Raspberry Pi based Smart Sensing Platform for Drinking Water Quality Monitoring System: A Python Framework Approach

Punit Khatri¹, Karunesh Kumar Gupta¹, Raj Kumar Gupta²

¹Department of Electrical and Electronics Engineering, Birla Institute of Technology & Science (BITS), Pilani-333031, India
²Department of Physics, Birla Institute of Technology & Science (BITS), Pilani-333031, India

Correspondence to: Punit Khatri (punit.khtr@gmail.com)

Abstract. Drinking or potable water quality monitoring is essential for mankind as it affects the human health directly or indirectly. This work reports a smart sensing platform for potable water quality monitoring. Five water quality parameters (pH, Dissolved Oxygen, Oxidation Reduction Potential, Electrical Conductivity, and Temperature) have been selected to monitor the water quality. The selection of water quality parameters is made based on guidelines of the Central Pollution and Control Board, New Delhi, India. A Graphical User Interface (GUI) is developed to provide an interactive Human Machine Interface for the end user. Python programming language is used for GUI development, data acquisition and for data analysis. Fuzzy computing technique is employed for decision making to categorize the water quality in different classes like bad, poor, satisfactory, good and excellent. The system has been tested for various water resources and results have been displayed.

1 Introduction

Drinking water quality monitoring is very essential before consumption in daily life as it affects directly or indirectly human health (Bhardwaj et al., 2018). The water crisis has become a global problem in recent years, it is not limited to a particular region or country. By the end of 2025, half of the world population will be living in water-stressed areas (World Health Organization (WHO), 1996). In developing countries, as much as 80% of illnesses are linked to poor quality water and sanitation conditions (Anan, 2003). India is one of the most water-challenged countries, among developing countries in the world. Groundwater levels are decreasing day by day as farmers, city residents and industries are regularly draining wells and aquifers. The available water is severely polluted and it may create the worst situation in the future. Drinking water with pollutant concentration exceeding BIS (Bureau of Indian Standard) limits, is considered unsafe. The groundwater quality was measured by IWT (India Water Tool) 2.1 tool (World Resources Institute, 2016) among 632 districts in India. Among 632 districts, 59 are above BIS limits. The yellow and red areas in figure 1 indicate places where chlorine, fluoride, iron, arsenic, nitrate and/or electrical conductivity exceed national standards. This is one of the most critical reasons why drinking water quality monitoring is essential.

Traditionally, water quality measurement involves sample collection on sites and subsequent laboratory chemical based analysis, which is both the labor and cost intensive (Korostynska et al., 2013). Therefore it is the need of the hour to have a real-time monitoring of water quality for drinking applications (Bhardwaj et al., 2015). The work presented in this paper aims
to provide an efficient real-time monitoring of ground-water in drinking applications and can address the issues related to drinking water quality for various species.

The proposed work consists of Multi-Sensor Array (MSA), hardware platform along with a software platform (i.e. python framework). In the first stage, MSA is designed that consists an array of commercially available individual sensors for pH, Dissolved Oxygen (DO), Oxygen Reduction Potential (ORP), Electrical Conductivity (EC) and temperature. The Total Dissolved Solid (TDS) is derived from Electrical Conductivity (EC). The selection of the parameters is based on the guidelines of the Central Pollution and Control Board (CPCB), New Delhi, India (Central Pollution Control Board, 2007). In the second stage, the hardware platform was designed based on the Raspberry Pi board. An additional serial port expander is used since Raspberry has only one I2C channel and it is required to communicate with all the sensors simultaneously. In the third stage, a software platform is designed which consists of Graphical User Interface (GUI) for interactive Human Machine Interface (HMI) and fuzzy modeling. A python framework is utilized in software and GUI development. The water standard is defined based on calculated water quality parameters using Fuzzy Inference System (FIS) since FIS has the capability to imbibe vagueness related to observed parameters. The fuzzy approach reported in this work has been used and discussed widely in many environmental applications and helps in decision making in many real-life complex problems (Lermontov et al., 2009). Many researchers (S. Ponsadailakshmi (Ponsadailakshmi et al., 2018), A. Tiri (Tiri et al., 2018), B. V Raman (Raman et al., 2009), A. M. Jinturkar (Jinturkar et al., 2010), Y. Icaga (Icaga, 2007), etc.) have implemented the fuzzy modeling in MATLAB for water quality index calculation. That process of modeling is offline e.g. researchers have collected the data and later done the modeling in MATLAB. Whereas in this paper, we are trying to implement the fuzzy in real-time calculation spontaneously after the data collection with the help of libraries in the Python framework.

