IOT Based Smart Water Management System with Machine Learning

G.S.Nandini,
Assistant Professor,
Electronics and communication Engineering,
Shree Venkateshwara HiTech Engineering College,
Gobi,Erode-638455
Tamil Nadu, India
nandhinigsece@gmail.com

P.Bhuvaneshwari,
Electronics and communication Engineering,
Shree Venkateshwara HiTech Engineering College,
Gobi,Erode-638455,
Tamil Nadu, India
2003bhuvanapooja10@gmail.com

S.Gowthami,
Electronics and communication Engineering,
Shree Venkateshwara HiTech Engineering College,
Gobi,Erode-638455,
Tamil Nadu, India
gowthamis1125@gmail.com

S.Elango
Electronics and communication Engineering,
Shree Venkateshwara HiTech Engineering College,
Gobi,Erode-638455,
Tamil Nadu, India
navaelango4@gmail.com

***Abstract*— This paper presents an Internet of Things (IoT)-based smart water meter with machine learning (ML)-aided water quality assessment capability. A flow rate sensor is utilized to measure water consumption while pH and turbidity sensors are employed for water quality assessment. The collected data is transmitted to a remote server via a cellular network, where it is utilized for monitoring purposes by both the utility company and customers. The system evaluates the collected data against relevant thresholds and issues appropriate notifications to the service provider and the customers. The thresholds for water quality are based on the national standards for potable water, while those for consumption are determined by the average monthly water consumption. This paper considers National Water and Sewerage Corporation (NWSC), the largest water utility company in Uganda, as a case study. A total of 1,760 samples collected by NWSC in the Kampala service area in 2022 were assessed using the feature selection algorithm of ML. The most dominant parameters were determined as residual chlorine, pH, turbidity, conductivity, and apparent color. In this paper, only pH and turbidity are considered.**

Keywords: Internet of Things (IoT), smart water meter, machine learning (ML), water quality assessment, flow rate sensor, pH sensor, turbidity sensor, remote monitoring, cellular network, threshold evaluation, notification system, potable water standards, consumption thresholds, National Water and Sewerage Corporation (NWSC), Uganda, feature selection algorithm

I.INTRODUCTION

Water quality monitoring involves the regular assessment of water supply [1], while water consumption monitoring focuses on measuring water usage [2]. Both processes are vital for safeguarding public health and preventing water wastage. Inadequate water quality monitoring heightens the risk of waterborne diseases such as malaria, typhoid, and cholera [1]. Similarly, a lack of water consumption monitoring leads to wastage through leakage. Consequently, the World Health Organization (WHO) mandates water monitoring, particularly in urban areas [1]. In Uganda, the National Water and Sewerage Corporation (NWSC) oversees water supply and routine quality monitoring [3]. Although NWSC has digitalized its services, customer complaints and manual billing remain challenges.

The IoT aspect, as described in [4] and [5], entails a network comprising physical devices, sensors, and software interconnected to collect and exchange data. This infrastructure facilitates seamless integration and communication among various devices, enabling real-time monitoring and control of water-related parameters. Through this system, both water consumption and quality are monitored in real-time using cost-effective sensors, facilitating effective monitoring by NWSC and customers alike. Additionally, ML algorithms are integrated to analyze collected data, pinpointing key parameters for water quality assessment. Thresholds for water quality and consumption adhere to national potable water standards [6] and average monthly usage, respectively. Incorporating alert messages enhances fault response time and customer safety.

The main contributions of the paper are outlined as follows:

1. Integration of both water quality and consumption monitoring into a unified system through the utilization of IoT technology. This methodology offers a comprehensive perspective on water management.

2. Adoption of data-driven analysis of water quality through the utilization of feature scaling ML algorithms for assessing water quality and consumption in Uganda. This data-centric approach enables targeted interventions addressing the underlying causes of water pollution.

