



Correlation and linear regression as a water quality monitoring tool for Serlui River in Kolasib district, Mizoram, India

Sangeeta Sunar, BP Mishra

Department of Environmental Science, Mizoram University, Aizawl, Mizoram, India

Abstract

The Serlui River is one of the important rivers in Kolasib, Mizoram. The present investigation was carried out for a period of two years (March 2015- February 2017) to study the water quality of Serlui River in vicinity of Serlui-B hydel project using statistical tools for significance and validity of findings. Altogether, three sampling sites were selected from upstream to downstream of river along the hydel project. The water samples were collected at monthly interval water quality attributes namely, pH, Acidity, Total Alkalinity, Dissolved Oxygen and Biological Oxygen Demand were analyzed following the standard methods. A correlation matrix test ($p < 0.05$) and linear regression analysis were computed to check the significant relationship between the water quality parameters. The findings reveal that the values were within the prescribed limits given by various scientific agencies, but intensity of pollutants increased from Site 1 (Control Site) to Site 3 (Downstream), indicating impact of hydel project on river water quality. From the statistical analysis, a highly positive and significant correlation was observed between pH and DO ($r = 0.940$, $p < 0.01$), Acidity and BOD ($r = 0.733$, $p < 0.05$), whereas, a highly negative and significant correlation was recorded between pH and Acidity ($r = -0.841$, $p < 0.01$), pH and BOD ($r = -0.938$, $p < 0.01$), Acidity and DO ($r = -0.764$, $p < 0.05$), DO and BOD ($r = -0.950$, $p < 0.01$). Total alkalinity showed a moderate correlation with pH ($r = -0.619$), Acidity ($r = -0.543$), and BOD ($r = -0.642$) and a weak correlation with DO ($r = -0.472$). The findings reveal that methodical calculation of correlation coefficient and regression analysis proved to be a convenient mean for rapid monitoring of the water quality of Serlui river.

Keywords: correlation coefficient, regression, Serlui-b dam, Serlui River

1. Introduction

Water is very important for the survival and sustainability of life on the earth. For most of our domestic purposes (including drinking, irrigation, power supply etc.), is mostly available as surface water in the form of rivers, ponds, and lakes. But, the rapid urbanization and developmental activities have led to severe environmental stress and deterioration of these surface water systems. The problematic degradation of our river ecology have requisite the monitoring of the numerous rivers throughout the country to assess their utility potential, production capacity and to plan proper management and restoration strategies [1, 2]. For determination of the water quality parameters, analytical methods are used to produce dependable outcomes but usually the laboratory approaches are time consuming and very expensive.³ However, in recent years, an easier and advanced method based on statistical tools have been developed to gather data and provide necessary information of the various parameters using mathematical relationships between the parameters [4, 5, 6]. The degree of association that exists between two variables, taking one as dependent variable and other as independent is measured by the correlation coefficient (r). A direct correlation exists between two variables, when the increase or decrease in the value of one is associated with the corresponding increase or decrease of the other.⁷ If the calculated correlation coefficient between two variables is nearer to +1 or -1, there is a probability of linear relationship between the two variables in such case a linear regression equation is established [8]. The

regression equation is a statistical tool used to predict the unknown values of one variable from the known value of another variable. The statistical analysis attempt to establish the nature of the relationship between the variables using the correlation coefficient and linear regression method and thus provides an implement for prediction and forecasting of the variables [9].

Therefore, in the present research an approach have been made to establish correlation and regression between water quality attributes of the Serlui River through physicochemical characteristics analysis.

2. Materials and Methods

The Serlui river flows through Kolasib district of Mizoram and is impounded by the Serlui-B Dam, 12 km from Bilkhawthlir Village in the Kolasib district. The Serlui-B has a 293m (961 feet) long, 51m (167 feet) high earthfill embankment dam, a 135m pressure tunnel and a semi-ground power house [10]. The hydel project has 3 units each with a capacity to generate 4MW power. The dam creates a reservoir catchment area of 53 square kilometers with life storage capacity of 453.59 cubic million [10]. Altogether three sampling sites along the hydel project were selected to study the water quality of Serlui River.

