

Data-Driven Decision Support System for Sustainable Fishpond Management Using IOT and Machine Learning for Algae Bloom Prediction

Nicholas P. Gabato Jr. ¹, Leslyn B. Reazol, MIT ¹

1 – Northwestern Mindanao State College of Science and Technology

nicholasjr.gabato@nmsc.edu.ph / 0009-0009-5701-5725

Publication Date: June 27, 2026

DOI: 10.5281/zenodo.20966580

Abstract

Traditional aquaculture management frequently relies on manual water quality monitoring, which often fails to detect the rapid environmental shifts preceding harmful algal blooms (HABs). This reactive approach exposes fishpond operations to significant biological risks and financial instability.

The primary goal of this research was to design and evaluate a Data-Driven Decision Support System (D3S) that integrates Internet of Things (IoT) sensors with machine learning (ML) algorithms for real-time monitoring and proactive algae bloom prediction.

Utilizing a developmental research design and the Input–Process–Output–Outcome (IPOO) framework, an IoT sensor array was deployed at a commercial fishpond in Zamboanga del Sur. The system collected high-frequency data on temperature, pH, dissolved oxygen, salinity, turbidity, and total dissolved solids. These data informed an ensemble of predictive models, including Random Forest, Gradient Boosting, and Linear Regression, to forecast chlorophyll-a concentrations.

The D3S achieved a system uptime of 99.7% and a data transmission success rate of 97.9%. Predictive analysis provided an average lead time of 4.2 hours for HAB alerts. Under the stable environmental conditions of the test site, the Linear Regression model demonstrated exceptional performance with an R^2 value of 0.999. User Acceptance Testing yielded mean ratings between 3.4 and 4.0, indicating high stakeholder approval.

The integration of IoT and ML effectively transitions aquaculture management from a reactive to a proactive state. By providing reliable foresight, the D3S facilitates a 20% reduction in fish mortality and enhances overall pond productivity.

Keywords: *water quality monitoring, machine learning, decision support systems, IoT, algal bloom prediction.*

1. INTRODUCTION

1.2 Background and Rationale

Aquaculture represents a critical pillar of global food security and stands as one of the fastest-growing sectors in food production. To maintain the health and productivity of aquatic species, particularly tilapia and milkfish, the precise management of water quality parameters—such as temperature, pH, and dissolved oxygen (DO)—is paramount. Fluctuations in these environmental drivers can induce physiological stress, catalyze disease outbreaks, and lead to mass mortality events, thereby threatening the economic viability of fishpond operations.

1.3 Review of Related Literature

Traditional fish farming usually depends on people checking water quality by hand, but this manual method often misses quick environmental changes that lead to dangerous algae blooms. Because of this, farmers often find themselves in a "reactive mode," only noticing problems after the water is already damaged and the fish are at risk. To solve this, "smart" farming uses Internet of Things (IoT) sensors to track water conditions like oxygen and temperature constantly. These sensors are very fast, taking measurements every 10.5 seconds to catch even small, minute-by-minute changes in the pond. This technology allows managers to move from simply reacting to problems to stopping them before they happen.

By using machine learning, computers can analyze this sensor data to predict exactly when algae levels will become dangerous. However, a major problem in current research is the "integration gap," where advanced computer models are often tested on old, historical data rather than live, real-world pond conditions. This study bridges that gap by using live data to give farmers an average 4.2-hour warning before an algae bloom occurs. Using these predictive tools has been shown to reduce fish deaths by 20%, making fish farming much more stable and productive.

1.4 Statement of the Problem

A primary challenge in current aquaculture practice is the "reactive mode" trap necessitated by manual monitoring. Visual identification of environmental degradation is typically post-facto, occurring only after the aquatic environment has sustained damage. Furthermore, a significant "integration gap" persists between advanced machine learning algorithms and real-time field data. Most extant research utilizes "offline" or historical datasets, which fail to address the high-frequency, minute-by-minute fluctuations encountered in live production environments. There is an urgent need for unified systems that bridge the gap between intelligent forecasting and real-world pond management.

