INTRODUCTION

Now-a-days, the rapid increase in human population coupled with expanding urbanization, industrialization motorisation as well as high input of fertilizers and pesticides in agriculture is polluting the water bodies (Ajayi and Osibanjo, 1981; Osibanjo, 1996; Waziri et al., 2009). As a result, the aquatic ecosystems we depend are under stress and the availability of fresh water is at a very high risk (Osibanjo, 1996; Kulshreshtha, 1998). Therefore, it is empirical to analyze and monitor water quality to determine its suitability for various purposes like drinking, domestic, industrial, agricultural etc.

The physico-chemical and microbiological parameters of different fresh water systems (river, stream, ocean etc.) have been studied by various researchers (e.g. Panda et al., 1991; Sahu et al., 1991; Singh et al., 2004; Ghose et al., 2009; El-Korashey, 2009) in different areas. Recently, Kar et al. (2010) have undertaken a detailed investigation to monitor the physico-chemical and microbiological parameters of Mahanadi river water in and around Hirakud. These detailed investigations involving many parameters very often take lot of time and therefore, it becomes difficult to provide results and implement immediate management measures in emergency situations like an outbreak of contagious water born disease. Under the above circumstances, statistical investigation in environmental science can provide some satisfactory results to meet the requirements at emergency situations (Nemade and Shrivastava, 2004). With this backdrop, the present work is undertaken to study the applicability of regression analysis in estimation of Mahanadi river water quality parameters in and around Hirakud, Orissa India, using predictor (surrogate) parameters.

MATERIALS AND METHODS

Study area

The present study is confined to Mahanadi river stretch, in and around Hirakud, an industrial hub in the district of Sambalpur, located in the western part of Orissa (Fig. 1). A detailed description of the study area has been highlighted in the introductory section of the research article of Kar et al., (2010) published in this issue.

Sampling and analysis

Water samples were collected from four different water stretches in Mahanadi river namely reservoir stretch, upstream stretch, midstream stretch and down stream stretch at monthly intervals between October 2007 and September 2009.
covering four seasons i.e. monsoon (June-September), post monsoon (October-November), winter (December-February) and pre monsoon (March-May).

Standard methods (APHA, 1985; ISI, 1984) were used during collection, preservation and estimation of different physico-chemical and microbiological parameters like temperature, turbidity, dissolved oxygen, pH, conductivity, total hardness, total alkalinity, total dissolved solids, chemical oxygen demand, biological oxygen demand, fluoride, chloride, cyanide, sulphate, arsenic, mercury, faecal and total coliform bacteria.

**Statistical procedure**

A simple linear regression equation is usually computed to draw the relationship between two continuous variables: one response (y) and one predictor (x) and on the basis of this, it is also possible to predict the value of a response variable from a predictor (surrogate) variable with greater accuracy. Regression provides the line that “best” fits the data which is based on the least-squares criterion. Helsel and Hirsch (1995) and El-Korashey (2009) explained the rationale and methodology for expressing water quality constituent concentrations in terms of other surrogate constituents or physical properties using linear regression equation.

In the present study, Minitab 15.10 software was used to examine the applicability of regression equations between the selected constituents (parameters) i.e. conductance, total dissolved solids, biological oxygen demand, dissolved oxygen, mercury, arsenic, chloride and sulphate using water-quality assayed data from above mentioned four different stretches during 2007-2009. To test the developed regression equations from the assayed data of 2007-08, the equations were applied to the assayed data for surrogate parameters of 2008-09 to calculate the concentrations of response parameters. The surrogate parameters used for biological oxygen demand, mercury, arsenic and sulphate were dissolved oxygen, total dissolved solids, conductance and chloride, respectively.

To measure the effectiveness of regression relation, the measured and estimated constituent concentrations of response parameters for the year 2008-09 were compared by calculating median Relative Percentage Difference (RPD) using the following equation.

\[
\text{RPD} = \frac{E-M}{M} \times 100
\]

Where:  
E is the constituent concentration estimated from the regression equation  
M is the measured constituent concentration.

**RESULTS AND DISCUSSION**

The measured data for selected constituents or parameters (conductance, total dissolved solids, biological oxygen demand, dissolved oxygen, mercury, arsenic, chloride and sulphate) for the period 2007-08 and 2008-09 are given in Table 1. Regression relations between constituents were examined and regression equations were developed for some particular constituents (biological oxygen demand, mercury, arsenic and sulphate) using the surrogate variables (dissolved oxygen, total dissolved solids, conductance and chloride) as
per the methods mentioned above. The regression equations and associated $R^2$, adjusted $R^2$, Pearson correlation $r$ and median Relative Percentage Difference (RPD) for biological oxygen demand, mercury, arsenic and sulphate are given in Table 2. The regression equations were evaluated using common measures of variability, including $R^2$ and Pearson corr. $r$. $R^2$ values ranged from 89.9 % for sulphate to 70.3 % for arsenic, while $r$ values ranged from 0.916 for mercury to -0.843 for biological oxygen demand.

