Fuzzy Based Water Quality Index Modeling for Groundwater Contamination

Jessy Mol. I[1]
Assistant Professor, Department of Civil Engineering
St.Xavier's Catholic college of Engineering
Nagercoil, India
jessymol@gmail.com
(Corresponding author)

Baskaran. T[2]
Professor, Department of Civil Engineering
Thiagarajar College of Engineering
Madurai, India
tbaskaran70@gmail.com

Justin Jose.D[3]
Assistant Professor, Department of Information Technology
Marthandam College of Engineering and Technology
Kuttakuzhi, India
mail2djose@gmail.com

***Abstract*—A basic problem in the case of water quality monitoring is the complexity associated with analysing the large number of variables. The groundwater quality management is one of the major environmental challenges. Monitoring pollutant from the different sources of groundwater is quite difficult and laborious work, which sometimes leads to analytical errors too. To predict the WQI of the study area, Fuzzy based Water Quality Index (FWQI) is developed. The physico-chemical parameters such as TDS, pH, NO_3, DO, EC and PO_4 of the area have been taken for deciding the characteristics of the study area. The membership functions of the input and the outputs are created based on standard ranges of WQI. The result shows that 55% of the samples under the category are "Good" and the remaining 45% samples are "Fair".**

Keywords— Fuzzy logic, FWQI, Water Quality Index, Groundwater

I. INTRODUCTION

Assessing the water quality is very important so as to identify the level of impurities present in water. Protecting it from contamination and carefully managing its use will ensure its future as it an important part of our ecosystems and human activity. A basic problem in the case of water quality monitoring is the complexity associated with analysing the large number of variables [1]. Recently, many statistical techniques, such as cluster analysis, factor analysis, and correlation analysis are available to analyse the quality of water. Prediction can be used for resource planning and management in case they are of acceptable accuracy [2]. Many intelligent tools, such as linear and nonlinear tools are successfully used to predict the water quality. Regression analysis is the linear modelling tool used in prediction. It works based on the mathematical equations. If linear model fails to accurate prediction, can be used nonlinear prediction such as artificial neural network, fuzzy logic, etc. Nowadays artificial neural networks and fuzzy logics are used to find a solution for many cases of water quality problems. Many groundwater quality models are developed by USGS.

Nowadays fuzzy logic integrated tool is used in many of the water quality problems for creating a relation between the variables. Fuzzy system is a good prediction tool for imprecise and uncertainty information, the approach is the most appropriate technique for modelling the prediction of water quality. On account of its ability to handle the uncertainties in Geosciences, water resources and particularly in water quality management, fuzzy logic has undergone an explosive development in application in almost all the areas of research. Besides it has been easily accepted by both researchers and decision makers. Consequently, great attention has been paid to develop the environmental indices using fuzzy logic. For a new indexing method of water quality, fuzzy synthetic evaluation has been proposed [3-4]. This same method of fuzzy synthetic evaluation has been applied to identify river water quality [5-6], also the identification of groundwater quality is proposed by using the fuzzy synthetic evaluation [7-8]. Some authors have developed cascade fuzzy integrated system for Indian River water quality prediction [9]. The proposed model is developed to predict the water quality index of five rivers in India. In this case, there are three types of water quality criteria, such as Indian criteria, Malysian criteria and USA criteria and they have been developed manually. The current study aims to develop Fuzzy Water Quality Index (FWQI) to assess the quality of groundwater for drinking purpose, including its physical and chemical characters.

2. MATERIALS AND METHODS

2.1 STUDY AREA

The area selected for this study is Mondaikadu. It is a panchayat town in Kannyakumari District, Tamil Nadu. It lies between 8^{0}9'47" latitude and 77^{0}16'48.09" longitude. It is a coastal area of Kanyakumari District. It is 75 km from Thiruvananthapuram, 95km from Tirunelveli and 18 km from Nagercoil. Coir making and fishing are the main occupations of this area. In the coir making process, the coconut husk is soaked in a pond for retting purposes. During this process, the retting pond produces a lot of organic, inorganic and biological components and it has a bad smell. On account of this process, the groundwater quality may be affected.

To represent the water quality, 20 groundwater sampling points were selected from the study area. The samples were collected from three Indian seasons such as winter, summer and monsoon. The sampling work was carried in the study area for three years, from May 2013 to April 2016 so that the total numbers of collected samples are 180. The physical, chemical and biological properties of the groundwater of all the three seasons were evaluated in the standard laboratory.