2 Materials and methods

2.1 Water Quality Parameter Selection Criteria

The Central Pollution and Control Board (CPCB) have suggested the criteria of water quality parameters for different usage of the water which is shown in table 1. In this work, category ‘C’ parameters have been considered since we are targeting drinking water source after conventional treatment and disinfection. Experimental work has proven that water quality parameters such as pH, EC, and DO are indirect indicators of nicotine, arsenic trioxide and Escherichia coli (Hall et al., 2007) (Power and Nagy, 1999). That’s why for the detection of contamination, these water quality parameters have been considered as promising criterion.

2.2 Defining Water Quality

The fuzzy inference system (FIS) mimics the way human thinks in his day to day life. Hence fuzzy logic based techniques proved to be very effective since it is less mathematically intensive and supports approximate reasoning. In FIS, the knowledge is presented as linguistic rules. The inputs are converted from crisp value to linguistic variable by the process called
fuzzification and these variables are fed to inference system. This inference system gives a new set of a linguistic variable which is then converted to crisp value with the help of defuzzification (MathWorks, n.d.). The basic process to design a fuzzy logic involves three (3) basic steps shown in figure 2.

The proposed fuzzy logic was implemented in python with the help of library known as scikit-fuzzy library (Anon, n.d.) to define the water quality from the groups of five linguistic variables defined as bad, poor, satisfactory, good and excellent. The fuzzy system uses Mamdani implication model, which takes five inputs pH, Electrical Conductivity (EC), Oxidation Reduction Potential (ORP), Dissolved Oxygen (DO) and temperature. The Mamdani Fuzzy Inference System produce a more accurate response as compared to Takagi-Sugeno type model since it exerts the centroid method of defuzzification. The defuzzified output of the model is water quality which corresponds to five inputs of the model. In this paper, Mamdani-type FIS model is implemented in a decision support system since it possesses spontaneous and interpretable nature of the rule base capability.

To decide the water quality, five inputs and one output are selected and modeling is performed based on these parameters. The selection of the membership function is done based on the complexity of the system considered for decision making. Triangular membership function (MF) is most commonly used membership functions because of its linear nature and easy implementation ability (Zhao and Bose, 2002) (Kosko, 1993), hence we have selected triangular MF to fuzzify the crisp variable into linguistic one. The triangular membership function as shown in figure 3 depends on three parameters \( l, m \) and \( n \) and are given by equation 1.

\[
 f(x;l,m,n) = \begin{cases} 
 0 & \text{for } x < l \\
 \frac{x-l}{m-l} & \text{for } l \leq x \leq m \\
 \frac{n-x}{n-m} & \text{for } m \leq x \leq n \\
 0 & \text{for } x > n 
\end{cases}
\]  

(1)

The logic operations used in the fuzzy logic are \( \text{min} \), \( \text{max} \) and \( \text{compliment} \) and these are defined by the equation (2), (3) and (4) respectively. Let A and B are two subsets.

\[
 \mu_{A\cap B}(x) = \max[\mu_A(x), \mu_B(x)] 
\]  

(2)

\[
 \mu_{A\cap B}(x) = \min[\mu_A(x), \mu_B(x)] 
\]  

(3)

\[
 \mu_{\overline{A}}(x) = 1 - \mu_A(x) 
\]  

(4)

After the logic operations, ‘if-then’ rule is applied. All the rules are applied in parallel and the rule which does not effect the output is fired. The outputs of all rules are then aggregated and all fuzzy sets that affect the output, are combined into a single fuzzy set. Finally, the fuzzy set is converted into a crisp set by means of defuzzification in which a single number is generated. There are several methods for defuzzification such as centroid, maximum, mean of maxima, height and modified height. In
this work, the centroid defuzzification method is used which is the most popular method. The output is calculated by averaging individual centroids, weighted by their heights as given by equation 5 (Zadeh, 1988).

\[ U_o = \frac{\sum_{i=1}^{n} u_i \mu(u_i)}{\sum_{i=1}^{n} \mu(u_i)} \]

(5)

Where \( \mu(u_i) \) is the min/max value of the membership degree of the input values (depends on min/max operator). The overall fuzzy inference system is shown in figure 4.

