www.cafetinnova.org

ISSN 0974-5904, Volume 10, No. 02

DOI:10.21276/ijee.2017.10.0236

International Journal of Earth Sciences and Engineering

April 2017, P.P. 383-387

Use of Multivariate Analytical Methods in Assessment of River Water Quality

SIDDHARTHA SHARMA

Department of Civil Engineering, Maharishi Markandeshwar University, Mullana, Ambala-133207, Haryana, India

Email: *siddharthas79@gmail.com*

Abstract: *This study is focused on the assessment of water quality of river Satluj in North Indian state of Punjab and evaluation of 34 physico-chemical variables monitored during the period 2015–2016, at 3 different sampling locations. Multivariate analytical techniques, such as Principal Component Analysis (PCA)/ Factor Analysis (FA) were applied to the water quality data set to identify characteristics of water quality in the studied catchment. PCA/FA was applied for source identification to data sets pertaining to 3 spatial groups (upper catchment, middle catchment and lower catchment) responsible for the data structure. These factors are conditionally named soil structure and soil erosion; domestic, municipal and industrial effluents; agricultural activities (fertilizers, livestock waste etc.) and seasonal effect factors.In the current study usefulness of multivariate analysis for evaluation of river Satluj water quality assessment and identification of dominant factors and pollution sources for effective water quality management and determination of spatial and temporal variations in water quality illustrated.*

Keywords: *Surface water quality, Multivariate statistical analysis, Satluj river*

1. Introduction

The surface water quality is a matter of serious concern today [1]. Rivers carry the municipal and industrial wastewater and run-off from agricultural land in their vast drainage basins and are the most vulnerable to pollution [2]. Human activities significantly decrease water quality and river inflows contribute pollutants to most riverine wetlands, thereby creating serious ecological problems [3]. For control of pollution effectively and water resources management, interpretation of a large number of monitoring data is required [1]. For data structure determination, its classification, time trends and contribution of pollution sources identification etc. envirometric methods are recommended [4].

In recent years accelerated development of industries and urbanization activities in the Satluj river catchment of Punjab (India) region caused the major water quality deterioration problems in the middle and lower catchments in the river Satluj basin.

A large monitoring data, obtained from the monthly monitoring of river Satluj during the period (2015-16) is subjected to multivariate analytical techniques in the present study for source identification and determination of variables responsible for spatial and temporal variations. The final results are expected to be helpful for the optimization of river monitoring plan.

2 Study Area

The river Satluj is one of the five great rivers of Punjab. It has its catchment in the hills of Himachal Pradesh. The river Satluj originates near Dharma Pass as the river Zaskar and flow through Tibetan Plateau. It then cuts through the Himalayan Range, entering plains near Nangal in Punjab. The present study has been conducted on the 200 km. stretch of river Satluj between Nangal Dam and Harike wetland, lying between 31°45'N 74°57'E and 30°45' N 76°50' E. The watershed area is approximately 10,882 Km^2. Swan, Sirsa, Siswan, East Bein, BudhaNala etc. are some of the important tributaries which drain water as well as pollution loads to the river Satluj in this stretch. Nangal and Ropar drain in upper catchment, Phillaur and Ludhiana drain in middle catchment, Phagwara, Jallandhar, Nawashahar and Hoshiarpur drain in lower catchment and are the major urban settlements of the watershed that contribute pollution loads in the form of effluents to the river Satluj.

The canal system exploits most of freshwater and provides irrigation. From the confluence of Satluj and Beas at Harike headworks, two canals, Rajasthan Feeder and Sirhind Feeder take off. This headworks is basically responsible for Harike wetland system, which is a man-made riverine freshwater Ramsar wetland site of international importance. Deteriorating water quality and water hyacinth infestation have been the major problems associated with this wetland system. Heavy pollution loads delivered by the river Satluj have been mainly responsible for this rapid deterioration.

