(Online) ISSN 2712-0562
Original Research

Sustainable Engineering and Innovation
Vol. 7, No. 2, 2025, pp.521-536
https://doi.org/10.37868/sei.v7i2.id606

Smart water quality regulation in sustainable aquaponics using PID control and long-term performance analysis

Jorge Saul Fandiño Pelayo[1], Luis Sebastian Mendoza Castellanos[2], Rocio Cazes Ortega[3], Luis Guillermo Hernández Rojas[4], O. Lengerke[5]
[1,3,5] Department of Natural Sciences and Engineering, Unidades Tecnológicas de Santander (UTS), Bucaramanga, Colombia
[2] Department of Engineering, Universidad Autónoma de Bucaramanga (UNAB), Bucaramanga, Colombia
[4] Tecnológico de Monterrey, School of Engineering and Sciences, Monterrey, México

*Corresponding author E-mail: jfandino@correo.uts.edu.co

Received Aug. 1, 2025
Revised Sep. 13, 2025
Accepted Sep. 18, 2025
Online Oct. 10, 2025

Abstract

This study evaluates the reliability and suitability of long-term water quality data collected from a sustainable aquaponics system equipped with a pH, dissolved oxygen, and temperature regulation control strategy based on a PID algorithm. Although PID control was implemented to maintain parameters within optimal biological ranges, natural fluctuations and out-of-range measurements were recorded, particularly in pH. Rather than being considered anomalies, these deviations represent realistic environmental variations that must be captured for comprehensive system analysis. A thorough data validation process was conducted, including descriptive statistics, outlier detection, correlation analysis, principal component analysis (PCA), and temporal stability evaluation. Results confirmed the absence of missing data, the presence of controlled variability in dissolved oxygen and temperature, and meaningful correlations between parameters, with pH showing the highest variability. Autocorrelation and long-term trend analyses indicated stable measurement patterns that reflect real-world aquaponic dynamics. The validated dataset provides a robust foundation for future studies, particularly for the development of artificial intelligence (AI)-based predictive models aimed at early detection of fish distress or mortality.

Keywords: Aquaponics, PID control, Water quality monitoring, Data validation, Sustainable aquaculture

1. Introduction

Aquaponics, which integrates aquaculture and hydroponics into a recirculating system, is increasingly recognized as a sustainable food production approach that maximizes resource efficiency and reduces environmental impact [1], [2], [3]. In these systems, fish waste provides nutrients for plants, while roots purify the water, reducing fertilizer inputs and water use. To sustain fish health, plant growth, and microbial biofiltration, precise regulation of water quality parameters such as pH, temperature, and dissolved oxygen (DO) is essential [4], [5], [6]. Moreover, fish welfare in aquaponic systems is closely related to water quality, particularly in connection with feed and waste management [5].

Automatic control strategies have been introduced to stabilize these parameters. Proportional–Integral–Derivative (PID) controllers, in particular, are widely applied due to their robustness and simplicity. Supriadi et al. (2019) demonstrated Arduino-based proportional control of pH and temperature in aquaponics [7], while more advanced Smart Aquaponics implementations combined automatic PID tuner tools with manual fine-tuning of K_p, K_i, and Kd gains across interconnected units [8]. However, system nonlinearities and environmental variability still generate deviations from setpoints, even under PID regulation [9], underscoring the need to analyze system performance under both controlled and uncontrolled conditions. Beyond PID, other strategies have been explored to improve resilience. Fuzzy logic-based controllers have been successfully applied to regulate pH and temperature in aquaponics biofilters [10], while IoT-based monitoring and machine learning frameworks have been integrated for predictive control and real-time decision support [11], [12]. Despite these advancements, one critical aspect remains insufficiently addressed: whether collected data are reliable, representative, and sufficiently variable for advanced modeling. Data validation is crucial for long-term monitoring, where sensor drift, noise, or missing records can compromise analytical outcomes [13]. In aquaponics, datasets capturing both stable ranges and outliers caused by stress or control limitations are indispensable for predictive tools capable of early error detection and decision support [14], [15].

