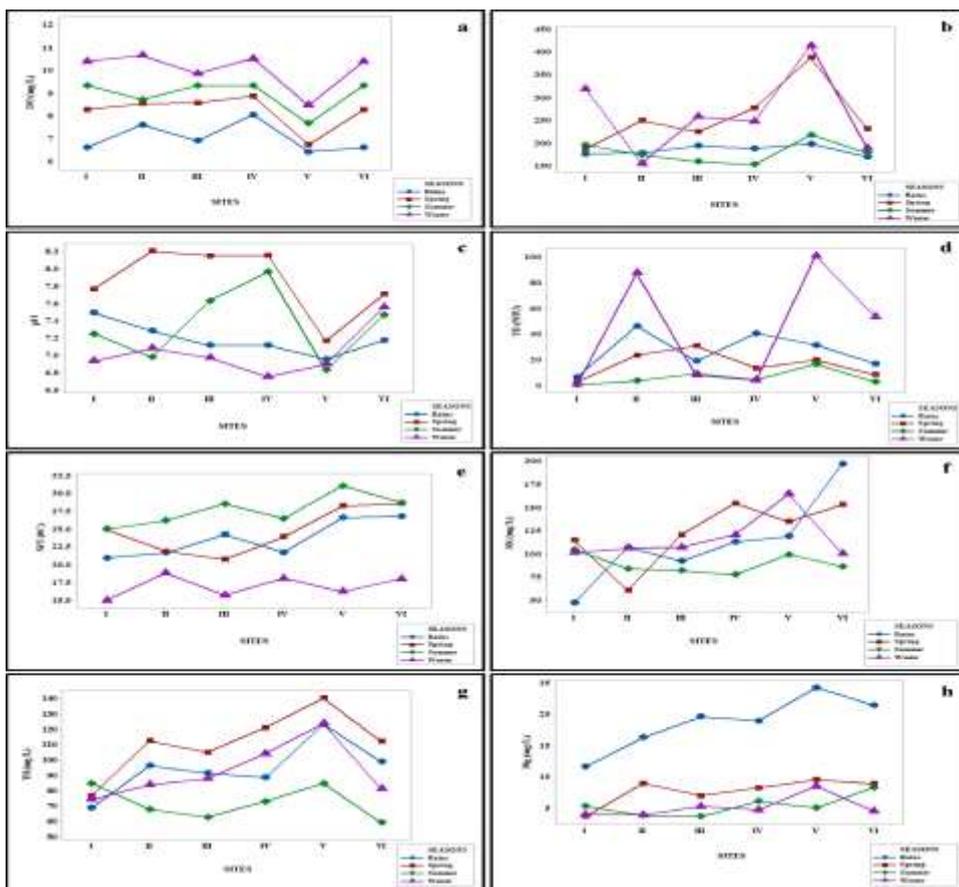
Figure 2 a-h Interaction plots of six sampling sites during four seasons
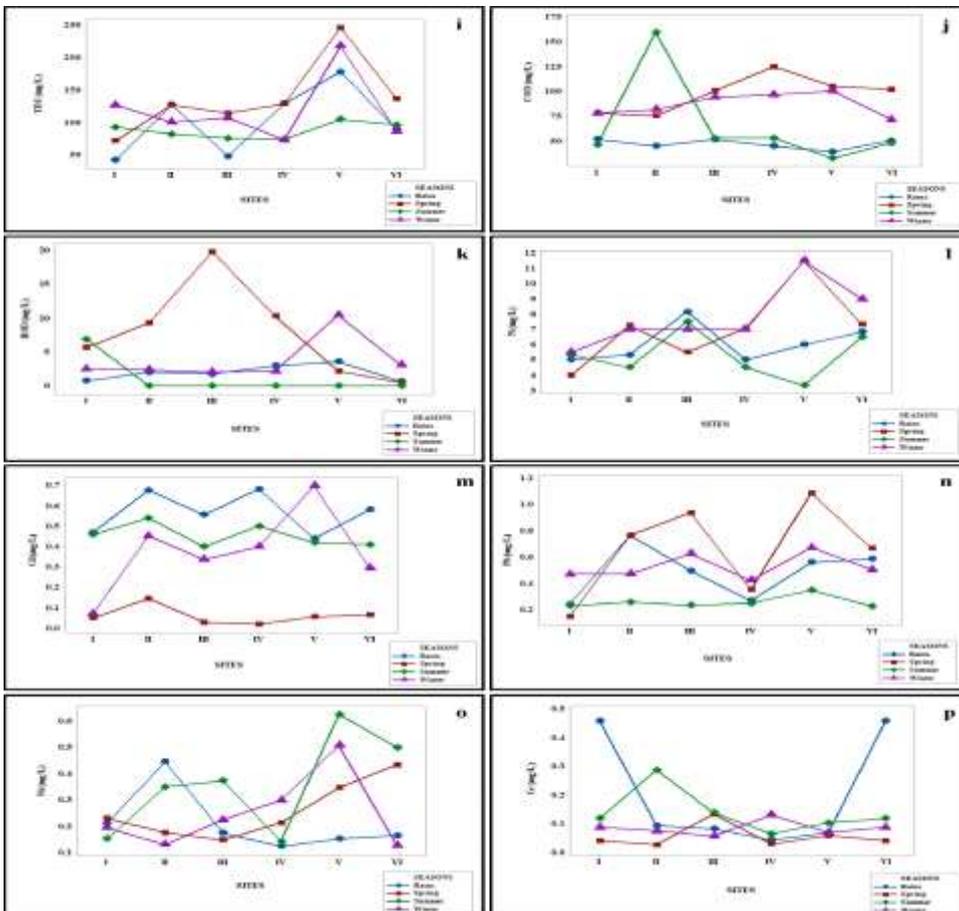