2.1Monitoring Sites

In the present study total 3 stations, namely Nangal Head Works (SAT-1), Phillaur d/s (SAT-2) and

Received: December 01, 2016; Accepted: February 28, 2017; Published: April 26, 2017
International Journal of Earth Sciences and Engineering, 10(02), 383-387, 2017, DOI:10.21276/ijee.2017.10.0236

Yusufpur (SAT-3) were selected on the river Satluj. The first site (SAT-1) is representing relatively lesser river pollution. Next site (SAT-2) is upstream to famous industrial city of Ludhiana and moderate in river pollution. The last site (SAT-3) is in the downstream region to Ludhiana and severely polluted. Monitoring stations are seen at Figure 1.

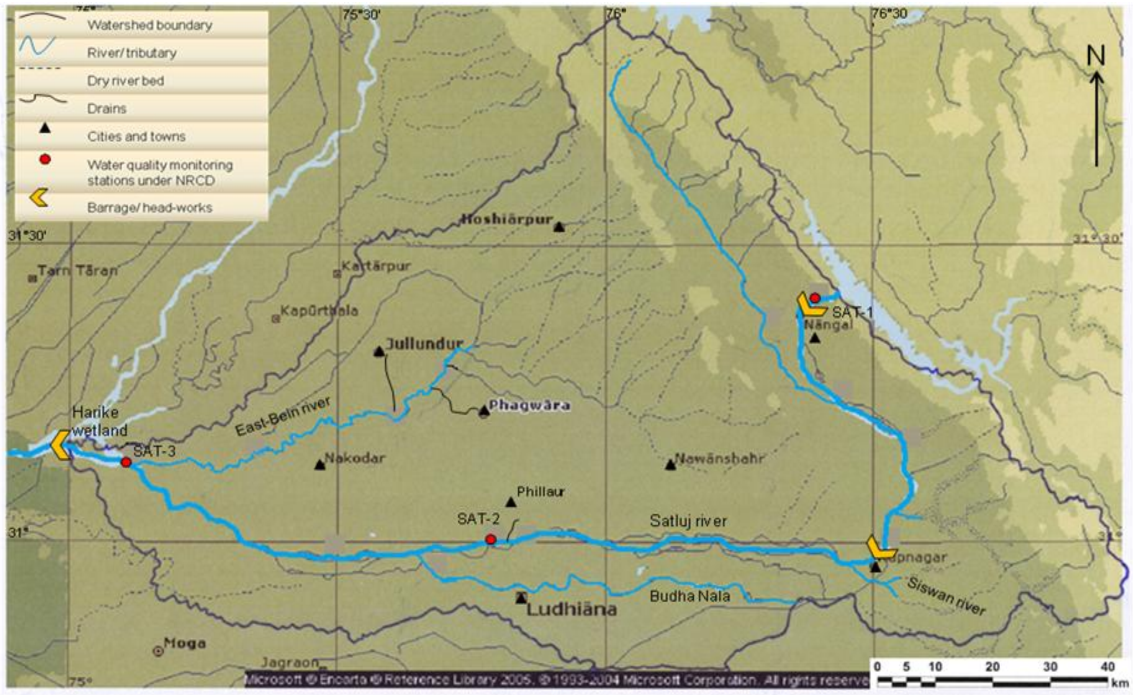

Figure 1: *Map of Satluj river watershed with monitoring stations indicated on it. (Source: Sharma S., 2017, Ph.D. Thesis, Thapar University, Patiala, India)*

2.2 Sampling and Chemical Analysis

Water samples were collected on monthly basis at the 3 sites during the (2015-16). Sampling procedures, preservation techniques and transportation of the water samples to the laboratory were followed as per APHA standard methods [5]. River flow data was collected from Punjab Water Supply and Sewerage Board office, Mohali. Water temperature and DO were measured on-site. Standard protocols [5] were followed for analysis of all the 34 parameters, namely Temperature of Water, DO, pH, Electrical Conductivity, Total Solids, Total Dissolved Solids, Suspended Solids, Turbidity, Flow, Total Hardness, Calcium Hardness, Magnesium Hardness, Total Alkalinity, Chloride, Sulphate, Total Phosphorous, Nitrate+Nitrite Nitrogen, Ammonical Nitrogen, Organic Nitrogen, BOD_5, COD, Total Coliform, Fecal Coliform, Enterococcus, Sodium, Calcium, Magnesium, Iron, Manganese, Copper, Zinc, Nickel, Lead and Chromium.

