

Internet of Things in Fish Ponds

Lecturer

Dr. Hager Saleh

Student

Ahmed Mohammed Alnahari

251003035

Introduction:

There are challenges facing aquaculture such as the availability of water, how to monitor and manage water quality, and more seriously, the problem of absence of dataset with which the farmer can use as a guide for fish breeding. Aquaponics is a system that combines conventional aquaculture with hydroponics (the method of growing plants in water i.e. soilless farming of crops). It uses these two technologies in a symbiotic combination in which the plant uses the waste from the fish as food while at the same time filtering the water for immediate re-use by the fish. This helps to solve the problem of frequent change of water. An Internet of Things (IOT) system consisting of an ESP-32 microcontroller which controls water quality sensors in aquaponics fish ponds was designed and developed for automatic data collection. The sensors include temperature, pH, dissolved oxygen, turbidity, ammonia and nitrate sensors. The IOT system reads water quality data and uploads the same to the cloud in real time. The dataset is visualized in the cloud and downloaded for the purposes of data analytics and decision-making. We present the dataset in this paper. The dataset will be very useful to the agriculture, aquaculture, data science and machine learning communities. The insights such dataset will provide when subjected to machine learning and data analytics will be very useful to fish farmers, informing them when to change the pond water, what stocking density to apply, provide knowledge about feed conversion ratios, and in predict the growth rate and patterns of their fishes. Here are some research papers in this area.

Some Research Paper Summaries:

Study 1

MS-CAGRU for Water Quality Prediction

The authors proposed ensemble models: bagging, boosting, and stacking, and applied different ML models: random forest (RF), support vector machine (SVM), decision tree (DT), K-nearest neighbor (KNN), and logistic regression (LR) to classify different types of fish using water quality parameters: pH, temperature, and turbidity. The dataset was collected from AquaBot, an IoT-based system. The results showed the ensemble model recorded the best performance 94% accuracy, F1-score of 94%, and a 93% MCC. In addition, Results showed MS-CAGRU achieved 4.25% better accuracy than RNN and 2.45% over GRU. Regularization improved performance further.

Study 2

IoT & Advanced ML Algorithms (Tilapia Ponds, 2025)

The authors applied advanced machine learning techniques to six months of IoT sensor data from tilapia aquaculture ponds. They monitored key water quality parameters (temperature, DO, pH, turbidity) using a custom IoT system and evaluated multiple predictive models including tree ensembles and neural networks. By leveraging techniques like bagging, boosting, and temporal feature analysis, they significantly improved prediction performance for parameters like dissolved oxygen. The results showed that ensemble models (combining decision trees and boosting methods) achieved very high accuracy (often ~99%) in forecasting water quality trends. The best model was a custom ensemble approach, demonstrating that integrating IoT data with ML can enable near real-time water quality predictions with minimal error.

Study 3

IOT + Random Forest with Quantum Optimization (Tropical Pond, 2025)

The researchers integrated IOT sensors and machine learning to continuously monitor and predict water quality in a tropical fish pond. They collected continuous data on temperature, dissolved oxygen, pH, and turbidity via low-cost IOT devices and used Random Forest regression to forecast key parameters. A quantum-inspired optimization algorithm (QAOA) was applied to reduce model training time by ~50% and fine-tune parameters. The Random Forest model achieved a near-perfect fit for dissolved oxygen ($R^2 \approx 0.999$, RMSE ~ 0.1 mg/L), indicating extremely high predictive accuracy. By using (IOT) data and QAOA optimization, the system enabled rapid adjustments to maintain > 90% fish survival rates. The best-performing model was the Random Forest regressor, augmented by quantum optimization for efficiency.