2.2 FUZZY LOGIC SYSTEM

Fuzzy inference is defined as the process of mapping a set of input data set into a set of output data, using an approach based on fuzzy logic and it falls under the category of black box models [10-11]. Fuzzy logic has rapidly become one of the most successful of today's technologies for developing sophisticated control systems. The reason is very simple. Fuzzy logic addresses such applications perfectly as it resembles human decision making with an ability to generate precise solutions from certain or approximate information. It fills an important gap in engineering design methods, left vacant by purely mathematical approaches and purely logic-based approaches in system design. While the other approaches require accurate equations to model real-world behaviours, fuzzy design can accommodate the ambiguities of a real-world human language and logic. It provides both an intuitive method for describing systems in human terms and automates the conversion of those system specifications into effective models [12]. The membership functions work on fuzzy sets of variables.

The generalized structure of a fuzzy system is presented in Fig. 1. There are two inference systems are commonly used, i.e. Mamdani Fuzzy Model and Suguno Fuzzy Model. They use different types of fuzzy reasoning and formulation of fuzzy IF-THEN rules. Typically, the consequence is a polynomial function of the desired input variables. Fuzzy system comprises four blocks; they are fuzzifier, knowledge base, inference engine and defuzzifier. The real world input of the fuzzy system is applied in the fuzzzifier. This input is called crisp input since it contains precise information about the specified information of the parameter. The fuzzifier converts this precise quantity to the form of imprecise quantity like 'large', 'medium', 'high', etc. with a degree of belongingness to it. Typically, the value ranges between 0 and 1.

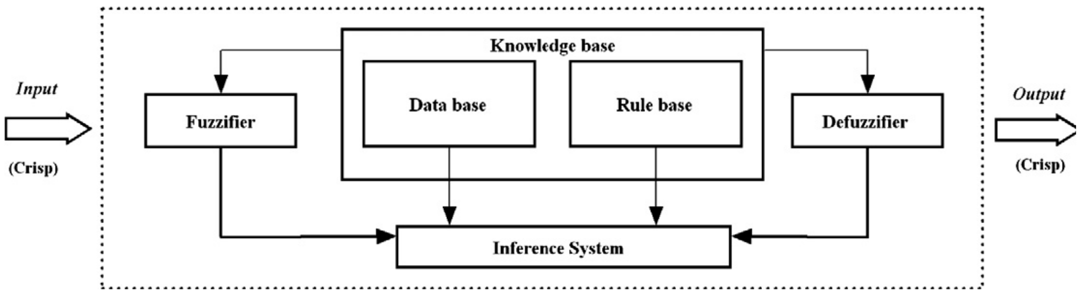


Fig. 1 Generalised structure of fuzzy system [10]

The knowledge base is the main part of the fuzzy system, in which both the rule base and database are jointly referred. The database defines the membership functions of the fuzzy sets used in the fuzzy rules whereas the rule base contains a number of fuzzy if-then rules. The inference system or the decision-making unit performs the inference operations of the rules. It handles the way in which the rules are combined. The output generated from the inference system is always fuzzy in nature. A real world system will always require the output of the fuzzy system to the crisp or in the form of real world input. The role of the defuzzifier is to receive the fuzzy input and provide the real world output. In operation, it works opposite to the input block [13].

2.2.1 Membership Function

Membership functions allow quantifying the linguistic terms and representing a fuzzy set graphically. The membership function for a fuzzy set A on the universe of discourse X is defined as $\mu_A:X \rightarrow [0,1]$. Here, each element of X is mapped to a value between 0 and 1. It is called membership value or degree of membership. It quantifies the degree of membership of the element in X to the fuzzy set A.

- x axis represents the universe of discourse.
- y axis represents the degrees of membership in the [0, 1] interval.

There can be multiple membership functions applicable to fuzzify a numerical value. Simple membership functions are used as use of complex functions does not add more precision in the output. The triangular and trapezoidal membership function are most common among the various other membership function shapes, such as singleton, and Gaussian. The triangular and trapezoidal fuzzy membership functions are specified as the Equations of (1) and (2) and the models shown in fig. 2.

Triangular: $$f(x,a,b,c) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x \le b \\ \frac{c-x}{c-b}, & b \le x \le c \\ 0, & c < x \end{cases} \quad (1)$$

Trapezoidal: $$f(x,a,b,c,d) = \begin{cases} 0, & x \le a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & b \le x \le c \\ \frac{d-x}{d-c}, & c \le x \le d \\ 0, & d < x \end{cases} \quad (2)$$

where a, b, c, d are the parameters of the linguistic value and x is the range of the input parameters [14].