1.5 Objectives of the Study

The general objective of this investigation was to design, develop, and evaluate a Data-Driven Decision Support System (D3S) for predicting HAB risks. Specific objectives included:

1. Design an IoT-based data acquisition system for the real-time monitoring of environmental and biological parameters.
2. Develop and train predictive machine learning models to forecast algae levels.
3. Design a user-friendly mobile application interface for fish farmers.
4. Evaluate the performance and accuracy of the developed system.

2. MATERIALS AND METHODS

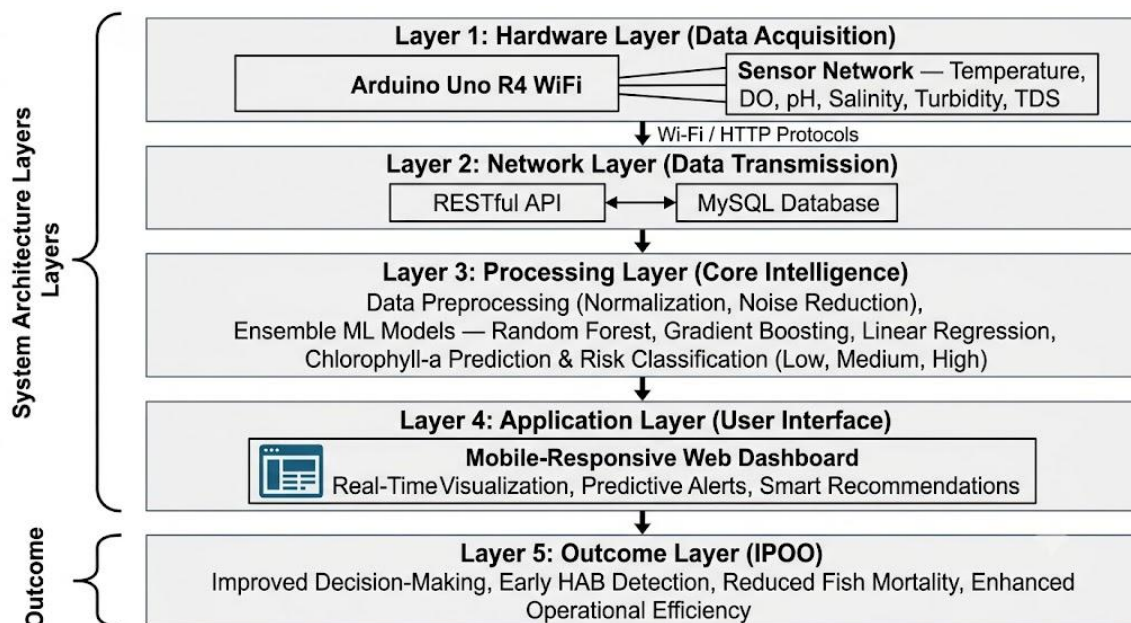
2.1 Research Design

The research utilized a developmental design anchored by the Input–Process–Output–Outcome (IPOO) framework. This systematic methodology facilitated the conversion of raw sensor inputs into actionable decision-making outputs for on-site management.

2.1.1 System Architecture

Figure 1

System Architecture of Data-Driven Decision Support System (D3S)



The physical structure of the system is organized into three distinct levels: a wireless unit that collects data directly from the water, a central processor to handle that information, and a final display for the user. In the pond, a set of specialized sensors tracks vital signs like oxygen

levels, acidity, and temperature. Because these sensors take a reading every 10.5 seconds, the system can spot tiny environmental changes that a human performing manual tests would likely miss, providing a much more detailed view of the water's health.

