CHEMICAL ENGINEERING TRANSACTIONS

VOL. 113, 2024

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.cetjournal.it

Guest Editors: Jeng Shiun Lim, Bohong Wang, Guo Ren Mong, Petar S. Varbanov
Copyright © 2024, AIDIC Servizi S.r.l.
ISBN 979-12-81206-14-4; **ISSN** 2283-9216

DOI: 10.3303/CET24113028

Integrated Internet of Things and AI for Real-Time Water Quality Monitoring in Laguna Lake

Jonalyn G. Ebron*, Gabriel T. Rivero, Joshua Victor R. Ang

Mapúa Malayan Colleges Laguna, College of Computer and Information Science
jgebron@mcl.edu.ph

This study presents a low-cost IoT-based real-time water quality monitoring in Laguna Lake. By strategically deploying sensors and leveraging Artificial Intelligence, the systems provide accurate temperature, pH, and ORP data, enabling early detection of water quality issues. The inclusion of ORP, often overlooked due to cost and complexity, distinguishes this research. Artificial Intelligence optimizes sensor performance and provides timely alerts. A user-friendly web application facilitates data visualization and stakeholder engagement. Key findings include alarmingly low ORP levels (Average of 207.52 mV), indicating potential water contamination and low oxygen levels. This integrated approach empowers decision-makers with actionable insights for sustainable water resource management and protecting Laguna Lake's ecosystem. The system contributes to UN SDG 6: Clean Water and Sanitation by providing a cost-effective and efficient solution.

1. Introduction

Monitoring water quality is essential for protecting public health and ecosystems. Almetwally et al. (2020) proposed practical solutions to address global challenges and ensure sustainable water resource management. The 2030 UN Agenda and Sustainable Development Goals (SDGs) prioritize water quality issues by providing a cost-effective and efficient solution for clean water. The Laguna Lake Development Authority (LLDA) was established in the Philippines to oversee the lake's well-being. However, the LLDA faced problems in the traditional collection, analysis, and visualization of data in the smaller streams or rivers that flow into a larger body of water, such as a lake or river. Integrating ecosystem health and water poverty underscores the importance of solving the problems in this area with the support of technology, as mentioned by Leones et al. (2023). Traditional monitoring methods, relying on manual sampling and laboratory analysis, often need to provide timely data for effective water quality management (Jan et al., 2021). Remote monitoring and IoT technologies are crucial for addressing this challenge (Wang et al., 2021). Real-time data access and integration, enabled by IoT technologies, is crucial for monitoring the lake (DERN, 2021). Reliant on manual sampling and laboratory analysis, they often fail to provide the timely data essential for effective water quality management (Jan et al., 2021), necessitating innovative solutions using remote monitoring and IoT by real-time data access can be done (Wang et al., 2021). The study aimed to develop a low-cost IoT system for real-time water quality monitoring in the local region by providing real-time data, identifying trends, and enabling initiative-taking management by alerting the stakeholders. Anthropogenic activities, particularly wastewater dumping, significantly deteriorate wetlands by increasing organic matter in these ecosystems, as mentioned in Acosta et al. (2023). The report of LLDA (n.d.) found that technological advancements have emerged to overcome the limitations of traditional manual monitoring methods. IoT-based systems offer a promising solution by enabling continuous monitoring of water quality parameters. For example, in the case of the lake's South Point and Canal Rivers areas, IoT sensors can be strategically placed to monitor water quality changes caused by urban runoff or industrial discharges. The progress in the contamination of rivers, as highlighted by Mamani et al. (2023), shows that the increasing contamination of rivers poses a significant health risk to those who use their surface water. Hui et al. (2020) utilized AI techniques to revolutionize water quality management using IoT sensors. Deep learning models can predict water quality trends, detect anomalies, and identify potential pollution sources (Randrianiaina et al., 2019). Data management and visualization are essential for informed decision-making