While prior works have primarily emphasized controller performance or short-term monitoring [7], [10], [11], few studies have validated extended time-series datasets that reflect the real variability and limitations of PID regulation under operational conditions. Addressing this gap, the present study introduces an eight-month dataset from a domestic aquaponics system regulated with PID control. Comprehensive validation was conducted—including descriptive statistics, outlier analysis, correlation, and principal component analysis (PCA), and temporal stability tests—to assess data quality and relevance. The results confirm strong internal consistency, realistic variability, and meaningful parameter relationships. By establishing a validated dataset that captures authentic deviations, this work provides a robust foundation for future development of AI-based predictive control systems aimed at enhancing stability and sustainability in aquaponic operations [12], [15].

2. Materials and methods

2.1. System description

The experimental aquaponic system operated in a closed-loop recirculation configuration, integrating the rearing of red tilapia (*Oreochromis niloticus*) with the cultivation of crisp lettuce (*Lactuca sativa var. crispa*). The main rearing tank (1) had a volume of 800 L (1 m^2 surface area) and was stocked with 30 juvenile tilapia, corresponding to a density consistent with small-scale aquaponic trials. The tank was instrumented with submerged sensors for dissolved oxygen (Model AR8406), water temperature (DS18B20 submersible thermocouple), and pH (Arduino-compatible PH45 probe). Two internal aerators were installed to maintain dissolved oxygen within the optimal range for both fish and plants. Water was circulated through a PVC outlet connected to a submersible JUIYU pump (800 L/h, 12 V DC, 5 m head) located at the tank bottom. The pump was initially fitted with narrow slits to prevent the escape of fingerlings, later widened as the fish grew to allow the passage of suspended solids.

The first filtration stage consisted of a cylindrical clarifier (2) for gravitational sedimentation of coarse particulate matter, including uneaten feed and feces. The second stage was a mechanical filtration unit (3) with a fine-mesh screen to capture finer suspended particles. The third stage was a biofilter (4), packed with synthetic polyester fiber and thoroughly cleaned plastic bottle caps. These materials were selected for their high surface area and structural heterogeneity, favoring colonization by nitrifying bacteria responsible for oxidizing ammonia into nitrite and subsequently into nitrate, thereby supplying plants with assimilable nitrogen.The nutrient-enriched effluent entered a floating raft hydroponic bed (5), where approximately 25–30 lettuce plants were cultivated in net pots inserted into perforations in a polystyrene sheet. Plant roots were in direct contact with the circulating water, enabling nutrient uptake and further water polishing. After this stage, the treated water returned by gravity to the fish tank (1), completing the recirculation loop.

Throughout the system, a Raspberry Pi 4 served as the supervisory control unit, acquiring sensor inputs and managing actuators in real time. It activated the aerators when dissolved oxygen dropped, a cold-water injection pump when temperature exceeded the target range, an aquarium glass heater during temperature drops, and a peristaltic pump for dosing apple vinegar or sodium bicarbonate in response to pH deviations. A schematic of the system is presented in Figure 1.

