Scientific
Research
Publishing

Detection and Mapping of Water Quality Variation in the Godavari River Using Water Quality Index, Clustering and GIS Techniques

Indrani Gupta[1], Awkash Kumar[2*], Chandrakant Singh[1], Rakesh Kumar[1]

[1]CSIR-NEERI, Mumbai Zonal Laboratory, Worli, Mumbai, India
[2]Centre for Environmental Science and Engineering, IIT Bombay, Mumbai, India
Email: *awkash.narayan@gmail.com

Abstract

The objective of this research is to develop a tool for planning and managing the water quality of River Godavari. This is achieved by classifying the pollution levels of Godavari River into several categories using water quality index and a clustering approach that ensure simple but accurate information about the pollution levels and water characteristics at any point in Godavari River in Maharashtra. The derived water quality indices and clusters were then visualized by using a Geographical Information System to draw thematic maps of Godavari River, thus making GIS as a decision support system. The obtained maps may assist the decision makers in managing and controlling pollution in the Godavari River. This also provides an effective overview of those spots in the Godavari River where intensified monitoring activities are required. Consequently, the obtained results make a major contribution to the assessment of the State's water quality monitoring network. Three significant groups (less polluted, moderately and highly polluted sites) were detected by Cluster Analysis method. The results of Discriminant Analysis revealed that five parameters *i.e.* pH, Dissolved Oxygen (DO), Faecal Coliform (FC), Total Coliform (TC) and Ammonical Nitrogen (NH$_3$-N) were necessary for analysis in spatial variation. Using discriminant function developed in the analysis, 100% of the original sites were correctly classified.

Keywords

---

*Corresponding author.

1. Introduction

Water Pollution is a major environmental issue in India. The largest source of water pollution in India is untreated sewage [1]. Other sources of pollution include agricultural runoff and unregulated small scale industry. Most rivers, lakes and surface water in India are polluted [2]. The Godavari is the second longest river in India after the river Ganges. It starts in Maharashtra and flows for 1465 kilometers into the Bay of Bengal. It forms one of the largest river basins in India. The catchment area of the river is 312,812 sq·km and is spread in the states of Maharashtra (48.6%), Andhra Pradesh (23.4%), Madhya Pradesh (10.0%), Chattisgarh (10.9%), Orissa (5.7%) and Karnataka (1.4%) [3]. In this study, the large database comprising of 34 monitoring stations of the Godavari River obtained during 5 years (2007-2011) was analyzed. water quality index was calculated based on four parameters viz. pH, DO, BOD and FC. The derived Water quality indices were then visualized by using Geographical Information System (GIS) to draw thematic maps. The data was also subjected to cluster analysis (CA) and discriminant analysis (DA) based on seven water quality parameters, with a view to extract information about the similarities or dissimilarities among the sampling sites. Cluster and discriminant analysis have been successfully applied in a number of studies, till date [3]-[12]. The clusters were then visualized by using GIS to draw thematic map of Godavari River in Maharashtra State.

2. Methodology

The data sets of the 30 water quality monitoring stations, which comprised of 7 water quality parameters monitored monthly over 5 years (2007-2011), were obtained from the Maharashtra Pollution Control Board. The selected water quality parameters included pH, dissolved oxygen (DO), Biochemical oxygen demand (BOD), Ammonical Nitrogen (NH_3-N), Nitrate (NO_3), Faecal Coliform (FC) and Total Coliform (TC).

2.1. Cluster Analysis (CA)

CA is an exploratory data analysis tool for solving classification problems. Its objective is to sort cases (monitoring points) into groups, or clusters, so that the degree of association is strong between members of the same cluster and weak between members of different clusters. Each cluster thus describes, in terms of the data collected, the class to which its members belong; and this description may be abstracted through use from the particular to the general class or type. Different types of cluster analysis have been applied to view water quality data for both surface and ground waters.

2.2. Discriminant Analysis (DA)

DA is a method to classify cases into categories by one or more continuous or binary independent variables. DA is performed with prior knowledge of membership of objects to a particular group or cluster. This is in contrast to CA where there is no prior knowledge of the group membership. It operates on raw data. The DA technique builds up a discriminant function by creating one or more linear combinations of independent variables. The total number of discriminant functions is either $Ng - 1$ where Ng = number of groups, or p (the number of independent variables), whichever is smaller. The first function maximizes the differences between groups on that function. The second function maximizes differences, but also must not be correlated with the previous function. The percentage of correct classification given in classification table, or prediction matrix, measures the performance of DA. DA is useful in determining whether all independent variables are necessary for predicting category membership and also whether the variables being monitored are sufficient for predicting group membership.