Paper Received: 17 May 2024; Revised: 7 September 2024; Accepted: 10 November 2024
Please cite this article as: Ebron J.G., Rivero G.T., Ang J.V.R., 2024, Integrated Internet of Things and AI for Real-Time Water Quality Monitoring in Laguna Lake, Chemical Engineering Transactions, 113, 163-168 DOI:10.3303/CET24113028

(Hui et al., 2020). GIS-based platforms can represent water quality data, facilitating analysis and identifying pollution hotspots (Boubakri et al., 2017). While previous research has shown that IoT-based systems are suitable for real-time water quality monitoring, there is a need to combine IoT and AI to improve how sensors work, analyze data, and show information in a way that is easy to understand (Chowdhury et al., 2023). This study aims to solve this gap by designing a low-cost IoT system that uses AI to combine data from different tributaries.

2. Materials and methods

2.1 **IoT Sensor Network**

The researcher selected the DFROBOT sensors because they are small, work well with Arduino, and are affordable and durable. Although these sensors offered a suitable balance of accuracy and affordability, with a unit cost of $50-$180, Integrating these sensors with Arduino boards and communication modules added to the overall system cost of $350, it remained significantly lower than comparable commercial monitoring systems. We calibrated the sensors to minimize costs according to the DENR AO 2016-08 Water Quality Guidelines and General Effluent Standards of 2016.

2.2 **Data Transmission Module**

Sensor data was transmitted to a cloud server using a secure wireless communication protocol. The chosen protocol considered data transmission rate, power consumption, and network coverage. An Arduino Uno R3 microcontroller was employed at each sensor node to preprocess data before transmission.

2.3 **Data Processing and Visualization Platform**

We transmitted the data to a cloud-based platform to manage and interpret water quality. We developed this centralized system to store, process, and visualize the collected water quality data. This centralized system incorporated the cloud incorporated Beebotte, a third-party tool, for initial data handling and transmission to an AWS server for advanced analytics. GIU presented real-time and historical water quality information, empowering stakeholders with data-driven insights for informed decision-making.

2.4 Data Collection

Data collection involved two primary sources: historical data and real-time sensor measurements. Historical water quality data from 2017 to 2022 was a comparison and model training baseline. Building upon the findings of Ebron et al. (2020), which identified three positively correlated water quality parameters, we focused on these parameters for our analysis. The researcher continuously collected real-time data from the deployed IoT sensors at the two selected stations. The researcher cleaned up the data to fix problems like missing information, incorrect values, and noise.

2.5 Data Analysis

The researcher employed cloud-based, incorporated AI-powered tools for data processing, pattern recognition, or anomaly detection. Descriptive statistics and correlation analysis using Rapid Miner and Python programming to assess parameter relationships. The research computed a Relative Weight Index (RWI) to assess water quality conditions comprehensively. This numerical value assigned to each parameter reflects its relative importance in overall water quality assessment. Parameters with higher RWIs contribute more significantly to the overall WQI. The researcher determines RWIs through expert opinion, statistical analysis, or both. RWI values usually range from 1 to 5, with 5 indicating the highest importance. RWI helps balance factors' influence on water quality assessment. Machine learning algorithms were also employed to develop predictive models for identifying patterns and trends Randrianiaina et al. (2019).

Table 1: Sensor Deployment and Data Collection for Prototype Testing

Location	Sensor ID	Deployment Date	Data Collection Period
South Point	Station 1	January 15, 2022	May 1, 2022 – May 14, 2023
Canal River	Station 2	Jan 15, 2022	May 1, 2022 – May 14, 2023

3. Results and Discussion

3.1 **IoT Sensor Network**

The map in Figure 1 shows the monitoring stations in Cabuyao City. We set up two IoT sensor stations at South Point and Canal River, following the advice of an expert engineer from LLDA. Each station had sensors to measure pH, temperature, and ORP. The system collected data every hour from May 1 to May 14, 2023.

