Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 6, Issue, 02, pp.5005-5008, February, 2014

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

ASSESSMENT OF SEASONAL AND POLLUTING EFFECTS ON THE QUALITY OF RIVER INDUS WATER BY USING MULTIPLE LINEAR REGRESSION ANALYSIS

*[1]Aziz-ur-Rahman, [2]Pervez Akhter and [3]Muhammad Mansha

[1,2]Hamdard University, City Campus-II, Karachi, Pakistan
[3]Pakistan Space and Upper Atmosphere Research Commission (SUPARCO), Karachi, Pakistan

ARTICLE INFO

Article History:
Received 15th November, 2013
Received in revised form
09th December, 2013
Accepted 26th January, 2014
Published online 21st February, 2014

Key words:
Water Quality,
Multiple linear regressions Analysis,
Indus River,
Monsoon.

ABSTRACT

In this study, statistical techniques such as Multiple linear regression analysis is applied to water quality data set monitored during Pre and Post monsoon 1982 -2012 to investigate in the extent of pollution and seasonal variation in river Indus waters. The variables were pH, BOD (Biochemical Oxygen Demand), COD (Chemical Oxygen Demand, TSS (Total Suspended Solid), TDS (Total Dissolve Solid), alkalinity, Cl^-, HCO_3^-, SO_4^{-2}, Ca^{+2}, Mg^{+2}, Na^+ and K^+. Multiple linear regression equations established between above parameters and dependent variable such as electrical conductivity, which caused to predict the value of one parameter, if value of other is known. The above study provides us a tool to find the value of physico-chemical parameters and extent of pollution theoretically and seasonal variation, which is time saving as well as cost effective. The result obtained provides a method to characterize river water quality using statistical multiple linear regression analysis.

INTRODUCTION

Water is vital to the existence of living organisms but this valued resource is increasingly being threatened as human population grown which increases the demand for more water of high quality for domestic purposes and economic activities (UNESCO 2003). The river Indus and its tributaries are the major source of water for drinking, agricultural and industrial desires in province of Sindh, Pakistan. The river cover a drainage area of exceeding 1,165,000 km^2, flows through densely populated areas from the North in a Southerly direction along the entire length of the country and finally merge into Arabian Sea near port city of Karachi, Pakistan. In the bank of Indus River in the province of Sindh, Pakistan big cities like Sukkur, Nawabsha, Dadu and Hyderabad are situated along the bank of the river with numerous small towns and villages. Additionaly, industrial estates like Sukkur Industrial Estate, Hyderabad and Kotri Industrial Estates are also situated on the side of the Indus River. These industrial estates are releasing their effluents into River Indus without treatment. Therefore, the River water is getting contaminated by both domestic and industrial waste as well as agricultural return flow. In addition, hydrology of Indus River altered significantly by the construction of channels, barrages, embankments, dykes and dams for hydroelectric power generation, irrigation and flood control. These have reduced the discharge significantly in the lower Indus region. The water quality in the region depends upon the water flow. There are concerns that extensive deforestation, industrial pollution, agriculture return flow and global warming are affecting the vegetation and wildlife of the Indus delta, while affecting agriculture products as well, in addition to quality of water. Environmental targets for water quality are implicitly biological in broad sense. First, there is consideration of human health. Are water contact hazardous and how much treatment is necessary before it is safe to drink or use in various ways? Second, are fish and other desirable 'living' things abundant? Is conservation further thread? Chemistry is important in that it enables environmental targets to be met and because various standard chemical determents have proved useful in surveillance because they are easy to measure and monitor. Compared with other parts, relatively little hydrological research has been conducted in the area of the impact of seasonal limnological variation, nutrient load and urbanization on the downstream of River Indus system, especially in the territories of the Sindh Province. This study addresses these gaps by assessing the water quality of River Indus.

***Corresponding author:** *Aziz-ur-Rahman*
Hamdard University, City Campus-II, Karachi, Pakistan.

Therefore, various Physico-chemical parameters were analyzed by many workers to understand the quality of surface water. Some researchers suggested the empirical relationship to measure the quality of water. Recently, multiple linear regression analysis technique has been established to evaluate the river water quality, which is cost effective as well as reliable (Agarwal and Sexena 2011; Mulla *et al.*, 2007).

Multiple linear regressions analysis has become one of the most widely used statistical tools for both experimental and non-experimental research by analyzing multifactor data. It is appealing because it provides a theoretically simple method for investigating functional relationships among variables. In recent decades, this statistical technique has been employed to extract significant information from physico-chemical datasets in compound systems (Pathak 2011, 2012a, b, 2013). Multiple linear regressions (MLR) are an accurate tool to evaluate River water quality, since it generates a minimum data set of indicators. Thus, an equation is explored for the sampled River water. Using E-views software, the structural equation modeling allows, to test in simultaneous analysis the entire system of variables (Physico-chemical parameters), in order to determine the extent to which it is consistent with the data. For this purpose, it should investigate simultaneously the interactions between the different Physico-chemical parameters of River water and their relationship with Electrical Conductivity.