2.3. Water Qualiy Index (WQI)

An index is a mean device to reduce a large quantity of data down to a simplest form. The water quality indices help to evaluate the water quality profile of a river in its entire stretch as well as to identify the reaches where the gap between the desired and the existing water quality is significant enough to warrant urgent pollution control measures. In India the National Sanitation Foundation WQI (NSFWQI) is being used by Central Pollution Control Board, with a slight modification in weights [14]. The NSF WQI is expressed mathematically as:

$$\text{NSFWQI} = \sum_{i=1}^{p} W_i I_i, \tag{1}$$

where, I_i = sub index for ith water quality parameter, W_i = weight (in terms of importance) associated with water quality parameter, and p = number of water quality parameters.

The modified weights (W_i) and the equation for the sub-indices (I_i) as per CPCB, are given in **Table 1** and **Table 2** respectively. The range of the NSF WQI corresponding to various designated best use classification is given in **Table 3**.

Table 1. Original and modified weights for the computation of NSF WQI (CPCB 2001).

Water quality parameters	Original weights from NSF WQI	Modified weights by CPCB
DO	0.17	0.31
FC	0.15	0.28
pH	0.12	0.22
BOD	0.1	0.19
Total	0.54	1.00

Table 2. Sub-index equations for water quality parameters (NSF WQI)[*].

Water quality parameters	Range applicable	Equation
DO (Percent saturation)	0% - 40% saturation	$IDO = 0.18 + 0.66 \times (\text{\% saturation DO})$
	40% - 100% saturation	$IDO = -13.55 + 1.17 \times (\text{\% saturation DO})$
	100% - 140% saturation	$IDO = 163.34 - 0.62 \times (\text{\% saturation DO})$
BOD (mg/l)	0 - 10	$IBOD = 96.67 - 7\,(BOD)$
	10 - 30	$IBOD = 38.9 - 1.23\,(BOD)$
	>30	$IBOD = 2$
pH	2 - 5	$IpH = 16.1 + 7.35 \times (pH)$
	5 - 7.3	$IpH = -142.67 + 33.5 \times (pH)$
	7.3 - 10	$IpH = 316.96 - 29.85 \times (pH)$
	10 - 12	$IpH = 96.17 - 8.0 \times (pH)$
	<2, >12	$IpH = 0$
FC (counts/100ml)	$1 - 10^3$	$IFC = 97.2 - 26.6 \times \log(FC)$
	$10^3 - 10^5$	$IFC = 42.33 - 7.75 \times \log(FC)$
	$>10^5$	$IFC = 2$

[*]Abbasi 2002.

Table 3. NSF WQI for various designated best use[*].

Serial No	NSF WQI	Description of quality (1978)	Class by CPCB	Remarks
1	63 - 100	Good to excellent	A	Non polluted
2	50 - 63	Medium to good	B	Non polluted
3	38 - 50	Bad	C	Polluted
4	38 & less	Bad to very bad	D, E	Heavily polluted

[*]CPCB 2001, Abbasi 2002.

2.4. GIS Application

GIS are well-established software tools that combine cartographic and database technologies [15] [16]. These information systems capture, store, manage, and present various data sets such as maps of cities and rivers that are linked to location. Users can create interactive queries, analyze spatial information, and thus better understand relationship and interactions between human activities and their surroundings. The results are viewed on maps, and this visualization which is advantageous over traditional records and tables used in spreadsheets and databases. The GIS data mapping used in this study were on the software package ArcGIS 9.1 Desktop.

3. Results and Discussion

Figure 1 gives the box plots of pH, DO, BOD, FC, TC, NH_3-N and Nitrate. A box plot or a box-and whisker diagram or plot is a convenient way of graphically depicting groups of numerical data through their five-number summaries: the minimum, lower quartile, median, upper quartile, and sample maximum. A box plot may also indicate which observations, if any, might be considered outliers. The ends of the whiskers represent the minimum and maximum of all the data. The bottom and top of the box are the 25th and 75th percentiles. The center represents the 50th percentile. Any data not included between the whiskers is plotted as an outlier with a small circle, or star. If the data are normally distributed, the locations of the five marks on the box plot will be equally spaced.

3.1. Percent Exceedances

Percent exceedances of water quality parameters viz. pH, DO, BOD, TC, NH_3-N and Nitrate with respect to MPCB A-II standard are given in **Table 4**. DO, BOD, NH_3-N and pH are not meeting the MPCB A-II standards at 27, 43, 30 and 63% of the locations. At all locations, Nitrate concentrations are within the standard. At Tapovan, Nashik BOD and DO levels do not meet the standard 93 and 23% of the time. Similarly, at Kapila Nashik, BOD and DO levels do not meet the standard 90 and 26% of the time. At Pathegaon, NH_3-N do not meet the standard 47% of the time. TC also do not meet the standard at 1 location viz, Kopargaon and percent exceedance is 50% at this station.