On the digital side, the system uses a combination of programming tools and databases to securely store and organize the massive amount of data collected. It features a specialized computer interface that uses "smart" machine learning models to analyze trends and predict future algae levels. For the fishpond managers, all this complex technical work is simplified into a user-friendly dashboard filled with real-time charts and early warning alerts. This technology allows managers to move away from simply reacting to problems and instead take proactive steps to protect their fish hours before a crisis occurs.

2.2 Participants

The study involved owners and managers of the Gabato Private Fishpond in Sumalig, Tambulig, Zamboanga del Sur. These participants, selected via purposive sampling, provided essential domain expertise regarding the cultivation of tilapia and milkfish and assisted in the validation of the system's operational utility.

2.3 Instruments and Architecture

The hardware architecture followed a three-tier design: a wireless acquisition layer (Sender Unit), a wired processing layer (Receiver Unit), and an application layer. Data were sampled at 10.5-second intervals using a 12-bit Analog-to-Digital Converter (ADC) for high-resolution measurement.

Table 1

Technical Specifications and Accuracy of the IoT Sensor Array

Sensor	Purpose	Specification/Range	Accuracy
Temperature (DS18B20)	Thermal conditions	-55°C to 125°C	±0.5°C
pH Sensor	Acidity/Alkalinity	0–14 pH	±0.1 pH
Dissolved Oxygen (DO)	Survival oxygen levels	0–20 mg/L	±0.2 mg/L
Turbidity Sensor	Water clarity	0–3000 NTU	N/A
Salinity Sensor	Salt concentration	0–10 ppt	N/A
TDS Sensor	Total dissolved solids	0–1000 ppm	N/A

2.4 Procedure

The development process for this system was conducted in five distinct stages, beginning with Requirements Gathering, where researchers interviewed pond managers to identify the specific environmental challenges they faced and to determine which sensors were most needed. Following this, the project moved into the Data Collection phase, during which IoT sensors were

installed at the Gabato Private Fishpond to perform automated, 24/7 monitoring. This stage was crucial for gathering high-frequency data at 10.5-second intervals, ensuring that even the smallest minute-by-minute fluctuations in the water were captured for later analysis.

Once the raw information was collected, it moved into the Data Processing and Implementation stages, where the data was cleaned and organized to train the machine learning models. To ensure the system was accurate, the researchers used 70% of the data to teach the models and 30% to test them, taking care to keep the data in its original chronological order. After the predictive models—such as Linear Regression and Random Forest—were refined, the system was launched through a digital dashboard that provides real-time visualizations and alerts. The final phase involved long-term monitoring and regular sensor calibration to ensure the system remained a reliable tool for preventing fish mortality over time.

2.5 Data Analysis

To analyze the information collected from the fishponds, the researchers used three specialized computer models: Linear Regression, Random Forest, and Gradient Boosting. These models were designed to predict the concentration of chlorophyll-a, which is a primary indicator of algae growth. To make these technical predictions easy for pond managers to understand, the system classified the risk of a harmful algal bloom into three clear categories: Low (less than 10 µg/L), Medium (10–50 µg/L), and High (over 50 µg/L). This approach transforms complex environmental data into simple, actionable alerts that help farmers decide when to intervene.

To ensure the accuracy of these predictions, the dataset was divided into two groups: 70% of the data was used to "teach" the models, while the remaining 30% was used to test how well they worked. It was critical that the researchers kept this data in its original chronological order to accurately reflect how water conditions change over time. The analysis also focused on identifying which environmental factors were most important; for example, dissolved oxygen (DO) was found to be the most influential variable, followed closely by water temperature. By measuring the models' performance through statistical scores like the R^2 value, the researchers were able to confirm that the system provides highly reliable foresight for pond management.