(a) (b)

Figure 1: (a) Map of the study for Station 1, (b) Map of the study Station 2

Table 2: Sensor Deployment and Data Collection for Prototype Testing

Location	Sensor ID	Deployment Date	Data Collection Period
South Point	Station 1	January 15, 2022	May 1, 2022 – May 14, 2023
Canal River	Station 2	Jan 15, 2022	May 1, 2022 – May 14, 2023

3.2 Data Transmission Module

Table 2 shows that the DS18B20 temperature sensor, operating at a fixed rate of 16.3 kbps, demonstrated consistent performance with an average packet loss of 2 % under optimal conditions (distance of 10 meters, minimal interference). While this rate is moderate, the sensor's low power consumption of 20 mA is advantageous for battery-powered applications.

Table 3: Data Transmission Metrics

Parameters	Average	Max	Min	Data Pocket Loss Rate	Average Power Consumption
Temperature (DS18B20)	750 ms	16.3 kbps	750 ms	2 %	20mA
pH (SENO161V2)	9600 to 19200 baud	115200 baud	1200 baud	< 1%	100–200 mW
ORP	9600 to 19200 baud	115200 baud	1200 baud	< 1%	100-300 mW

ms – milliseconds *mW-* milliwatts

The study focused on monitoring the water quality intended for using freshwater or the Class C classification to grow fish and other aquatic resources based on the DENR-Water Quality Guidelines (WQG).

Table 4: DENR Water Quality Parameters

Parameter	DAO-2016-08 Standard (Class C)	Compliance
Temperature (°C)	25 – 31	Acceptable
pH	6.5 – 9.0	Acceptable
ORP (mv)	600 – 800 mV	Low value

Note:

- The samples were taken from 9:00 am – 4:00 pm
- The natural background temperature, as determined by EMB shall prevail if the temperature is lower or higher than the WQG, provided that the maximum increase is only up to 10 % and will not cause any risk to human health.

The notably low ORP value of 207.52, as indicated in Table 2, is significantly lower than the typical ORP range of 200-700 mV in freshwater environments. Low-value ORP suggests a pronounced reducing environment, potentially due to elevated organic matter, reducing substances, or stagnant conditions. A low ORP generally indicates a lack of oxygen and can be associated with poor water quality, anaerobic conditions, and the potential for harmful bacteria growth. Conversely, a high ORP often suggests a well-oxygenated environment, which is generally desirable but can be excessive, harming aquatic life. To fully comprehend the implications of this low

ORP, it is crucial to correlate it with other water quality parameters such as organic matter, nutrient, and bacterial levels. Examining surrounding land use and historical water quality data will provide valuable context.

3.3 Data Processing Visualization

The GUI features a dashboard that utilizes Google Maps to display the on-site location, providing a visual context for the data. Circular gauges display numerical values, such as pH, temperature, or ORP. Bar gauges are used to compare values, and dial gauges will visually represent a range of values. These interactive gauges allow users to hover over them for details, zoom in or out, and view real-time updates. Users can quickly identify trends or anomalies by observing changes in the gauges over time. Additionally, gauges can compare data from separate locations or periods, facilitating effective monitoring and analysis. The dashboard's functionality allows users to interpret and analyze the data quickly. It provides data filtering and customization, and collected data is integrated into reports to inform policy decision-makers. Users can filter data by time and view different water quality parameters individually or in combination. Furthermore, the dashboard allows users to download data in CSV or PDF format for further analysis or integration with other systems. The analytics page offers additional visualizations, such as scatter plots and boxplots, for in-depth data analysis. By providing these features, the dashboard effectively supports decision-making by offering a user-friendly interface for accessing and analyzing water quality data.