3.2. Water Quality Index

Yearly and average WQI of five years is given in **Table 5**. However, there were number of missing values and WQI could not be calculated for these years. In many cases, data was available for few months of the year and the yearly WQI have been calculated based on the available data. Average Water Quality Indices of 30 sitesvaried between Good to Excellent, Medium to Good and Bad. The average WQI was "good" at 27 sites, "medium to good" at 2 sites and "bad" at 1 site. Straight line fit was carried out on these 5 years data and slopes were calculated. Positive values imply an increasing trend and negative values imply a decreasing trend. At 11 locations
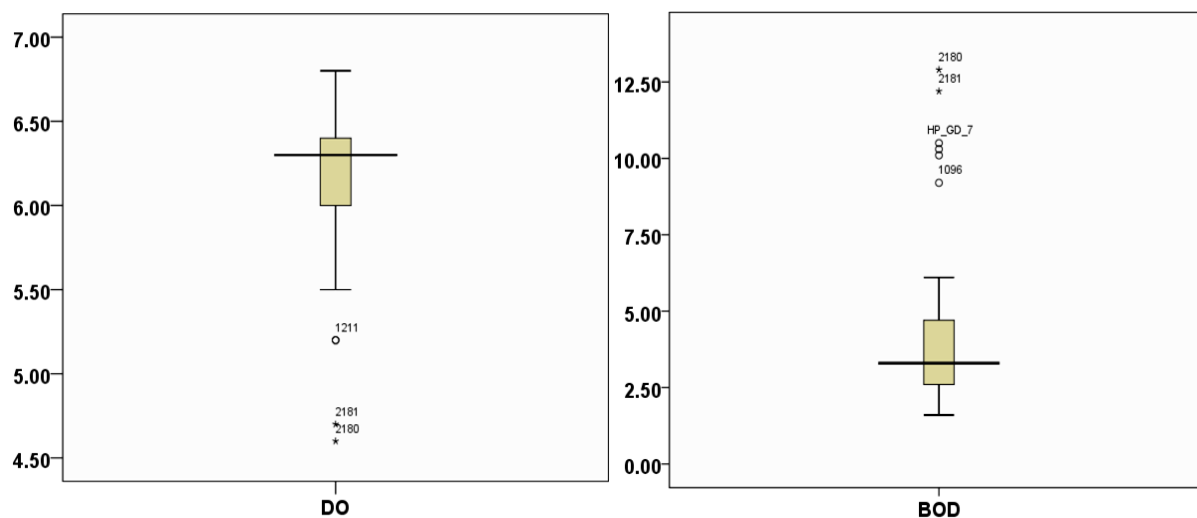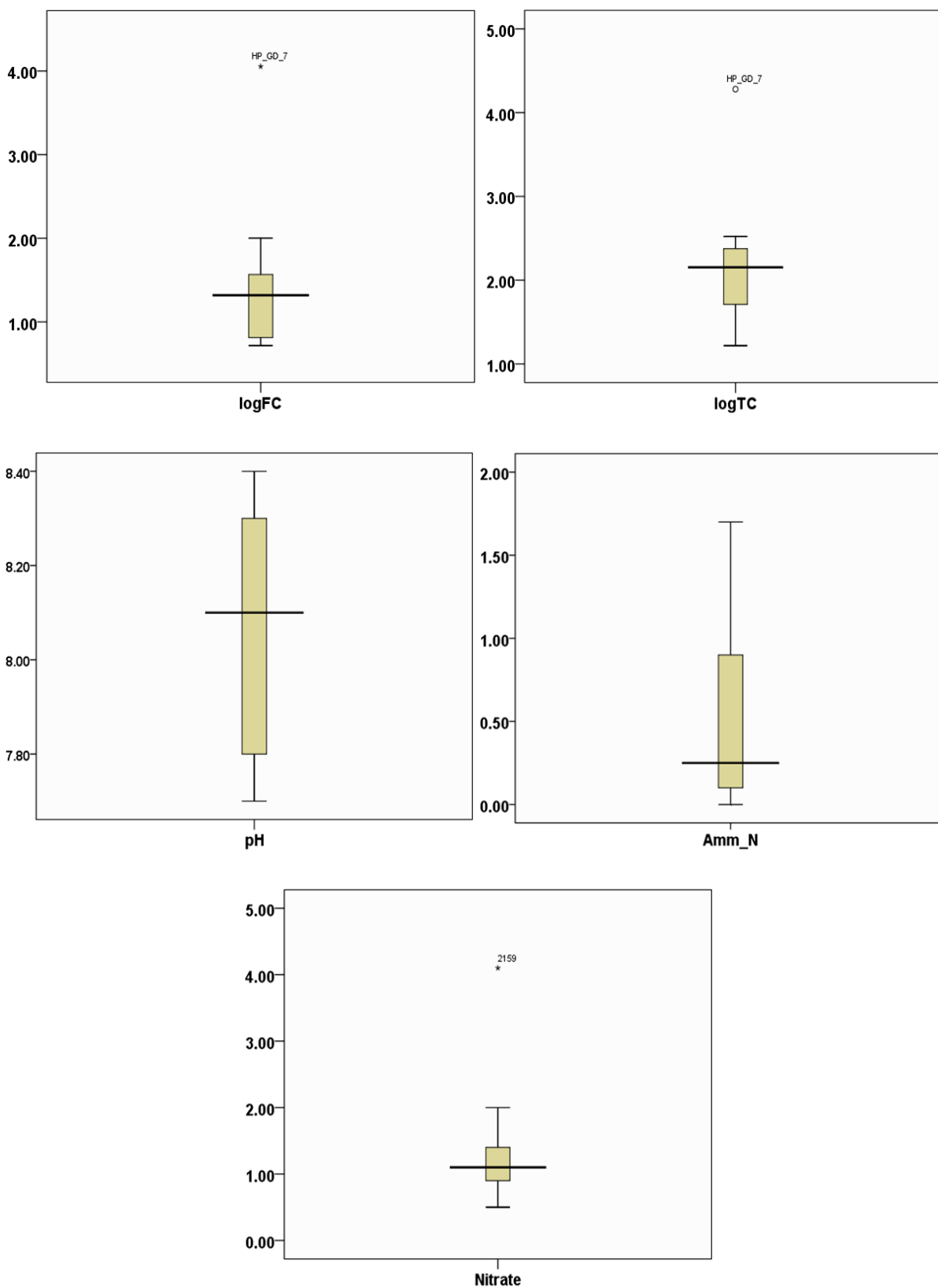
Figure 1. Boxplots of water quality parameters.