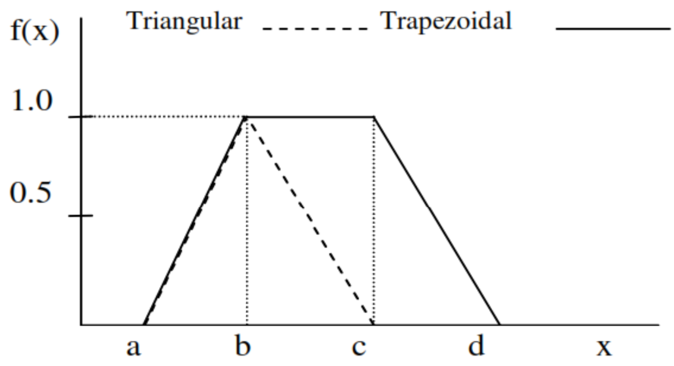


Fig. 2 Triangular Membership Function

2.2.2 Fuzzy Rules

The relationship between the input and the output is represented in the form of IF-THEN rules. If the number of each parameter membership function is $\mu(x)$ and the number of input parameters is n, then the number of rules R is,

$$R(Rule) = \mu(x_1)\mu(x_2) \dots \mu(x_n) \quad (3)$$

Where $x_1, x_2 \dots\dots x_n$ is the range of input parameters [15].

3. RESULT AND DISCUSSION

3.1 FUZZY WATER QUALITY INDEX (FWQI)

Fuzzy Water Quality Index (FWQI) model involves the following steps:

1. Selection of the input and output variables.
2. Selection of the membership functions for the input and output variables.
3. Formation of the linguistic rule base.
4. Defuzzification

3.1.1 Selection of the Input and Output Variables.

The first step in system modeling is the identification of the input and output variables called the system's variables. Only those inputs that affect the output largely have been selected. The most important input variables are pH, Dissolved Oxygen (DO), Total Solids (TS), Phosphate, Nitrogen, and Electric Conductivity (EC). The universe of discourse is also decided, based on the physical nature of the problem. In the selection procedure, the above-mentioned inputs and the outputs are also taken in the form of linguistic format that displays an important role in the application of fuzzy logic.]

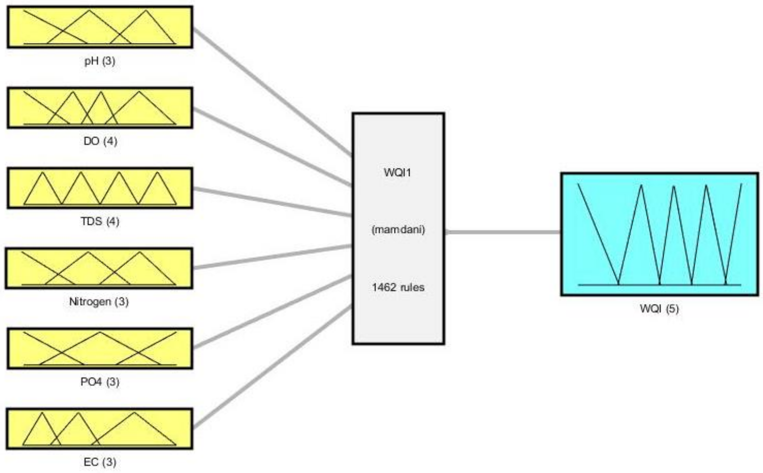


Fig. 3 Fuzzylogic Network for WQI

The above fig. 3 shows the fuzzy logic network between the input and output variables. The input variables are classified as pH = {low, medium, high}, DO = {low, medium, medium high, high}, TS = {low, medium, medium high, high}, Nitrogen = {low, medium, medium high, high}, Phosphate = {low, medium, high}, and EC = {low, medium, high}. The output variable is similarly divided into WQI = {A-Excellent, B-Desirable, C-Acceptable}.

3.1.2 Selection of Membership Functions

Linguistic values are expressed in the form of fuzzy sets. A fuzzy set is usually defined by its membership functions. In general, a triangular membership function is used to normalize the crisp inputs because of its simplicity and computational efficiency. It is described mathematically in the following manner

$$\text{Triangle } (x; a, b, c) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right) \quad (4)$$

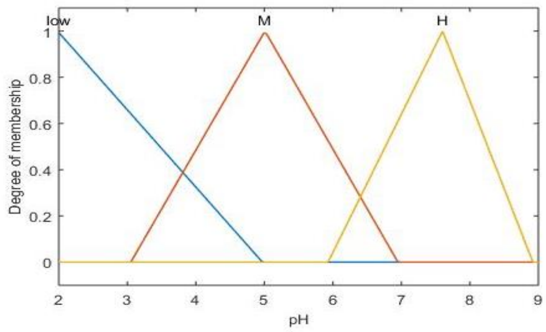
Where a, b, c are the parameters of the linguistic value and x is the range of the input parameters. This triangular membership functions as described in the above expressions (4) convert the linguistic values to a range of 0–1. The membership function of inputs is represented in Fig. 4 and the output is represented in Fig. 5 and the membership ranges are given in the table 1 and 2.