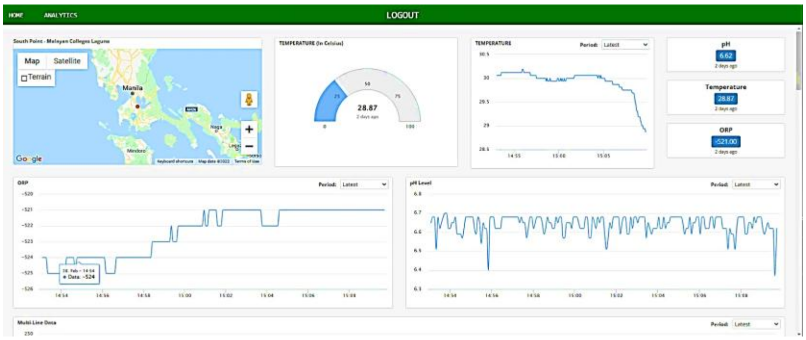


Figure 2: Analytics dashboard of the website

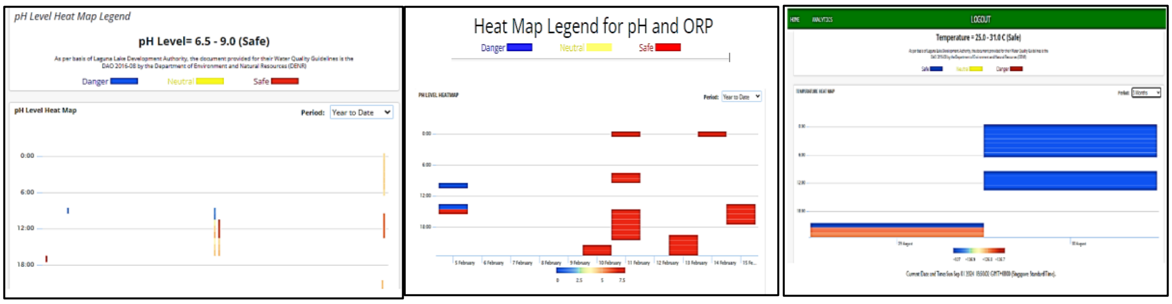


Figure 3: Heat Map for each parameter for decision-making

The heat maps in Figure 3 in the dashboard are integrated with Google Maps to show the location of different water quality measurements. By looking at these maps, users can easily see where the temperature, pH, and ORP levels are high or low to help identify areas with water pollution problems. Users can find connections between water quality and location by analyzing the heat maps and the time series data (Jan et al., 2021). By looking at these maps, users can easily see where the temperature, pH, and ORP levels are high or low to help identify areas with water pollution problems. Users can find connections between water quality and location by analyzing the heat maps and the time series data (Jan et al., 2021). For example, Station 2, with more factories or residents, has different water quality levels compared to areas with less human activity, like Station 2. Stakeholders can collect data to decide on ways to improve water quality in specific areas (Chowdhury et al., 2023).

3.4 Data Collection

The study utilized historical water quality data from LLDA from 2017 to 2022. This dataset, encompassing 712 data points from various monitoring stations, served as the training dataset for model development. The dataset included parameters such as pH, temperature, and other relevant water quality indicators. The researcher performed preprocessing to manage missing values and outliers. Data validation was conducted by cross-referencing LLDA's published reports and internal quality control procedures. Ebron et al. (2020) adapted data collection and analysis procedures to enhance the dataset.

3.5 Data Analysis

The provided Table 1 presents real-time water quality data collected from two (2) pilot stations.

Table 5: Station 1 and 2 Real-time Data Sample

Reading ID	ORP Reading	Ph Reading	temp	% Diff. from Reading Mean	Calibrated pH Reading	Date /Time	Timestamp in milliseconds
66431ce22e2dc747c42071b8218		6.98	25	-0.35 %	7.05	02/14/22 3:12 pm	1.71567E+12
66431cdd2e2dc747c42971a4218		7.06	25	0.80 %	7.13	02/14/22 3:12 pm	1.71567E+12
66431cd82e2dc747c42071a4218		6.98	25	-0.35 %	7.05	02/14/22 3:12 pm	1.71567E+12
66431cd82e2dc747c429718c218		7.06	25	0.80%	7.13	02/14/22 3:12 pm	1.71567E+12
66431cd32e2dc747c420716b218		6.98	25	-0.35 %	7.05	02/14/22 3:11 pm	1.71567E+12

