Determination of physico-chemical parameters of coal mines water samples and their correlation establishment with WQI using python programming

Satyajeet Parida¹, Abhishek Kumar Tripathi², D. P. Tripathy³, Purabi Bora⁴

¹,²Department of Mining Engineering, Aditya Engineering College, Surampalem, Andhra Pradesh, India-533437
E-mail: ¹parida.satyajeet@aec.edu.in, ²abhishekkumar@aec.edu.in
³Department of Mining Engineering, National Institute of Technology, Rourkela, Odisha, India-769008
E-mail: debi_tripathy@yahoo.co.in
⁴Petroleum Engineering Department, Chandigarh University, Chandigarh, India-140413
E-mail: purabiboraduiet@gmail.com

Abstract. The quality assessment of water is the need of the hour as water pollution has reached to an alarming level. The pollution of natural water bodies due to mine drainage system and mining activities is a major environmental concern worldwide. There are many potential reasons of water pollution such as agricultural, sewage, oil, radioactive materials, dumping & mining activities etc. Mining activities are responsible for the contamination of watercourses with metal and increment of sediment levels in it, however acid mine drainage can be viewed as the most lethal means of polluting watercourse. In this study an analysis was done on the water samples collected from different coal mines of Jharkhand and Telangana State. The WQI for each sample were calculated and correlated with their physico-chemical parameters. The lowest grades of the water samples are mainly due to the presence of the strongest correlated parameters. It was observed that the iron content in the samples has the strongest correlation with a Pearson coefficient of 0.9977 and highest significance with a P value lower than 0.001.

Keywords: WQI, Mine Water, Water Pollution, Physico-chemical parameters, Coal mines, Python, NumPy, Pandas, Seaborn.

1. Introduction

There are many natural resources available for mankind such as air, coal, natural gases, minerals, natural oil, forest & timber, soil and water. Among all-natural resources, water could be considered as one of the most vital and valuable natural resources for mankind. Water is one of the important basic needs of all life forms on earth ranging from micro-organisms to human beings. It is a vital natural resource to mankind's existence on the earth, without it there would be no life on the earth. Exponential increase in the demand of water brought water resources around the world under pressure, which is needed to be managed in a war like emergency for human survival [1]. The quality assessment of water is the need of the hour as water pollution has reached to an alarming level. The pollution of natural water bodies due to mine drainage system and mining activities is a major environmental concern worldwide [1].
There are many potential reasons of water pollution such as agricultural, sewage, oil, radioactive materials, dumping & mining activities etc. Mining activities are responsible for the contamination of watercourses with metal and increment of sediment levels in it, however acid mine drainage can be viewed as the most lethal means of polluting watercourse [2]. The degree and significance of contamination are affected by complex biochemical reactions which are result of the interaction between water, ore bodies and mine waste material [3]. Apart from all these chemical oriented contaminations, water bodies in and around mine area are significantly affected by the regular activities done for mining such as shaft sinking, opening pits, excavation of overburden and ore etc. [3].

It is well known that the water outflow from the mines is severely polluted. Any amount of pollution in the water bodies changes its water quality. Therefore, the quality of outflow water from the mines needs to be appropriately understood, quantified and communicated to stakeholders, which in turn can provide an appropriate guidance for the environment water management system. In this paper, an attempt has been made to assess the water quality parameters of samples collected from various mines in India and quantify their cumulative effects by estimating the water quality index [4].

2. Collection of Water Samples from Mines

The collection of water samples was carried out from five different mines which are situated at different parts of India. The collected water samples are designated by different sample Ids as mentioned in Table number 1. Among 9 water samples, the first 5 samples were collected from Jharkhand and the remaining samples were collected from Telengana state of India. The details of collected water samples are tabulated in Table 1.