- 1. Site 1-** This site is situated at the upstream of the dam with least anthropogenic activities and maintains its natural flow, and is demarcated as reference (control site).
- 2. Site 2-** This is the reservoir of the river where the flow of

the water recedes with the development of the dam.

- 3. Site 3-** This is where the confluence of diversion outlet, spillover and power house meets towards the downstream of the river.

2.1 Sampling and Methods

For analysis of water quality, the samples were collected from selected study sites at monthly interval (in triplicates) for two years (March 2015- February 2017). Wide mouth bottles were used to collect the samples with necessary precautions.

The water samples were analyzed for various parameters namely; pH, Acidity, Total Alkalinity, Dissolved Oxygen (DO) and Biological Oxygen Demand (BOD) following the methods as outlined in the 'Standard Methods for Examination of Water and Wastewater' as prescribed by APHA^[11] and compared with standards given by as USPH^[12], BIS^[13], WHO^[14] and ICMR^[15].

The pH was determined at the site using electronic pH meter. For Dissolved Oxygen content, the water samples were fixed immediately after collection. The samples were stored in 4°C for further analysis of various parameters.

2.2. Statistical Analysis

All the obtained data were subjected to statistical analysis. To check the significant relationship among all the physico-chemical parameters (significance level 0.05), the correlation coefficient (r) was calculated using the computer software SPSS 16.0. The linear regression model was developed for water quality parameters having highly significant correlation coefficients (r) using SPSS 16.0 and MS-Excel 2013. The findings on water quality attributes were expressed as average mean values for two years data (March 2015 to February 2017).

3. Results and Discussion

The average mean values, minimum and maximum values, standard deviation and standard error for selected water quality attributes are presented in Table 1, and compared with various scientific agencies (Table 2). The minimum pH (6.1) was recorded at Site 3 during the monsoon season and the maximum value (7.5) at Site 2 during the post-monsoon season (Fig 1a). Acidity was minimum (37mgL^{-1}) at Site 1 during the pre-monsoon season and maximum (69mgL^{-1}) at Site 2 during the post-monsoon season (Fig 1b). Total alkalinity was minimum (34mgL^{-1}) at Site 1 during the post-monsoon season and maximum (63mgL^{-1}) at Site 2 during the monsoon season (Fig 1c). The DO content is an indicator of the extent of pollution in the water bodies. During the present study, the minimum DO (5.4mgL^{-1}) was recorded at Site 3 during the monsoon season and maximum (7.8mgL^{-1}) at Site 2 during the post-monsoon season (Fig 1d). The BOD content was minimum (0.8mgL^{-1}) at Site 2 during the post-monsoon season and maximum (2.4mgL^{-1}) at Site 3 during the monsoon season (Fig 1e). A similar trend of results was reported by earlier workers^[16, 17, 18, 19, 20].

From the statistical analysis (Table 3) it is evident that a highly positive and significant correlation was observed between pH and DO ($r=0.940$, $p<0.01$), Acidity and BOD ($r=0.733$, $p<0.05$), whereas, a highly negative and significant correlation was established between pH and Acidity ($r=-0.841$,

$p<0.01$), pH and BOD ($r=-0.938$, $p<0.01$), Acidity and DO ($r=-0.764$, $p<0.05$), DO and BOD ($r=-0.950$, $p<0.01$). Total alkalinity showed a moderate correlation with pH ($r=-0.619$), Acidity ($r=-0.543$), and BOD ($r=-0.642$) and a weak correlation with DO ($r=-0.472$). If the range of the correlation coefficient is from +0.8 to 1.0 and -0.8 to -1.0 it is termed as strong, from +0.5 to 0.8 and -0.5 to -0.8 as moderate and weak when it is in the range of +0.0 to 0.5 and -0.0 to -0.5.²¹ In the present study we have established four strong correlations (pH-Acidity, pH-DO, pH-BOD and DO-BOD), five moderate correlations (pH-Alkalinity, Acidity- Alkalinity, DO-Acidity, BOD-Acidity and Alkalinity-BOD) and only one weak correlation (Alkalinity-DO). Altogether six regression equations were derived for the parameters having moderate to strong correlation coefficients. The regression equation $\text{DO}=1.413\text{pH}-2.876$, $\text{BOD}=-0.947\text{pH}+7.995$ (Fig: 2a) and $\text{Acidity}=-17.692\text{pH}+173.4$ (Fig: 2b) can be used to measure the values of DO, BOD and Acidity respectively by substituting the known values of pH in the equations. Similarly, from the regression equations $\text{DO}=0.054\text{Acidity}+9.66$ and $\text{BOD}=0.035\text{Acidity}-0.3353$ (Fig 2c) the values of DO and BOD can be measured respectively by substituting the known values of Acidity in the equation and from the equation $\text{BOD}=-0.638\text{DO}+5.845$ (Fig 2d) the BOD values can be estimated by substituting the known values of DO in the equation. As DO exhibited negative correlation with most of the parameters, therefore, it can be considered as an useful pollution index of water quality as with the increase in the values of the other parameters there in a significant decrease in the DO content^[22]. A negative correlation between DO and BOD indicates that the decrease in DO content is linked with the accelerated decomposition and oxidation of organic matter^[22].