MATERIALS AND METHODS

Seven (07) different sampling locations (Fig.1) are selected to study the physico-chemical parameters of Indus River; these are receiving bulk quantity of effluent either by municipal/ industrial/ agriculture return flow, resulting in reasonably high pollution.

Fig. 1. A Layout and sampling point of study area

Chemical analysis

Most of the samples were collected from the middle of the flow using boat, two to four sub samples of equal volume were collected from vertical section. The water samples were collected within 3-9 inches from the surface of water. The samples were mixed well and a sample of 2.5 L was transferred to clean glass bottle. The samples were analyzed for pH, electrical conductivity, total dissolved solid, and total suspended solid, alkalinity, chloride, hydrogen carbonate, chemical oxygen demand, biological oxygen demand, sulphate, sodium, potassium, calcium and magnesium using standard methods (APHA, 2005). Results obtained were subjected to multiple linear regression statistical analysis using Statistical Package E-Views.

Statistical Analysis

Multiple linear regressions - Selecting the Best Equation When fitting a multiple linear regression model, a researcher will likely include independent variables that are not important in predicting the dependent variable Y(Weisberg 1980). In the analysis one will try to eliminate these variables from the final equation. The objective in trying to find the "best equation" will be to find the simplest model that adequately fits the data. This will not necessarily be the model explains the most variance in the dependent variable Y (the equation with the highest value of R^2). This equation will be the equation with all of the independent variables in the equation. The objective will be to find the equation with the least number of variables that still explain a percentage of variance in the dependent variable that is comparable to the percentage explained with all the variables in the equation. Multiple linear regressions method provides equation linking a dependant variable Y to the independent variable X using the multiple linear regression analyses; examine the relationships between Electrical conductivity and various physico-chemical parameters.

$$Y_D = \beta_o + \beta 1\, X_i1 + \beta 2X_i2 \text{------------} + \beta nX_in + \varepsilon_t$$

The intercept β_o and the regression coefficients of descriptors (βi) are determined by least square method. Xi descriptors are used to describe Physico-chemical parameters and (n) is the number of water samples while ε_t disturbance term. The reduction in the number of descriptors (variables) is included in the study to minimize the information overlap in variables. The best equation is selected while being based on the highest multiple correlation coefficients (R), lowest standard deviation (SD). Relationships between variables were established using the forward stepwise regression method. The MLR modeling method was performed by the E-Views.

RESULTS AND DISCUSSION

Multiple linear Regression analysis was established to find out the relationships between Electrical Conductivity (EC) and physico-chemical parameters. EC was taken as Criterion and 13 physico-chemical parameters were considered as Predictors. After developing various equations with river water samples, an analysis of residuals was established and R^2 values were obtained. The 13 predictors' variables account for 99.9% of the variability of EC (Adjusted R-squared=0.99) in pre and post-monsoon season. The positive sign of the beta coefficients pertaining to these variables indicates that there is a positive relationship between EC and physico-chemical parameters.

Multiple Linear Regression Analysis for different Parameters

Table 1-A (Pre-Monsoon)

Estimation Equation:
=====================
EC = β (1) + β (2)*Al + β (3)*BOD + β (4)*Ca + β (5)*Cl + β (6)*COD + β (7)*HCO_3 + β (8)*K + β (9)*Mg + β (10)*Na + β (11)*pH + β (12)*SO_4 + β (13)*TDS + β (14)*TSS

Substituted Coefficients:
=====================
EC = -0.380 - 0.0099*Al + 0.0155*BOD + 0.0005*Ca + 0.0135*Cl - 0.0043*COD + 0.007*HCO_3 - 0.017*K - 0.0128*Mg - 0.0069*Na - 0.001*pH - 0.004*SO_4 + 0.0025*TDS + 4.35e-05*TSS

Dependent Variable: EC
Method: Least Squares
Date: 01/01/13 Time: 02:23
Sample (adjusted): 1982 2012
Included observations: 31 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	-0.380	0.546	-0.695	0.496
Al	-0.0099	0.011	-0.879	0.392
BOD	0.0155	0.013	1.193	0.249
Ca	0.0005	0.0016	0.339	0.739
Cl	0.0135	0.0218	0.620	0.543
COD	-0.0043	0.0027	-1.561	0.137
HCO_3	0.007	0.0106	0.631	0.536
K	-0.017	0.025	-0.687	0.502
Mg	-0.0128	0.0096	-1.341	0.197
Na	-0.0069	0.023	-0.302	0.766
pH	-0.001	0.026	-0.042	0.967
SO_4	-0.004	0.024	-0.174	0.864
TDS	0.0025	0.004	0.655	0.521
TSS	4.35E-05	0.0004	0.098	0.923
R-squared	0.999	Mean dependent var		0.602
Adjusted R-squared	0.999	S.D. dependent var		0.173
S.E. of regression	0.003	Akaike info criterion		-8.256
Sum squared resid	0.0002	Schwarz criterion		-7.608
Log likelihood	141.902	F-statistic		6149.711
Durbin-Watson stat	1.255	Prob(F-statistic)		0.000