Table 4. Percent exceedance of water quality parameters.

Se. No	Site details	Station No	pH	DO	BOD	TC	NH₃-N	Nitrate
1	Anantwadi	HP_GD_26	4.8	0	0	0	0	0
2	Damrencha	HP_GD_27	0	0	0	0	0	0
3	Deori	HP_GD_28	7.7	0	0	0	0	0
4	Dhaba	HP_GD_29	9.1	2.2	0	0	0	0
5	Drugwada	HP_GD_30	0	0	0	0	0	0
6	Gangapur, Nashik	1095	7.4	0	20.4	0	0	0
7	Jaikwadi Dam	1312	5.9	5.9	5.9	0	25.8	0
8	Kamtheekhairi	HP_GD_31	0	0	0	0	0	0
9	Kapila-Nashik	2181	9.3	25.6	90.7	0	22.9	0
10	Kolgaon	HP_GD_33	0	0	0	0	0	0
11	Kopargaon	HP_GD_7	0	0	62.5	50	0	0
12	Mahagaon	HP_GD_34	0	0	0	0	0	0
13	Mathani	HP_GD_35	0	0	0	0	0	0
14	Pathegaon bridge	2159	14.3	14.3	23.8	0	46.7	0
15	Paithan Aurangabad	2158	4.8	14.3	33.3	0	50	0
16	Panchavati at Ramkund	1096	3.6	0	85.7	0	1.9	0
17	Petta	HP_GD_36	0	0	0	0	0	0
18	Raher. Nanded	1209	9.3	5.9	10.9	0	17	0
19	Saikheda, Nashik	2182	4.7	2.5	60	0	7.5	0
20	Saiphal	HP_GD_37	0	0	0	0	0	0
21	Someshwar Temple	2177	4.3	0	56.5	0	19	0
22	Taklidhangar	HP_GD_21	7.7	0	0	0	0	0
23	Tapovan. Nashik	2180	9.3	23.3	93	0	28.2	0
24	Tembhurdoh	HP_GD_39	0	0	0	0	0	0
25	Toka	HP_GD_22	4.8	0	13.6	0	0	0
26	Vishnupuri, Nanded	1210	8.3	0	16.2	0	0	0
27	Wadvali	HP_GD_23	27.3	0	0	0	0	0
28	Wagholi-Butti	HP_GD_41	7.1	0	0	0	0	0
29	WarudBagaji	HP_GD_42	6.8	0	0	0	0	0
30	Yelli	HP_GD_24	0	0	0	0	0	0

Table 5. Water quality indices of Godavari River during 2007-2011.