II. LITERATURE REVIEW

This section delves into various methodologies utilized to address issues concerning both water quality and water consumption, examining alternative approaches to tackle these challenge

TABLE 1
MAXIMUM PERMISSIBLE NATIONAL STANDARDS FOR POTABLE WATER(WHO)

Parameter	Maximum Permissible	Unit
pH	6.5 - 8.5	-
Turbidity	5.0	NTU
Total Dissolved Solids (TDS)	700	mg/L
Hardness (as $CaCO_3$)	500	mg/L
Chlorine Residual	0.2 – 5	mg/L
Nitrate (as NO_3)	5.0	mg/L
Fluoride	1.5	mg/L
Arsenic	0.01	mg/L
Lead	0.05	mg/L
Mercury	0.001	mg/L
Calcium	75	mg/L
Magnesium	50	mg/L
Bi-carbonate	500	mg/L
Manganese	0.2	mg/L
Chloride	250	mg/L
Iron	<0.3	mg/L
Sulphate	200	mg/L
Ammonia	0.5	mg/L
Cadmium	<0.001	mg/L
Copper	1.0	mg/L
Zinc	5.0	mg/L
E. Coli	0	CFU/100mL
Total Coliforms	0	CFU/100Ml
Fecal Coliforms	0	CFU/100mL

The method introduced by the authors in [7] involves utilizing specific laboratory equipment to gather water samples, which are subsequently analyzed in the laboratory. The determination of water quality entails testing multiple parameters, which are then assessed according to the national standards established by the WHO [6], as outlined in Table 1

Authors in [8] discuss a methodology wherein various sensors, including those for conductivity, pH, and turbidity, are employed to measure different water quality parameters. Each sensor is responsible for measuring its respective parameter, and the data collected is stored in a remote server. Subsequently, the server evaluates these values against thresholds established in accordance with the national standards for potable water set by the WHO [6].

In the ML-based water quality assessment method presented in [9], the process is structured into three primary steps: data collection and processing, model training, and model evaluation. Throughout the data collection and processing phase, various water quality parameters are consolidated to create a standardized measure referred to as the Water Quality Index (WQI). This index forms the foundation for establishing the Water Quality Class (WQC), facilitating the application of ML models for predictive tasks. The computation of the WQI is accomplished using equation (1), and the respective WQC can be identified by consulting Table

$$WI = \sum qn\ Wn / \sum Wn \tag{1}$$

As indicated in (1), Wn is the unit weight, expressed as:

$$Wn = 1 / \sum Sn \tag{2}$$

TABLE II
WATER QUALITY CLASS (WQC)

WQI Range	Class
0-25	Excellent
25-50	Good
50-70	Poor
70-90	Very poor
90-100	Unfit for consumption

where Sn is the standard desirable value of the given water parameter under evaluation. Qn is the sub-index value and is expressed as:

$$Qn = (Vn - Vo)/(Sn - Vo) \times 100\% \tag{3}$$

where Vn is the mean concentration of the water quality parameter and Vo is the actual value of the parameters in pure water

. In the model training phase, feature selection, classification, and regression algorithms are utilized to pinpoint the most significant water quality parameters. Following this, during the evaluation stage, various Classification and Regression ML models are scrutinized to pinpoint those that demonstrate superior performance in predicting the WQC.

Water Consumption Monitoring

Water metering involves quantifying the volume of water utilized. Consequently, the water meter emerges as a crucial instrument for both the utility provider and its clientele to gauge the precise water consumption. Various types of water meters exist, including positive displacement meters, velocity water meters, mechanical water meters, and digital water meters [10]. Additionally, an IoT-based smart water meter can be deployed to monitor water usage via wireless flow sensor nodes. These nodes possess the capability to transmit consumption data to a remote server in real-time via the Internet. Subsequently, the consumption data can be analyzed to detect instances of excessive water usage, prompting the issuance of notifications to customers for proactive monitoring of their water consumption.

III. METHODOLOGY

This section outlines the comprehensive procedure for the development, construction, and testing of the prototype for water quality assessment and consumption monitoring.

A. Water Quality Assessment Subsystem

The subsequent steps were undertaken in crafting the prototype for the water quality assessment system:

- Identification of water quality parameters through ML techniques.
- Conceptualization of the prototype for the water quality system
- .Execution of the prototype for the water quality system.

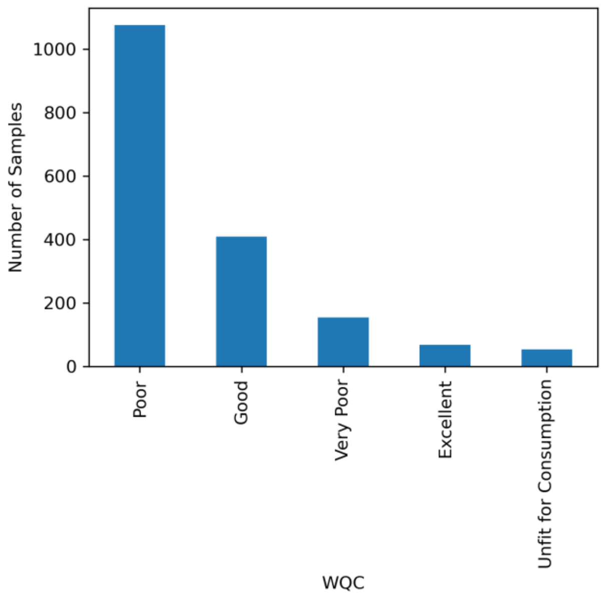

Fig. 1. WQC distribution

Determining the water quality parameters: The relevant water quality parameters to be measured were determined using a method of Classification ML called feature selection. The steps involved include: data processing, model training, and finally model evaluation.