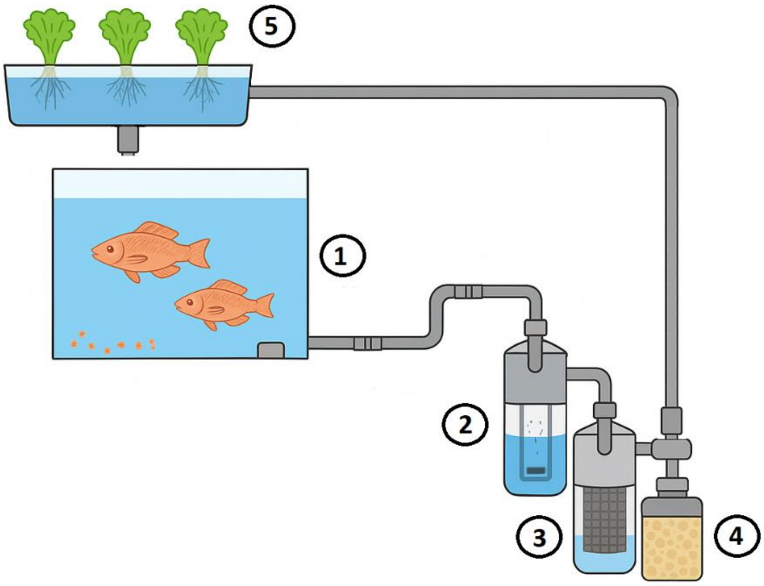

Figure 1. Tilapia tank (1), mechanical filter (2), biofilter (3), hydroponic beds (4), and water pump (5) in the aquaponic system

2.2. Control strategy

The unit operated off-grid using a photovoltaic array rated at 400 Wp, which charged a 12 V, 120 Ah $LiFePO_4$ battery through a 40 A MPPT controller. During daytime operation, the PV array directly supplied power to the DC instrumentation and actuators while simultaneously charging the battery. When irradiance was insufficient (nighttime or cloudy conditions), the battery bank provided an uninterrupted supply to all loads. This automatic exchange between photovoltaic generation and stored energy ensured backup capacity and continuity of supply, thereby guaranteeing stability in actuator responses and uninterrupted acquisition of sensor data. The energy configuration allowed the system to operate autonomously for approximately 48 hours without solar input, securing data continuity even under adverse weather conditions.

A Raspberry Pi 4 functioned as the supervisory controller and data hub, acquiring real-time signals from the dissolved oxygen (AR8406), temperature (DS18B20), and pH (PH45) sensors. Based on these measurements, a Proportional–Integral–Derivative (PID) algorithm computed corrective actions relative to predefined setpoints (pH = 7.0, DO = 6 mg/L, temperature = 25 °C) and their allowable biological ranges.

Temperature regulation was bidirectional; when water temperature exceeded the upper threshold, a 12 V submersible pump injected cold water from a sealed auxiliary reservoir; conversely, when temperature dropped below the lower bound, a 300 W thermostat-controlled heater was activated to restore the setpoint. pH stability was maintained via a dual-channel peristaltic pump, which dosed apple vinegar or diluted sodium bicarbonate according to the PID output. Dissolved oxygen was controlled through two 12 V DC aerators (15 W, 800 L/h each), automatically engaged when levels approached the minimum threshold for tilapia welfare. In parallel with these conditional actions, a solids-lift submersible pump operated continuously to sustain water recirculation and solids removal. All measurements, actuator states, and PID outputs were locally logged and synchronized to a Firebase Realtime Database when internet connectivity was available, ensuring continuous monitoring and full traceability of system performance. A schematic of the control architecture is presented in Figure 2.