Study 4

Hybrid EEMD-MLR-LSTM Model (Salmon Farm, 2023)

This study examined a hybrid deep learning approach for multivariate water quality forecasting in an offshore salmon aquaculture farm. The authors developed a model combining Ensemble Empirical Mode Decomposition (EEMD) for feature extraction, Multivariate Linear Regression (MLR), and an LSTM neural network. Using time-series sensor data (e.g., chlorophyll-a, turbidity) from a salmon farm in Scotland, the hybrid EEMD-MLR-LSTM model was trained to predict harmful algal bloom indicators. The hybrid model achieved lower prediction errors than single models, with RMSE dropping to around 0.05 in forecasting chlorophyll levels. It outperformed comparison models (e.g., SAE-LSTM or standalone LSTM) by effectively capturing multi-scale features. The best model was the proposed EEMD-MLR-LSTM, which demonstrated the highest accuracy and the least error among all methods tested.

Study 5

Pearson Correlation + LSTM with Attention (Freshwater Aquaculture, 2025):

The authors developed a water quality prediction model for a freshwater fish farm by integrating feature selection and attention mechanisms into LSTM. They first applied Pearson correlation analysis to select the most relevant water quality variables for predicting dissolved oxygen. An LSTM model with an attention module (LSTM-AM) was then trained on the selected features to forecast DO levels five days ahead. The enhanced model (Pearson-LSTM-AM) showed markedly better performance than baseline LSTM, CNN-LSTM, and Bi-LSTM models. It achieved an R^2 of 0.8592 and reduced RMSE to 0.3791 for DO prediction, improving RMSE by ~8–10% over competing models. The best-performing model was the Pearson-LSTM-AM, which combined correlation-based feature filtering and attention to yield the most accurate DO forecasts.

Study 6

CNN + Self-Attention + Bi-SRU (Intensive Pond, 2025):

This study proposed a hybrid deep learning framework to predict dissolved oxygen in an intensive aquaculture pond using a CNN, self-attention, and a bidirectional Simple Recurrent Unit. The model applied 1D convolution to extract features from time-series sensor data, then a self-attention mechanism to weight important features, followed by a Bi-SRU recurrent layer for sequence prediction. Using an intensive fish farm dataset from Guangzhou (China), the hybrid CNN-SA-BiSRU model achieved very high accuracy in DO estimation. It achieved a mean absolute error around 0.03–0.04 and an RMSE of 0.0471, with an R^2 of 0.9765 in predicting DO levels. Short-term prediction accuracy improved significantly compared to standard RF or basic neural nets. The best model was the proposed CNN-SA-BiSRU, which outperformed traditional methods and provided highly accurate real-time DO predictions.

Study 7

Support Vector Machine vs Neural Networks (Industrial Fish Farm, 2022):

The authors compared several machine learning models for predicting multiple water quality parameters in an industrial aquaculture system. They used historical data (both published and measured) of an intensive fish farm, including parameters such as DO, pH, ammonia, nitrate, and nitrite. Four models were tested: backpropagation neural network (BPNN), radial basis function neural network (RBFNN), support vector machine (SVM), and least-squares SVM. The SVM consistently yielded the most accurate and stable predictions across all five water quality parameters. In particular, the SVM achieved about 99% accuracy for both the literature dataset and actual farm data, significantly outperforming the neural network models (which had more variable and lower accuracy). Thus, the SVM was recommended as the best model, demonstrating excellent predictive performance (e.g., near-perfect correlation coefficients of 0.99 for key parameters) in simulating water quality dynamics.

Study 8

AquaSense IoT Framework + Dilated CNN-BiLSTM (Smart Aquaculture, 2024):

The authors presented “AquaSense,” an autonomous sensor network framework for continuous water quality monitoring in aquaculture ponds. The system deployed multiple IoT sensors to measure temperature, pH, dissolved oxygen, and salinity, feeding data to a cloud-based analysis platform. They implemented a dilated CNN-BiLSTM deep learning architecture to predict water quality status and issue timely alerts. The model was evaluated using metrics like accuracy, precision, recall, and F1-score. AquaSense achieved an overall accuracy of 96.98% in classifying water quality conditions, substantially higher than existing methods (which ranged from ~88% to 93% accuracy). In particular, the dilated CNN-BiLSTM model outperformed prior frameworks (such as

ISAS and AquaStat) by 3–8 percentage points in accuracy. The best model was the proposed AquaSense deep learning framework, which reliably predicted water quality trends and enabled proactive farm management.