Table 1 Parameter and Ranges of Input Membership Functions (Mf)

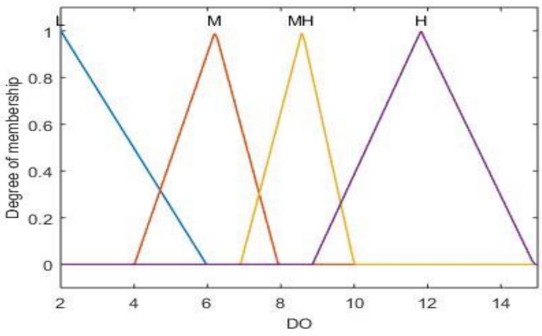
Parameter	Units	Range			
		Low (L)	*Medium (M)*	*Medium High (MH)*	*High (H)*
pH	-	2-5	3-7	-	6-9
DO	Mg/l	2-6	4-8	7-10	9-15
TDS	Mg/l	0-100	100-200	200-300	300-400
Nitrogen	Mg/l	1-5	2-10	-	7-14
Phosphate	Mg/l	0-0.4	0.1-0.9	-	0.6-1
EC	µ/cm	0-1000	750-2000	-	1750-4000

Table 2 Membership Function and their Ranges of Output MF

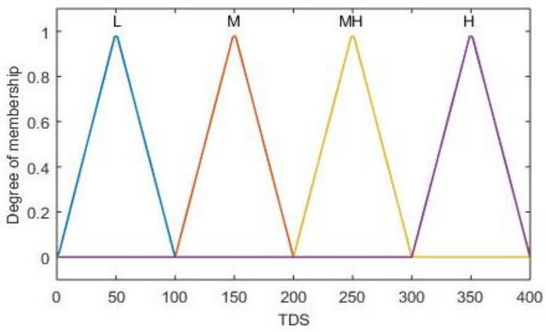
Type	Range	Classification
A	0-25	Excellent
B	25-50	good
C	50-70	Fair
D	70-90	poor
E	90-100	Very poor



