Cyber Times International Journal of Technology & Management

Vol. 11 Issue 1, October 2017 – March 2018

TABLE OF CONTENTS

WATER QUALITY PARAMETERS ESTIMATION THROUGH DATA DRIVEN MODEL

Ms. Swapnali Mahadik
Research Scholar, TMV Pune
Asst. prof. MCA Dept. DES's NMIT, Dadar, Mumbai
swapn.mahadik30@gmail.com

Dr. Anup Girdhar
CEO-Founder
Sedulity Solutions & Technologies
anup@sedulitygroups.com

ABSTRACT

In the 21st century monitoring water quality is rising challenges because of the large number of elements are used in everyday lives. Monitoring water provides the necessary evidence to make sound decisions on managing water quality. This research paper deals with the assessment of water quality of Krishna River with the help of data model based on the predefined major principal water quality parameters. Electrical Conductivity and Total Dissolved Solids are considered as important parameters in determining the quality of drinking water because it shows the direct salt concentration in the water. In this research, estimation of these two parameters is studied using the K-Nearest neighbour algorithm and also used for the water quality estimation.

KEYWORDS: *Electrical Conductivity, Decision Making, Total Dissolved Solids, K-Nearest neighbor algorithm, Water Quality.*

I. INTRODUCTION

Water quality refers to the Chemical, Physical and biological characteristics of water. The most common standards used to access water quality relate to health of environment, safety. The parameters of water quality are determined by the intended use. Such as water quality inclines to be focused on water that is treated for human consumption, industry purpose and environment. The complexity of water quality as subject is reflected in the many types of measurements of water quality indicator. Numerous models have been developed for management of water quality, such as QUAL2E, Water Quality Analysis Simulation, and the US Army Corps of Engineers' Hydrologic Engineering Center-5Q (Chen et al. 2003). Using these models is time-consuming and expensive; therefore, development of cost-effective models is encouraged [1]. Data-driven models are computer tools that are capable of modeling water quality parameters by using various algorithms [13]. The quantity of Ammonia, dissolved oxygen, chlorine, copper, lead, suspended solids zinc, sodium, Electrical Conductivity, Total dissolved solids and pH is frequently measured at monitoring stations. EC and TDS levels in water are two major parameters of water because they represent directly total concentration of salt in water.

As per the N.R. Prasad, J.M. Patil , physio-chemical parameters of Krishna river water was studied and compared with the standards of ICMR and Who and results has been shown in the year 20118 as maximum parameters are in permissible limit.[6]

II. LITERATURE REVIEW

The growing use of techniques in water resource engineering is seen at recent literature. As the author is suggesting development model with a minimal number of chemical parameters but with acceptable accuracy to estimate EC and TDS values reduces the cost of water quality monitoring [1]. Taeyun Kim et. al (2010) has evaluated a good capability water quality model simulations using model parameters produced by modified Gausss- Newton Method and also some modeled parameters water quality model simulation employing model parameters obtained by the trial-and-error method [2]. S.P. Gorde and M.V. Jadhav has also suggested water quality is dependent on the type of the pollutant added and the nature of self-purification of water [14]. The majority of the existing techniques are limited to most of the substantial features of water quality [15].

In general, water quality model have many kinetic equations which includes various unknown parameters. The purpose of these unknown parameters is to analysis the complex physical and biochemical factors, such as geological characteristics of temperature, freshwater inflows, nutrient loads, algae, nutrient etc.

III. AREA STUDIED AND DATA USED

The data set used in this research is obtained from the Regional Water Authority "National Water Quality Monitoring Programme (NWMP)". Real datasets are used as sample shown in table 1, where the sample data has been used from Central Pollution Control Board (CPCB) year 2014 of Krishna River in Maharashtra, India. The Krishna River originates in Western Ghats near Mahabaleshwar at an elevation of about around 1300 Km in length. Krishna Basin occupies an area of 2, 58,948 sqkm which is nearly 8% of the total geographical area of the country. Largest part of the basin, nearly 44% lies in Karnataka. 26% of the basin falls in Maharashtra, about 15% in Telangana and another 15% in Andhra Pradesh [7][8]..

Figure 1: Path Directions of Krishna River [9]

- **Total Dissolved solids**

Total dissolved solids (TDS) are a term which defines the inorganic salts and small amounts of organic matter present in solution in water. The principal ingredients are usually calcium, magnesium, sodium, and potassium cations and carbonate, chloride, sulfate, and nitrate anions. The presence of TDS may affect its test. The method of identifying TDS in water supplies most commonly used is the measurement of specific conductivity which detects the presence of ions in water [3]. In this research is the flow of river is not considered as an input parameter.