3 Data Analysis and Multivariate Statistical Methods

In this study, multivariate statistical analysis of the river water quality data set was performed through Principal Components Analysis/ Factor Analysis (PCA/FA) techniques [6]. The variables were standardized first using z-transformation in order to avoid wide difference in data dimensionality [7] and so PCA/FA was calculated further on it.

Excel 2003 used for all the general mathematical calculations. GraphPad Prism 4 used for normalization and statistiXL 1.7 used for multivariate calculations.

3.1 Principal Component and Factor Analysis

PCA/FA was performed on correlation matrix of transformed data set while rotating with Varimax rotation method. The PCA of the monitored data set indicated that, most of the data variance is contained in the first 2 components. Hence, the Biplots were developed between PC1 and PC2 for pattern recognition and for discussion on probable sources and types of pollution.

Interpretation of PCs involved evaluation of the loadings which may be associated with any geochemical and biological processes and/or may indicate some sources of water body contamination. High loading of any variable means that the concerned variable may be important to a given process or contamination source and low loadings for any variables means they are not that much important. In this way one logically associates the given factor with the given source. Hence, pollution sources can be identified with some subjective logics.

Any factor having eigenvalue greater than 1 was accepted as significant [3]. Any variable having loading value >0.5 was considered as the most representative of that factor.

Table 1: *Results of water quality parameters at different locations of the Satluj river (2015-2016)*

Parameters	SAT-1	SAT-2	SAT-3
Water Temperature	18.8±0.9	21.8±0.5	24.1±0.3
DO	8.31±0.3	7.57±0.5	2.21±1.3
pH	7.62±0.18	7.72±0.12	7.51±0.15
Conductivity	197±27	278±53	556±61
TS	186±20	346±43	617±102
TDS	149±17	200±29	401±66
SS	38±14	146±46	215±67
Ammonical N	0.64±0.34	1.05±0.78	5.87±1.88
Organic N	0.60±0.39	0.71±0.30	1.66±1.03
Nitrate+Nitrite N	0.61±0.17	0.68±0.17	1.68±0.83
Total Phosphorous	0.12±0.11	0.27±0.3	0.82±0.21
BOD_5	1.2±0.3	2.7±1.3	23.4±11.7
COD	9.5±3.1	24.6±.7	119.7±48.1
Chloride	6.6±1.5	14.5±7.0	70.2±14.1
Turbidity	12.9±1.0	51.9±19.3	57.9±23.4
Total Hardness	99.7±6.8	110.2±10.1	161.4±8.2
Calcium Hardness	69.5±1.7	78.8±1.6	105.4±5.5
Magnesium Hardness	32.7±4.7	34.9±8.0	58.5±2.5
Total Alkalinity	63.2±3.9	78.3±4.1	156.1±22.3
Sulphate	36.5±5.4	47.3±22.4	72.2±34.4
Flow	511±54	83±45	115±102
Total Coliform	61,336±93,007	197,979±99,623	4,642,683±5,058,941
Fecal Coliform	1,682±1,142	26,710±36,085	324,504±216,618
Enterococcus	394±212	6,405±7,766	117,947±153,131
Na	7.7±3.5	8.0±4.3	20.4±5.2
Ca	26.8±2.7	26.7±5.9	36.8±12.4
Mg	8.1±3.2	9.2±4.5	12.8±4.1
Fe	1.1±0.9	3.3±2.0	3.7±2.4
Mn	0.19±0.09	0.30±0.08	0.28±0.12
Cu	0.24±0.25	0.19±0.20	0.17±0.18
Zn	0.32±0.26	0.31±0.24	0.38±0.31
Ni	0.11±0.09	0.13±0.12	0.12±0.11
Cr	0.11±0.12	0.14±0.17	0.21±0.22
Pb	0.16±0.12	0.17±0.14	0.19±0.18

Note: Results are written as Value ± Standard Deviation. All the parameters are in mg/l except pH; Water Temperature in °C; EC in µS/cm; Turbidity in NTU; TC, FC and Ent in MPN/100 ml; Flow in m^3/s.

4 Results and Discussion

The basic statistics of 34 parameters of water quality at 3 monitoring sites of the river Satluj are given in Table-1.

