Access Quality of Water Parameters by Suggestion and Correlating of Water Parameters

L Roshini and C.R.K. Reddy

July 10, 2023

Access Quality of Water Parametres by suggestion and correlating water parameters

Roshini L[1] [0009-0001-9709-0289] and C.R.K.Reddy[2] [0000-0003-2246-3930]

[1] Research Scholar, Osmania University, Hyderabad, India-500044
[1] Assistant Professor, CSIT Department, CVR College of Engineering, Hyderabad, India-501510
[2] Department of Computer Science & Engineering, MGIT, Hyderabad, India-500044,
roshini1604@gmail.com,crkreddy_cse@mgit.ac.in

Abstract: One of the key issues the globe has recently faced is the evaluation of water purity and the location of sources for safe potable water. Since the manner of water poisoning is unpredictable, it is difficult to analyze and maintain, the key problem is to preserve water and maintain the purity of the water from extreme environmental degradation. As a result, improving water purity through resource management is particularly effective. The determination of the water quality indicator (WQI) depends on a few specific water-related elements. However, the conventional approach to calculating the WQI requires a lot of work, and occasionally errors are found while the computation is being done. In this study, we assess the water's quality and predict its grade using a machine learning methods. Our algorithms will identify the chemical components of the water if it is not pure and will suggest the necessary combination to bring the pH level of the water into balance and make it stable enough to drink. In order to improve water purity, both manually and through the application of machine learning algorithms, good management of water supply is essential. The main goals of this study are the gathering of water data, the analysis of chemical combinations, correlations between combinations, the analysis of water samples for their physico-chemical properties, the calculation of the WQI using this information, and the use of machine learning techniques to forecast water purity. To assess the effectiveness of the forecast algorithms, several statistical indicators were employed.

Keywords: Correlations, Machine Learning, Water purity, WQI, chemical compositions

1. INTRODUCTION

On Earth, neither a living being nor a plant can exist without water. Water is regularly contaminated, though, as a result of the industry's yearly expansion in response to rising demand and because these enterprises constantly discharge hazardous waste into rivers and lakes. Millions of deaths are caused by contaminated water.

Each year, there are significant financial losses and damage to farming areas. Numerous studies have found that groundwater quality has recently deteriorated dramatically in the majority of countries. As a result, the groundwater's quality is deteriorating everyday [1]. Finding the "quality determining factors" of water, which help to identify water pollution, is a quick, easy, and sensible technique to assess the quality of water for various purposes.
For the highest water quality, monitoring and analyzing water purity is a crucial task.

For the best possible administration of various water supplies, it is crucial to monitor and assess the cleanliness of the water. Either directly or indirectly related to the continual increase of human activity is the rise in demand for clean, unpolluted water. The demand for pure water drives researchers to investigate the physical properties of groundwater in the area in order to develop novel, cutting-edge models for water quality forecasts. Potable water supply availability has recently raised significant issues on a global level[2]. Therefore, it is now crucial to assess water purity and predict its future status.For the best possible administration of various water supplies, it is crucial to monitor and assess the cleanliness of the water. Either directly or indirectly related to the continual increase of human activity is the rise in demand for clean, unpolluted water. The demand for pure water drives researchers to investigate the physical properties of groundwater in the area in order to develop novel, cutting-edge models for water quality forecasts. Potable water supply availability has recently raised significant issues on a global level[2]. Therefore, it is now crucial to assess water purity and predict its future status.

2. RELATED WORKS

The examination of the quality of the groundwater is crucial and necessary, hence it is carried out frequently throughout the country. Since groundwater is utilized for drinking, a thorough physico-chemical examination is required due to its connection to human health. Physico-chemical and other parameters have been defined by many organizations (such as UNESCO, WHO, BSI, etc.) to evaluate the quality of water. It is important to note the reports of similar investigations conducted by other experts in the subject.

The results of physico-chemical variables like pH, EC, TDS, total hardness, nitrate, and sulfate have been examined by Prasad et al. [4]. They concluded that EC and TDS displayed a significant positive correlation. Some metals, including iron, copper, zinc, and nickel, are required for the proper division of some compounds, claim Garg et al. [5].

3. METHODOLOGIES

The standard method is used to assess the various physiochemical characteristics [6]. The collected water samples are retained for examination of dangerous heavy metals using the advised method[7], which involves adding 5 milliliters of pure HNO_3 to a liter of water samples to maintain a pH lower than 4.0. After carefully filtering the samples with Whitman 40-filter paper, measurements can be made with an atomic absorption spectrometer (AAS).