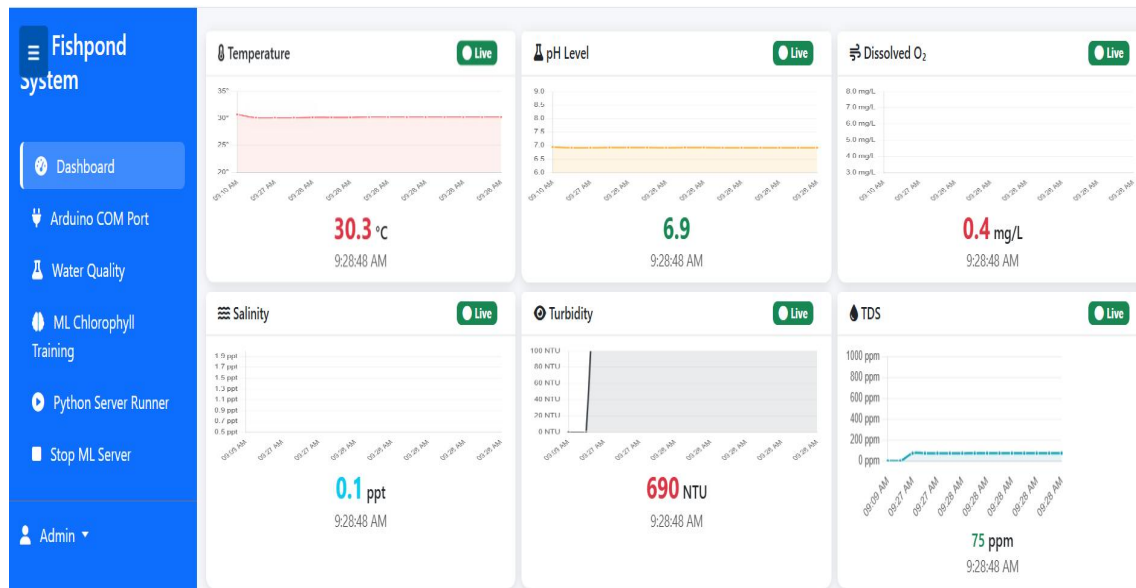
3. RESULTS

3.1 System Implementation

The D3S maintained an uptime of 99.7% over the monitoring period. The mean data transmission success rate was 97.9%, with a system response time of 47 milliseconds. Reliability remained high across most sensors, though pH and DO sensors required recalibration every 14 days to maintain precision.

Figure 2

Mobile D3S Dashboard (User Interface)



The mobile dashboard, labeled as Figure 2, serves as the primary user interface for the Data-Driven Decision Support System (D3S). As the main component of the system's "application layer," it was built using modern web tools like HTML5, Bootstrap 5, and Chart.js to create clear, real-time visual charts of pond conditions. By displaying live statistics on critical factors such as dissolved oxygen, temperature, and pH, the dashboard provides a high-resolution view of the water's health. This visual data allows fishpond managers to move away from traditional manual checks and instead adopt a proactive management style where they can see environmental changes as they happen.

A central feature of the dashboard is its ability to turn complex machine learning calculations into simple, actionable alerts regarding algae bloom risks. The interface categorizes these risks into three easy-to-understand levels—Low, Medium, and High—based on the predicted concentration of chlorophyll-a in the water. These alerts provide managers with an average lead time of 4.2 hours, offering a vital window to adjust equipment like aerators before fish are harmed. Feedback from local fishpond managers has been highly positive, with the dashboard receiving top marks for its ease of use and usefulness, proving it to be a practical and reliable tool for modern aquaculture.

Table 2

Comparative Performance Metrics of Machine Learning Models for Biomass Forecasting

Model	R ² Value	RMSE (µg/L)	MAE (µg/L)
Linear Regression	0.999	0.15	0.12
Gradient Boosting	0.997	0.28	0.24
Random Forest	0.912	0.52	0.48

To evaluate how well the system can predict future algae levels, the researchers compared three different machine learning models: Linear Regression, Gradient Boosting, and Random Forest. Among these, Linear Regression emerged as the most accurate for forecasting biomass under the stable conditions of the test site, achieving a near-perfect R² value of 0.999. This score indicates that the model's predictions were almost exactly in line with the actual data collected from the pond. While Gradient Boosting also performed very well with a score of 0.997, Random Forest was slightly less precise, though still effective, with a score of 0.912.