4.1 Principal Component Analysis/ Factor Analysis for Source Identification

PCA/FA was applied on the standardized data (34 variables) for each of the sampling location individually. Table 2, 3 and 4 are having summaries of the PCA/ FA results. Here variance contributed by each Factor and their cumulative variance contribution are mentioned along with accepted eigenvalues. The significant parameters are indicated with their corresponding loadings.

PCA/FA of SAT-1 data set evolved that first 2 factors are explaining 49.13% of the total variance in respective water quality data as seen in Table 2. Factor-1 indicating strong relation on DO & Total Hardness with positive loading values; negative loadings on Total Coliform, Fecal Coliform & Enterococcus; moderate relation with positive loadings on Conductivity, Chloride & Calcium; Water Temperature, Turbidity, Calcium Hardness, Flow and Nickel are having negative loadings. This factor is explaining 34.9% of total variance. This factor basically represents seasonal impacts, soil structure of the catchment and municipal waste water discharges into the river. This station being the water reservoir, the water quality is less influenced by riverine flow as this factor is having moderate negative loading on flow.

Table 2: *Factor Analysis for SAT-1*

Explained Variance	Factors				
	Fa1	**Fa2**	**Fa3**	**Fa4**	**Fa5**
	Fecal Coliform (-0.94)	Cr(0.93)	TS(0.95)	Na(0.90)	Organic N(0.98)
	DO(0.87)	Mg(-0.81)	SS(0.94)	Fe(0.81)	COD(0.71)
	Total Coliform (-0.81)	Ammonical N (-0.63)	Pb(0.81)	Sulphate(0.77)	
	Enterococcus (-0.80)	$NO_2 + NO_3$ N(0.59)	Magnesium Hardness(-0.64)		
	Total Hardness (0.75)	pH(0.59)			
	Water Temperature (-0.73)				
	Flow (-0.71)				
	Turbidity (-0.71)				
	EC (0.69)				
	Chloride (0.67)				
	Calcium Hardness (-0.59)				
	Ca (0.58)				
	Ni (-0.58)				
Eigenvalue	11.86	4.84	3.99	3.01	2.54
% of Var.	34.90	14.23	11.74	8.85	7.47
Cum. %	34.90	49.13	60.86	69.71	77.18

Note: Variables are indicated with their loadings in brackets.

Factor-2 accounting for 14.23% of total variance and this factor is showing strong relation with positive loadings on Chromium and negative loadings on Magnesium; moderate relation with positive loadings on pH & Nitrate+Nitrite Nitrogen and negative loadings on Ammonical Nitrogen. This factor indicated the impact of agricultural run-off activity.

PCA/FA of SAT-2, revealed that first 2 factors are explaining 47.79% of the total variance (Table-3). Factor-1 explaining 30.81% of total variance and has strong positive relation with Total Coliform; moderate positive relation with Fecal Coliform and negative relation with Chloride. This factor shows the severe impact of municipal sewage.

Table 3: *Factor Analysis for SAT-2*

Explained Variance	Factors				
	Fa1	Fa2	Fa3	Fa4	Fa5
	Total Coliform(0.84)	DO(-0.92)	Mg (-0.93)	Enterococcus (-0.88)	Flow(0.88)
	Fecal Coliform(0.68)	Water Temperature(0.90)	Zn(0.83)	Cu(0.78)	Calcium Hardness(-0.76)
	Chloride(-0.63)	Ni(0.71)	$NO_2 + NO_3$ N(0.70)	BOD(0.71)	Ca(0.44)
		Total Phosphorous(0.69)	Mn(0.58)		
		Organic N(0.58)	Fe(-0.58)		
		Total Hardness(-0.57)			
Eigenvalue	10.47	5.77	4.27	3.09	2.49
% of Var.	30.81	16.98	12.56	9.08	7.33
Cum. %	30.81	47.79	60.35	69.43	76.75

Note: Variables are indicated with their loadings in brackets.

Factor-2 explaining 16.98% of total variance. This factor was positively related to Water Temperature; negatively to DO; moderately related to Org-N, Total P, Total Hardness & Nickel. This factor shows the impact of seasonal variations and agricultural run-off.PCA/FA of SAT-3, is explaining 55.67% of the total variance as seen in Table-4. Factor-1 has accountedfor 40.27% of total variance. It has strong variable relation on Flow; moderate relation on DO, Cr, Ca & Mg. This factor is revealed mineral composition in the river system which had been influenced by urban run-off.