To determine the parameters to be determined for water evaluation, operational tracking must be exact and dependent on detecting relevant biological, hydromorphological, and physico-chemical data. Three different sorts of criteria are provided by these technologies for worldwide environmental surveillance to gauge water quality

Water Purity

Indicators of water purity include a variety of physical, molecular, biological, and bacteriological characteristics. These characteristics are sometimes known as factors or markers. According to physical, molecular, biological, and bacteriological characteristics, water can be categorized and used for many purposes.

Depending on the planned usage, different water purity criteria must be met. Therefore, any poisons or bacteria that could be harmful to a person's health are not allowed in drinkable water. The use of water for irrigation that contains significant amounts of sodium ions, significant amounts of nitrate, or significant amounts of other contaminants is not permitted. Businesses are subject to less rigorous water

usage rules than those that apply to potable water [3].
Water purity also varies depending on the type of water source, which is another aspect.

4. DATA SET

The data has 21 parameters with salinity indicators, Heavy metals, other relevant indicators and the target parameter.

Table 1: Parameters of the data

ph	Hardness	Solids	Chloramines	Sulfate	Conductivity	Organic_carbon	Trihalomethanes	Turbidity	Potability	
0	4.500000	204.890456	20791.31898	7.300212	368.516441	564.308654	10.379783	86.990970	2.963135	0
1	3.716080	129.422921	18630.05786	6.635246	NaN	592.885359	15.180013	56.329076	4.500656	0
2	8.099124	224.236259	19909.54173	9.275884	NaN	418.606213	16.868637	66.420093	3.055934	0
3	8.316766	214.373394	22018.41744	8.059332	356.886136	363.266516	18.436525	100.341674	4.628771	0
4	9.092223	181.101509	17978.98634	6.546600	310.135738	398.410813	11.558279	31.997993	4.075075	

Figure 1: Sample data set with parameters

4.1 Correlating Parameters:

Correlation reveals a relationship between two variables. The assessment of the association between two such parameters is referred to as the correlation coefficient. It is also known as the degree to which the two components are correlated.

Covariance

Two variables in a set of data, x and y, are correlated to determine how linearly related they are. A positive covariance would imply a positive linear relationship between the elements, whereas a negative covariance would imply the opposite.

The correlation coefficient between two variables in a data collection is computed by dividing the covariance by the total of the individual standard deviations for the two variables. It performs a standardized measurement of their straight relationship.

$$S_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

The correlation coefficient between two variables in a data collection is computed by dividing the covariance by the total of the individual standard deviations for the two variables. It performs a standardized measurement of their straight relationship.
The following formula, where sx and sy are the sample standard deviations and sxy is the sample covariance, serves as a general definition of the sample correlation coefficient..

$$r_{xy} = \frac{s_{xy}}{s_x \quad s_y}$$

If the correlation value is close to 1, the scatter plot will almost always display a straight line with a positive slope, indicating that the variables are positively linearly connected. The scatter plot almost exactly follows a downward-inclining straight line when the value is 1, indicating a negative linear relationship between the variables. Furthermore, a value of zero indicates that there is just a weak linear relationship between the variables.
The following algorithm is used to determine the correlation coefficient, where x represents the values of the independent variable (height in this case) and y represents the values of the dependent variable (anatomical dead space in this case). The calculation appears as follows.:

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{[\sum(x - x^{-2})(y - y^{-2})]}}$$