**Biological oxygen demand (BOD)**

Biological oxygen demand is a measure of the quantity of dissolved oxygen used by microorganisms in the oxidation of organic matter (Metcalf and Eddy, 2003). Urban sewage, industrial effluents and agricultural runoff increase oxygen demand. Organisms that are highly tolerant of lower dissolved oxygen levels may replace a diversity of more sensitive organisms leading to imbalance in the system (Ryberg, 2007). Dissolved oxygen (DO) was chosen as explanatory variable for estimating biochemical oxygen demand (BOD) because these two parameters directly or indirectly depend on oxygen budget of the water sample. The range of dissolved oxygen values was 6.60 mg/L to 7.67 mg/L, whereas the range of biological oxygen demand values was 16.98 mg/L to 40.90 mg/L during 2007-08. The derived regression equation is:

$$\text{BOD} = 168 - 19.4 \text{DO} \quad \text{Equation (1)}$$

Where, $R^2$ for this equation is 76.6% with Pearson correlation coefficient of -0.843 (Fig. 2).

The measured BOD concentrations of 2008-09 were compared with estimated BOD concentrations calculated from equation -1 using the DO values of 2008-09 (Fig. 3) and the obtained RPD value was 13.9.

**Mercury**

The source of mercury in water may be attributed to burning of coal in captive power plants in the study area. Indian coal contains 0.17-0.32 ppm of mercury (Wankhade and Agrawal, 2003). The use of coal in steel plants also leads to mercury pollution of water through different pathways (Koshle et al., 2008). There is a coal belt in the catchment of the reservoir where mining activities are going on. The presence of captive power plants and steel industries in and around Hirakud may also be contributing mercury to the water. The pesticides used in the agricultural fields in the region may also contribute mercury load to the river water. Total Hg in water is sometimes strongly correlated with total dissolved solids (Berzas Nevado et al., 2008). The total dissolved solid (TDS) was chosen as explanatory variable for estimating mercury concentration in Mahanadi river water. The range of TDS values was 218.70 mg/L to 289.10 mg/L, whereas the range of mercury values was from 0.0008 mg/L to 0.0015 mg/L during 2007-08. The computed regression equation is as follows:

$$\text{Mercury} = -0.00141 + 0.000010 \text{TDS} \quad \text{Equation (2)}$$

Where, $R^2$ for this equation is 76.7% with Pearson correlation coefficient of 0.916(Fig. 4).

The measured mercury concentrations in 2008-09 were compared with estimated mercury concentrations calculated from equation-2 using TDS values of 2008-09 (Fig. 5). The calculated RPD value for this comparison was 6.3.

**Arsenic**

Arsenic is introduced into the aquatic environment from both natural and man-made sources. Typically, arsenic occurrence in water is caused by the weathering and dissolution of arsenic-bearing rocks, minerals and ores. Industrial effluents also contribute arsenic to water body. Combustion of fossil fuels is a source of arsenic in the environment through atmospheric deposition. The greatest threat to public health arises from
arsenic in drinking water. Arsenic occurs as a major constituent in more than 200 minerals, including elemental arsenic, arsenides, sulphides, oxides, arsenates and arsenites. Although arsenic exists in both organic and inorganic forms, the inorganic forms are more prevalent in water and are considered more toxic. Arsenic concentration is sometimes strongly correlated to electrical conductivity (EC) suggesting that evapoconcentration is an important process regulating total As concentrations (Ryu et al., 2002). Conductance was chosen as explanatory variable for estimating arsenic. The range of conductance values was 0.68 mS/cm to 0.88 mS/cm, whereas the range of arsenic values was from 0.0021 to 0.0035 mg/L during 2007-08. The derived regression equation is:

\[
\text{Arsenic} = -0.00105 + 0.00519 \text{Conductance}
\]  

### Table 2: The regression equations and associated \( R^2 \), adjusted \( R^2 \), Pearson correlation \( r \) and median Relative Percentage Difference (RPD) for biological oxygen demand, mercury, arsenic and sulphate