4. Conclusion

The present investigation involves a new approach to evaluate the water quality parameters of Serlui River using correlation coefficient and linear regression method. Using correlation coefficient and linear regression tools in this work, a significant relationship has been obtained between different pairs of water quality parameters to describe the realistic water quality status of Serlui River. The application of mathematical equation modeling to evaluate the water quality of Serlui River showed that most of the parameters are either positively or negatively correlated to each other. Most of the correlated variables are influenced by one or more other variables. DO and pH are important parameter for determination of water quality as they are significantly correlated with most of the parameters. Implementing linear regression method can be a major approach to get an indication of the water quality of the water by determining a few factors experimentally. The regression equation obtained from the study can be widely used to estimate the unknown values by substituting the known values in the equation and therefore it can be a helpful technique to predict the water quality status of Serlui river prior to the detailed monitoring as this techniques are precisely more convenient to get an accurate idea of the water quality and for the proper planning and sustainable management of the river system throughout its length. Based on the experimental results it can be concluded that the water

quality at Site 3 showed a higher pollution index than other Sites thereby it clearly shows that the Serlui river water is comparatively polluted at the downstream indicating the negative impact of the dam. Seasonally, the monsoon season showed higher values, this may be due to excess runoff from the nearby agricultural fields and villages carrying more organic loads indicating the extent of anthropogenic disturbance on the downstream areas. Even though all the parameters are within the permissible limits, there is an ample scope of appropriate management measures, as long term use of such water may adversely affect the lives of human as well as aquatic lives. The finding of the present study maybe a base line for further study on maintenance of water quality of

Serlui, through formulating appropriate management strategies.

Table 1: Descriptive Statistics of the water quality characteristics for two year data (Mar 2015- Feb 2017)

	Minimum	Maximum	Range	Mean	Std. Deviation	Std. Error
pH	6.1	7.5	1.4	6.9	0.4927	0.1642
Acidity	37	69	32	52	10.392	3.464
Alkalinity	34	63	29	49	10.259	3.420
DO	5.4	7.8	2.4	6.8	0.7407	0.2469
BOD	0.8	2.4	1.6	1.5	0.4975	0.1658

Table 2: Findings of present investigation in relation to the water quality standards given by various scientific agencies

Parameter	Water Quality Standards				Range of water quality characteristics during the study period
	USPH	BIS	WHO	ICMR	
pH (nano mole L ⁻¹)	6 - 8.5	6.5- 8.5	6.5-8.5	7- 8.5	6.1-7.5
Acidity(mgL ⁻¹ CaCO ₃)	-	-	-	-	37 mgL ⁻¹ -69 mgL ⁻¹
Total Alkalinity (mgL ⁻¹ CaCO ₃)	-	200	-	120	34 mgL ⁻¹ -63 mgL ⁻¹
DO (mgL ⁻¹)	>4	>5	-	-	5.4mgL ⁻¹ -7.8mgL ⁻¹
BOD (mgL ⁻¹)	-	<3	-	-	0.8 mgL ⁻¹ -2.4 mgL ⁻¹