Figure 3 i-p Interaction plots of six sampling sites during four seasons

4. Conclusion

With increasing anthropogenic activities and climatic changes in recent years, river water quality has become an issue of concern and there is need to incorporate multivariate statistical techniques for analysis of quality of the water. For this study principal component analysis/ Factor analysis (PCA/FA) were applied that reduced huge data in small valuable data set. Spatial PCA revealed that water of site V was a little bit more polluted than other sites mainly due to strong loadings of TD, TDS, COD, AK, Mg, TH, WT, BOD and EC. Temporal PCA indicated that water quality of summer and rains was different from winter and spring seasons and variation was contributed by N, WT, DO, BOD, COD, TD, AK, TH, TDS and Mg. Two way ANOVA revealed that all the physicochemical water quality variables showed significant variance in all the sources of variations viz. seasons, sites and interaction except DO and Mg that did not had significant variance for interaction. Results indicated that water of River Beas is of good quality and fit for human consumption and aquatic life except for the site V which should be taken into consideration for improvement of water quality and pollution therein. This study suggests that by applying multivariate statistical techniques we can interpret the large cluttered chemometric data into cost effective sampling model approach to predict relationship between water quality and other parameters.

References
1. APHA/AWWA/WEF 2005. Standard methods for examination of water and wastewater. Vol. 21, American Public Health Association, American Water Works Association, Washington, USA.
2. BIS 1991. Indian Standard Water Specification. Bureau of Indian standard.
3. Bu, H., Wan J., Meng, W. and Zhang, Y. 2014. Relationships between land use patterns and water quality in the Taizi River basin, China. Ecological Indicators, 41: 187–197.
4. Chumkiew, S., Jaroensutasinee, C.S. and Jaroensutasinee, M. 2015. Assessment of spatio-temporal variations in water quality of Bandon Bay, Thailand. Indian Journal of Marine Sciences, 44(7), 1000-1010.
5. Das, J. and Acharya, B.C. 2003. Hydrology and assessment of lotic water quality in Cuttak city, India. Water Air and Soil Pollution, 150, 163-175.
6. Dunteman, G.H. 1989. Principal Components Analysis. Sage, Thousand Oaks, ISBN: 9780803931046.
7. EPAR Report 2012. General standards for discharge of environmental pollutants. Part A: Effluents into Inland waters. Environmental (Protection) Amendment Rules, G.S.R. 152 (E). Ministry of Environment and Forest, India.
8. Gholizadeh, M.H., Melesse, A.S. and Redddi, L. 2016. Discriminant analysis application in spatiotemporal evaluation of water quality in South Florida. Journal of Hydroinformatics, 18 (6), 1019–1032.
9. Gupta, S.P. 2007. Statistical Methods. Sultan Chand & Sons. India. pp 144-200.
10. Hamzah, F.M., Jaafar, O., Jani, W.N.F.A. and Abdullah, S.M.S. 2016. Multivariate Analysis of Physical and Chemical Parameters of Marine Water Quality in the Straits of Johor, Malaysia. Journal of Environmental Science Technology, 9, 427-436.
11. Hart, A.I. and Zabbey, N. 2005. Physico-chemical and benthic fauna of Woji Creek in the Lower Niger Delta, Nigeria. Environment and Ecology, 23(2), 361-368.
12. Isiyaka, H.A. and Juahir, H. 2015. Analysis of surface water pollution in the Kinta River using multivariate technique. Malaysian Journal of Analytical Sciences, 19 (5), 1019-1031.
13. Kaur, R. and Dua, A. 2014. Application of Multivariate Statistical Analysis in Assessing Surface Water Quality of Chanera-1 Reservoir on River Ravi. International Journal of Sciences and Research, 3(10), 1966-1971.
14. Kaur, R. and Dua, A. 2016. Assessment of Water Quality of Tung Dhab Drain-An International Water Channel-Using Multivariate Statistical Techniques. Asian Journal of Water Environment and Pollution, 13(3), 85-93.
15. Liu, C.W., Lin, K.H. and Kuo, Y.M. 2003. Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. Science of Total Environment, 313(1), 77-89.
16. Manjare, S.A. Vhanalakar, S.A. and Muley, D.V. 2010. Analysis of Water Quality Using Physico-Chemical Parameters Tamdalge Tank in Kolhapur District, Maharashtra. International Journal of Advanced Biotechnology and Research, 1(2), 115-119.
17. Mazurek, M., Paluszkiewicz, R., Rachlewicz G. and Zwolinski Z. (2012). Variability of Water Chemistry in Tundra Lakes, Petuniabukta Coast, Central Spitsbergen, Svalbard. Scientific World Journal, 2012, 1-13.
18. Mustapha, M. K. 2008. Assessment of the Water Quality of Oyun Reservoir, Offa, Nigeria, Using Selected Physico-Chemical Parameters. Turkish Journal of Fishery and Aquatic Science, 8, 309-319.
19. Najar, I.A. and Khan, A.B. 2012. Assessment of water quality and identification of pollution sources of three lakes in Kashmir, India, using multivariate analysis. Environmental Earth Science, 66(8), 2367-2378.
20. Ogwueleka, T.C. 2014. Assessment of the water quality and identification of pollution sources of Kaduna River in Niger State (Nigeria) using exploratory data analysis. Water and Environment, 28(1), 31-37.
21. Priya, S., Das, S.S.M. and Vareethiah, K. 2016. Analysis of Water Quality in Selected Stations alone River, Tambaraparani Kanyakumari District, Tamilnadu, India. International Journal of Innovative Science Engineering and Technology, 3(8), 319-323.