Study 9

AquaBot IOT System + Ensemble ML (Smart Fish Pond, 2024):

The authors developed an integrated smart pond management system called AquaBot, which automates water monitoring and fish species recommendations. An IoT-enabled robotic platform was used to collect water parameters (pH, temperature, turbidity) throughout the pond, and the data were fed into various machine learning models. They trained several algorithms – including a custom ensemble model, Random Forest, SVM, decision tree, k-NN, and logistic regression – to analyze water quality and suggest suitable fish species for that water. Among these, their proposed ensemble model delivered the best performance, with ~94% accuracy, precision, and recall in multi-class classification (evaluating fish species recommendations). This ensemble (which stacked multiple classifiers) outperformed individual learners (which achieved ~88–93% accuracy). The top model was thus the custom ensemble integrated into AquaBot, which provided accurate fish farming recommendations based on real-time water quality data.

Study 10

IoT Sensors + Decision Tree Classifier (Fish Behavior Analysis, 2023):

This study monitored ornamental fish behavior using IoT sensors and machine learning to infer stress or health issues from environmental changes. Goldfish in a controlled tank were observed under varying water temperatures and dissolved oxygen levels while recording their

behavioral changes. The researchers collected behavioral data (e.g., resting, erratic swimming, gasping at surface) and trained four classifiers – decision tree, Naïve Bayes, k-NN, and LDA – to predict fish behavioral states from the sensor inputs. Cross-validation showed the decision tree to be the most accurate model, with a prediction error of ~13.78% (i.e., ~86.2% accuracy), outperforming k-NN and Naïve Bayes which had higher errors (~30% and ~28%). Using the decision tree model, they identified specific thresholds: for example, fish became erratic or showed gasping behavior when water temperature exceeded ~40.5°C or when DO fell below ~6.6 mg/L. The best-performing model was the decision tree, enabling automated detection of fish stress behaviors from IOT sensor data with high accuracy.

Study 11

Dilated CNN + BiGRU with Attention (Indoor Mariculture, 2024):

The authors introduced a prediction method for dissolved oxygen in an indoor recirculating aquaculture system using a hybrid model based on variational mode decomposition (VMD), dilated CNN, and Bi-GRU with attention. They partitioned the water quality time-series using VMD to capture different frequency components, then applied a dilated convolutional neural network for feature extraction, followed by a Bi-GRU recurrent network enhanced with an attention mechanism to predict DO levels. Testing on an indoor fish tank dataset showed that this VMD-CNN-BiGRU-Attention model met the required accuracy for practical DO prediction. It provided a reliable multi-step forecast of DO with substantially lower error than conventional LSTM or GRU models. The study reported that the hybrid model's predictions closely matched observed DO trends, satisfying the needs of real-time monitoring. The best model was the VMD-CNN-BiGRU with attention, which delivered robust predictive performance for short-term DO fluctuations in closed aquaculture environments.

Study 12

Dual-Attention ResNet + LSTM (Offshore Aquaculture, 2025):

In this work, researchers developed a dual-channel neural network to improve early warning of water quality in an offshore aquaculture farm. The model, called DAM-ResNet-LSTM, consists of two parallel feature extraction streams: one using a deep residual network (ResNet) and another using LSTM, each augmented with an attention mechanism to emphasize important features. The model was trained on an environmental monitoring dataset from a coastal fish farm, including water parameters (pH, DO, salinity, COD, ammonia, etc.) and meteorological data. Results showed that the dual-attention structure significantly enhanced prediction accuracy. For instance, the model's Nash-Sutcliffe efficiency coefficients for pH, DO, and salinity were ~ 0.93 – 0.94 (indicating around 93–94% relative accuracy), whereas for more challenging parameters (COD, $\text{NH}_3\text{-N}$, etc.) it achieved ~ 0.83 – 0.86 . This was an improvement of roughly 10–18% in predictive skill over single-channel attention models. The best model was the proposed dual-attention ResNet-LSTM, which demonstrated superior accuracy in forecasting multiple water quality factors simultaneously.