Descriptive statistics summarize the collected data, including mean, median, standard deviation, and range from water quality parameters. Table 1 below shows that the water quality parameters analysis revealed that both stations' temperature and pH levels were within acceptable limits for Class C water bodies. However, the consistently low ORP values at both sites are cause for concern, as they indicate potential water quality issues. The researcher used the Relative Weight Index (RWI) method. This method calculates the water quality rating scale, relative weight, and overall WQI using the function formula below, suggested by an expert from LLDA. The relative weight per parameter is distributed equally to 100 % in the research; however, if more parameters participate in future studies, the percentage per parameter will differ.

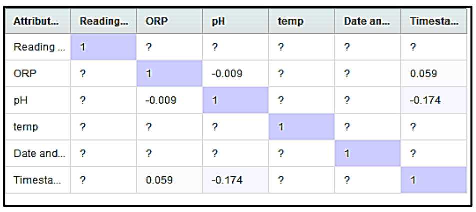
f (Overall relative weight of each parameter) = Average of the parameters) x (rating of each parameter) x (Assigned weight of each parameter)
Overall relative weights of all parameters = (relative weights of parameter 1) + (relative weights of parameter 2)+ (relative weights of parameter 3) (1)

Table 6: Descriptive Statistics Summary

Parameters		Mean	Min	Max	Standard (Class C)	Compliance
Temperature	Station 1	25	25	25	25 – 31	Acceptable
	Station 2	27	27	25		
pH	Station 1	7.004	6.87	7.12	6.6 -9.0	Acceptable
	Station 2	Station7.22	6.93 194	7.35		
ORP	1	211.53	220	227	600 -800 mV	NOT Acceptable
	Station 2	270.63		2952		

ms – milliseconds *mW*- milliwatts

To draw more conclusions, the provided correlation matrix offers preliminary insights into the relationships between temperature, pH, and ORP in the dataset. The researchers observed the following results below.



Attribut...	Reading...	ORP	pH	temp	Date an...	Timesta...
Reading ...	1	?	?	?	?	?
ORP	?	1	-0.009	?	?	0.059
pH	?	-0.009	1	?	?	-0.174
temp	?	?	?	1	?	?
Date and...	?	?	?	?	1	?
Timesta...	?	0.059	-0.174	?	?	1

Figure 4: Correlation Analysis of parameters

The correlation matrix shows a weak -0.009 connection between the ORP and pH levels. However, both parameters' levels change over time. The collected parameters provided valuable insights into potential relationships with algal blooms, even though not included in this study. Seasonal changes in pH levels may correlate with rainfall, temperature, and organic matter decomposition, impacting aquatic life and water quality. Rising temperatures can create favourable conditions for algal growth. Data analysts can analyze the system's data to identify potential correlations with algal bloom occurrences, as the DENR (2021) suggested. Additionally, low ORP levels, often indicating pollution, can be monitored to detect potential pollution and take appropriate action. If the system detects a significant drop in ORP levels, the water is not good at cleaning itself, indicating reduced dissolved oxygen. Local authorities or environmental agencies can immediately increase water aeration, investigate potential pollution sources, and issue public advisories. The system triggers alerts for industrial plants around the lake when ORP levels fall below a predetermined threshold. Local governments can use the data to inform decisions regarding water allocation, pollution control measures, and public advisories. Moreover, if the system detects elevated pollutant levels, authorities can implement restrictions on industrial discharges or initiate cleanup efforts. Additionally, the data can raise community awareness about water quality issues and engage them in conservation efforts through the dashboard for visualizing real-time and historical data. This dashboard will facilitate stakeholder communication. Local governments can leverage the data to adjust water allocation policies, implement pollution control measures, and issue public advisories. This integration will enable decision-makers to utilize the data effectively.

