ASSESSMENT OF SURFACE WATER QUALITY OF CHAMBAL RIVER A MULTIPLE LINEAR REGRESSION ANALYSIS

RAKESH PARMAR ,P.B.REDDY

Abstract: Recently the Biomathematics is an emerging field which has enabled the development of a range of new theories with significant problems of pollution management and control. Based on the earlier concepts, existing mathematical models would be utilized as well as more advanced mathematical models would be developed, for better understanding of the environmental problems.

In this study multiple linear regressions was applied on the surface water quality data with the aim of identifying the pollution sources and their contribution toward water quality variation. Surface water samples were collected from four different sampling points along Chambal River at Nagda. Eleven physico-chemical water quality parameters were selected for analysis: dissolved oxygen (DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), suspended solids (SS), pH, conductivity, temperature, nitrogen in the form of ammonia (NH_3), turbidity, total solids (TS), nitrates (NO_3), chloride (Cl) and sulphates. The correlation of BOD with hydro-meteorological parameters was computed by the linear regression analyses. The correlation coefficient was positive in all water samples and the confidence interval was statistically significant. Biodegrading activity indicated by the BOD activity was affected by the hydrological regime, by the irregularities of anthropogenic management and waste waters fluctuation, as well.

Keywords: Stepwise regression, correlation coefficient BOD, Water pollution

Introduction: With the growth of human populations, commercial and industrial activities, surface water has received large amount of pollutants from variety of sources [1].The quality of surface water provides significant information about the available resources for supporting life in the ecosystem [2].The physical, chemical and biological compositions of surface water is controlled by many factors such as natural (precipitation, geology of the watershed, climate and topography) and anthropogenic (domestic, industrial activities and agricultural run-off). Increasing surface water pollution causes not only deterioration of water quality, but also threatens human health, balance of aquatic ecosystem, economic development and social prosperity [3].It is very important to prevent and control the surface water pollution and to have reliable information on its quality for effective management [4].Characterization of the spatial variation and source apportionment of water quality parameters can provide an improved understanding of the environmental condition and help policy makers to establish priorities for sustainable water management [5]. One of the major challenges in surface water quality assessment is identifying the sources of pollutants and the contribution of the parameters/variables in explaining water quality variation.

Several researchers used BOD to identify water quality sources apportionment. [6], [5]. Juahir et al. [7] studied spatial variability of surface water quality. Reddy and Baghel [8], [9] in their studies, they discovered that Chambal River in Nagda was heavily polluted as a result of industrial discharge and municipal waste (anthropogenic source of pollution). Similarly, study was conducted by Koklu et al. [10]. They revealed that, multiple regressions analysis identified important and effective parameters that contributed to the water quality variation.

Therefore, this study aims to evaluate the surface water pollution sources of Chambal River at Nagda (Western Madhya Pradesh) through BOD and estimating the contribution of the significant parameters towards water quality variation using multiple linear regressions model.

Expected outcome : The Chambal River in M.P. is affected by various effluents of industrial complex located in this area offers an opportunity to further qualify the impact of such effluents on the water quality which may lead to better understanding of pollution process in the River that may lead to improved regulation and policy development.

Materials and Methods

Study Area. About Nagda: Nagda is very close to tropic of cancer at 23'27N and 75'25 and 517 meters above MSL.Nagda is a city and municipality in Ujjain district in the Indian state of Madhya Pradesh. It is an Industrial town in the Malwa region of western M.P and is situated in the bank of Chambal River. Presently Nagda is a major industrial town having manufacturing unit of viscose fiber, thermal plant and chemical plant were developed by Grasim Industries Ltd. It is a major railway junction on Delhi-Mumbai railway line. The region has rapid population growth and industrial development with increase the mass of sewage discharge. River receives many inputs both anthropogenic and natural in origin that may cause deterioration of the river water quality.

River Chambal River Chambal Originates from Barnagar (M.P) and joins River Yamuna after Udi at Jahika (U.P).From its origin onwards, tributaries, Khan, and Kshipra join river Chambal before Nagda, M.P. More than one lakh of residents in and around the Nagda area rely on water from Chambal River for public use, industrial supplies, power plant cooling and waste water treatment. The river receives water from different units of Grasim Industries and sewage from Nagda town. Waste after coming from the factory complex runs in a channel for about 3km and joins River Chambal near Juna Nagda.