Table 3: Correlation coefficient (r) between the water quality attributes for two years data

		pH	Acidity	Alkalinity	DO	BOD
pH	Pearson Correlation	1	-.841**	-.619	.940**	-.938**
	Sig. (2-tailed)		.004	.075	.000	.000
Acidity	Pearson Correlation	-.841**	1	.543	-.764*	.733*
	Sig. (2-tailed)	.004		.131	.017	.025
Alkalinity	Pearson Correlation	-.619	.543	1	-.472	.642
	Sig. (2-tailed)	.075	.131		.200	.062
DO	Pearson Correlation	.940**	-.764*	-.472	1	-.950**
	Sig. (2-tailed)	.000	.017	.200		.000
BOD	Pearson Correlation	-.938**	.733*	.642	-.950**	1
	Sig. (2-tailed)	.000	.025	.062	.000	

** . Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

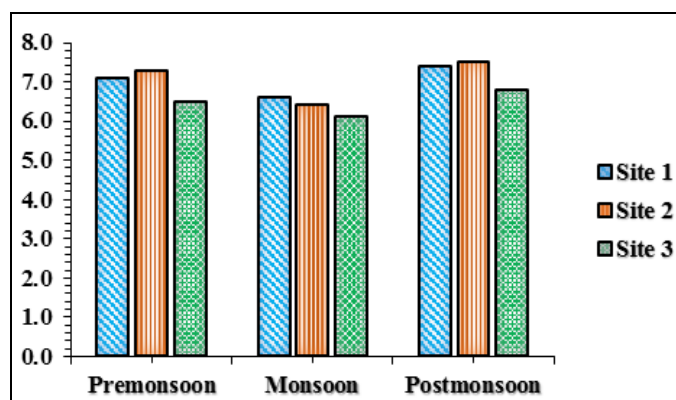


Fig 1a: Seasonal variation in pH values at selected study sites for two years average data

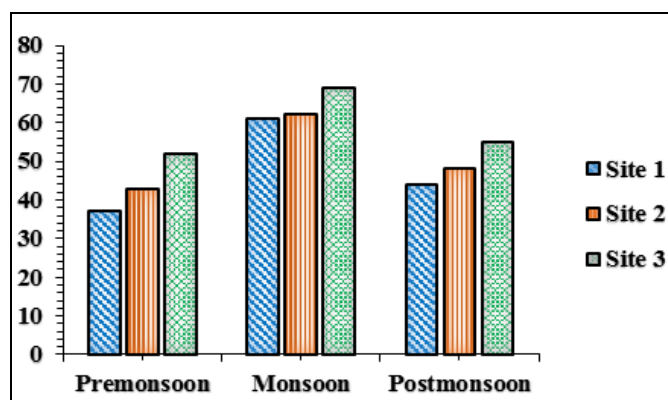


Fig 1b: Seasonal variation in Acidity values at selected study sites for two years average data

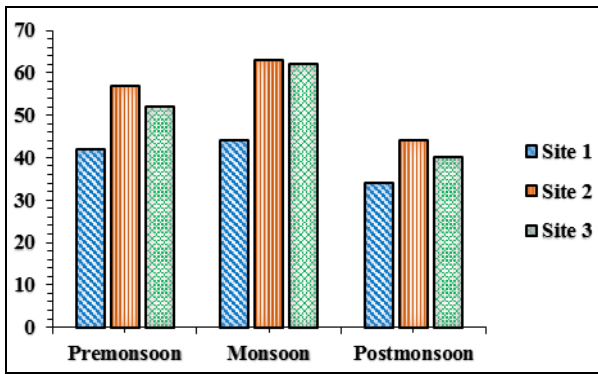


Fig 1c: Seasonal variation in Total Alkalinity values at selected study sites for two years average data

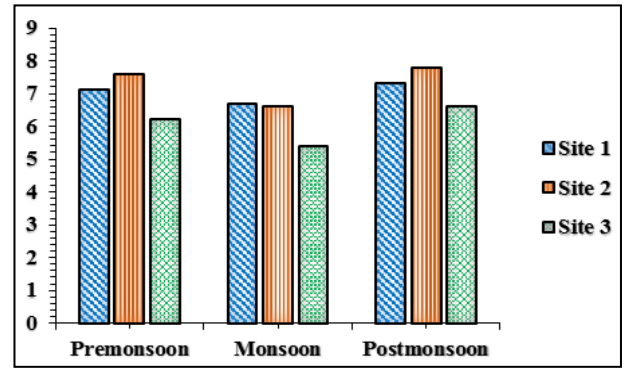


Fig 1d: Seasonal variation in DO values at selected study sites for two years average data