Se. No	Station name	Station No.	2007	2008	2009	2010	2011	Avg.	Avg. WQI	Slope (m value)
1	Anantwadi	HP_GD_26	68	73	65	66	74	69.2	GE	0.423
2	Damrencha	HP_GD_27	79	80	70	71	80	76	GE	0.839
3	Deori	HP_GD_28	74	73	67	74		72	GE	0.554
4	Dhaba	HP_GD_29	65	64	57	63	69	63.6	GE	0.797
5	Drugwada	HP_GD_30		73	69	72	79	73.25	GE	1.96
6	Gangapur Dam	1095	74	78	74	80	78	76.8	GE	0.821
7	Jaikwadi Dam	1312	75	78	73	76	80	76.4	GE	0.756
8	Kamtheekhairi	HP_GD_31	65	73	70	71	77	71.2	GE	2.222
9	Kapila	2181		58	59	67	58	60.5	MG	0.748
10	Kolgaon	HP_GD_33	68	67	66	68	73	68.4	GE	0.984
11	Kopargaon	HP_GD_7		49		38		43.5	Bad	10.693
12	Mahagaon	HP_GD_34	78	77	68			74.3	GE	−5.38
13	Mathani	HP_GD_35	72	73	64			69. 7	GE	−4.166
14	D/S of Paithan	2159		76	72	78	76	75.5	GE	0.527
15	Paithan Pump House	2158		76	74	75	79	76	GE	0.995
16	Panchavati	1096	64	70	62	59	63	63.6	GE	−1.272
17	Petta	HP_GD_36	76	76	67			73	GE	−4.64
18	Raher	1209		79	70	76	76	75.25	GE	−0.331
19	Saikheda	2182		76	69	71	69	71.25	GE	−1.638
20	Saiphal	HP_GD_37	70	65	69	73	74	70.2	GE	0.46
21	Someshwar Temple	2177	63	72	69	70	71	69	GE	1.474
22	Taklidhangar	HP_GD_21	83	71	68	77		74.75	GE	−2.12
23	Tapovan, Nashik	2180		58	57	59	57	57.75	MG	−0.23
24	Temburdoh	HP_GD_39	65	72	71	69	79	71.2	GE	2.5
25	Toka	HP_GD_22	83	80	74	70		76.75	GE	−4.435
26	Vishnupuri Nanded	1210	76	78	75	75	78	76.4	GE	0.182
27	Wadvali	HP_GD_23	82	77	77	69		76.25	GE	−3.96
28	WagholiButti	HP_GD_41	71	67	71		76	71.25	GE	1.79
29	WarudBagaji	HP_GD_42	68	73	67	68	78	70.8	GE	1.414
30	Yelli	HP_GD_24	82	73	74	79		77	GE	−0.61

slopes are negative implying a deterioration of water quality. At 19 locations slopes are positive implying improvement of water quality. Yearly WQI are shown in **Figure 2** to **Figure 6**.

3.3. Cluster Analysis

Spatial cluster analysis was carried out with average concentrations for five years. Spatial cluster analysis produced a dendrogram with three groups (**Figure 7**). Dendrogram gives an idea of how great the distance was between cases or groups that are clustered in aparticular step, using a 0 to 25 scale along the top of the chart. The bigger the distances before two clusters are joined, the bigger the differences in these clusters. The branching-type nature of the dendrogram allows one to trace backward or forward to any individual case or cluster at any level.

Group 1 consists of 20 sites which were less polluted. Group 2 comprised of 9 sites which were moderately polluted. Group 3 comprised of 1 monitoring location which was bad. **Table 6** presents the means and standard deviations for the clusters for the examined parameters. The clusters were well separated by the water quality

Figure 2. WQ classification of Godavari River in 2007.