Sampling and Analytical Procedures: The sampling network and strategy were designed to cover wide range of determinant at the key sites, which reasonably represent the surface water quality in the area. Sampling was carried out every day from 1st November ,2011 to 30th November 2011 at three different sampling locations .samples were collected at 30cm below the water level using a water sampler and acid washed container to avoid unpredicted changes. The samples were immediately transported to the laboratory under low temperature conditions in ice-boxes and stored in the laboratory at 4° C until analysis.

All samples were analyzed for eleven various physiochemical parameters. Water temperature, DO, pH, conductivity, BOD, turbidity, TS, SS and NH_3 of the water samples were carried out using the standard method [11]. The dissolved oxygen content was determined before and after the incubation. Sample incubation was for 5 days at 20°C in BOD bottle and BOD_5 was calculated after the incubation period. COD was determined after oxidation of organic matter in strong tetraoxosulphate VI acid medium by $K_2Cr_2O_7$ at 148° C with blank titrations. Cl was determined using 100 mg/l of the water sample which was measured into 250 mg/L conical flask and pH was adjusted with 1 M NaOH. 1 ml/g of $K_2Cr_2O_4$ indicator was then added and titrated with $AgNO_3$ solution. A blank titration was carried out using distilled water. Cl mg/L was then calculated. NO_3 was determined using calorimetric method [11].

Multiple Linear Regressions: Multiple linear regressions is a statistical tool for understanding between an outcome variable and several predictors (independent variables) that best represent the relationship in a population [10]. The technique is used for both predictive and explanatory purposes within experimental and non experimental designed. Multiple linear regressions can be expressed using the equation below:

$Y = bo + b_1X_1 + b_2X_2 + \times\times\times\times\times + b_mX_m + e$ (2) Where Y represent the dependent variable, $X_1 \times\times\times X_m$ represent the several independent variables bo × × × ×bm represent the regression coefficient and e represent the random error. The present data was analyzed by using computer soft ware . http://www.stattools.net.

Results and Discussion:

Descriptive Statistics: Table 1 present range, minimum, maximum, mean, standard deviation and variance of the parameters under study. It is clear that, DO, SS, TS, Cl and conductivity are the dominant parameters with high mean concentration of 5.4mg/L,133.3 mg/L, 432. mg L, 105.56 mg/L, 1261.66mg/L and 427.36 mg/L respectively. This showed that, these variables have common source of origin. The mean value of pH ranged from 6.27 to 8.7 mg/L which the average value of 7.79 mg/L which is slightly above neutral level. The concentration of BOD was 23.1 to 54.1 and COD ranged from 23.1 to 47.21.00 mg/L respectively. The order of abundance is COD > BOD_5 > DO, showing very high anthropogenic pressure on the surface water.

Surface water quality prediction using multiple linear regressions model: To find out the best predictor of water quality variation in the Chambal River at Nagda, a stepwise multiple linear regressions model was used (http://www.stattools.net). Before interpreting the result, classical assumptions of linear regressions was checked. An inspection of normal p-p plot of regression standardized residuals revealed that all the observed values fall roughly along the straight line indicating that the residuals are from normally distributed population. Moreover, the scatter plot (standardized predicted values against observed values) indicated that, the relationship between the

Tab.1Changes in surface water parameters of Chambal River at Nagda in November, 2011.