a) Data processing: In this step, a dataset provided by NWSC consisting of 1,760 water samples collected from tertiary points within the Kampala service area in 2022 was processed. This involved cleaning the data and subsequently using the data to calculate the WQI. The WQI values were then utilized to determine the WQC based on the reference provided in Table 2. The distribution of the number of samples with respect to the WQC is shown in Fig. 1.

We then conducted a correlation analysis to examine the association between various water quality parameters and their corresponding WQCs. Their findings indicated that residual chlorine exhibited the strongest correlation with the WQC, as depicted in Fig. 2. The elevated levels of residual chlorine can be attributed to its initial high concentration at the distribution point. This practice is commonly employed to ensure effective chlorine treatment reaches all customers, especially those located at the end of the distribution network.

Feature scaling was then performed to ensure the data is both normalized and fits into a single scale from 0 to 1 for all water quality parameters. This helps improve the accuracy of the ML algorithms when predicting the WQI.

Finally, the dataset was partitioned into an 80% training set and a 20% testing set which was used to train the ML models.

a)A Model training and evaluation: The training phase involved utilizing the training set, which comprised 80% of the collected data, to train various classification ML models.

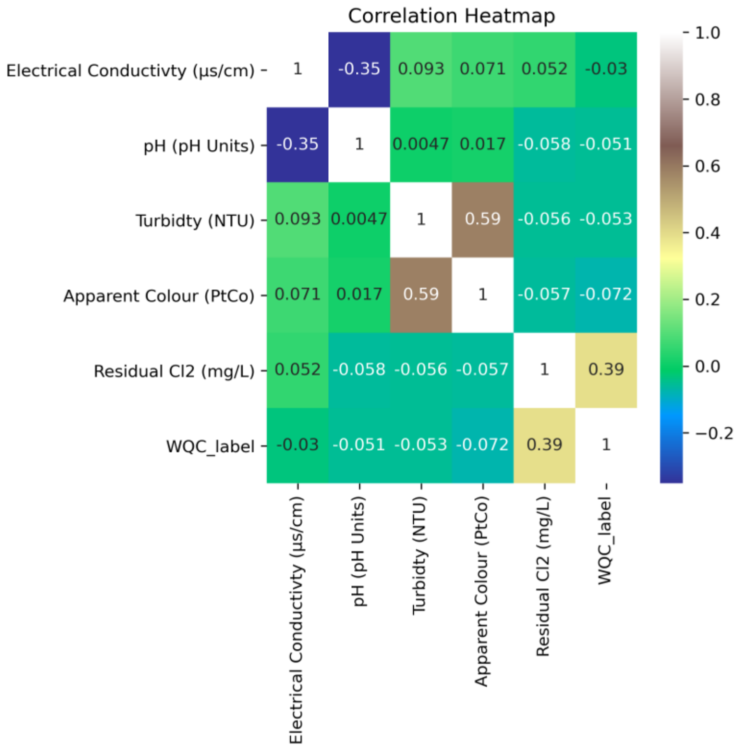

Fig. 2. Correlation of the different water quality parameters with WQC

TABLE III
PERFORMANCE OF ML MODELS

ML MODEL	ACCURACY	PRECISION	RECALL
Random Forest	89.2%	85%	88%
Gradient Boosting	89.20%	83%	86%
Extra Trees	88.63%	86%	85%
Decision Tree	86.93%	87%	87%

Additionally, to assess the performance of the trained models, the remaining 20% of the data, known as the testing set, was used for evaluation. The models were evaluated based on four metrics: accuracy score, precision, recall, and F1 score, to determine their effectiveness in predicting the WQC. The results of model evaluation are shown in Table 3. Notably, the Random Forest Classifier exhibited the best performance, achieving an accuracy score of 89%, precision of 85%, recall of 88%, and an F1-score of 88%.