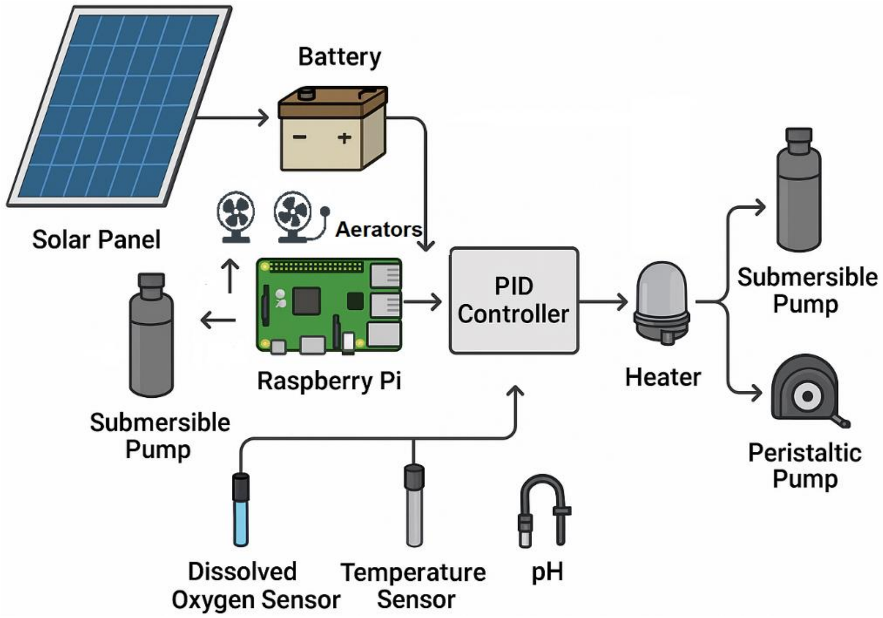

Figure 2. Solar-powered control architecture. Raspberry Pi 4 reads dissolved oxygen, temperature, and pH and drives actuators: continuous solids-lift submersible pump; cold-water injection pump; two 12 V, 15 W aerators; dual-channel peristaltic dosing; and a 300 W thermostat heater. Power is supplied by a 400 Wp PV array, 12 V 120 Ah $LiFePO_4$ battery, and 40 A MPPT controller; an inverter is used only when the AC heater is operated from the battery

2.3. Experimental design and procedure

The aquaponic system was configured as a closed-loop arrangement integrating the rearing tank, biofiltration units, and the hydroponic grow bed. Tilapia (*Oreochromis niloticus*) were selected as the experimental species due to their robustness and adaptability to recirculating aquaculture systems. The rearing tank (800 L, 1 m² surface area) was stocked with 30 juvenile tilapia, while approximately 25–30 lettuce plants (*Lactuca sativa* var. *crispa*) were cultivated in the floating raft bed. This configuration provided a balanced nutrient exchange between aquaculture and hydroponics, maintaining water quality parameters within biologically acceptable thresholds.

Water quality monitoring was conducted continuously using the sensors integrated into the Raspberry Pi 4, which recorded pH, temperature, and dissolved oxygen at predefined intervals. These measurements were automatically logged and served as inputs for the PID-based control system. The closed-loop feedback ensured that deviations from setpoints were immediately corrected through actuation.

Thresholds for favorable operation were established based on literature values for tilapia welfare and lettuce growth. Water temperature was maintained within 24–28°C, with corrective actions triggered when it exceeded 28°C (activation of the auxiliary cold-water pump) or dropped below 24°C (engagement of the 300 W thermostat-controlled heater) [5], [10]. The acceptable pH range was 6.5–7.5, with the setpoint fixed at 7.0, consistent with both tilapia tolerance and nutrient availability for plants [5], [10]. Dissolved oxygen (DO) was kept above 5 mg/L, with aerators activated automatically when DO approached this threshold, thereby ensuring adequate oxygenation for tilapia welfare [5], [15].

Sensor calibration was performed weekly in accordance with manufacturer guidelines and widely adopted protocols. The pH probe (PH45) was adjusted using the manufacturer's potentiometric procedure (2.5 V in pH 7 buffer) and calibrated with standard buffer solutions at pH 4.0 and 7.0. The dissolved oxygen sensor (AR8406), factory-calibrated and certified, was periodically verified in air-saturated water, checked with sodium sulfite solution for zero reference, and occasionally cross-validated using a YSI meter. The temperature sensor (DS18B20) was verified against a laboratory-grade glass thermometer at the start of the experiment and during periodic maintenance. According to manufacturer specifications, accuracies were ±0.1–0.2 pH units, ±0.5 °C for temperature, and ±0.3 mg/L for dissolved oxygen, ensuring reliable and traceable measurements throughout the study.

The experimental phase lasted eight months and encompassed multiple production cycles. Data collected during this period were later used for system validation and for the development of predictive artificial intelligence models. In this way, the aquaponic installation functioned not only as a life-support system for fish and plants but also as a reliable experimental platform for acquiring validated datasets.