Days	TDSmg/L	TSSmg/L	EC,umho/cm	DO mg/L	COD mg/L	Chl. mg/L	Sulpha. mg/L	TH mg/L	pH	Temp. C	BOD mg/L
1	448.1±6.3	148.1±.4.3	512±7.1	6.1±0.4	27.1±. 1.2	1250± 24.5	724± 11.2	1423± 13.6	6.3± 0.1	24.1±. 0.1	24.1± 1.2
2	392.1±5.3	140.2± 8.7	492.1±5.9	6.4±0.3	34.2±. 4.1	1265± 18..5	658± 8.9	1411± 14.1	6±0.2	24.2±. 0.1	25.43± .3.3
3	412.1±7.1	132.2± 5.2	452.1±7.4	6.3±0.1	29.2±. 3.1	1389± 14.2	852± 8.8	1121.± 6.3	8.5± 0.1	23.1±. 0.1	23.1± 2.1
4	411.8±4.9	134.12± 4.4	348.1.8±4.	5.8±0.2	32.1±. 4.3	1289± 22.4	874± 14.2	869.1± 5.3	7.8± 0.2	24.±.0. 1	54.1.± .6.1
5	398.1±8.3	152.2± 3.4	398.1±8.3	5.6±0.2	36.1±. 3.1	1453± 23.5	769± 12.3	968.1± 6.3	8.6± 0.2	24.1±. 0.12	47.5 ±.5.2
6	442.2.±8.3	111.3± 5.4	345.1.±8.3	5.7±0.4	29.8±. 2.7	1369± 21.3	963± 13.1	789.1± 5.3	7.7± 0.4	231±.0 .1	49.8± 4.8
7	409.7±3.8	137.2± 6.1	439.7±6.8	4.8±0.2	38.1±. 4.1	1458± 18.5	689± 11.2	689.1± 8.2	8.8± 0.2	24.±.0. 1	52.4± 6.3
8	387.2±8.1	138.3± 5.3	361.2±4.4 1	6.1±0.3	33.6.± .3.1	1530± 24.5	735± 11.1	894±6. 1	8.1± 0.3	22.6.±. 0.11	39.2± 4.2
9	434.2±5.3	139.3± 8.2	289.2±3.2	5.8±0.2	23.1.± .2.11	1485± 21.8	1114± 9.6	1125.1 ±8.3	8.8± 0.2	23.1.±. 0.11	53.3± 5.1
10	389.2±5.9	142.2± 5.4	376.3 ±4.2	6.4±0.4	34.3.± .3.1	1369± 19.8	968± 12.3	1224.1 ±8.8	7.4± 0.4	24.±.0. 1	43.8.±. 305
11	444.1±8.3	129.3± 7.3	298.3 ±3.6	6.2±0.2	32.1.± .3.1	1480± 16.9	1123± 14.5	1005 ±6.8	8.2± 0.2	22.1.±. 0.1	52.1± 3.4
12	423.1±7.2	137.2± 9.4	402.1 ±3.5	5.9±0.1	31.7.. ±.2.1	1469± 15.8	1121± 14.3	1133 ±11.1	7.9± 0.1	23.7..± .0.1	48.3.± 6.2
13	367.2±4.9	142.1± 6.7	343.2±3.9	6.3±0.3	29.8.± .2.1.	1458± 13.7	879± 13.7	1254± 12.2	8.3± 0.3	22.8.±. 0.1.	52.1± 3.8
14	359.2±6.4	137,3± 6.9	311.1 ±4.3	6.4±0.1	31.3.± .3.1.	1463± 17.8	963± 14.2	1008± 9.2	8.4± 0.1	24.3.±. 0.1.	47.2± 4.1
15	434.3±7.3	144.2± 8.3	410.1 ±3.3	6.6±0.5	34.1.± .3.1	1453± 18.4	1003± 13.5	1186 ±9.9	8.6± 0.5	23.5.±. 0.1	44.5± 3.4
16	411.2±6.6	129.5± 7.3	361.1 ±5.3	6.7±0.3	29.9.± 3.2	1550± 16.8	890± 11.5	1325± 11.2	8.7± 0.3	22.9.±. 0.1	29.8± 2.5
17	408.2±5.8	152.1± 8.3	412.3 ±5.1	6.4±0.4	34.2.± .2.3	1367± 18.3	1008± 11.5	1425± 11.2	6.4± 0.4	24.1.±. 0.1	52.1± 4.3
18	434.2±7. 1	136,4± 7.8	345.6 ±4.8	5.9±0. 1	31.3. ±.2.1	1411± 18.2	1241± 14.5	1354 ±12.3	7.9± 0.1	24.3. ±.0.1	46.6± 6.3
19	389.3±6. 8	138.4± 6.2	334.3 ±4.1	5.8±0. 1	32.1. ±.2.3.	1362± 21.3	1234± 12.8	1325 ±11.3	7.8± 0.1	23.6. ±.0.1	51.1± 3.3
20	389.3±7. 3	137.2± 6.8	321.2±2. 9	5.2±0. 2	42.1. ±.3.2	1380± 15.8	1100± 13.1	1085 ±9.9	8.2± 0.2	23.8. ±.0.1	47.2± 4.2
21	373.2±4. 6	141.3± 5.9	383.2 ±4.6	5.9 ±0.2	24.1. ±.0.2	1563± 14.5	1123± 14.1	1425 ±11.2	7.9 ±0.2	24.1. ±.0.2	44.6± 4.3
22	458.1±6. 3	137.4± 6.9	443.2±5. 3	6.1±0. 1	36.5. ±.2.4	1689± 17.8	1169± 13.6	1369 ±9.6	8.1± 0.1	23.5. ±.0.1	49.7± 5.2
23	367.8±3. 9	138.5± 5.8	398.2±7. 3	5.7±0. 1	34.2. ±.3.2	1568± 15.6	1243± 13.1	1423 ±11.2	7.7± 0.1	24.1. ±.0.1	29.5± 3.3