- **Electrical Conductivity**

Electrical Conductivity is the ability of water to conduct electricity depends on the concentration of Ions, Types of Ions and temperature of solution.[4] The most desirable limit of Electrical Conductivity in drinking water is prescribed as 1.500 µs/cm [5]. Electrical conductivity at this point is more than permissible. The reason behind such increase in conductivity is a large amount of religious activities at upstream side of the river, it produces solid waste which is directly discharged into river water due to religious belief. This part of the river has 8 ghats and more than 100 of temples in Wai [11]. Table 1 has represented the mean value of EC and TDS of 9 locations.

Table 1: Details of EC and TDS Dataset Krishna River [10]

SR NO	STATION CODE	LOCATIONS	STATE	PARAMETERS	
				EC	TDS
1	1194	Krishna at Mahabaleshwar Dhom dam near Koina dam	Maharashtra	160	110
2	1153	Krishna at Rajapur Weir	Maharashtra	449	222
3	36	Krishna at Krishna bridge, karad	Maharashtra	262	180
4	1906	Krishna d/s of Islampur	Maharashtra	441	336
5	37	Krishna at Sangli	Maharashtra	711	579
6	1310	Krishna at Kurundwad in Kolhapur	Maharashtra	400	320
7	2187	Krishna river at Kshetra Mahuli.	Maharashtra	551	389
8	2188	Krishna river at Krishna-Venna sangam at Mahuli.	Maharashtra	193	127
9	2190	Krishna river at Wai.	Maharashtra	413	249

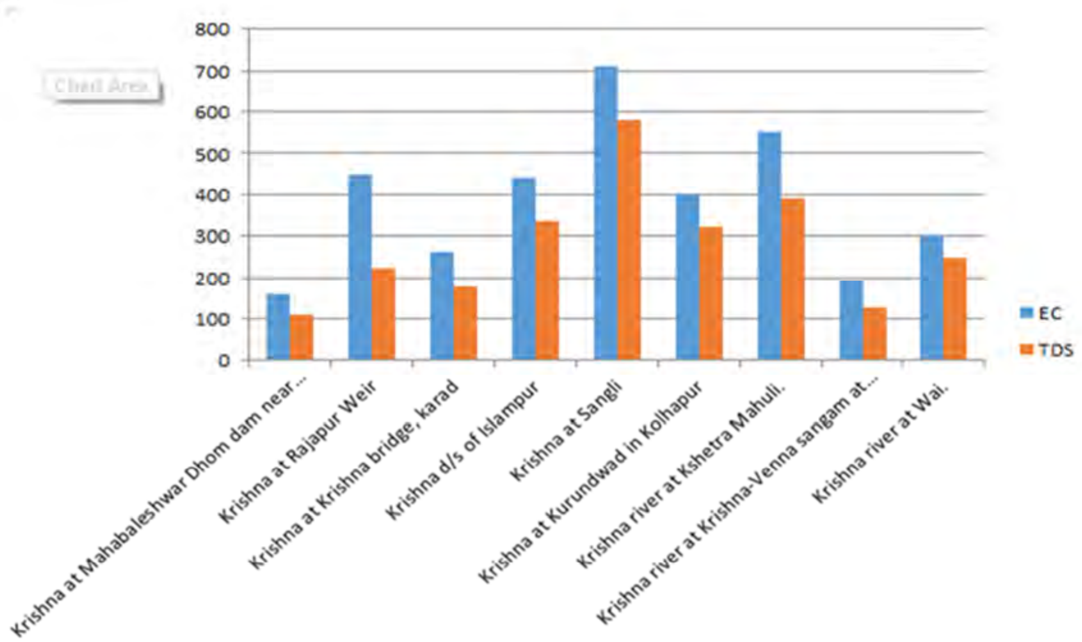

Figure 2: EC and TDS concentration of Krisha at different location

IV. K NEAREST NEIGHBOR MODEL

In Knowledge discovery process, data cleaning and preprocessing is an important step before choosing the data mining algorithms and data mining. This cleaning and preprocessing includes basic operations, like deciding strategies for appropriately handling missing data fields, removing noise [12].

KNN model is examined using the correlation efficient, Root Mean square Error and mean absolute Error. In particular, evaluate the predictive performance with respect to its readings at different time and the performance is evaluated in terms of their root-mean-square-error (RMSE).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{M}(y_i - \hat{y}_i)^2}. \quad (1)$$

Here Root Mean Square Error can be used for measuring the difference between values which are actually observed between different locations.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i| \quad (2)$$

Based on this two methods KNN model examined using the Correlation Efficient , Root Mean Square Error (RMSE) in equation 1 and in equation 2 Mean Absolute Error (MAE).