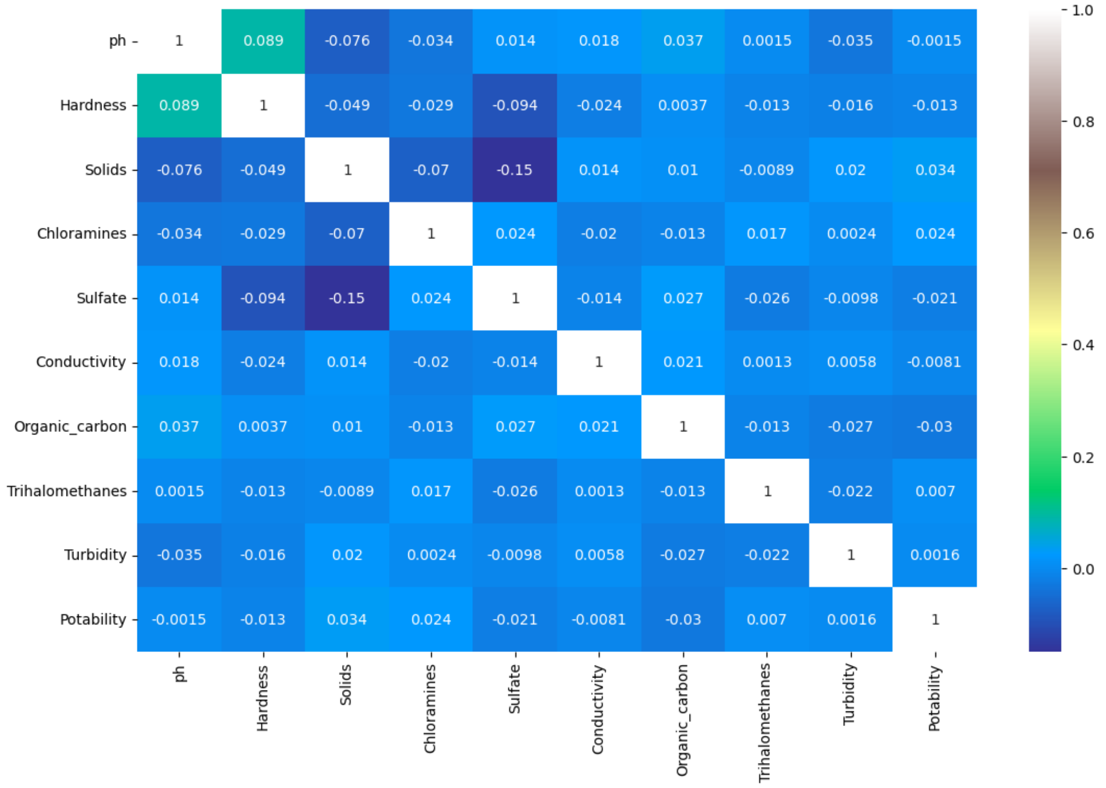

Figure 2: The correlation between every parameter used in the data set

5. RESULTS AND EVALUATION

Water serves a variety of purposes and is essential to all known living things. Therefore, everyone should be concerned about the water's quality. The state of water can be characterized by a set of physical, molecular, biological, and bacteriological characteristics. With the use of these characteristics, water can be divided into different categories and used for diverse purposes.
This chapter establishes the connections between external influences and how they affect different characteristics of water purity. Data on water purity can be provided using a variety of Water Quality Index (WQI) techniques. The links between the various water quality measurements have been investigated using a number of correlation variables. Using Pearson and Spearman correlations, we have created

```
[ ] actual = [2.8,0.01,2,0.005,4,0.1,1.3,1.5,0,0,0.015,10,1,0.002,56,5,0.5,0.1,0.3]
```

Figure 4: The sample data for testing the water purity

```
water is unsafe
Please decrease value of aluminium from : 3.0 to : 0.20000000000000018
Please decrease value of arsenic from : 0.05 to : 0.04
Please decrease value of barium from : 4 to : 2.0
Please decrease value of cadmium from : 0.01 to : 0.005
Please decrease value of chloramine from : 6 to : 2.0
Please decrease value of chromium from : 0.5 to : 0.4
Please decrease value of copper from : 2.0 to : 0.7
Please decrease value of flouride from : 2.0 to : 0.5
Please decrease value of bacteria from : 0.01 to : 0.01
Please decrease value of viruses from : 0.01 to : 0.01
Please decrease value of lead from : 0.02 to : 0.005000000000000001
Please decrease value of nitrates from : 15 to : 5.0
Please decrease value of nitrites from : 2 to : 1.0
Please decrease value of mercury from : 0.005 to : 0.003
Please decrease value of perchlorate from : 60 to : 4.0
Please decrease value of radium from : 10 to : 5.0
Please increase value of selenium from : 0.1 to : -0.4
Please decrease value of silver from : 0.5 to : 0.4
Please decrease value of uranium from : 0.5 to : 0.2
```

Figure 5: The recommended features to be managed to make the water pure.
The Parameters are described as increasing or decreasing in order to make the water pure enough to drink. The whole model works on the principle of correlation of the features of the data set. The model seems to be accurate and efficient in recommending the values for the features.