Designing the water quality subsystem prototype: The basic system architecture of the water quality subsystem is a range of components dedicated to seamless data collection and transmission. In this comprehensive setup, the turbidity sensor and pH sensor play pivotal roles in acquiring essential water quality data. Once collected, the data is efficiently transmitted to the water consumption subsystem via a radio frequency (RF) module. Furthermore, the integration of a Global System for Mobile Communications (GSM) module is utilized to efficiently upload this critical data to a remote cloud server, facilitating accessibility and data analysis. Lastly, to provide an intuitive user interface, a Liquid Crystal Display (LCD) is used to offer a visual representation of the real-time water quality information. The equivalent schematic diagram derived from the system architecture is presented in Fig. 3

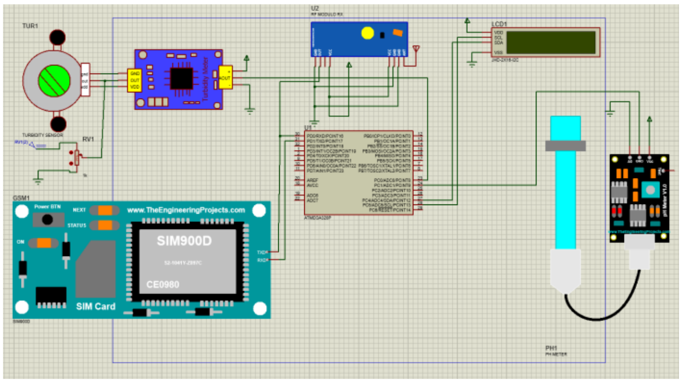

Fig. 3. Water quality system schematic diagram

B. Water Consumption Monitoring subsystem

The water consumption monitoring subsystem was developed following a similar approach to the water quality assessment subsystem.The design incorporates a water flow sensor for collecting consumption data, an RF module for transmitting the consumption data to the water quality assessment subsystem, an SD card module for storing the data collected from all the sensors, and a Real-Time Clock (RTC) module for providing the date and time of data collection. Additionally, an LCD is included to offer a visual representation of the water consumption in liters. Furthermore, the system architecture allows for scalability, enabling additional sensors to be integrated for measuring other water quality parameters. The equivalent schematic diagram derived from the water consumption architecture is presented in Fig.4

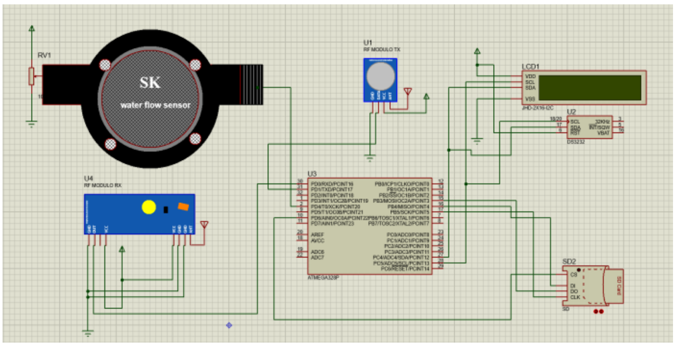

Fig. 4. Water consumption system schematic diagram

IV. RESULTS

This section discusses the results obtained from testing the system. The tests carried out included:

- Measurement of water consumption and water quality.
- Transmission of data between the two systems.
- Storage of data on the SD card.
- Uploading of data and receipt of email alerts from a remote Thing Speak cloud server.

Measurement of water consumption and water quality

The water consumption test was conducted to evaluate the accuracy of the water consumption subsystem by measuring the volume of water in liters (L). According to the results, which are depicted in Fig. 8, the water consumption was recorded as 0.91 L on the LCD screen. Similarly, the water quality test aimed to assess the accuracy of the water quality assessment subsystem. This assessment relied on the compliance of both pH and turbidity values with national standards for potable water, which specify that pH should range between 5.5 and 8.5, and turbidity should be between 0 and 5 [4]. Accordingly, water quality is classified as eithergood or bad based on its adherence to these thresholds. Figs. 5 and 6 illustrate the classification of water quality according to these criteria, with the results conveniently displayed on on an LCD screen