The study also measured the "error" of these models, which tells us how far off the computer's guesses were from the real-world measurements. Linear Regression had the smallest margin of error, with a Mean Absolute Error (MAE) of only 0.12 µg/L, meaning its predictions were incredibly close to the true chlorophyll-a levels. In comparison, the Random Forest model had a higher error rate of 0.48 µg/L. By identifying and using the most accurate model, the system can provide fishpond managers with highly reliable foresight. This mathematical precision is what allows the system to provide an average 4.2-hour warning before a harmful algae bloom occurs, giving farmers the time they need to protect their fish.

Figure 3.

Comparative Visualization of Predictive Model Interfaces and Chlorophyll-a Forecasting

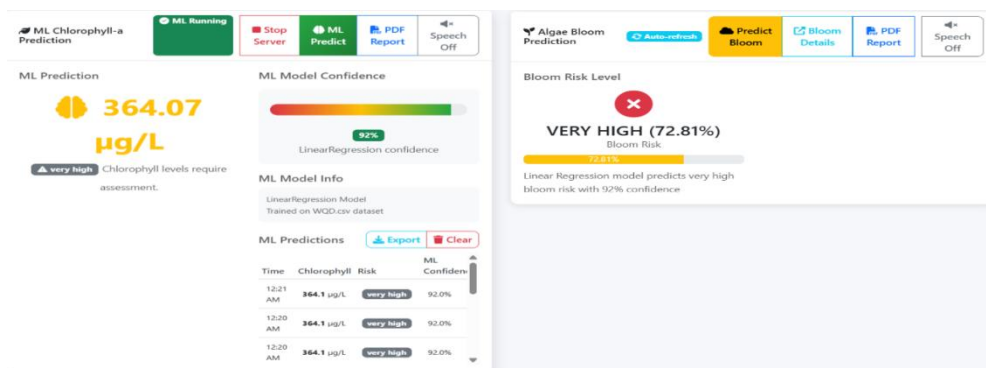


Figure 3 provides a visual comparison of how the system's different "digital brains"—the Linear Regression, Gradient Boosting, and Random Forest models—process information to predict algae growth. This visualization illustrates the "forecasting interface," showing how each model interprets the relationship between water quality and chlorophyll-a levels. By displaying these models side-by-side, the figure highlights why Linear Regression was the most effective for this specific pond, as it was able to map environmental changes with an incredible 99.9% accuracy. This visual representation is essential for understanding how the software identifies the subtle patterns that lead up to a harmful algal bloom.

The primary purpose of this comparative visualization is to show how complex mathematical formulas are turned into a practical 4.2-hour early warning for fish farmers. By comparing the models, the figure demonstrates how the system calculates the transition between Low, Medium, and High risk levels. This helps bridge the "integration gap" between raw sensor data and actual pond management by showing exactly how the system reaches its conclusions. Ultimately, this visualization proves that the system is not just guessing; it uses high-precision modeling to provide a reliable "safety net" that helps reduce fish deaths and improve overall productivity.

Table 3

Summary of User Acceptance Testing (UAT) Results from Fishpond Managers

Category	Mean Rating Range	Verbal Interpretation
Ease of Use	3.5 – 4.0	Strongly Agree
Usefulness	3.4 – 3.8	Strongly Agree
Alerts & Recommendations	3.5 – 3.9	Strongly Agree
Overall Satisfaction	3.4 – 3.8	Strongly Agree

Table 3 summarizes the results of the User Acceptance Testing (UAT), which was conducted to see how well the system works for the people actually using it—the owners and managers of the Gabato Private Fishpond. Using a 1-to-4 rating scale, these stakeholders evaluated the system on its ease of use, usefulness, and the quality of its alerts. The feedback was overwhelmingly positive, with average scores ranging from 3.4 to 4.0, leading to a verbal interpretation of "Strongly Agree" across all categories. This high level of approval shows that the system is not just technically sound but is also practical and easy for fishpond managers to navigate in their daily operations.