As per the acceptable range of water quality parameters, we have assigned two groups to each parameter which are desirable (DES) and undesirable (UNDES) as described in table 2. If the parameters are in the desirable range than only fuzzy has been applied otherwise the sample has been rejected. After checking the acceptable range, we have assigned individual membership functions to each parameter as shown in figure 5. Also, we have defined the membership function for water quality on the scale of 0 to 100. After assigning the membership functions, “if-then” rule is applied and overall quality is defined on the basis of adopted rule base formulation.

3 Hardware Platform Design

The hardware platform plays a vital role in any system development since data acquisition and data processing is done with the help of the hardware platform itself. The main task of the hardware platform design is MSA design and its interfacing with the raspberry pi board followed by python framework. The details of the design are given below.

3.1 Multi-Sensor Array (MSA) Design

For the proposed work, Multi-Sensor Array (MSA) is designed using the industrially manufactured sensors from Atlas Scientific, USA. The individual sensors are arranged in an array form to make the MSA. The sensors used are pH sensor, Electrical Conductivity (EC) sensor, Dissolved Oxygen (DO) sensor, Oxidation Reduction Potential (ORP) sensor and a temperature sensor. Total Dissolved Solids (TDS) parameter was derived from EC. The block diagram of MSA and its interfacing with hardware followed by software framework is shown in figure 6.

3.2 Integration of MSA with Raspberry Pi board and interfacing with Python

Once the MSA is designed, it has to be integrated with the Raspberry Pi for data acquisition and further data processing. Raspberry Pi is a single board credit card size micro-computer with ARM cortex A-53 processor (Raspberry Pi Foundation, 2014). It is an open hardware with many on-board running components like CPU, graphics, memory, USB controller etc.
Nowadays, Raspberry Pi board is being used in many real-time applications (for e.g. real-time video surveillance system (MathWorks, n.d.), real-time paper currency recognition of new Indian notes after demonetization (Anilkumar and Srikanth, 2018), automatic traffic control system (Talukder et al., 2017), smart traffic system (Kumar et al., 2017), energy management system based on real-time electricity pricing model (Qureshi et al., 2017) etc.), self-driving system (Sumardi et al., 2018), 3D wavelet transform (Bernabé et al., 2018), air quality monitoring system (Alkandari and Moein, 2018), cyber physical system (CPS) based water quality monitoring (Bhardwaj et al., 2018) using Arduino Uno board which has got some limitation with computational capabilities. Although Arduino is user-friendly, the reason for selecting the Raspberry Pi for this work is its computational capabilities that cannot be done in the Arduino platform. Apart from the computational capabilities, Raspberry Pi is bundled with inbuilt Bluetooth module, Wi-Fi module, HDMI interface, camera interface, display slot, SD card slot, USB slot, etc. which has to be interfaced externally in case of Arduino depending upon application.

For stand-alone applications, we can use a display panel and interface it with the python framework. We have used a 7” touch screen and interfaced it with the python programming. Python is an open source programming platform and supports the data processing and computing. The only drawback that we are facing with raspberry Pi board is a limited number of I2C channel, so we have interfaced external serial port expender with the board to enhance the number of I2C channels. The experimental setup of MSA integrated with Raspberry Pi board is shown in figure 7. It includes MSA connected to serial port expender and Raspberry board followed by python programming.

4 Results and Discussion

4.1 Experimental Procedure

The proposed system has been tested for water samples from five different locations. In order to get accurate readings, each sensor node has been calibrated before going for measurement. The calibration was performed with the reference solution given for each sensor. Initially, the measurement iteration was carried out for at least five minutes so that sensor reading gets stabilized because the original readings have to be recorded only after the sensor attains stability in order to make any conclusive illation out of data. The data readings obtained from MSA were fed to the fuzzy decision support system implemented in python with the help of available libraries. The complete procedure of data acquiring and fuzzy decision support system was implemented in python for real-time measurement.