Study 13

Hybrid Multi-Scale Feature Model (DO Prediction, 2018):

The study presented a hybrid modeling approach to predict dissolved oxygen in aquaculture ponds by combining multi-scale feature extraction with deep learning. The authors decomposed the original water quality time-series into multiple scales (e.g., using wavelet or empirical mode decomposition) to capture both short-term fluctuations and long-term trends in DO levels. These features were then fed into a deep neural network (such as an LSTM or hybrid CNN-LSTM) to forecast future DO concentrations. The hybrid multi-scale model achieved higher precision than models using single-scale data. In evaluations on aquaculture pond data, it reduced prediction error significantly – for example, cutting RMSE

by around 20–30% compared to a standard LSTM without multi-scale inputs. By integrating information across frequency scales, the model more accurately tracked DO dynamics. The best-performing approach was this multi-scale hybrid model, which outperformed conventional single-resolution predictors in both accuracy and stability of aquaculture DO forecasts.

Study 14

Deep Belief Network (Recirculating Aquaculture, 2020):

This work applied a deep belief network (DBN) to predict dissolved oxygen in a recirculating aquaculture system (RAS). The authors designed a DBN-based regression model, leveraging its layered learning to model complex, nonlinear relationships between input variables (such as feed rate, aeration, temperature) and DO output. They trained and tested the model on data from multiple marine fish ranches, comparing its performance to traditional approaches. The DBN model demonstrated faster convergence and higher accuracy in DO prediction than conventional mechanistic models. For instance, it achieved a high coefficient of determination in validation tests and was able to predict DO levels with low error for the RAS dataset. By capturing underlying feature representations through its deep layered structure, the DBN outperformed simpler neural networks in both stability and precision. The best model was the DBN-based predictor, which reliably forecasted short-term DO variations in the recirculating system, aiding automated control of aeration.

Study 15

Ensemble LSTM-SVM with Optimization (Coastal Aquifer Case, 2023):

In this study, researchers aimed to enhance water quality parameter prediction by combining recurrent neural networks with support vector regression. They proposed two hybrid models – one integrating a unidirectional LSTM with SVM, and another integrating a bidirectional LSTM (BiLSTM) with SVM – and introduced a novel “Ali Baba and the Forty Thieves” (ABFT) optimization algorithm to fine-tune model parameters. The case study focused on forecasting total dissolved solids (TDS) and electrical conductivity (EC) in a coastal aquaculture-adjacent aquifer using historical water chemistry data. Results indicated that the BiLSTM-SVM hybrid model outperformed all other tested models, achieving very low MAPE (around 2% for TDS) and a high R^2 (~ 0.94 for EC) for one-day-ahead predictions. The accuracy of all models declined slightly for longer horizons, but the BiLSTM-SVM consistently maintained the lowest error. Thus, the best model was the BiLSTM-SVM optimized with ABFT, which provided the most precise and reliable water quality predictions among the methods compared.

Comparison Table :