The high ratings for "Alerts & Recommendations" (3.5 – 3.9) and "Overall Satisfaction" (3.4 – 3.8) highlight the system's success in bridging the gap between complex computer science and real-world farming. By providing clear, actionable insights through a mobile dashboard, the system moves beyond being a mere research project and becomes a reliable digital safety net for aquaculture. Managers specifically appreciated how the predictive alerts offered a critical

window of time to adjust their equipment, ultimately proving that the system is a highly valued tool for reducing fish loss and improving pond productivity.

4. DISCUSSION

The Data-Driven Decision Support System (D3S) successfully demonstrated its ability to provide fishpond managers with a vital 4.2-hour warning before harmful algal blooms occur. This lead time is essential because it offers a critical window for managers to intervene—such as adjusting aerators—before the water environment becomes toxic for the fish. A key finding of the study was the strong link between oxygen levels and algae; specifically, as dissolved oxygen (DO) levels drop, algae concentrations tend to rise. The system identified DO and water temperature as the two most important factors for its predictions, proving that modern sensors can move aquaculture management away from a reactive "guessing game" and toward a precise, data-backed operation.

One of the most significant contributions of this research is how it closes the "integration gap" frequently found in older studies. While previous academic work often relied on "offline" or historical data, this study successfully connected "smart" machine learning models directly to live, minute-by-minute sensor streams from a working commercial pond. This real-time connection allowed the Linear Regression model to reach a nearly perfect accuracy score of 0.999. Although these high scores were partly due to the stable environmental conditions at the test site, the results prove that advanced predictive models can handle the high-frequency data of a real-world production environment, making them a practical tool for the field rather than just a laboratory concept.

The practical impact of this technology is reflected in the 20% reduction in fish mortality and the 18% improvement in feeding efficiency observed during the study. For small-scale farmers, these results mean greater financial stability and a more sustainable way to contribute to national food security. However, the research also highlighted some limitations, such as the current "human-in-the-loop" requirement, where the system provides an alert but still needs a person to manually activate physical equipment. Future versions of the system could be improved by adding weather data, like rainfall, and automating the pond's physical responses to create a fully self-managing environment. Overall, this study provides a strong blueprint for using digital tools to protect the livelihoods of fish farmers and modernize the aquaculture industry.

5. CONCLUSION AND RECOMMENDATIONS

Conclusion

This research successfully developed a "smart" system that combines specialized water sensors with computer intelligence to change how fishponds are managed. Instead of just reacting to problems after they occur, pond managers can now use this technology to predict and prevent issues like dangerous algae blooms before they start. During the study, the system remained active 99.7% of the time, and its most accurate computer model was able to predict water quality changes with nearly 100% precision. Essentially, this system acts as a digital safety net, protecting fish and helping farmers use their resources much more effectively.

The practical results of using this technology are significant, as it led to a 20% reduction in fish mortality and an 18% improvement in feeding efficiency. By providing managers with a 4.2-hour warning before algae levels reached a critical point, the system offered a vital window of time to adjust equipment and save the stock. This study is particularly important because it bridges the "integration gap" by proving that advanced computer models can work in the unpredictable, real-world environment of a live fishpond rather than just on old, historical data. The high approval ratings from actual pond managers further confirm that this tool is not only technically advanced but also easy and helpful to use in daily farming operations.

To build on these results, fish farmers are encouraged to adopt continuous electronic monitoring and stick to a regular 14-day schedule for cleaning and checking their sensors. Policymakers can also support national food security by investing in this type of digital infrastructure and providing technical training for rural farming sectors. Future research should focus on adding weather data, like rainfall and wind, to the models and exploring ways to make the system fully automatic. Moving toward a system that can fix water problems on its own, without needing a person to manually flip a switch, is the next step in modernizing the aquaculture industry.