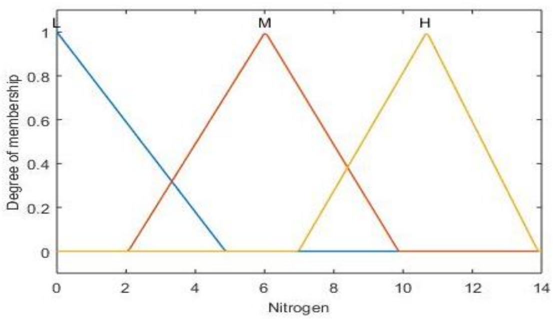
pH



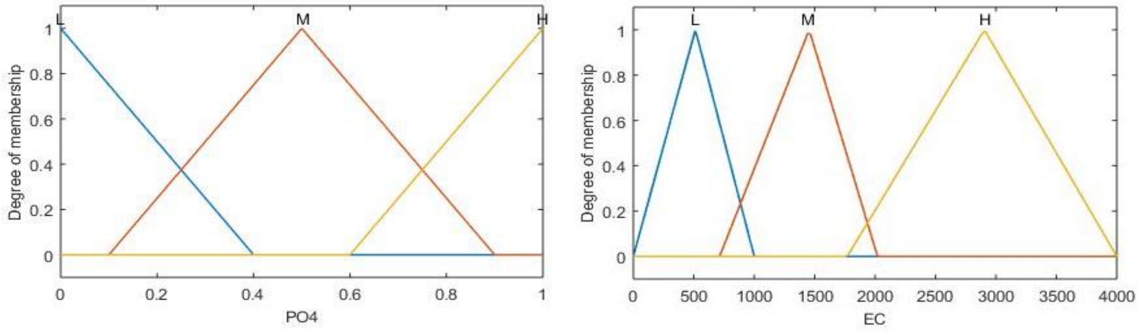
DO



TDS



Nitrogen



PO₄ **EC**

Fig. 4 Membership Functions for Inputs

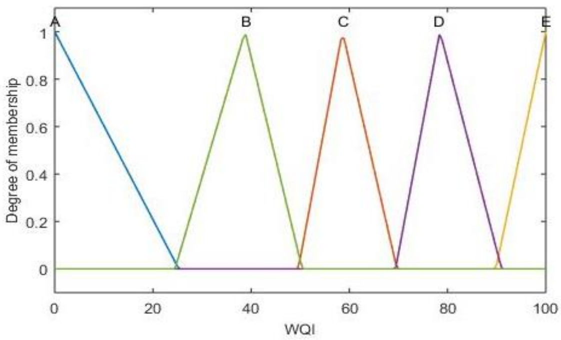


Fig. 5 Membership Functions for Output

3.1.3 Formation of Linguistic Rule-Base

The relationship between the input and the output are represented in the form of IF-THEN rules. The inputs, pH, DO, TDS, Nitrogen, Phosphate, EC are taken as X_1, X_2, X_3. ...X_6 respectively. The output of the system WQI is taken as y. In this system, X_1 have three MF, X_2 input have four MF, X_3 input have four MF, X_4 input have four MF, X_5 have three MF, and X_6 have three MF. The fuzzy system is designed for predicting WQI and it has six inputs as X_1, X_2, X_3, X_4, X_5 and X_6.

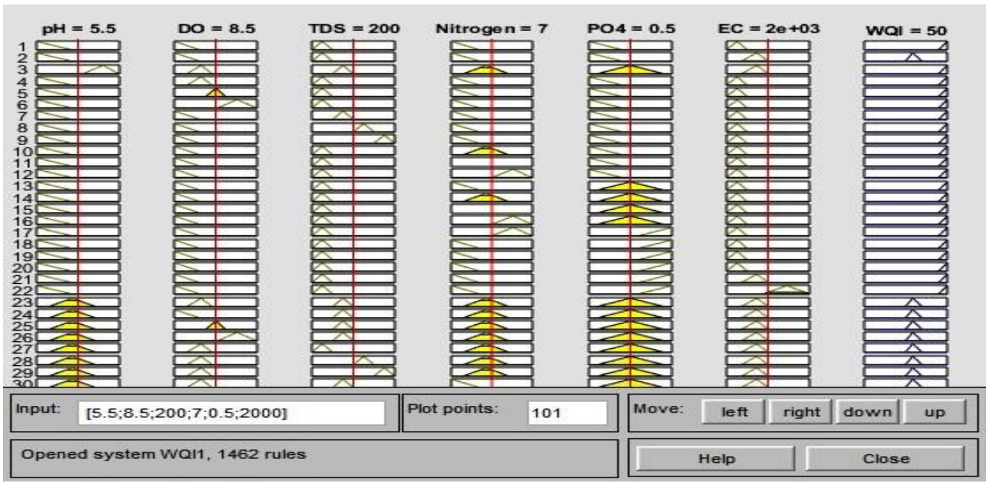


Fig. 6 Rule View of the Fuzzy Water Quality Index

The rules of the Mamdani fuzzy system are generated in the following ways and the rule views are shown in fig. 6.

R_1: **IF** X_1 is ${}^1\mu_{x1}$= "low" AND X_2 is ${}^1\mu_{x2}$= "low" AND X_3 is ${}^1\mu_{x3}$= "low" AND X_4 is ${}^1\mu_{x4}$= "low" AND X_5 is ${}^1\mu_{x5}$= "low" AND X_6 is ${}^1\mu_{x6}$= "low" THEN WQI(y) is y= ${}^3\mu_{y1}$ = ''C-Acceptable'';

R_2: **IF** X_1 is ${}^1\mu_{x1}$= "low" AND X_2 is ${}^1\mu_{x2}$= "low" AND X_3 is ${}^1\mu_{x3}$= "low" AND X_4 is ${}^1\mu_{x4}$= "low" AND X_5 is ${}^1\mu_{x5}$= "low" AND X_6 is ${}^1\mu_{x6}$= "Medium" THEN WQI(y) is y= ${}^2\mu_{y1}$ = ''B-desirable'';