22. Razmkhaha, H. Abrishamchi, A. and Torkian, A. 2010. Evaluation of spatial and temporal variation in water quality by pattern recognition techniques: A case study on Jajrood River (Tehran, Iran). Journal of Environment and Management, 91: 852–860.

23. Septong, C. and Chuersuwan, N. 2015. Evaluation of Spatial Variation in Water Quality of the Lower Lam Takhong River, Thailand. 3rd International Conference on Biological, Chemical & Environmental Sciences. Kuala Lumpur (Malaysia), 106-109.

24. Shinde, S.E., Pathan T.S., Raut, K.S. and Sonawane, D.L. 2011. Studies on the Physio-Chemical Parameters and Correlation coefficient of Harsool-Savangi Dam, District Aurangabad, India. Middle East Journal of Scientific Research, 8(3), 544–554.

25. Shrestha, S. and Kazama, F. 2007. Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. Environmental Modelling Software, 22(4), 464-475.

26. Simpi, B., Hiremath, S.M., Murthy, K.N.S., Chandrashekhrappa, K.N., Patil, A.N. and Puttiah, E.T. 2011. Analysis of water quality using Physico-Chemical parameters in Hosahalli Tank in Shimoga District, Karnataka, India. Global Journal of Science Frontier Research, 3(11), 31-33.

27. Singh, K.P., Malik, A., Mohan, D. and Sinha, S. 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) a case study. Water Research, 38, 3980-3992.

28. Solanas, A., Manolov, R. and Leiva, D. 2011. Retaining principal components for discrete variables. Anuario de Psicologia Juridica, 41(1-3), 33-50.

29. Trivedi, R.K. and Goel, P.K. 1986. Chemical and Biological Methods for Water Pollution Studies. Environmental Publications, Karad, India. pp-56-78.

30. Varol, M., Gokot, B., Bekleyen, A. and Sen, B. 2012. Spatial and Temporal Variations in Surface Water Quality of the Dam Reservoirs in the Tigris River Basin, Turkey. Catena, 92: 11-21.

31. VishnuRadhan, R., Zainudin, Sreekanth, G.B., Dhiman, R., Salleh, M.N. and Vethamony, P. 2017. Temporal water quality response in an urban river: a case study in peninsular Malaysia. Applied Water Science, 7, 923-933.

32. Wang, X., Cai, Q., Ye, L. and Qu, X. 2012. Evaluation of Spatial and Temporal variation in stream quality by multivariate statistical techniques: A case study of the Xiangxi River basin, China. Quaternary International, 282: 137-144.

33. Wang, Y., Wang, P., Bai, Y., Tian, Z., Li, J., Shao, X., Mustavich, L.F. and Li, B.L. 2013. Assessment of surface water quality via multivariate statistical techniques: A case study of the Songhua River Harbin region, China. Journal of Hydro-environment Research, 7, 30-40.

34. Yang, Y.H., Zhou, F., Guo, H.C., Sheng, H., Liu, H., Dao, X. and He, C.J. 2010. Analysis of spatial and temporal water pollution patterns in Lake Dianchi using 231 multivariate statistical methods. Environment Monitoring and Assessment, 170: 407–416.