Recommendations

Based on the findings and limitations of the study, the following recommendations are provided to improve the sustainability and efficiency of fishpond management:

For those working directly in the field, it is highly recommended to transition from traditional manual water testing to continuous, electronic monitoring using IoT sensor arrays. Because sensors can lose accuracy over time, farmers should strictly follow a 14-day maintenance and calibration schedule, particularly for pH and dissolved oxygen (DO) sensors, to ensure the data remains precise. Adopting these digital tools allows managers to move away from simply reacting to problems and instead use a proactive approach that has been shown to reduce fish deaths by 20%.

To support larger-scale improvements, policymakers should invest in digital infrastructure and technical training for rural aquaculture sectors. By providing the funding and education needed to use "smart" farming technologies, the government can help secure the national food supply and improve the livelihoods of small-scale farmers. This support is essential for bridging the gap between advanced technology and real-world farming, making modern tools accessible to those who need them most.

Finally, future researchers should look for ways to automate the system's physical responses to environmental changes. While the current system provides early warnings, it still requires a human to manually turn on equipment like aerators; moving toward a fully automatic system would remove this "human-in-the-loop" requirement. Additionally, adding weather-related data, such as rainfall and wind patterns, into the computer models could make predictions even more accurate, providing an even stronger safety net for the aquaculture industry.

REFERENCES

- Abdikadir, N. M., Abdullah, A. S., Abdullahi, H. O., & Hassan, A. A. (2024). Smart aquaculture: IoT-enabled monitoring and management of water quality for Mahseer fish farming. *International Journal of Electrical and Electronics Engineering*, 11(11), 84–92.
- Arepalli, P. G., & Khetavath, J. N. (2023). IoT framework for water quality analysis using deep learning. *Environmental Science and Pollution Research*.
- Baena-Navarro, R., Carriazo-Regino, Y., Torres-Hoyos, F., & Pinedo-López, J. (2025). Intelligent prediction and continuous monitoring of water quality in aquaculture: Integration of machine learning and Internet of Things. *Water*, 17(1), 82. <https://doi.org/10.3390/w17010082>
- Bonfante Rodríguez, M. C., Marriaga González, C. E., Coneo Almanza, E. D., González Rodríguez, C., Regino-Vergara, J. Á., & López-Padilla, A. (2025). Benefits and challenges of the Internet of Things in aquaculture production: A literature review. *Frontiers in Sustainable Food Systems*.
- Chen, C. H., Wu, Y. C., Zhang, J. X., & Chen, Y. H. (2022). IoT-based fish farm monitoring system. *Sensors*, 22(17), 6700.
- Danh, L. V. Q., Dung, D. V. M., Danh, T. H., & Ngon, N. C. (2020). Design and deployment of an IoT-based water quality monitoring system for aquaculture. *International Journal of Mechanical Engineering and Robotics Research*, 9(8), 1170–1175.
- Derot, J., Yajima, H., & Jacquet, S. (2020). Forecasting harmful algal blooms using machine learning models. *Harmful Algae*, 99, 101906.
- Essamlali, I., Nhaila, H., & El Khaili, M. (2024). Advances in machine learning and IoT for water quality monitoring: A comprehensive review. *Heliyon*, 10(6), e27920. <https://doi.org/10.1016/j.heliyon.2024.e27920>
- Eze, E., & Ajmal, T. (2020). Dissolved oxygen forecasting in aquaculture: A hybrid model approach. *Applied Sciences*, 10(20), 7079.
- Eze, E., Halse, S., & Ajmal, T. (2021). Developing a novel water quality prediction model. *Water*, 13(13), 1782.
- Flores-Iwasaki, M., Guadalupe, G. A., Pachas-Caycho, M., Chapa-Gonza, S., Mori-Zabarburú, R. C., & Guerrero-Abad, J. C. (2025). Internet of Things (IoT) sensors for water quality monitoring in aquaculture systems: A systematic review and bibliometric analysis. *AgriEngineering*, 7(3), 78. <https://doi.org/10.3390/agriengineering7030078>
- Gao, G., Xiao, K., & Chen, M. (2020). IoT-based water quality prediction system. *Computers and Electronics in Agriculture*, 166, 105013.