Factor-2 accounted for 15.39% of total data variance which has strong relation on TS, SS, Org-N & COD. Organic pollution load from industrial effluent are the main constituents of this factor. The summarized sources identified for river Satluj water pollution are indicated in Table-5.

Table 4: *Factor Analysis for SAT-3*

	Factors			
Explained Variance	**Fa1**	**Fa2**	**Fa3**	**Fa4**
	Flow (-0.82)	SS (-0.90)	Enterococcus (0.91)	Turbidity (0.92)
	Ca (-0.71)	TS (-0.88)	Fecal Coliform (0.90)	Zn (-0.73)
	Mg (-0.68)	Organic N (-0.85)	Total Coliform (0.80)	
	DO (0.67)	COD (-0.76)	Mn (0.67)	
	Cr (0.57)		Water Temperature (-0.64)	
Eigenvalue	13.69	5.23	4.29	3.12
% of Var.	40.27	15.39	12.62	9.17
Cum. %	40.27	55.67	68.29	77.45

Note: Variables are indicated with their loadings in brackets.

Table 5: *Summary of identified sources for different spatial locations*

Spatial Locations	Factor 1	Factor 2
SAT-1 (Upper Catchment)	seasonal impacts, soil structure of the catchment and municipal waste water discharges	agricultural run-off
SAT-2 (Middle catchment)	severe impact of municipal sewage	seasonal variations and agricultural run-off
SAT-3 (Lower Catchment)	mineral component of river influenced by urban run-off	organic pollution load from industrial effluent

5 Conclusions

This study revealed that multivariate analysis was applied to data set obtained from river Satluj water quality monitoring program. PCA/FA analysis was quite helpful in identifying the factors which are responsible for the variations in Satluj river water quality at all 3 spatial groups of sampling sites. Factors indicated that the variables responsible for Satluj river water quality influences are mainly belongs to municipal, industrial waste water discharges; agricultural run-off, organic pollution; nutrients and soluble salts of soil. In this way this study revealed usefulness of multivariate statistical analysis techniques in water quality assessment and identification of pollution sources with a view to get better information about the water quality and design of monitoring program for effective management of rivers.

References

[1] Koklu, R., Sengorur, B., Topal, B., “Water quality assessment using multivariate statistical methods – A case study: Melen river system (Turkey)”, Water Resources Management, 24 (5), PP. 959-978, 2010, DOI:10.1007/s11269-009-9481-7

[2] Singh, K.P., Malik, A., Sinha, S., “Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical technique: a case study”, Analytica Chimica Acta, 538 (1-2), PP. 355-374, 2005, DOI:10.1016/j.aca.2005.02.006

[3] Tanriverdi, C., Alp, A., Demirkiran, A. R., Uckardes, F., “Assessment of surface water quality of the Ceyhan river basin, Turkey”, Environmental Monitoring and Assessment, 167 (1-4), PP. 175-184, 2010, DOI:10.1007/s10661-009-1040-4

[4] Simeonov, V., Sarbu, C., Massart, D. C., Tsakovski, S., “Danube river water data modeling by multivariate data analysis”, Microchimica Acta, 137(3-4), PP. 243-248, 2001, DOI:10.1007/s006040170017

[5] APHA, “Standard methods for the examination of water and wastewater”. 22nd ed. Washington DC, 2012.

[6] Zhang, Y., Guo, F., Meng, W., Wang, X., “Water quality assessment and source identification of Daliao river basin using multivariate statistical methods”, Environmental Monitoring and Assessment, 152 (1-4), PP. 105-121, 2009, DOI:10.1007/s10661-008-0300-z.

[7] Simeonov, V., Stratis, J.A., Samara, C., Zachariadis, G., Voutsa, D., Anthemidis, A., Sofoniou, M., Kouimtzis, T., “Assessment of the surface water quality in Northern Greece”, Water Research, 37 (17), PP. 4119-4124, 2003, DOI:10.1016/S0043-1354(03)00398-1.