Figure 3. Water quality classification of Godavari River in 2008.
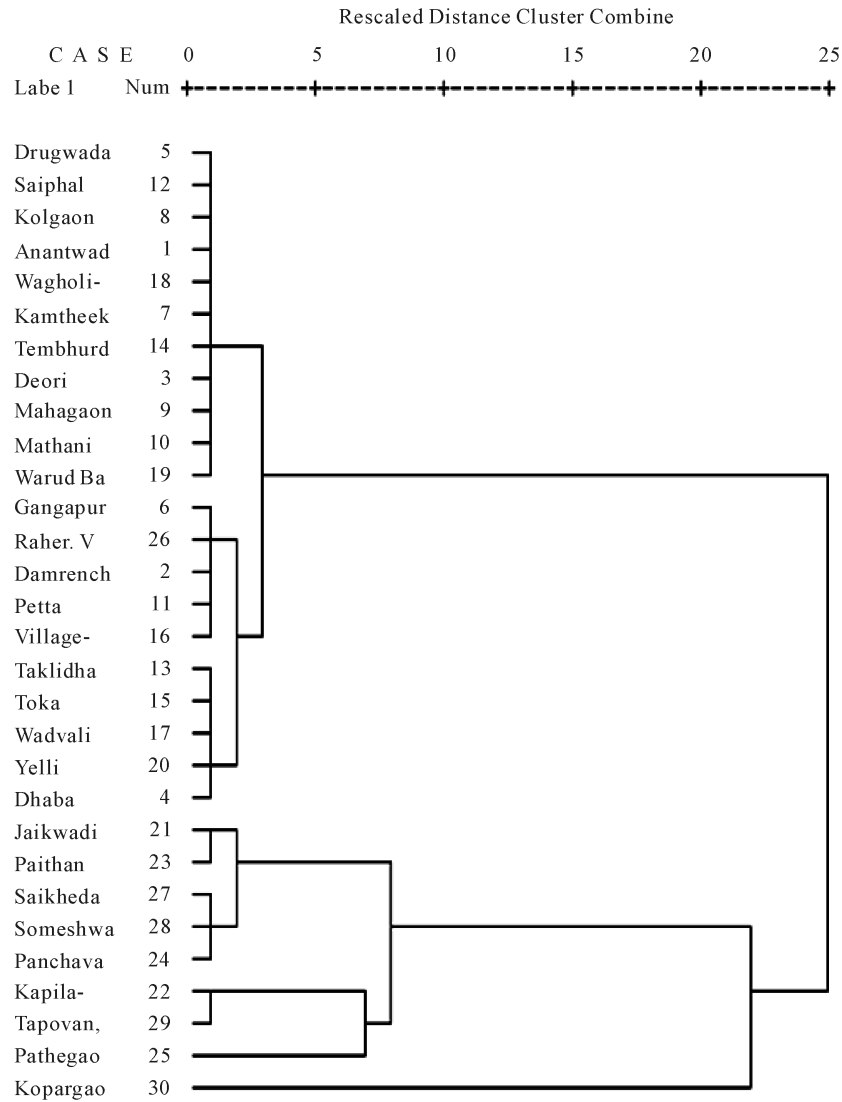
Figure 4. Water quality classification of Godavari River in 2009.

Figure 5. Water quality classification of Godavari River in 2010.

Figure 6. Water quality classification of Godavari River in 2011.

Rescaled Distance Cluster Combine

```
    C A S E      0         5        10        15        20        25
    Labe l   Num +---------+---------+---------+---------+---------+

    Drugwada    5
    Saiphal    12
    Kolgaon     8
    Anantwad    1
    Wagholi-   18
    Kamtheek    7
    Tembhurd   14
    Deori       3
    Mahagaon    9
    Mathani    10
    Warud Ba   19
    Gangapur    6
    Raher. V   26
    Damrench    2
    Petta      11
    Village-   16
    Taklidha   13
    Toka       15
    Wadvali    17
    Yelli      20
    Dhaba       4
    Jaikwadi   21
    Paithan    23
    Saikheda   27
    Someshwa   28
    Panchava   24
    Kapila-    22
    Tapovan,   29
    Pathegao   25
    Kopargao   30
```

Figure 7. Dendrogram of 30 monitoring locations using ward method.

Table 6. Mean and standard deviations for water quality parameters of 3 clusters in Godavari River.

	Cluster 1		Cluster 2		Cluster 3	
	Mean	Stdev.	Mean	Stdev.	Mean	Stdev.
pH	8.2	0.2	7.8	0.1	7.9	0.0
DO	6.4	0.2	5.7	0.6	6.3	0.0
BOD	2.8	0.7	7.1	3.3	10.5	0.0
FC	33.8	26.7	12.8	7.9	11326.0	0.0
TC	135.6	102.7	153.4	105.9	19075.0	0.0
Am-N	0.2	0.2	1.0	1.0	0.1	0.0
Nitrate	1.1	0.3	1.5	1.58	1.1	0.0

parameters. Highest DO, lowest BOD and lowest TC were observed in cluster 1. However, moderate FC and Nitrate were observed in cluster 1. In cluster 2 moderate BOD and TC were observed. In cluster 3, highest BOD, FC and TC were observed. However, moderate DO and Nitrate were observed in cluster 3. The results were also viewed in a GIS layer (**Figure 8**).

The results of water quality index and those of cluster analysis more or less match since water quality index considers only 4 parameters, viz. pH, BOD, DO and FC while cluster analysis was performed using 7 parameters, viz. pH, BOD, DO, FC and TC, NH$_3$-Nand Nitrate. The results of water quality index and those of cluster analysis match at 23 locations.