4.2 Interactive User Interface

The graphical representation is being provided for real-time data obtained from various sensors with the help of GUI platform for the interactive user interface. An interactive Graphical User Interface (GUI) has been designed in python framework where the user can select the individual parameter to be measured as well as check overall water quality. A screenshot of GUI is
shown in figure 8. The acquired data was kept for future use by means of saving in memory drive provided with Raspberry board. The live plotting of data is shown in figure 9. The X-axis represents time and Y-axis represents the sensor node reading.

4.3 Validation and Performance Comparison

The validation of results acquired from the proposed system was done by comparing the results obtained from available benchmark equipment YSI EXO-1 sonde monitoring system and calculating the percentage relative error (PRE) (Lee, 2016). PRE expresses the error in percentage to determine the accuracy and is given by

\[
PRE = 100 \times \left( \frac{\text{actual} - \text{observed}}{\text{actual}} \right)
\]  

The system was tested for the total duration of 21 hours over seven days. The results obtained from MSA were accumulated through fuzzy implemented in python framework. The average values of the experiment are shown in table 3. The calculated PRE plot is shown in figure 10. Based on the results of the parameters obtained from MSA, water quality has been defined for all the locations using fuzzy libraries as shown in table 4.

5 Conclusion

Water quality monitoring is essential before consumption and its real-time monitoring will reduce the risk of illness in the human being. This paper reported a smart sensing platform for real-time water quality monitoring and to collect a large database. The designed platform is compatible with IoT networks as the Raspberry Pi used here has a built-in Wi-Fi module and will be implemented in future looks. The work presented here has both academic and practical importance. Currently, the calibration of the sensor is time-consuming and requires a certain time period to get stabilized. In future looks, the focus will be on auto-calibration of the sensors and also the drift analysis and compensation by means of algorithms.

Acknowledgments

The authors are pleased to acknowledge the Department of Science and Technology, Govt. of India, New Delhi for their financial support (DST/TM/WTI/2K16/103) and Birla Institute of Technology and Science for providing an enabling environment to carry out the research work.

Conflict of Interest

The authors declare that they have no conflict of interest.
References


Kosko, B.: Fuzzy thinking: the new science of fuzzy logic, Hyperion. [online] Available from:


Table 1. Central Pollution and Control Board criteria for Water Quality (Central Pollution Control Board, 2007)

<table>
<thead>
<tr>
<th>Type of Water</th>
<th>Category</th>
<th>Quality Parameter Criteria</th>
</tr>
</thead>
</table>
| Drinking Water Source without conventional treatment but after disinfection  | ‘A’      | 1. Total Coliforms Organism MPN/100ml shall be 50 or less  
2. pH between 6.5 and 8.5  
3. Dissolved Oxygen 6mg/l or more  
4. Biochemical Oxygen Demand 5 days 20°C 2mg/l or less  |
| Outdoor bathing (Organised)                                                  | ‘B’      | 1. Total Coliforms Organism MPN/100ml shall be 500 or less  
2. pH between 6.5 and 8.5  
3. Dissolved Oxygen 5mg/l or more  
4. Biochemical Oxygen Demand 5 days 20°C 3mg/l or less  |
| Drinking water source after conventional treatment and disinfection           | ‘C’      | 1. Total Coliforms Organism MPN/100ml shall be 5000 or less  
2. pH between 6 to 9  
3. Dissolved Oxygen 4mg/l or more  
4. Biochemical Oxygen Demand 5 days 20°C 3mg/l or less  
5. TDS 2000 mg/l  |
| Propagation of Wildlife and Fisheries                                         | ‘D’      | 1. pH between 6.5 to 8.5  
2. Dissolved Oxygen 4mg/l or more  
3. Free Ammonia (as N) 1.2 mg/l or less  |
| Irrigation, Industrial Cooling, Controlled Waste disposal                    | ‘E’      | 1. pH between 6.0 to 8.5  
2. Electrical Conductivity at 25°C micromhos/cm Max.2250  
3. Sodium absorption Ratio Max. 26  
4. Boron Max. 2mg/l  |
Table 2. Groups defined for water quality parameters