| Sample Id | Mine Type | Date of Collection |
|-----------|-----------|--------------------|
| S1        | Underground Coal Mines Water pumped from ground sump | 22/02/2019 |
| S2        | Open cast coal mines water | 22/02/2019 |
| S3        | Underground Coal Mines Water pumped from ground sump | 23/02/2019 |
| S4        | Underground Coal Mines Water pumped from ground sump | 23/02/2019 |
| S5        | Underground coal mines settled sump water | 13/02/2019 |
| S6        | Underground coal mines | 21/12/2019 |
| S7        | Underground coal mines | 21/12/2019 |
| S8        | Underground coal mines drinking water supply unit | 21/12/2019 |
| S9        | Underground Coal Mines Water pumped from ground sump | 21/12/2019 |

3. Analysis of Samples

In this section the analysis of samples are presented in detail. In order to analyze the samples a two-step approach has been followed. The first step involves the experimentation and WQI calculation, for which, first the physico-chemical parameters were determined and using the determined parameters WQI was
calculated for each sample [4].

In the second step a correlation between the determined parameters and the calculated WQI was established and some visual analysis of the strongest correlating parameters with calculated WQI are done.

3.1 Determination of Physico Chemical Parameters

The physico-chemical parameters of the samples needed to be determined for any further assessment of WQI [3]. Laboratory tests in line with APHA (1995) [5] standards were conducted to determine the following parameters: pH, Dissolved Oxygen(DO), Chloride(Cl), Calcium(Ca), Total Dissolved Solid(TDS), Sulphate(SO4), Nitrate(NO3), Iron(Fe), Alkalinity(CaCO3), BOD (Biological Oxygen Demand), Magnesium(Mg) and Total Hardness(TH). The determined values were noted and tabulated in Table 2.

Table 2. Estimated Physico-Chemical Parameters of Water Samples

| Sample ID | pH   | DO   | Cl  | Ca  | TDS  | SO4  | NO3 | Fe  | CaCO3 | BOD  | Mg  | TH |
|-----------|------|------|-----|-----|------|------|-----|-----|-------|------|-----|----|
| S1        | 6.93 | 0.20 | 566.4 | 54 | 10   | 1781.4 | 4.29 | 1.2 | 0.2  | 2.5  | 100.6 | 285 |
| S2        | 7.04 | 0.37 | 637.2 | 76 | 4.58 | 410   | 3.95 | 1.1 | 2.8  | 1.89 | 102.48 | 280 |
| S3        | 7.39 | 0.19 | 778.8 | 54 | 3.8  | 681.4  | 3.86 | 1.2 | 3.2  | 1.16 | 84.49 | 247 |
| S4        | 7.41 | 0.21 | 708   | 48 | 4.5  | 1338.54 | 4.63 | 1.02 | 2.6  | 1.72 | 89.49 | 235 |
| S5        | 6.98 | 0.24 | 2194.8 | 90 | 4    | 767.14  | 8.4  | 0.9 | 0.8  | 0.83 | 31.51 | 120 |
| S6        | 7    | 1    | 156   | 112.22 | 2.2 | 25.625 | 6.24 | 0.31 | 0.72 | 0.86 | 23.02 | 106 |
| S7        | 7    | 1.2  | 138.2 | 114.2 | 3  | 69.25  | 7.28 | 0.28 | 0.91 | 0.76 | 24.24 | 92  |
| S8        | 8.23 | 1.6  | 109.9 | 94.18 | 2.1 | 69.25  | 8.28 | 0.17 | 0.82 | 0.69 | 18.18 | 90  |
| S9        | 7.93 | 1.8  | 74.4  | 94.18 | 2.4 | 69.25  | 7.88 | 0.13 | 0.62 | 0.79 | 8.48  | 88  |

After determining the parameters from the samples, the composite effect due to all parameters was needed to be evaluated and indexed in a manner, which can not only represent the overall quality of the water sample but also will be simpler for communication to the potential end users and policy making authorities. That’s why Water Quality Index was estimated using the determined parameters [6].