Where ∑ is the sum y denotes the observed values and y^ denotes the computed values. The KNN is a non-parametric method used for classification and regression. The input is considered as K closest training example. The KNNs are determined on the basis of RMSE.

Generally, EC is highly correlated with TDS, the scenario must include the input parameters of sodium, magnesium chloride, bicarbonate, iron, Calcium which highly affect the estimated values of EC and TDS also.

V. CONCLUSION

This research shows, the Electrical Conductivity and Total Dissolved Solid, these two parameters are commonly used with the help of the KNN, as water quality estimation technique because they represent direct total concentration of salt in water. In this research, the study shows that, there is no need to use the costly process of basic sampling hydro chemical parameters that affect river water quality. But also, this research has shown that the best scenario to estimate the TDS in water and EC should involves a combination of the parameters like sodium, Mg, Ca, chloride, and bicarbonate ion. Without these parameters estimation of TDS can affect the quality value. Overall the approach discussed (KNN) in this article reduces the cost of the water quality monitoring.

REFERENCES

[1] Mohammad Taghi Sattari, 1 Ali Rezazadeh Joudi,2 And Andrew Kusiak3, April 2016 , “Estimation of Water Quality Parameters With Data-Driven Model”, **http://dx.doi.org/10.5942/jawwa.2016.108.0012**, Volume / Number: 108, Number 4, Page(s): E232-E239.

[2] Taeyun Kim, Y.P. sheng, May 2019, “Estimation of Water quality model parameters”, KSCE Journal of Civil Engineering, Volume 14, Issue 3, PP–437.

[3] World Health Organization, Guidelines available at http://who.int/water_sanitation_health/dwq/chemicals/tds.pdf, [Accessed 12th Feb 2018].

[4] Hydrology Project, available at <http://www.cwc.nic.in/main/HP/download/08%20Understanding%20EC.pdf>, [Accessed 12th Feb 2018].

[5] Sajitha V., Smitha Asok Vijayamma, March 2016 , “Study of Physico-Chemical Parameters and Pond Water Quality Assessment by using Water Quality Index at Athiyannoor Panchayath, Kerala, India”, Emer Life Sci Res (2016) 2(1): 46-51, E-ISSN: 2395-6658.

[6] N.R. Prasad, J.M. Patil, 2008, "A Study Of Physico-Chemical Parameters Of Krishna River Water Particularly In Western Maharashtra", Rasayan J Chem, Vol.1, No.4 (2008), 943-958.

[7] Krishna River Report, May 2014 http://mpcb.gov.in/ereports/pdf/KrishnaRiverReport.pdf, [Accessed 14th Feb 2018].

[8] S. Mahadik , Feb 2018, "Analysis of Principal Water Quality Parameters by Data Driven Technique", ISSN No: 2249-3867.

[9] www.mapsofIndia.com Copyright at 2014 [Accessed 18th Feb 2018].

[10] Water quality data CPCB ENVIS <https://data.gov.in/catalog/water-quality-india-2013> [Accessed 18th Feb 2017].

[11] Mr. Prashant R. Narayane, Dr. Sunil B.Thakare and Prof. Sagar M. Gawande, may 2016 , "the analysis of Physico-chemical characteristics of water in krishna river at bhuinj, satara", International Journal for Research in Emerging Science and Technology, Volume-3, Issue-5, e-ISSN: 2349-7610.

[12] Hui Zou 1,2, Zhihong Zou 1,* and Xiaojing Wang, November 2015 , "An Enhanced K-Means Algorithm for Water Quality Analysis of The Haihe River in China", International journal of environmental research and public health, ISSN 1660-4601.

[13] H. Orouji1 ; O. Bozorg Haddad2; E. Fallah-Mehdipour3 ; and M. A. Mariño, Dist.M.ASCE4, June 2013 "Modeling of Water Quality Parameters using Data-Driven Models", Ascelibrary.org by University of Texas at San Antonia.

[14] S.P. Gorde, M.V. Jadhav , Nov- Dec 2013, "Assessment of water quality parameters : A review" , Int. Journal of Engineering Research and Applications, ISSN : 2248-9622, Vol. 3, Issue 6, pp.2029-2035.

[15] Priya Sing, Dept of Computer Science and Pankaj Kaur, Dept. Of Computer Science, May June 2017,"Review on Data Mining Techniques for Prediction of Water Quality" International Journal of Advanced Research in Computer Science, Volume 8 No.5 ISSN No: 0976-5697.