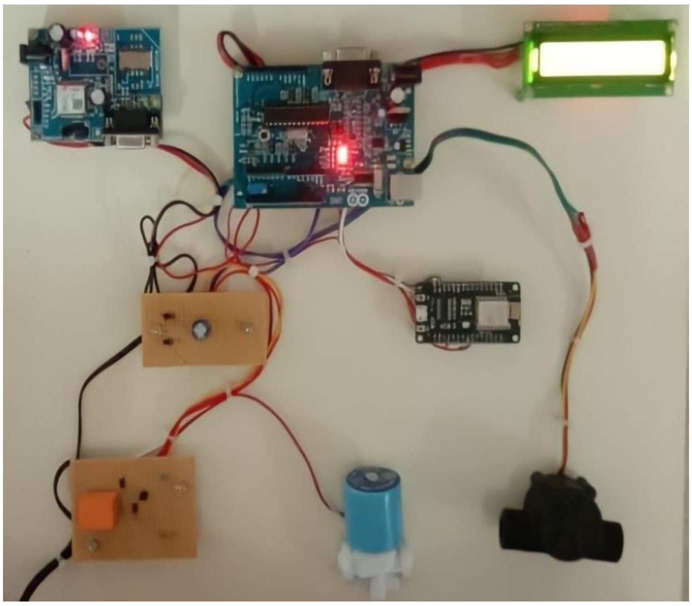

Fig 5 Hardware implementation

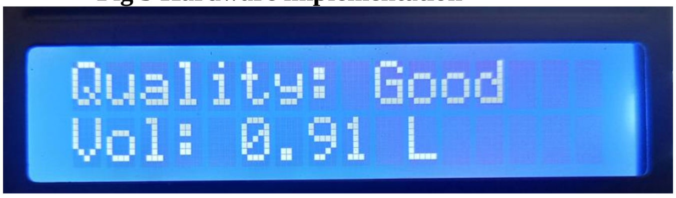

Fig. 6. LCD display showing that water quality is "Good".

V. CONCLUSION

In conclusion, it can be observed that the development of a combined water quality assessment and consumption monitoring system yields significant advantages in water management. This system enables precise real-time monitoring of both water quality and consumption, thereby ensuring the safety and reliability of the water supply. Through the integration of low-cost sensors, the system allows for cost-effective measurement and monitoring of various water parameters. Additionally, the utilization of the Thing Speak server and SD card storage facilitates remote monitoring and on-site access to data, thereby promoting efficient management and analysis. The incorporation of email alerts enhances response times, enabling timely actions to be taken in addressing water quality issues. Furthermore, graphical data representation assists in on-site fault investigation, reducing the necessity for physical assessment and enhancing customer safety An essential aspect of the system involves the application of feature scaling ML techniques to ascertain key parameters for water quality assessment. Through this approach, efficient evaluation of the water's chemical and physical characteristics is achieved, thereby enabling targeted investigation of water quality parameters in the event of pollution. Our assessment indicated that the most dominant parameters include residual chlorine, pH, turbidity, conductivity, and apparent color. In summary, this combined system makes a substantial contribution to efficient water management, highlighting the significance of integrating multiple functionalities and utilizing ML techniques for accurate assessment and monitoring.

REFERENCES

[1] K. Khatri, K. Vairavamoorthy, and M. Porto, "Challenges for urban water supply and sanitation in developing countries," in Water for a Changing World-Developing Local Knowledge and Capacity. CRC Press, 2008, pp. 93–112

[2] T. Boyle, D. Giurco, P. Mukheibir, A. Liu, C. Moy, S. White, and R. Stewart, "Intelligent metering for urban water: A review," Water, vol. 5, no. 3, pp. 1052–1081, 2013.

[3] M. Robert Biti and A. N. Gisele, "An online web based e-billing water system: A case study national water and sewerage corporation (nwsc)," 2012.

[4] S. Li, L. D. Xu, and S. Zhao, "The internet of things: a survey, " Information systems frontiers, vol. 17, pp. 243–259, 2015.

[5] A. Ghasempour, "Internet of things in smart grid: Architecture, applications, services, key technologies, and challenges," Inventions, vol. 22, March 2019.

[6] S. Nayebare, L. Wilson, D. Carpenter, and D. Dziewulski, "A review of potable water accessibility and sustainability issues in developing countries - case study of uganda," Reviews on environmental health, vol. 0, 06 2014.

[7] J. Pritt and J. W. Raese, Quality Assurance/quality Control Manual: National Water Quality Laboratory. US Geological Survey, 1995, vol. 95, no. 443.

[8] V. Lakshmikantha, A. Hiriyannagowda, A. Manjunath, A. Patted, J. Basavaiah, and A. A. Anthony, "Iot based smart water quality monitoring system," Global Transitions Proceedings, vol. 2, no. 2, pp.181–186, 2021.