3.2. Determination of Water Quality Index

Water Quality Index is the cumulative representation of the determined physico-chemical parameters of the water samples. It indicates the overall quality of water represented in terms of an index number. WQI can be determined for various intended use such as drinking, industrial use, irrigation etc. In the present study, WQI was determined for finding suitability of mine water as potential source of drinking water by using the Weighted Arithmetic Index method [7]. The first step of WQI estimation involved assigning weights to the determined parameters [8]. The weights of the parameters can be calculated by using equation (1). The constant of proportionality can be calculated by using equation (2).

\[
W_i = \frac{K}{S_i}
\]  
\[
K = \frac{1}{\sum S_i}
\]

where, \(W_i\) = Weight of the ith parameter.

\(K\) = Constant of proportionality

\(S_i\) = Standard permissible limit of the ith parameter.
The categorization of WQI with respect to water quality status is presented in Table 3. Grades A, B, C, D and E were assigned to different levels of water quality which represent the status of water quality as excellent, very good, good, poor, very poor and unsuitable for drinking respectively [9]. The permissible limits and calculated weights for each parameter are presented in Table 4.

### Table 3. WQI Level And Grading[9]

| WQI Level | Grade |
|-----------|-------|
| 0-25      | A     |
| 26-50     | B     |
| 51-75     | C     |
| 76-100    | D     |
| >100      | E     |

### Table 4. Drinking Water Standards, Recommending Agency & Unit Weights [10]

| Sl. No. | Parameters                  | Permissible Limit | Recommended Agency | Unit Weight |
|---------|-----------------------------|-------------------|--------------------|-------------|
| 1       | pH                          | 8.5               | BIS                | 0.0753      |
| 2       | Dissolved Oxygen            | 5                 | BIS                | 0.1280      |
| 3       | Chloride                    | 1000              | BIS                | 0.0006      |
| 4       | Calcium                     | 200               | BIS                | 0.0032      |
| 5       | Total Dissolved Solid       | 2000              | BIS                | 0.0003      |
| 6       | Sulphate                    | 400               | BIS                | 0.0016      |
| 7       | Nitrate                     | 45                | BIS                | 0.0142      |
| 8       | Iron                        | 0.3               | BIS                | 0.6401      |
| 9       | Alkalinity                  | 600               | BIS                | 0.0010      |
| 10      | BOD (Biological Oxygen Demand) | 5             | BIS                | 0.1280      |
| 11      | Magnesium                   | 100               | BIS                | 0.0064      |
| 12      | Total Hardness              | 600               | BIS                | 0.0010      |

A common observation was that the weight for the parameters with low permissible limit is high and vice versa as seen in Table 4. After calculating weight, a quality rating was assigned to each determined parameter. The quality rating is the relative representation of the determined parameter with respect to the standard permissible value of that parameter. It can be calculated by using equation (3). Further, from the calculated weights (Wi) and quality rating (Qi), the WQI can be estimated by using equation (4). The calculated quality rating and estimated WQI of each sample is presented below in Tables 5 and 6 respectively [9].

\[
Q_i = \frac{V_i - V_P}{S_i - V_P} \tag{3}
\]

\[
WQI = \frac{\sum W_i \times Q_i}{\sum W_i} \tag{4}
\]

where, \(Q_i\) = Quality Rating of the \(i_{th}\) parameter.
\(V_i\) = Determined value of the \(i_{th}\) parameter.
\(S_i\) = Standard permissible of the \(i_{th}\) parameter.
\(V_P\) = Value of the \(i_{th}\) parameter in pure water.
Table 5. Calculated Quality Rating

| Sl.No. | Parameters                  | Qi    |
|--------|-----------------------------|-------|
| 1      | pH                          | -4.6667 |
| 2      | Dissolved Oxygen            | 149.9740 |
| 3      | Chloride                    | 56.6400 |
| 4      | Calcium                     | 27.0000 |
| 5      | Total Dissolved Solid       | 0.5000 |
| 6      | Sulphate                    | 445.3500 |
| 7      | Nitrate                     | 9.5333 |
| 8      | Iron                        | 120.0000 |
| 9      | Alkalinity                  | 0.0333 |
| 10     | BOD(Biological Oxygen Demand)| 50.0000 |
| 11     | Magnesium                   | 23.0200 |
| 12     | Total Hardness              | 15.3333 |