<table>
<thead>
<tr>
<th>Range Parameters</th>
<th>UNDES</th>
<th>DES</th>
<th>UNDES</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>&lt; 6.5</td>
<td>6.5-8.5</td>
<td>&gt; 8.5</td>
</tr>
<tr>
<td>EC</td>
<td>&lt; 300</td>
<td>300-1000</td>
<td>&gt; 1000</td>
</tr>
<tr>
<td>ORP</td>
<td>&lt; 200</td>
<td>200-800</td>
<td>&gt; 800</td>
</tr>
<tr>
<td>DO</td>
<td>&lt; 3</td>
<td>3-11</td>
<td>&gt; 11</td>
</tr>
<tr>
<td>Temperature</td>
<td>&lt; 2</td>
<td>2-35</td>
<td>&gt; 35</td>
</tr>
</tbody>
</table>

* UNDES = undesirable  
  DES = desirable

Table 3 (a). Average values of samples for pH, DO & EC and their calculated PRE

<table>
<thead>
<tr>
<th>Location</th>
<th>pH</th>
<th>DO (mg/l)</th>
<th>EC (µS/cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSA</td>
<td>Commercial System</td>
<td>PRE (%)</td>
</tr>
<tr>
<td>1</td>
<td>7.45</td>
<td>7.51</td>
<td>0.8</td>
</tr>
<tr>
<td>2</td>
<td>7.62</td>
<td>7.68</td>
<td>0.78</td>
</tr>
<tr>
<td>3</td>
<td>6.95</td>
<td>6.99</td>
<td>0.57</td>
</tr>
<tr>
<td>4</td>
<td>8.1</td>
<td>8.2</td>
<td>1.2</td>
</tr>
<tr>
<td>5</td>
<td>7.8</td>
<td>7.87</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 3 (b). Average values of samples for ORP & Temperature and their calculated PRE

<table>
<thead>
<tr>
<th>Location</th>
<th>ORP (mV)</th>
<th>Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSA</td>
<td>Commercial System</td>
</tr>
<tr>
<td>1</td>
<td>213</td>
<td>212</td>
</tr>
<tr>
<td>2</td>
<td>212</td>
<td>210</td>
</tr>
<tr>
<td>3</td>
<td>185</td>
<td>187</td>
</tr>
<tr>
<td>4</td>
<td>206</td>
<td>208</td>
</tr>
<tr>
<td>5</td>
<td>191</td>
<td>194</td>
</tr>
</tbody>
</table>
Table 4. Fuzzy water quality for all locations

<table>
<thead>
<tr>
<th>Location</th>
<th>Fuzzy Water Quality (FWQ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Good</td>
</tr>
<tr>
<td>2</td>
<td>Good</td>
</tr>
<tr>
<td>3</td>
<td>Satisfactory</td>
</tr>
<tr>
<td>4</td>
<td>Good</td>
</tr>
<tr>
<td>5</td>
<td>Good</td>
</tr>
</tbody>
</table>

Figure 1 Groundwater Quality (Number of BIS standard breaches)
Figure 2 Fuzzy logic designing process

Figure 3 Triangular Membership Function

Figure 4 Fuzzy Inference System
Figure 5 Input and output membership functions
Figure 6 Block Diagram of Multi-Sensor Array (MSA) and It's Interfacing with Hardware and Software

(a)
Figure 7 (a) The proposed System (b) Multi-Sensor Array (MSA) (c) MSA integration with serial port expender and Raspberry Pi board

![Image of the proposed system with sensors and a Raspberry Pi board]

Figure 8 Graphical User Interface

![Image of the graphical user interface for drinking water quality measurement]

Drinking Water Quality Measurement

<table>
<thead>
<tr>
<th>Select Measurement</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. pH</td>
<td></td>
</tr>
<tr>
<td>2. Electrical Conductivity</td>
<td></td>
</tr>
<tr>
<td>3. Dissolved Oxygen</td>
<td></td>
</tr>
<tr>
<td>4. Oxidation Reduction Potential</td>
<td></td>
</tr>
<tr>
<td>5. Temperature</td>
<td>Temperature 32.259</td>
</tr>
<tr>
<td>6. Water Quality</td>
<td></td>
</tr>
</tbody>
</table>

Figure 8 Graphical User Interface
Figure 9 Temperature plot in GUI

Figure 10 Percentage Relative Error (PRE) Plot for different water quality parameters