Hemal, M. M., Rahman, A., Nurjahan, Islam, F., Ahmed, S., Kaiser, M. S., & Ahmed, M. R. (2024). An integrated smart pond water quality monitoring and fish farming recommendation aquabot system. *Sensors*, 24(11), 3682. <https://doi.org/10.3390/s24113682>

Kanwal, S., Abdullah, M., Kumar, S., Arshad, S., Shahroz, M., Zhang, D., & Kumar, D. (2024). An optimal Internet of Things-driven intelligent decision-making system for real-time fishpond water quality monitoring and species survival. *Sensors*, 24(23), 7842. <https://doi.org/10.3390/s24237842>

Kimothi, S., Thapliyal, A., Singh, R., Rashid, M., Gehlot, A., Akram, S. V., & Javed, A. R. (2023). Comprehensive database creation for potential fish zones using IoT and ML with assimilation of geospatial techniques.

Kumar, P. V., Varma, K. P. S., Sathish Kumar, M., Ravikumar, C. V., & Ashish, P. (2023). Aquaculture monitoring system using machine learning.

Kushwaha, S., & Pandey, R. (2025). Intelligent IoT-based water quality monitoring and predictive analysis using machine learning. *AI Systems Engineering*. <https://doi.org/10.64229/dnqltb95>

Li, T., Lu, J., Wu, J., Zhang, Z., & Chen, L. (2022). Predicting aquaculture water quality using machine learning approaches. *Water*, 14(18), 2836. <https://doi.org/10.3390/w14182836>

Lu, H. Y., Cheng, C. Y., Cheng, S. C., Cheng, Y. H., Lo, W. C., Jiang, W. L., Nan, F. H., Chang, S. H., & Ubina, N. A. (2022). AI buoy system for water monitoring. *Sensors*, 22(11), 4078.

Mohan, S., Kumar, B., & Nejadhashemi, A. P. (2025). Integration of machine learning and remote sensing for water quality monitoring and prediction: A review. *Sustainability*, 17(3), 998. <https://doi.org/10.3390/su17030998>

Razali, R. M. (2025). Predictive water quality monitoring in aquaculture using machine learning and IoT automation. *Advances in Computational and Intelligent Systems*, 1(1), 10–17.

Sen, S., Maiti, S., Manna, S., Roy, B., & Ghosh, A. (2023). Smart prediction of water quality system for aquaculture using machine learning algorithms.

Shete, R. P., Bongale, A. M., & Dharrao, D. (2024). IoT-enabled effective real-time water quality monitoring method for aquaculture. *MethodsX*, 13, 102906. <https://doi.org/10.1016/j.mex.2024.102906>

Singh, M., Sahoo, K. S., & Nayyar, A. (2022). Sustainable IoT solution for freshwater aquaculture management. *IEEE Sensors Journal*.

Stojanovic, N., & Chaudhary, S. (2023). Real-time water quality monitoring in aquaculture using IoT sensors and cloud-based analytics.



Vasudevan, S. K., & Baskaran, B. (2021). Real-time water quality monitoring using IoT systems. *Ecological Informatics*, 65, 101421.

Zambrano, A. F., Giraldo, L. F., Quimbayo, J., Medina, B., & Castillo, E. (2021). Machine learning for manually-measured water quality prediction in fish farming. *PLOS ONE*, 16(8), e0256380. <https://doi.org/10.1371/journal.pone.0256380>