Table 1-B (Post-Monsoon)

Estimation Equation:
=====================
EC = β (1) + β (2)*pH + β (3)*Al + β (4)*BOD + β (5)*COD + β (6)*TSS + β (7)*TDS + β (8)*Cl + β (9)*HCO_3 + β (10)*SO_4 + β (11)*Mg + β (12)*K + β (13)*Na + β (14)*Ca
Substituted Coefficients:
=====================
EC = -1.024 - 0.007*pH + 0.004*Al - 0.0098*BOD + 0.0025*COD + 0.0017*TSS - 0.0067*TDS - 0.007*Cl + 0.029*HCO_3 + 0.0088*SO_4 - 0.032*Mg + 0.074*K + 0.0345*Na - 0.017*Ca

Dependent Variable: EC
Method: Least Squares
Date: 01/01/13 Time: 02:23
Sample: 1982 2012
Included observations: 31

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	-1.024	3.786	-0.270	0.790
pH	-0.007	0.0167	-0.426	0.676
Al	0.004	0.011	0.389	0.703
BOD	-0.0098	0.016	-0.626	0.540
COD	0.0025	0.009	0.288	0.777
TSS	0.0017	0.003	0.555	0.586
TDS	-0.0067	0.002	-3.078	0.007
Cl	-0.007	0.014	-0.499	0.624
HCO_3	0.029	0.038	0.756	0.461
SO_4	0.0088	0.0145	0.606	0.553
Mg	-0.032	0.016	-1.957	0.068
K	0.074	0.017	4.376	0.0005
Na	0.0345	0.015	2.313	0.034
Ca	-0.017	0.018	-0.954	0.354
R-squared	0.999	Mean dependent var		-0.014
Adjusted R-squared	0.999	S.D. dependent var		0.271
S.E. of regression	0.0021	Akaike info criterion		-9.1405
Sum squared resid	7.41E-05	Schwarz criterion		-8.487
Log likelihood	151.108	F-statistic		35508.91
Durbin-Watson stat	1.096	Prob(F-statistic)		0.000

Conclusion

According to MLR the independent variables are indirectly related to dependant variable. Therefore in Pre-monsoon period, If independent variables such as BOD, Ca, Cl, HCO_3, TDS and TSS is increased by one unit then EC will be increased by 0.0155, 0.0005, 0.0135, 0.007, 0.0025, 0.0000435 respectively while if independent variable such as Al, COD, K, Mg, Na, pH and SO_4 is decreased by one unit then EC will be decreased by -0.0099, -0.0043, -0.017, -0.0128, -0.0069, -0.001, -0.004 respectively. In post-monsoon period, If Independent variable such as Al, COD, TSS, HCO_3, SO_4, K and Na is increased by one unit then EC will be Increased by 0.004, 0.0025, 0.0017, 0.029, 0.0088, 0.074, 0.0345 respectively while if Independent variable such as pH, BOD, TDS, Cl, Mg and Ca is decreased by one unit then EC will be decreased by -0.007, -0.0098, -0.0067, -0.007,-0.032 and -0.017. The F-Statistic value of both the tables indicates that overall model is statistically significant. Therefore, Durbin-watson Stat value is greater than R^2, means results are not Spurious.

REFERENCES

Agarwal, and Sexena, M. 2011. Assessment of pollution by physico-chemical parameters using regression analysis: A case study of Gagan River at Moradabad-India. Advances in applied science research., 2 (2):185-189.

APHA, 2005. Standard methods for the examination of water and waste water, 21st edition, American Public,Health Association, Washington, DC., USA.

Mulla, J. G., Farooqui, M., Zaheer, A. 2007. A correlation and regression equations among water quality parameters. *Int. J. chem. sci*., 5(2): 943-952.

Pathak H. 2012a, *J Environ Anal Toxicol*, 2(5), 01-05.

Pathak H., Limaye S. N. 2011. Interdependencey between Physicochemical Water Pollution Indicators: A Case Study of River Babus, Sagar, M.P., India analele Universităii din Oradea – Seria Geografie., 1: 23-29.

Pathak, H. 2012b. Evaluation of ground water quality using multiple Linear regression and athematical equation Modeling Analele UniversităŃii din Oradea – Seria Geografie., pp. 304-307.

Pathak, H. 2013. Indicators of the deteriorate water quality status of reservoir, Sagar city, MP, India by multivariate analysis. Analele Universitatii "Ovidius" constanta-Seria Chimie,23(2): 155–162, DOI: 10.2478/v10310-012-0026-1.

UNESCO. 2003. Water for people, water for life. UN World Water Development Report (WWDR).

Weisberg, S. 1980. Applied Linear Regression. Wiley and Sons, New York.

****