3.4. Discriminant Analysis

Discriminant analysis is carried after standardizing the independent variables. **Table 7** presents the Tests of Equality of Group Means. It indicates whether there is a statistically significant difference among the dependent variable means for each independent variable. The smaller the Wilks' Lambda, the more important is the independent variable. Here only Nitrate is not statistically significant and hence we drop the Zscore (Nitrate). **Table 8** gives the results of Box's Test of Equality of Covariance Matrices. The Box's M test tests the homogeneity of covariance assumption. This test is also sensitive to meeting the assumption of multivariate normality. **Table 10** shows that the test is significant so we conclude the groups have equal covariance matrices. **Table 9** gives the summary of Canonical Discriminant Functions. The larger the eigen value the more of the variance in the dependent variable is explained by that function. The magnitudes of the eigen values are indicative of the functions' discriminating abilities. Since there are three categories or groups there are two discriminant functions in the descending order of importance. Percent of variance is the proportion of discriminating ability of the continuous variables found in a given function. The squared canonical correlation is the percent of variation in the dependent discriminated by the independents in DA. In this analysis, the first function accounts for 100% of the discriminating ability of the discriminating variables. **Table 10** gives the Wilks' lambda which tests the significance of the eigen value for each discriminant function. In this case both are significant. Standardized discriminant function coefficients in **Table 11** indicate the relative importance of the independent variables in predicting the dependent. **Table 12** provides an indication of the success rate for prediction of membership of the grouping variable's categories using the discriminant function developed from the analysis. Percentage of correct classification for original cases was observed to be 100%. Percentage of correct classification for cross-validated grouped cases was observed to be 90%.

4. Conclusion

Average water quality index for five years *i.e.* for 2007 to 2011 was calculated with four parameters viz. DO, BOD, FC and pH. The results show that out of 30 locations on Godavari River in Maharashtra, 27 have "good to excellent" water quality, 2 have "medium to good" water quality and 1 has "bad" water quality. The slope of the

Figure 8. Water quality classification of Godavari River based on cluster analysis.

Table 7. Tests of equality of group means.

	Wilks' Lambda	F	df1	df2	Sig.
Zscore (FC)	0.000	1.18E5	2	27	0.000
Zscore (pH)	0.326	27.97	2	27	0.000
Zscore (DO)	0.612	8.54	2	27	0.001
Zscore (BOD)	0.412	19.30	2	27	0.000
Zscore (TC)	0.001	1.57E4	2	27	0.000
Zscore (NH$_3$-N)	0.340	26.24	2	27	0.000
Zscore (Nitrate)	0.877	1.90	2	27	0.169

Table 8. Box's test of equality of covariance matrices[a].

Box's M		67.38
	Approx.	3.30
F	df1	15
	df2	1.002E3
	Sig.	0.000

[a]Some covariance matrices are singular and the usual procedure will not work. The non-singular groups will be tested against their own pooled within-groups covariance matrix. The log of its determinant is −20.306.

Table 9. Summary of canonical discriminant functions.

Function	Eigenvalue	% of variance	Cumulative %	Canonical correlation
1	2.48E4[a]	100.0	100.0	1.00
2	3.84[a]	0.0	100.0	0.89

[a]First 2 canonical discriminant functions were used in the analysis.

Table 10. Wilks' Lambda.

Test of function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	0.000	292.42	10	0.000
2	0.21	39.44	4	0.000

Table 11. Standardized canonical discriminant function coefficients.

	Function	
	1	2
Zscore (pH)	−0.59	0.68
Zscore (DO)	1.23	0.34
Zscore (FC)	2.04	0.21
Zscore (TC)	−0.99	−0.55
Zscore (NH$_3$-N)	0.52	−0.43

Table 12. Classification results[bc].

	Group	Predicted group membership			Total
		1	2	3	
Original	Count 1	20	0	0	20
	2	0	9	0	9
	3	0	0	1	1
	% 1	100.0	0.0	0.0	100.0
	2	0.0	100.0	0.0	100.0
	3	0.0	0.0	100.0	100.0
Cross-validated[a]	Count 1	19	1	0	20
	2	1	8	0	9
	3	1	0	0	1
	% 1	95.0	5.0	0.0	100.0
	2	11.1	88.9	0.0	100.0
	3	100.0	0.0	0.0	100.0

[a]Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case; [b]100% of original grouped cases correctly classified; [c]90% of cross-validated grouped cases correctly classified.

straight line fit of the WQI of five years shows that water quality is deteriorating at 11 locations and there is improvement at 19 locations. Hierarchal clustering of the data based on seven parameters viz. DO, BOD, FC, TC, NH_3-N, Nitrate and pH for five years produce three clusters. Group 1 consists of 20 sites which were less polluted or good to excellent. Group 2 comprised of 9 sites which were medium to good. Group 3 comprised of 1 monitoring location *i.e.* at Kopargaon which was highly polluted. DA provided an important data reduction and used five parameters (pH, DO, FC, TC and NH_3-N) for 100% correct assignment in spatial analysis. Therefore, DA allowed a reduction in the dimensionality of data set, delineating a few indicator parameters responsible for variations in water quality. GIS is a very comprehensive tool which provides the decision makers with a simple and reliable overview of the status of surface water quality. The results of WQI for 5 years and clustering were viewed in GIS.