2.4. Data analysis and processing

The collected data underwent a comprehensive preprocessing phase to ensure quality and consistency. Measurements outside the biologically relevant ranges—pH between 5 and 9, dissolved oxygen (DO) between 0 and 15 mg/L, and temperature between 18 and 35 °C—were identified as anomalies. Specifically, 92 pH readings fell outside the optimal range and were treated as outliers for descriptive and trend analyses, following standard procedures for aquaponics data validation [15], [16].

Descriptive statistics were computed for each variable, including mean ($\bar{X}$), standard deviation (σ), median, and interquartile range (IQR), according to the following expressions:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_{i,} \qquad \sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{1}$$

$$Median = x_{\frac{N+1}{2}}, \qquad IQR = Q_3 - Q_1 \tag{2}$$

Where x_i represents the i-th measurement and N the total number of measurements. These statistics provided a summary of central tendency and variability within the system [15].

Long-term trends were evaluated using moving averages and moving standard deviations over a 30-measurement window [16], which enabled visualization of temporal variability and detection of deviations from the target ranges. Additionally, time series were decomposed into trend, seasonal, and residual components using an additive model [17]:

$$x_t = T_t + S_t + R_t \tag{3}$$

Where T_t represents the trend, S_t the seasonal variation, and R_t the residual (noise) at time t. This decomposition facilitated the identification of periodic fluctuations caused by environmental or operational cycles [17].

Autocorrelation analysis was performed to detect temporal dependencies, computing the autocorrelation function (ACF) for lags up to 50 measurements:

$$\rho_k = \frac{\sum_{t=1}^{N-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{N}(x_t - \bar{x})^2} \tag{4}$$

Where ρ_k is the autocorrelation at lag k. This provided information on the system's memory and the effectiveness of PID control loops [15], [18].

Finally, the data were visualized using line graphs, histograms, and bands of moving average ± standard deviation. Optimal operating ranges were shaded for clarity (e.g., pH 5–9, DO 4–8 mg/L, temperature 22–27 °C), and median lines with 95% confidence intervals were included to highlight stability and variability over time [16], [17]. These analyses provided a comprehensive characterization of system performance, supporting subsequent validation and predictive modeling of water quality parameters [15], [18].

2.5. System validation

The validation of the aquaponic system was performed by comparing measured pH, dissolved oxygen (DO), and temperature values with the predefined setpoints of the PID control system. Key performance indicators, including the mean absolute error (MAE) and root mean squared error (RMSE), were computed to quantify control accuracy:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|x_i - x_{set}|, \qquad RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - x_{set})^2} \tag{5}$$

Where x_i is the measured value at time i, x_{set} is the setpoint, and N is the number of measurements. These metrics quantified deviations from target values and allowed evaluation of system precision.

The computed errors demonstrated effective control performance across all monitored parameters. For pH, the MAE was 0.599 and the RMSE 0.846, indicating that deviations from the setpoint remained minimal. Dissolved oxygen exhibited an MAE of 0.758 and an RMSE of 0.927, while water temperature showed an MAE of 0.800 and an RMSE of 1.042. These results confirm that the PID controller maintained stable operational conditions within the ranges suitable for both tilapia welfare and lettuce cultivation.

Previous research has demonstrated the effectiveness of automated water quality control in aquaponics. Fuzzy logic-based controllers have been successfully applied to regulate pH and temperature, minimizing oscillations and improving overall system stability [19]. Similarly, modular IoT-based monitoring systems with edge computing have been shown to maintain optimal conditions for both fish and plants through real-time data acquisition and adaptive control [20].

The long-term stability of the present system was further evaluated through autocorrelation and trend decomposition analyses. Positive short-lag autocorrelation indicated smooth parameter adjustments, while the decomposition revealed minimal seasonal or environmental influence on overall dynamics [21]. Together, these results confirm that the PID controller sustained a stable operational environment suitable for both tilapia rearing and lettuce cultivation.