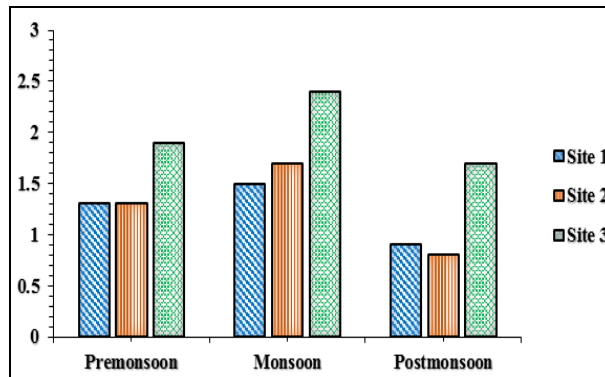


Fig 1e: Seasonal variation in BOD values at selected study sites for two years average data

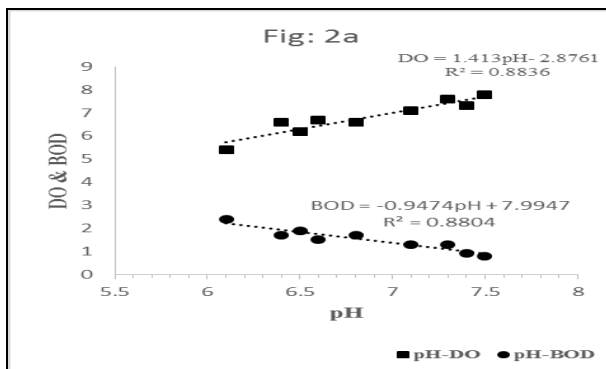


Fig 2a

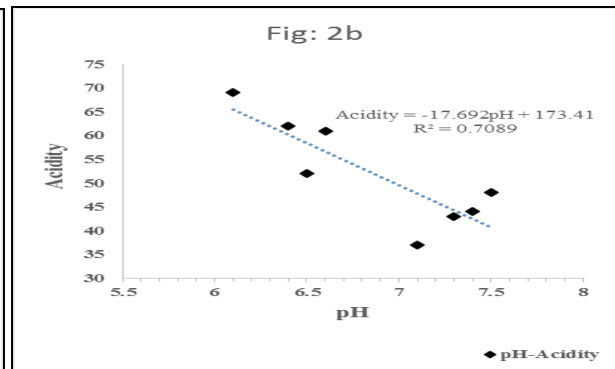


Fig 2b

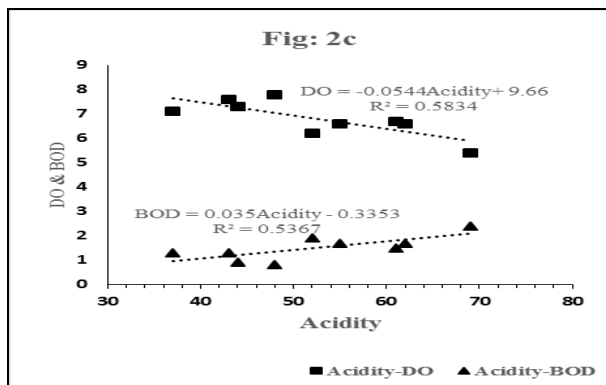


Fig 2c

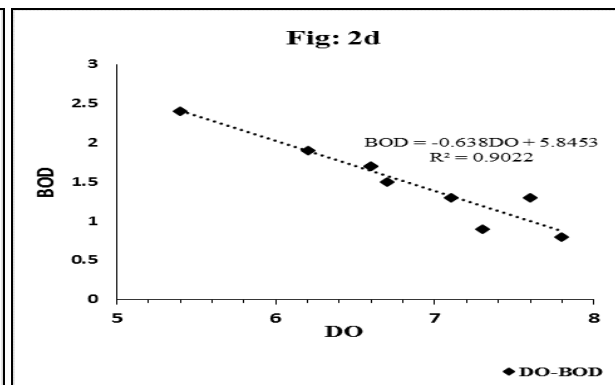


Fig 2d

Fig 2(a, b, c, d): Plots of water quality parameters as a linear regression model