Acknowledgements

References

[1] Central Pollution Control Board, Evaluation of Operation and Maintenance of Sewage Treatment Plants in India (2007). http://www.cpcb.nic.in

[2] Central Pollution Control Board, Annual Report (2008-2009). http://cpcb.nic.in/upload/AnnualReports//AnnualReport_37_ANNUAL_REPORT-08-09.pdf

[3] Central Pollution Control Board, Environmental Atlas of India (2001) New Delhi.

[4] Bierman, P., Lewis, M., Ostendorf, B. and Tanner, J. (2011) A Review of Methods for Analyzing Spatial and Temporal Patterns in Coastal Water Quality. *Ecological Indicators*, **11**, 103-114. http://dx.doi.org/10.1016/j.ecolind.2009.11.001

[5] Einax, J.W., Truckenbrodt, D. and Kampe, O. (1998) River Pollution Data Interpreted by Means of Chemometric Methods. *Microchemical Journal*, **58**, 315-324. http://dx.doi.org/10.1006/mchj.1997.1560

[6] Gazzaz, N.M., Yusoff, M.K., Ramli, M.F., Aris, A.Z. and Juahir, H. (2012) Characterization of Spatial Patterns in River Water Quality Using Chemometric Pattern Recognition Techniques. *Marine Pollution Bulletin*, **64**, 688-698. http://dx.doi.org/10.1016/j.marpolbul.2012.01.032

[7] Guler, C., Thyne, G.D., McCray, J.E. and Turner, K.A. (2002) Evaluation of Graphical and Multivariate Statistical Methods for Classification of Water Chemistry Data. *Hydrogeology Journal*, **10**, 455-474. http://dx.doi.org/10.1007/s10040-002-0196-6

[8] Gupta, I., Dhage, S., and Kumar, R. (2009) Study of Variations in Water Quality of Mumbai Coast through Multivariate Analysis Techniques. *Indian Journal of Marine Sciences*, **38**, 170-177.

[9] Gupta, I., Salunkhe, A., Rohra, N. and Kumar, R. (2013) Chemometrics Data Analysis of Marine Water Quality of Maharashtra, West Coast of India. *Indian Journal of Geo Marine Sciences*, **42**, 97-105.

[10] Kovacs, J., Kovacs, S., Magyar, N., Tanos, P., Hatvani, I.G. and Anda, A. (2014) Classification into Homogeneous Groups Using Combined Cluster and Discriminant Analysis. *Environmental Modeling & Software*, **57**, 52-59. http://dx.doi.org/10.1016/j.envsoft.2014.01.010

[11] Santos-Roman, M.D., Warner, S.G. and Scatena, F. (2003) Multivariate Analysis of Water Quality and Physical Characteristics of Selected Watershed in Puerto Rico. *Journal of the American Water Resources Association*, 829-839. http://dx.doi.org/10.1111/j.1752-1688.2003.tb04408.x

[12] Xu, H.S., Xu, Z.X., Wu, W. and Tang, F.F. (2012) Assessment and Spatiotemporal Variation Analysis of Water Quality in the Zhangweinan River Basin, China. *Procedia Environmental Sciences*, **13**, 1641-1652. http://dx.doi.org/10.1016/j.proenv.2012.01.157

[13] Zeng, X. and Rasmussen, T.C. (2005) Multivariate Statistical Characterization of Water Quality in Lake Lanier, Georgia, USA. *Journal of Environmental Quality*, **34**, 1980-1991. http://dx.doi.org/10.2134/jeq2004.0337

[14] Abbasi, S.A. (2002) Water Quality Indices State-of-the-Art, Pondicherry University, Centre for Pollution Control & Energy Technology, Pondicherry.

[15] Nemeth, T., Szabo, J., Pasztor, L. and Bakacsi, Z. (2002) Elaboration of a Complex GIS Application in a Catchment Area. *Water Science and Technology*, **45**, 133-140.

[16] Shaban, M., Urban, B., El-Saadi, A. and Faisal, M. (2010) Detection and Mapping of Water Pollution Variation in the Nile Delta Using Multivariate Clustering and GIS Techniques. *Journal of Environmental Management*, **91**, 1785-1793. http://dx.doi.org/10.1016/j.jenvman.2010.03.020