24	412.3±7.7	143.2± 6.7	414.1±4.7	6.3±0.2	32.1.±.2.6	1452±16.5	1156±13.1	1231±12.2	8.3±0.2	22.1.±.0.11	43.6±4.3
25	394.3±8.3	137.4± 7.4	437.2±6.3	6.1±0.4	33.2.±.2.6	1350±24.5	1231±12.4	1276±12.8	8.1±0.4	24.2.±.0.1	39.5.±4.1
26	428.1±6.2	142.2± 7.7	432.1±4.3	6.4±0.3	29.1.±.2.1	1337±16.2	1005±11.4	1358±11.3	7.4±0.3	24.3.±.0.1	41.8±.3.2
27	369.3±5.4	141.3± 4.9	467.2±3.9	5.9±0.1	31.1.±.2.2	1414±16.7	1489±11.9	1325±9.8	7.9±0.1	22.1.±.0.1	46.5±3.2
28	421.2±4.8	136.2± 6.8	456.7±6.6	6.2±0.1	42.1.±.3.1	1244±18.9	1366±11.2	1154±12.3	8.2±0.1	22.7.±.0.4	42.5±4.3
29	438.1±7.6	142.1± 4.8	434.5±4.1	6.3±0.2	41.1±.3.1	1487±14.4	1411±14.1	1235±10.2	8.3±0.2	.24.1±.0.1	38.5±.4.1
30	407.1±7.2	141.2 ±5.9	477.3±6.9	6.2±0.3	43.2.±.2.4	1432±18.1	1423±13.6	1369±11.2	8.2±0.3	22.7.±.0.2	51.2±4.5

Tables of output: Cor = Partial correlation coefficient, PSReg = Partial standardised regression coefficient
PReg = Partial regression coefficient, SE = Standard error of partial regression coefficient
p = alpha (α) probability of Type I error of PReg

Tab.3.Stepwise removal of the least effective independent variable

Var	PCor	PSReg	PReg	SE	t	p
1	0.3523	0.1784	0.1368	0.0834	1.6407	0.1182
2	-0.3166	-0.2233	-0.3765	0.2588	-1.4547	0.163
3	-0.4519	-0.3858	-0.1714	0.0776	-2.208	0.0405
4	-0.5739	-0.4435	-18.7313	6.1312	-3.0551	0.0068
5	0.3942	0.2363	0.8959	0.4791	1.8699	0.0779
6	-0.3673	-0.2078	-0.0402	0.0234	-1.7213	0.1023
7	-0.116	-0.0929	-0.0223	0.0437	-0.5089	0.617
8	0.3522	0.303	0.0353	0.0215	1.6401	0.1183
9	0.2254	0.1335	3.9988	3.9658	1.0083	0.3267
10	-0.1375	-0.0944	-2.3294	3.8491	-0.6052	0.5526

const=248.5634
R = 0.8676, R Sq = 0.7527.