Table 6. Calculated WQI for All Samples

| Sample ID | WQI    |
|-----------|--------|
| S1        | 103.7306 |
| S2        | 95.5914  |
| S3        | 102.0661 |
| S4        | 92.3643  |
| S5        | 79.8696  |
| S6        | 40.746   |
| S7        | 38.3606  |
| S8        | 36.7405  |
| S9        | 32.5857  |

4. Determination of Correlation

Correlation is a measure of interdependency between variables. In this study the Pearson correlation is calculated among the parameters and calculated WQI. Pearson correlation is the measure of linear interdependence between two variables. The linearity of dependence can be both positive and negative. The value of Pearson coefficient generally vary between -1 to +1. For a value of -1 the two variables are said to be in negative linear correlation and a value of +1 shows the variables are in positive linear correlation. If the value comes as 0 it indicates that the variables are not at all correlated [11].

Further to study the degree of significance of the correlation between the parameters and WQI the P values of each parameter with the WQI were determined. The P value is the probability of statistical significance of the correlation between the variables. The determined P values show the probability in the following convention

- If P Value < 0.001, Strongly correlated
- If 0.001 < P Value < 0.05, Moderately correlated
- If 0.05 < P Value < 0.1, Weakly correlated
- If P Value > 0.01, No evidence of correlation

To determine the Pearson coefficient and P values the Pandas, NumPy modules under the SciPy eco system of the Python programming language were used [12] [13]. The reason for using Python
programming language is its ease of calculating these complex mathematical parameters. We have used only 9 samples for determination, but by using these modules even data for thousands of samples can be analysed in a matter of few lines of codes independent of the size of the data. For visual analysis, the help of two other visualization modules Matplotlib and Seaborn were taken [14]. The correlation, along with its significance, between the parameters and WQI were determined and is presented in Table 7 in terms of Pearson correlation coefficient and P value.

Table 7. Pearson Correlation Coefficient and P Values

| Parameters | Pearson Correlation coefficient with WQI | P Values | Remarks                |
|------------|------------------------------------------|----------|------------------------|
| pH         | -0.4525                                  | 0.2212   | No Evidence of Correlation |
| DO         | -0.9365                                  | 0.0001   | Strongly Correlated    |
| Cl         | 0.5276                                   | 0.1442   | No Evidence of Correlation |
| Ca         | -0.8667                                  | 0.0024   | Moderately Correlated  |
| TDS        | 0.7111                                   | 0.0317   | Moderately Correlated  |
| SO4        | 0.8054                                   | 0.0087   | Moderately Correlated  |
| NO3        | -0.7880                                  | 0.0116   | Moderately Correlated  |
| Fe         | 0.9976                                   | 2.01E-09 | Strongly Correlated    |
| CaCO3      | 0.5639                                   | 0.1138   | No Evidence of Correlation |
| BOD        | 0.7793                                   | 0.0132   | Moderately Correlated  |
| Mg         | 0.9263                                   | 0.0003   | Strongly Correlated    |
| TH         | 0.9257                                   | 0.0003   | Strongly Correlated    |

The parameters DO, Fe, Mg, TH are having strong correlation with WQI, whereas parameters pH, Cl, CaCO3 have no evidence of correlation with WQI. All other parameters left are in moderate correlation with WQI. Here the point to ponder is that, the determined correlations cannot be generalized, for example, pH as a parameter may have very high significance in determining WQI but for the samples taken into consideration, it is not playing a major role in the increment or decrement of WQI. Similarly, the parameters strongly correlating with WQI are only significant for the samples taken into consideration. Observing the visual analysis of moderately correlated and not correlated parameters has no significance for the study, for further analysis the parameters which are strongly correlated with WQI were only taken into consideration and presented herewith. The line and regression plots of the strongly correlated parameters with WQI are presented below in Figures 1 to 8.

Figure 1. Line Plot DO-WQI
In figure 1 and 2 the line and regression plots of DO-WQI is presented, in which it can be observed that the variables are almost nearly in negative linear correlation, which is proved with a Pearson coefficient value of -0.9365, with changes in the value of DO, WQI is also changing, that means the correlation is significant which is also and evident from a P value of 0.000199 which is much below 0.001.