References

1. Datar MO, Vashistha RP. Physicochemical aspects of pollution in river Betwa. *Ind J Environ Prot.* 1992; 12:577-580.
2. Das NK, Sinha RK. Assessment of water quality of Ganga River, Patna, India. *Environ Ecol.* 1993; 11:829-832.
3. Agarwal M, Agarwal A. Linear Regression and Correlation Analysis of Water Quality Parameters: A Case Study of River Kosi at District Rampur, India. *International Journal of Innovative Research in Science, Engineering and Technology.* 2013; 2(12):7273-7279.
4. Joshi DM, Bhandari NS, Alok K, Namita A. Statistical analysis of Physico-Chemical parameters of water of river Ganga in Haridwar district. *Rasayan. J. Chem.* 2009; 2(3):579-587.
5. Khatoon N, Khan AH, Rehman M, Pathak V. Correlation Study for the Assessment of Water Quality and its Parameters of Ganga River, Kanpur, Uttar Pradesh, India. *IOSR Journal of Applied Chemistry.* 2013; 5(3):80-90.
6. Patel JY, Vaghani MV. Correlation Study for the Assessment of Water Quality and its Parameters of Par River Valsad, Gujrat, India. *International Journal of Innovative and Emerging Research in Engineering.* 2015; 2(2):150-156.
7. Jothivenkatachalam K, Nithya A, Mohan SC. Correlation Analysis of Drinking Water Quality in and Around Perur Block of Coimbatore District, Tamil Nadu, India. *Rasayan journal.* 2010; 3(4):649-654.
8. Heydari MM, Abasi A, Rohani SM, Hosseini SMA. Correlation Study and Regression Analysis of Drinking Water Quality in Kashan City, Iran. *Middle-East Journal of Scientific Research.* 2013; 13 (9):1238-1244.
9. Kumar N, Sinha DK. Drinking water quality management through correlation studies among various physicochemical parameters: a case study. *International Journal of Environmental Sciences.* 2010; 1(2):253-259.
10. PHE. Department of Public Health Engineering. Bilkhawthlir, Kolasib, Mizoram, 2012.
11. APHA. Standard methods for the examination of water and waste water, 21st edition as prescribed by American Public Health Association (American Water Works Association and Water Environment Federation), Washington, 2005, 387-388.
12. USPH. Drinking Water Standards, PHS. Pub. US. Department of Health, Education and Welfare, Washington, 1962, 965-978.
13. BIS. Manual for Specifications for Drinking Water, BIS, India, 1983, 23-56.
14. WHO. Guidelines for drinking water quality, World Health Organization 3rd Edition, Switzerland, 2004, 1-515.
15. ICMR. Guidelines for Drinking Water Manual, New Delhi: Indian Council of Medical Research, India, 1996, 456-463.
16. Sunar S, Mishra BP. Assessing the impact of Hydroelectric Power Project on the water quality of Serlui River in Kolasib district, Mizoram, North-East India. *Int. Res. J. Environment Sci.* 2016; 5(9):40-44.
17. Mishra BP, Tripathi BD. Sewage quality analysis: Pollutants removal efficiency of a sewage treatment plant. *J. Ind. Pollut. Contr.* 2000; 16(2):239-251.
18. Mishra BP, Tripathi BD. Impact of city sewage discharge in physico-chemical characteristics of river Ganga water. *Asian J. Micro, Biotech and Env. Sc.* 2001; 3(4):333-338.
19. Mishra BP, Tripathi BD. Seasonal variation in physico-chemical characteristics of Ganga water as influenced by sewage discharge. *Indian J. Ecol.* 2003; 30(1):27-32.
20. Lalpamawii S, Mishra BP. Seasonal variation in water quality of Tuirial River in vicinity of the hydel project in Mizoram, India. *Sci Vision.* 2012; 12(4):159-163.
21. Achuthan N, Abdullan G, Mohamad I, Fadiel MM. *Poll Res.* 2005; 24(1):1-6.
22. Rawatkar S, Doifode S, Kanojiya A. Correlation and linear regression analysis of water quality parameters of Wainganga River at Desaiganj (Wadsa) in Gadchiroli district Maharashtra, India. *International Journal of Academic Research and Development.* 2016; 1(9):39-44.