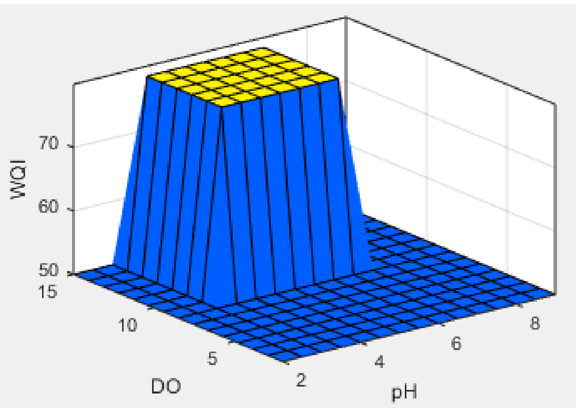
R_{1728}: **IF** X_1 is ${}^3\mu_{x1}$= "High" AND X_2 is ${}^4\mu_{x2}$= "High" AND X_3 is ${}^4\mu_{x3}$= "High" AND X_4 is ${}^3\mu_{x4}$= "High" AND X_5 is ${}^4\mu_{x5}$= "High" AND X_6 is ${}^3\mu_{x6}$= "High" THEN WQI(y) is y= ${}^3\mu_{y1}$ = ''C-Acceptable''.

3.1.4 Defuzzification

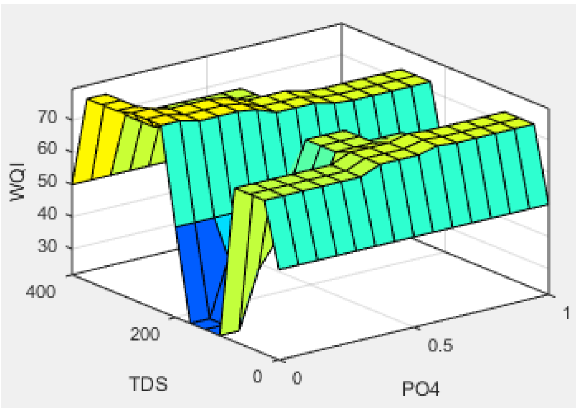
In this model, Centroid of area (COA) method of defuzzification is used for determining the output as expressed in Equation (6).

$$\text{centroid of Area COA} = \frac{\int \mu A(z) z dz}{\int \mu A(z) dz} \tag{6}$$

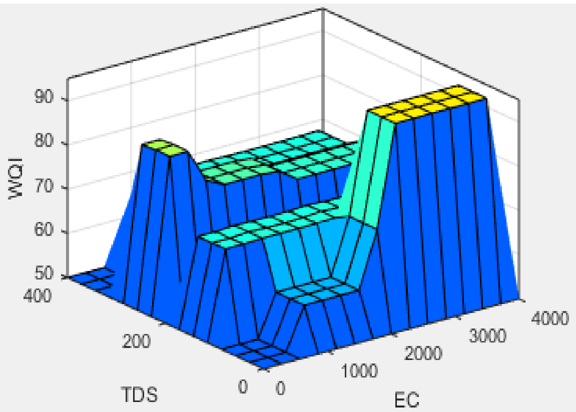
Where, µA (z) as the membership value of set A and z is the variable. The surface model is created between the two selected inputs and outputs. It shows the variations of the output data in respect to the input data. Fig. 7 shows the surface model of the different input data and the output data.