Fe has the highest linearity and significance of correlation with WQI which is evident from Figures 3 and 4 as well as from the determined Pearson coefficient and P value of 0.9976 and 2.01E-09 respectively. Similarly in the plots for Mg and TH which are presented in Figures 5 to 8 it can be observed that both the parameters follow a positive linear correlation with WQI and it is evident from the Pearson coefficient and P values presented in Table 8.
Figure 5. Line Plot Mg-WQI

Figure 6. Regression Plot Mg-WQI

Figure 7. Line Plot TH-WQI
Figure 8. Regression Plot TH-WQI

5. Conclusions

In this study, the WQI values of each water sample from different mines were calculated and compared with the standard grades assigned for WQI levels. The comparison has shown that the samples 1 & 3 are of grade E (Unsuitable for drinking), samples 2, 4 & 5 are of grade D (Very poor for drinking) and Samples 6, 7, 8 and 9 are of grade B (very good for drinking). Among all the samples, sample number 9 has the lowest WQI of 32.5857 whereas sample number 1 has the highest WQI of 103.7306. The lowest grades of the water samples are mainly due to the difference in the presence of the strongest correlated parameters, among which the stood-out parameter was iron which has the strongest correlation with a Pearson coefficient of 0.9976 and highest significance with a P value lower than 0.001 with WQI among all the determined parameters.

References

[1]. Das, S., & Rout, S. (2012). Assessment of Water Quality Around Mine Sites [PhD Thesis].
[2]. Akcil, A., & Koldas, S. (2006). Acid Mine Drainage (AMD): Causes, treatment and case studies. Journal of Cleaner Production, 14(12–13), 1139–1145.
[3]. Booth, C. J. (2006). Groundwater as an environmental constraint of longwall coal mining. Environmental Geology, 49(6), 796–803.
[4]. Egün, N. K., & Ogiesoba-Eguakun, C. U. (2018). Physico-chemical and Water Quality Index analysis of the Okhuanhe River, Edo State, Nigeria. African Journal of Aquatic Science, 43(4), 345–351.
[5]. methods for the examination of water and wastewater (Vol. 2). American Public Health Association.
[6]. Chande, M. M., & Mayo, A. W. (2019). Assessment of Groundwater Vulnerability and Water Quality of Ngwerere Sub-Catchment Urban Aquifers in Lusaka, Zambia. Physics and Chemistry of the Earth, Parts A/B/C.
[7]. Brown, R. M., McClelland, N. I., Deininger, R. A., & O’Connor, M. F. (1972). A water quality index—Crashing the psychological barrier. In Indicators of environmental quality (pp. 173–182). Springer.
[8]. Amalraj, A., & Pius, A. (2018). Assessment of groundwater quality for drinking and agricultural purposes of a few selected areas in Tamil Nadu South India: A GIS-based study. Sustainable Water Resources Management, 4(1), 1–21.
[9]. Chatterjee, C., & Raziuddin, M. (2002). Determination of Water Quality Index(WQI) of a degraded river in Asansol industrial area(West Bengal). Nature, Environment and Pollution Technology, 1(2), 181–189.
[10]. Rao, N. S., Rao, P. S., Reddy, G. V., Nagamani, M., Vidyasagar, G., & Satyanarayana, N. (2012). Chemical characteristics of groundwater and assessment of groundwater quality in Varaha River Basin, Visakhapatnam District, Andhra Pradesh, India. Environmental Monitoring and Assessment, 184(8), 5189–5214.

[11]. Kirch, W. (2008). Pearson’s Correlation Coefficient. Encyclopedia of Public Health; Springer: Dordrecht, The Netherlands, 1090–2013.

[12]. McKinney, W. (2010). Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference, 445, 51–56.

[13]. Walt, S. van der, Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: A structure for efficient numerical computation. Computing in Science & Engineering, 13(2), 22–30.

[14]. Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Computing in Science & Engineering, 9(3), 90–95.