4. Conclusion

This research successfully developed and deployed a low-cost IoT-based real-time water quality monitoring system. The system effectively collected and transmitted data to a cloud-based platform. The web application underwent user acceptance testing with positive feedback from the stakeholders. The system's ability to notify relevant stakeholders, such as local authorities, environmental agencies, and nearby communities, about the potential water quality issue. In conclusion, ORP can be a valuable tool for indirectly estimating dissolved oxygen levels, especially when combined with pH and temperature. Future research should focus on expanding the sensor network and exploring advanced machine-learning techniques to enhance the system's capabilities. By addressing the challenges of traditional monitoring methods, the developed IoT-based system provides a foundation for initiative-taking and data-driven approaches to water quality management in Laguna Lake.

References

Acosta Suasnabar, E., Aquino Soto, M.L., Benites-Alfaro, E., Banon Arias, J.V., 2023, Environmental Nanotechnology: Carbon Nanotubes as an Alternative to Reduce Organic Matter in Wetlands, Chemical Engineering Transactions, 101, 103-108.

Almetwally, S.A.H., Hassan, M.K., Mourad, M.H., 2020, Real Time Internet of Things (IoT) Based Water Quality Management System, Procedia CIRP, 91, 478–485.

Boubakri, S., Rhinane, H., 2017, The Contribution of GIS to display and analyze the water quality data collected by a wireless sensor network: Case of Bouregreg Catchment, Morocco ISPRS, IV-4/W4. 331-334.

Chowdury, M.S.U., Emran, T.B., Ghosh, S., et al., 2019, IoT Based Real-time River Water Quality Monitoring System, Procedia Computer Science, 155, 161-168.

DENR., 2021, Department of Environment and Natural Resources <denr.gov.ph/priority-program/clean-water-program> accessed 16.05.2023

Ebron, J.G., de Leon, R.I.D., Alejandro, A.D., Amoranto, B.A., 2020, Computational and numerical modelling for classification of water quality of lake. International Journal of Environmental Science and Development, 11(9), 425–431.

Cham, H., Malek, S., Milow, P., Ramli, M.R., 2020, Web-based system for visualisation of water quality index, All Life, 13(1), 426-432.

Jan, F., Min-Allah N., Düstegör D., 2021, IoT Based Smart Water Quality Monitoring: Recent Techniques, Trends and Challenges for Domestic Applications, Water, 13(13), 1729.

Leones, J., Restrepo, J., Velandia, K., Sanchez, E., Ojeda, K., 2023, Sustainability Assessment of Water Ecosystems Based on Modified Water Poverty Index (WPI) with Birds Biodiversity Parameters. Case Study: Juan Angola Creek- Cartagena, Colombia, Chemical Engineering Transactions, 98, 147-152.

LLDA (n.d.) Laguna de Bay and its Tributaries <llda.Gov.ph/ldb-and-its-tributaries> accessed 06.22.2022

Mamani Gómez, M., Lizarzaburu Aguinaga, D., Castaneda-Olivera, C.A., Ramirez Paredes, M., Tello Mendivil, V., Benites-Alfaro, E., 2023, Reduction of Organic Matter and Bacteria in Surface Waters of the Huatanay River using Chitosan Nanoparticles, Chemical Engineering Transactions, 101, 67-72.

Randrianiaina Jerry J.C.F., Rakotonirina Rija I., Ratiarimanana Jean R., Lahatra Razafindramisa Fils., 2019, Modelling of Lake Water Quality Parameters by Deep Learning Using Remote Sensing Data, American Journal of Geographic Information System, 8(6), 221-227.

Wang, Y., Ho, I. W.H., Chen, Y., Wang, Y., Lin, Y., 2022, Real-time Water Quality Monitoring and Estimation in IoT for Freshwater Biodiversity Conservation, IEEE Internet of Things Journal, 9(16), 14366 – 14374.