Table.4.Analysis of variance

	df	SSq	MSq	F	p
Reg	10	5970.6353	597.0635	12.4481	<0.0001
Res	19	911.3235	47.9644		
Tot	29	6881.9587			

dependent variable and the predictors are linear and the residuals variances are equal or constant. The water quality variation in winter season was explained by five predictor variables namely: DO, BOD5, SS, TS and Cl. The R-square of 0.94.2 revealed that 94.2% of the variation of water quality as explained by the mentioned five predictors. The estimate of coefficient of the model is presented in table 3. The Beta coefficient among the parameters calibrated by stepwise regressions analysis, TS makes the strongest unique contribution in water quality variation (0.668). The Beta value for DO (0.547) was the second highest, followed by Cl (0.545), BOD5 (-0.491) and the least contributor was SS with -0.292. The ANOVA (Table.4) showed that the *F*-statistics ($F = 12.4481$) was very large and the corresponding p value is highly significant ($p = 0.0001$) or lower than the alpha value (0.01). This indicated that, the slope of the estimated linear regression model is not equal to zero, confirming that, there is linear relationship between the predictors of the models.

Conclusion: In this study, Multiple linear regression supported the above result and identified the contribution of each variable with significant values R = 0.8676, R Sq = 0.7527. These statistical tools provide more objective interpretation of surface water quality variables. From the analysis, it is clear that DO, BOD5, SS, TS, Cl, and conductivity were found to be the most abundance parameters responsible for water pollution in Chambal River. Therefore, there is need to properly manage wastes in the city and monitor human activities, in order to ensure minimal negative effects on the rivers.

References:

1. P.B.Satheeshkumar, K.Anisa, Identification of mangrove water quality by multivariate statistical analysis methods in Pondicherry coast, India," Environmental Monitoring Assessment, doi: pp.10.1007/s10661-011-4. 15th July 2011.
2. R. Manikannan, S. Asokan, A.M. Samsoor-Ali, Seasonal variations of physic-chemical properties of the Great Vedaranyam Swamp, Point Calimere Wildlife Sanctuary, South-east coast of India," African Journal of Environmental Science & Technology, 5(9): pp.673-681.2011.
3. M.Milovanovic, Water quality assessment and determination of pollution sources along the Axios/Vardar River, Southeastern Europe," Desalinization, 213(1-3): pp.159-173.2007.
4. K.P.Singh, A. Malik, S. Sinha, K.Vinod, R.C Murthy, "Estimation of source of heavy metal contamination in sediments of Gomti River (India) using principal components analysis," Water Air and Soil Pollution, 166(1-4: pp.321-341.2005.
5. F.Huang, X.Wang, L.Liping, Z.Zhiqing, and W. Jiaping, Spatial variation and source apportionment of water pollution in Qiantang River (China) using statistical techniques," Water Research, 44:pp.1562-1572,2010.
6. Responses in fish due to pollution at Nagda. M.P.India. DAV Int. Jrnl.Sci. -1, (1) pp. 86-91.2012.
7. S.Shrestha, F. Kazama, Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan," Environmental Modelling & Software, 22(4): pp.464- 475.2007.
8. H.Juahir, M.S. Zain, M.K Yusoff, T.I.T.Hanidza, A.S Mohd Armi, M.E.Toriman, M. Mokhtar, Spatial water quality assessment of Langat River Basin (Malaysia) using environ metric techniques," Environmental Monitoring and Assessment, 173(1-4): pp.625-641.2011.
9. P.B. Reddy and B.S. Baghel, Impact of Industrial waste water on the Chambal River and Biomarker P.B.Reddy, and B.S.Baghel, Impact of the industrial effluent on physico Chemical characteristics of Chambal River, Nagda, M.P. Nature. Pollu. Tech 9, (3) pp.593-595. 2010.
10. R.Koklu, B. Sengorur, B. Topal, B, Water quality assessment using multivariate statistical methods, a case study: Melen River system," Water Resource management, (24)5: pp.959-978.2010.
11. APHA, Standard Methods for the Examination of Water and Wastewater, 19th edition. American Water Works Association, Washington, DC.1998.http://www.stattools.net.

***

PG Dept of Mathematics, Asst.Professor.Govt.PG College, Jaora (M.P)
PG Dept of Zoology, Assoc.Professor.Govt.Arts and Science PG College,Ratlam,
E.mail: reddysir@yahoo.co.in