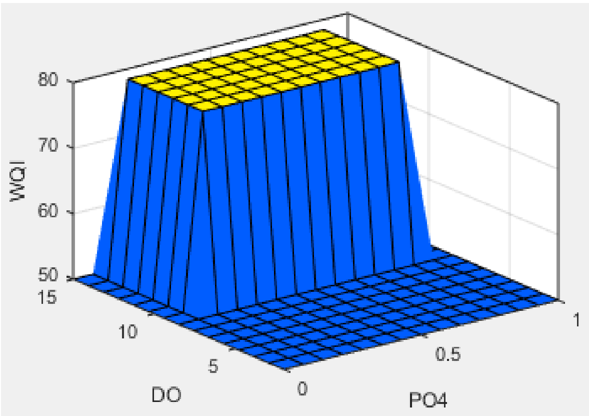
DO pH and WQI



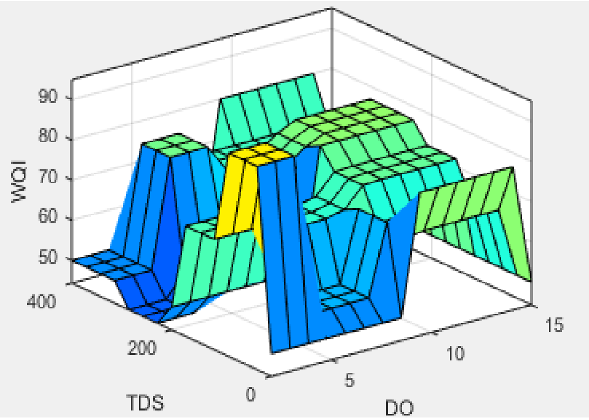
PO_4, TDS and WQI



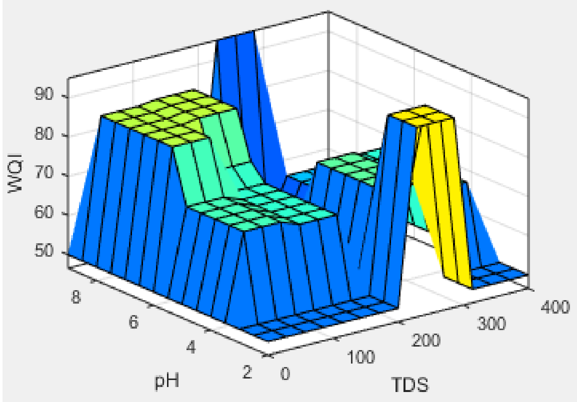
EC, TDS and WQI



DO, PO_4 and WQI



DO, TDS and WQI



TDS, pH and WQI

Fig. 7 Surface Models between WQI and Inputs

3.1.5 Simulation of WQI

In FWQI evaluation method, the potable water quality is classified in to five forms of "Excellent, Good, Fair, Poor and Very Poor" and we can easily suggest the final groundwater quality. In this study, 40 groundwater samples are used to simulate the Fuzzy Water Quality Index. When we consider the result, about 55% of the water samples fall under the category of "Good" and the remaining 45% samples are in the category of "Fair" as shown in table 3. According to FWQI, the water quality of the study area is "Good" or "Fair" no sample belongs to the "Excellent" category. pH level of the study area is less than the permissible level and it influences the groundwater quality as well. This is the reason for the degradation of the groundwater quality.

Table 3 Classification of Water Quality Based on FWQI

Sample No	pH	DO	TDS	Nitrogen	Phosphate	EC	FWQI	Quality
1.	5.66	0.12	112	0	0.05	170	50	Good
2.	5.6	0.2	104	0	0.15	157	42.5	Good
3.	5.5	0.24	109	0.01	0.1	165	52.6	Fair
4.	5.56	0.16	290	0.03	0.1	295	38.65	Good
5.	5.4	0.08	137	0.01	0.15	208	48.5	Good
6.	6.4	0.16	154	0	0.05	233	52.02	Fair
7.	6.2	0.24	76	0.01	0.03	115	48.02	Good
8.	6.2	0.32	68	0.01	0.2	103	46.65	Fair
9.	5.46	0.08	172	0	0.05	261	38.67	Good
10.	5.43	0.12	247	0	0.65	375	47.65	Fair
11.	5.3	0.2	185	0.01	0.05	162	50.26	Fair
12.	6.1	0.16	134	0.01	0.05	186	62.58	Fair
13.	5.5	0.32	190	0.22	0.1	384	39.12	Good
14.	5.72	0.12	225	0.16	0	156	42.65	Good
15.	5.9	0.28	127	0	0	178	38.96	Good
16.	5.1	0.16	121	0	0.05	325	42.36	Good
17.	5.65	0.24	72	0.01	0.1	246	59.14	Fair
18.	5.8	0.32	69	0.01	0	162	47.63	Good
19.	5.8	0.08	137	0	0	248	52.16	Fair

20.	6.2	0.12	204	0	0.05	178	61.65	Fair
21.	5.7	0.28	99	0.04	0	150	32.4	Good
22.	5.7	0.24	72	0.03	0.05	109	40.72	Good
23.	5.62	0.12	101	0.04	0	153	39.96	Good
24.	5.8	0.2	260	0.05	0	181	48.57	Good
25.	5.9	0.28	81	0.08	0.01	123	58.73	Fair
26.	6.2	0.12	110	0.03	0	158	62.57	Fair
27.	6	0.16	68	0.04	0	103	40.15	Good
28.	6	0.4	60	0.03	0.03	98	42.35	Good
29.	5.15	0.32	146	0.05	0	187	38.63	Good
30.	5.23	0.56	186	0.01	0.05	210	52.54	Fair
31.	5.5	0.32	111	0.01	0.1	121	42.23	Good
32.	5.3	0.08	89	0	0.15	143	54.27	Fair
33.	5.8	0.12	126	0	0.05	226	52.36	Fair
34.	5.3	0.2	128	0.04	0.03	156	41.28	Good
35.	5.6	0.16	60	0.03	0.2	125	41.36	Good
36.	5.12	0.32	94	0.04	0.05	202	51.13	Fair
37.	6.2	0.12	56	0.05	0.65	162	62.27	Fair
38.	5.1	0.28	58	0.08	0.05	125	42.18	Good
39.	5.3	0.16	68	0.03	0.05	156	51.18	Fair
40.	6.3	0.24	115	0.01	0.1	125	65.28	Fair