<table>
<thead>
<tr>
<th>Constituent</th>
<th>Regression Equation</th>
<th>( r )</th>
<th>( R^2 )%</th>
<th>( R^2 ) % (adj)</th>
<th>Median RPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOD</td>
<td>( \text{BOD} = 168 - 19.4 \text{DO} )</td>
<td>-0.843</td>
<td>76.6</td>
<td>74.2</td>
<td>13.9</td>
</tr>
<tr>
<td>Mercury</td>
<td>( \text{Mercury} = -0.001441 + 0.000010 \text{TDS} )</td>
<td>0.916</td>
<td>76.7</td>
<td>74.4</td>
<td>6.3</td>
</tr>
<tr>
<td>Arsenic</td>
<td>( \text{Arsenic} = -0.00105 + 0.00519 \text{Conductance} )</td>
<td>0.863</td>
<td>70.3</td>
<td>67.3</td>
<td>8.0</td>
</tr>
<tr>
<td>Sulphate</td>
<td>( \text{Sulphate} = 50.6 + 0.284 \text{Chloride} )</td>
<td>0.858</td>
<td>89.9</td>
<td>88.9</td>
<td>5.7</td>
</tr>
</tbody>
</table>

**Figure 4:** Fitted line plot for Mercury regression equation

**Figure 5:** Comparison of measured and estimated Mercury concentration in Mahanadi river water

**Figure 6:** Fitted line plot for Arsenic regression equation

**Figure 7:** Comparison of measured and estimated Arsenic concentration in Mahanadi river water

**Figure 8:** Fitted line plot for Sulphate regression equation

**Figure 9:** Comparison of measured and estimated Sulphate concentration in Mahanadi river water
The obtained RPD value was 5.7.

The measured arsenic concentrations on 2008-09 were compared with estimated arsenic concentrations calculated from equation – 3 using the conductance values of 2008-09 (Fig. 7). The RPD value estimated for this comparison was 8.0.

**Sulphate**

Sulphate ions are important constituents of the mammalian body and of natural food stuffs and commonly constitute between 25 and 50% of the solute load. The concentration of sulphate ion in drinking water that constitutes a laxative dose is related to intake, but it is generally accepted that cathartic effects are commonly experienced by people using drinking water with sulphate concentration above 600 mg per litre (USDHEW, 1962). If taste alone is considered, a maximum of 400 mg per litre of sulphate should not be exceeded (WHO, 1963) and preferably this figure should be no more than 250 mg/L (USDHEW, 1962). Sulphates are found in all body cells, with highest concentrations in connective tissues, bone and cartilage. Sulphates play a role in several important metabolic pathways, including those involved in detoxification processes. In anaerobic environments sulphate is biologically reduced to (hydrogen) sulphide by sulphate reducing bacteria, or incorporated into living organisms as source of sulphur, and thereby included in the sulphur cycle. There are water bodies in which the sulphate concentration is related to chloride concentration (Wilson et al., 2006). Chloride was therefore chosen as explanatory variable for estimating sulphate. The range of chloride values was 20.80 mg/L to 80.14 mg/L, whereas the range of sulphate values was from 53.74 mg/L to 73.19 mg/L during 2007-08. The derived regression equation is:

\[
\text{Sulphate} = 50.6 + 0.284 \text{ Chloride} \quad \text{Equation (4)}
\]

Where, \( R^2 \) for this equation is 89.9% with Pearson correlation coefficient of 0.858 (Fig. 8).

The measured sulphate concentrations on 2008-09 were compared with estimated sulphate concentrations calculated from equation- 4 using chloride values of 2008-09 (Fig. 9). The obtained RPD value was 5.7.

**CONCLUSIONS**

Regression equations can have wider applicability in estimating the water quality. In the present study, four different regression equations have been derived and the relative percentage difference (RPD) between measured constituent concentration and the estimated constituent concentration was found to be very low and ranged from 5.7 (for sulphate) to 13.9 (for biological oxygen demand). Therefore, these equations can be effectively used to estimate the concentration of biological oxygen demand, mercury, arsenic and sulphate in Mahanadi river water using the measured concentration of surrogate parameters namely dissolved oxygen, total dissolved solids, electrical conductivity and chloride, respectively. This statistical procedure would be very helpful in case of the requirement of large data set and particularly in case of emergency when immediate mitigation measures are required to maintain the water quality standard of the river.

**ACKNOWLEDGEMENT**

The authors acknowledge VSS University of Technology and Department of Environmental Sciences, Sambalpur University for providing Laboratory facilities during this work.

**REFERENCES**


Sahib, B. K., Panda, R. B., Sinha, B. K. and Nayak, A. 1991. Change of
Water Quality Index at different monitoring stations on the River Brahmani due to discharge of effluent of Talcher Industrial Complex of Orissa at Nandira. Sambalpur University joint of Sc-2 Tech. X: 19-23.