6 . CONCLUSION

The statistical analysis of the experimentally derived water quality indicators on water samples yielded the mean, standard deviation, and coefficient of variation. Because they show how distinct traits relate to one another, correlation values were computed. The correlation study's results show a strong and positive correlation between the traits of nitrites, chloramine, chromium, and copper. It may be concluded from the above talk that the nitrites, chloramine, chromium, and copper of water are essential drinking water parameters because they are strongly connected with eight out of seventeen factors in the research region.

References

1 Krishnan, K.S.D. Multiple Linear Regression-Based Water Quality Parameter Modeling to Detect Hexavalent Chromium in Drinking Water. In Proceedings of the 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 22–24 March 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 2434–2439.

2 Ighalo JO, Adeniyi AG (2020) A comprehensive review of water quality monitoring and assessment in Nigeria. Chemosphere 260:127569

3 Bhoi SK, Mallick C, Mohanty CR (2022) Estimating the Water Quality Class of a Major Irrigation Canal in Odisha, India: ASupervised Machine Learning Approach. Nat Environ Pollut. Technol 21(2):433–446

4 Prasad B.C. and Narayana T.S., Nat. Environ. Poll. Tech., 3, 10 (2004)

5 Garg A., Kripalani C. and Brighu U., Nat. Environ., Poll. Tech., 3, 1(2004) 73.

6 Ahmed AN, Othman FB, Afan HA, Ibrahim RK, Fai CM, Hossain MS, Elshafie A (2019) Machine learning methods for better water quality prediction. J Hydrol 578:1–18

7 Pant, R.R.Spatio temporal variations of hydrogeo chemistry and its controlling factors in the Gandakiriver basin, Central Himalaya Nepal. Sci. Total. Environ. 2018, 622–623, 770–782. [CrossRef]

8. Khatunm, M. Phytoplankton assemblage with relation to water quality in Turag River of Bangladesh. Casp. J. Environ. Sci. 2020, 18, 31–45.

9. Rose-Rodrígue, R. Water and fertilizers use efficiency in two hydroponic systems for tomato production. Hortic. Bras. 2020, 38, 47–52. [CrossRef]

10. Trombadore, O. Effective Data Convergence, Mapping, and Pollution Categorization of Ghats at Ganga River Front in Varanasi; Springer: Berlin/Heidelberg, Germany, 2020; Volume 27, pp. 15912–15924.

11. King, K.W.; Williams, M.R.; Macrae, M.L.; Fausey, N.R.; Frankenberger, J.; Smith, D.R.; Kleinman, P.J.A.; Brown, L.C. Phosphorus Transport in Agricultural Subsurface Drainage: A Review. J. Environ. Qual. 2015, 44, 467–485. [CrossRef]

12. Hupfer, M.; Lewandowski, J. Oxygen Controls the Phosphorus Release from Lake Sediments—A Long-Lasting Paradigm in Limnology. Int. Rev. Hydrobiol. 2008, 93, 415–432. [CrossRef]

13. Tammeorg, O.; Möls, T.; Niemistö, J.; Holmroos, H.; Horppila, J. The actual role of oxygen deficit in the linkage of the water quality and benthic phosphorus release: Potential implications for lake restoration. Sci. Total Environ. 2017, 599–600, 732–738. [CrossRef]

14. Søndergaard, M.; Jensen, J.P.; Jeppesen, E. Role of sediment and internal loading of phosphorus in shallow lakes. Hydrobiologia 2003, 506–509, 135–145. [CrossRef]

15. Cooke, G.D.; Welch, E.B.; Peterson, S.A.; Nichols, S.A. Restoration and Management of Lakes and Reservoirs, 3rd ed.; Taylor & Francis/CRC Press: Boca Raton, FL, USA, 2005; p. 591.

16. Klapper, H. Control of Eutrophication in Inland Waters; Ellis Hornwood: New York, NY, USA, 1991; pp. 116–266.

17. Lossow, K.; Gawro ́nska, H. Lakes—Review of restoration methods. Munic. Rev. 2000, 9, 91–106.

18. Kuha, J.K.; Palomäki, A.H.; Keskinen, J.T.; Karjalainen, J.S. Negligible effect of hypolimnetic oxygenation on the trophic state of lake Jyväsjärvi, Finland. Limnologica 2016, 58, 1–6. [CrossRef]