4. CONCLUSION

The study has demonstrated the utility of Fuzzy interference technology combined with laboratory analysis in evaluation of groundwater quality. In FWQI evaluation method, the potable water quality is classified in to five forms “Excellent, Good, Fair, Poor and Very Poor” and we can easily come to a conclusion regarding the quality of the groundwater. The result reveals that about 55% of the sample comes under “Good” category and remaining 45% comes under “Fair” category. Hence, it shows that 55% of the water can be directly used for drinking purposes and the remaining 45% of the water needs to be pre-treated before drinking.

References

[1] I. Radojevic, L. Comic, V.Rankovic, A.Ostojic and M.Topuzovic, “Applying Neural Networks for Predicting the Facultative Oligotrophic Bacteria in Two Reservoirs with Different Trophic State,” J. Env. Prot. & Ecol., vol. 14, pp. 55-63, 2013.

[2] A.M. Seyyed, A.K. Gholam, P. Zeynab, B.Zeynab and S.Mohammad, “Estimate the spatial distribution TDS the fusion method Geostatistics and artificial neural networks,” Int. J. Agri. Crop. Sci., Vol. 6(7), pp. 410–420, 2013.

[3] H. Gharibi, A.H. Mahvi, R. Nabizadeh, H.Arabalibeik , M.Yunesian and M.H. Sowlat, “A novel approach in water quality assessment based on fuzzy logic,” J. Environ. Manage, Vol. 112, pp. 87- 95, 2012.

[4] R. Sadiq and M.J. Rodriguez, “Fuzzy synthetic evaluation of disinfection by-products: a risk-based indexing system,” J. Envir. Mngmt., vol. 73, pp.1–13, 2004.

[5] W. Ocampo-Duque , N.Ferre-Huguet , J.L. Domingo and M. Schuhmacher, “Assessing water quality in rivers with fuzzy inference systems: A case study,” Envir. Inter., Vol. 32, pp. 733-742, 2006.

[6] A. Lermontov , L. Yokoyama , M. Lermontov and M.Augusta Soares Machado, “River quality analysis using fuzzy water quality index: Ribeira do Iguape river watershed,” Brazil, Ecol. Indicators, Vol.9, pp.1188-1197, 2009.

[7] S. Dahiya, B.Singh, S.Gaur , V.K. Garg and H.S. Kushwaha, "Analysis of groundwater quality using fuzzy synthetic evaluation," J. Hazardous Materials, Vol.147, pp.938-946, 2007.

[8] N. Venkat Kumar, S.Mathew and G.Swaminathan, "Fuzzy Information Processing for Assessment of Groundwater Quality," Inter. J. Soft Comp., Vol. 4 (1), pp. 1-9, 2009

[9] Amir Saberi Nasr, Mohsen Rezaei and Majid Dashti Barmaki, "Groundwater contamination analysis using Fuzzy Water Quality index (FWQI): Yazd province," Iran J. of Geope, Vol. 3(1), pp. 47-55, 2013.

[10] S.S. Mahapatra, Santosh Kumar Nanda and B.K. Panigrahy, "A Cascaded Fuzzy Inference System for Indian River water quality prediction," Adv. in Engg. Soft., Vol. 42(10), pp. 787–796, 2011.

[11] Z. Katambara , J. Ndiritu,, "A fuzzy inference system for modeling stream flow: Case of Letaba River, South Africa," Physics & Chemistry of the Earth, Vol. 34, pp. 688-700, 2009.

[12] Debasmita Pattnaik, Bonani Sahu Devadutta and Samantaray, "Microcontroller Based Implementation Of A Fuzzy Knowledge Based Controller," Department Of Electrical Engineering National Institute Of Technology Rourkela.

[13] M. Firat, M. ErkanTuran, M.A. Yurdusev, "Comparative analysis of fuzzy inference systems for water consumption time series prediction," J. Hydrology, Vol. 374, pp. 235-241, 2009.

[14] Renish M. Vekariya, Rakesh P. Ravani, "Investment Casting Process Using Fuzzy Logic Modelling," Inter. J. of Mech.l Engg. & Robotics Res., Vol. 2(1) , pp. 232-241, 2013.

[15] Z. Sen and A. Altunkaynak ," Fuzzy system modelling of drinking water consumption prediction, Expert Systems with Applications," vol. 36, pp. 11745-11752, 2009.