Papers	Year	Methods	Datasets	The best model	Results
Study [1]	2025	Applied Conv. Autoencoder + GRU (deep learning)	Aquaponic fish pond IOT sensors www.kaggle.com/dataset/e81da8b7666dc7af41cdc3aa5ef96c5547e4f412598a030f40d444550965e34f	MS-CAGRU (CAE+GRU hybrid)	AUC ~0.94; ~5% higher accuracy vs baseline RNNs
Study [2]	2025	Ensemble ML (bagging & boosting); time-series ML	Tilapia pond IOT data (6 months)	Ensemble (Bagging+ Boosting)	Accuracy (AUC)= 94%, ~5% higher accuracy vs baseline RNNs
Study [3]	2025	Random Forest regression; Quantum optimization	Tropical pond sensors (IOT) https://data.mendeley.com/data-sets/dgdr2kfby1	Random Forest + QAOA	$R^2 = 0.999$ (DO); RMSE ~0.0998 mg/L (very high)
Study [4]	2023	Applied different EEMD (signal decomposition) + LSTM + MLR	Salmon farm sensor time-series	EEMD-MLR-LSTM hybrid	Lowest error (RMSE ~0.05); outperformed LSTM
Study [5]	2025	Applied Pearson correlation + LSTM + Attention	Freshwater pond (DO focus)	Pearson-LSTM-AM (Attn.)	$R^2 \approx 0.859$; RMSE 0.379 (best among tested)
Study [6]	2025	CNN + Self-Attention + Bi-SRU (RNN variant)	Intensive pond (Guangdong, China)	CNN-SA-BiSRU (hybrid DL)	RMSE 0.047; $R^2 \approx 0.9765$ (excellent DO prediction)
Study [7]	2022	Applied different BPNN, RBFNN, SVM, LSSVM (model comparison)	Industrial fish farm data	SVM model (regression)	~99% accuracy; correlation ~0.99 on key params
Study [8]	2024	IOT sensor network; Dilated CNN + BiLSTM DL	Smart aquaculture pond (multi-sensor)	Dilated CNN-BiLSTM	96.98% accuracy; +3–8% vs prior methods
Study [9]	2024	Applied IOT robot + ML (RF, SVM, k-NN, ensemble)	Smart pond (Bangladesh) IOT data	Custom Ensemble (stacked)	~94% accuracy /precision/recall; best recommender
Study [10]	2023	IoT env. sensors + ML classifiers (DT, KNN, etc.)	Ornamental fish tank (behavior)	Decision Tree classifier	~86% accuracy; identified stress behavior triggers
Study [11]	2025	Physics-informed Bi-LSTM model	Integrated multi-trophic aquaculture	Bi-LSTM + physics constraints	Improved long-term biomass prediction accuracy
Study [12]	2024	VMD + CNN + BiGRU + Attention (hybrid deep model)	Indoor RAS (recirculation system)	VMD-CNN-BiGRU-Attn model	Met DO forecasting needs; more accurate
Study [13]	2025	Dual-channel ResNet + LSTM with dual attention	Offshore farm IoT (multi-param)	Dual-Attention ResNet-LSTM	~93–94% Nash efficiency for DO/pH; +10% vs single-channel
Study [14]	2018	Multi-scale feature extraction + deep network	Aquaculture pond (DO time-series)	Hybrid multi-scale DL model	~20–30% error reduction vs single-scale LSTM
Study [15]	2020	Deep Belief Network (DBN) regression	Recirculating aquaculture (marine)	DBN predictor	High R^2 in DO forecasts; outperformed legacy models

Conclusion :

Across these studies, several common trends emerge in applying IoT-based data and machine learning for aquaculture. **Deep learning models (LSTM, GRU, Bi-LSTM, etc.) have become dominant** for predicting key water quality parameters like dissolved oxygen, often outperforming traditional methods (e.g., SVM or simple neural nets) in accuracy. Many works incorporate **hybrid approaches** – for example, combining feature decomposition (EEMD, wavelets) or attention mechanisms with LSTM/GRU – to capture both short-term fluctuations and long-term trends in complex aquaculture datasets. **Ensemble and hybrid models** (such as stacking ML algorithms or integrating neural networks with SVMs or physics knowledge) frequently achieve the best performance, indicating that multi-model strategies benefit water quality prediction. The use of **IoT sensor networks** provides high-resolution, real-time data, enabling models to attain very high predictive precision (often with R^2 or accuracy well above 0.9 for critical parameters). Many studies report that their best models can predict water quality indicators with **low error (e.g., RMSE on DO in the 0.05–0.5 range)**, allowing proactive management (like timely aeration or feeding adjustments). In summary, the integration of IOT monitoring with advanced regression and deep learning techniques is driving a new level of predictive capability in aquaculture, leading to smarter fish pond management, early warning of unfavorable conditions, and improved fish health and yields.

References

1. Sundararajan S. C. M. et al. (2025). IoT-based prediction model for aquaponic fish pond water quality using multiscale feature fusion with convolutional autoencoder and GRU networks. Scientific Reports.
<https://www.nature.com/articles/s41598-024-84943-7>
<https://www.kaggle.com/datasets/bobsis/small-aquaculture-fishpond>
2. Hridoy, M. A. A. M. et al. (2025). Predictive modelling of aquaculture water quality using IOT and advanced machine-learning algorithms. Results in Chemistry 16, 102456.
<https://www.sciencedirect.com/science/article/pii/S2211715625004394>
3. 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 for Sustainable Management. <https://www.mdpi.com/2073-4441/17/1/82>

4. Eze, E., Kirby, S., Attridge, J., & Ajmal, T. (2023). Aquaculture 4.0: hybrid neural network multivariate water quality parameters forecasting model. *Scientific Reports* **13**, Article 16129. DOI: <https://www.nature.com/articles/s41598-023-41602-7>
5. Yu, X. et al. (2025). Research on a Pearson-LSTM-AM-based water quality prediction model for freshwater aquaculture. *Journal of the World Aquaculture Society* <https://onlinelibrary.wiley.com/doi/10.1111/jwas.70041?af=R>
6. Xu, L. et al. (2025). Hybrid deep learning framework for real-time DO prediction in aquaculture. *Scientific Reports* **15**, Article 24643. DOI: <https://www.nature.com/articles/s41598-025-10786-5>
7. Li, T., Lu, J., Wu, J., Zhang, Z., & Chen, L. (2022). Predicting Aquaculture Water Quality Using Machine Learning Approaches. *Water* **14**(18): 2836. DOI: <https://www.mdpi.com/2073-4441/14/18/2836>
8. Iniyan Arasu, M., Subha Rani, S., Thiyagarajan, K., & Appathurai, A. (2024). AQUASENSE: aquaculture water quality monitoring framework using autonomous sensors. *Aquaculture International* **32**(7): 9119–9135. DOI: <https://link.springer.com/article/10.1007/s10499-024-01606-0>
9. Hemal M. M. et al. (2024). An Integrated Smart Pond Water Quality Monitoring and Fish Farming Recommendation Aquabot System. *Sensors*. <https://doi.org/10.3390/s24113682>
10. Patro K. S. K. et al. (2023). IoT and ML approach for ornamental fish behavior analysis. *Scientific Reports*. <https://www.nature.com/articles/s41598-023-48057-w>
11. Kunapinun, A., Fairman, W., Wills, P., & Ouyang, B. (2025). Integrating Bi-LSTM and physics constraint models for improved biomass and water quality prediction in aquaculture systems. In *Proc. SPIE 13460: Machine Learning from Challenging Data* <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/13460/1346008/Integrating-Bi-LSTM-and-physics-constraint-models-for-improved-biomass/10.1117/12.3053764.short>
12. Arepalli, P. G., & Naik, K. J. (2024). A Lot-Based VMD-CNN-BiGRU indoor mariculture water quality prediction model with attention mechanism. (Research study reported in 2024; appeared in conference/journal literature via World Scientific, <https://www.worldscientific.com/doi/10.1142/S0129156425400105?srsIid=AfmBOoo3gRuGsR06XmS7MT-LxbA7dnS80dum1BsfA4rYn00U45yWWGV>
13. Liu, W., Wang, J., Li, Z., & Lu, Q. (2025). A Hybrid Improved Dual-Channel and Dual-Attention Mechanism Model for Water Quality Prediction in Nearshore Aquaculture. *Electronics* <https://www.mdpi.com/2079-9292/14/2/331>
14. Li, C., Li, Z., Wu, J., Zhu, L., & Yue, J. (2018). A hybrid model for dissolved oxygen prediction in aquaculture based on multi-scale features. *Information Processing in Agriculture* **5** <https://www.sciencedirect.com/science/article/pii/S2214317316301184>
15. Ren, Q., Wang, X., Li, W., et al. (2020). Research of dissolved oxygen prediction in recirculating aquaculture systems based on deep belief network. *Aquacultural Engineering* **90**: 102085. DOI: <https://www.sciencedirect.com/science/article/abs/pii/S0144860919301669>