2.6. Data acquisition and logging

Data acquisition was conducted using a modular IoT system that integrated sensors for pH, dissolved oxygen, and temperature. The system used edge computing to process data locally, minimizing latency and enabling real-time control decisions. Data were recorded at regular intervals, generating time-series datasets suitable for system performance analysis and AI-based predictive modeling.

Data integrity was ensured through cloud synchronization and local redundancy, allowing remote monitoring and post-processing analysis. Prior studies have confirmed that IoT-enabled aquaponic monitoring improves both operational efficiency and data reliability, particularly when combined with adaptive algorithms for real-time control [22]. Additionally, sensor validation techniques are crucial for maintaining accurate long-term measurements in aquaponics [23].

3. Results

3.1. Descriptive statistics and data validation

A preliminary analysis of the dataset was conducted to evaluate its completeness, central tendencies, variability, and compliance with recommended biological ranges for aquaponic systems. The dataset comprises 1,823 observations of three key water quality parameters: pH, dissolved oxygen (DO), and temperature. No missing values were detected across the dataset, ensuring consistency for further statistical and machine learning applications.

The descriptive statistics are summarized in Table 1. The mean pH was 6.60 with a standard deviation of 0.75, while DO averaged 5.57 mg/L with moderate variability (standard deviation = 0.82). Temperature showed the lowest variability, with a mean of 24.70 °C and a standard deviation of 0.99, remaining tightly clustered around the optimal range for aquaponic operation. Out-of-range values were primarily detected in pH measurements (92 instances), whereas DO and temperature measurements remained fully within the recommended ranges.

Table 1. Summary statistics of water quality parameters

Parameter	Count	Mean	Std	Min	25%	50%	75%	Max
pH	1823	6.60	0.75	4.50	6.20	6.72	7.10	8.90
Dissolved Oxygen (mg/L)	1823	5.57	0.82	3.10	4.93	5.60	6.30	6.97
Temperature (°C)	1823	24.70	0.99	22.00	24.19	24.75	24.94	27.19

To visually assess variability and the presence of outliers, Figure 3 presents boxplots for each parameter. Median values are highlighted in red, means are shown as black markers, and 95% confidence intervals are represented as error bars. Shaded areas correspond to the biologically recommended operating ranges for aquaponic systems (pH 5–9, DO 0–15 mg/L, and temperature 18–35 °C). The visualization reveals that while DO and temperature remained within acceptable bounds, pH measurements exhibited occasional deviations below the lower threshold, reflecting natural fluctuations in the system.

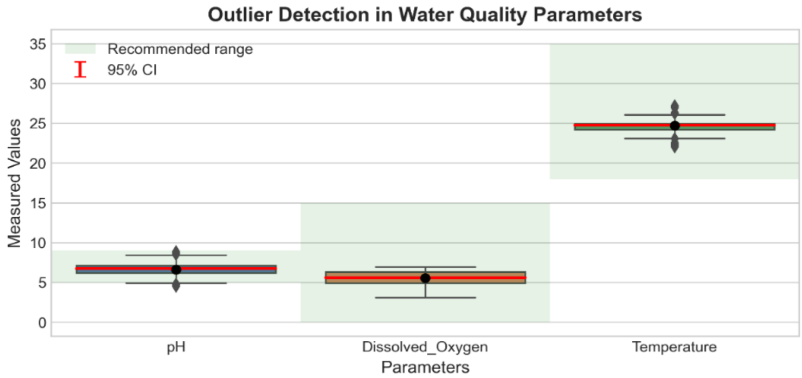

Figure 3. Boxplot of water quality parameters with recommended ranges, mean values, and 95% confidence intervals

3.2. Correlation analysis

To examine the interdependence among water quality parameters, a Pearson correlation analysis was conducted. The correlation matrix is presented in Figure 4, where colors indicate the strength and direction of pairwise associations. The results revealed a moderate positive correlation between pH and dissolved oxygen ($r = 0.55$), a moderate positive correlation between dissolved oxygen and temperature ($r = 0.67$), and a weak positive correlation between pH and temperature ($r = 0.20$). All correlations were statistically significant ($p < 0.001$).

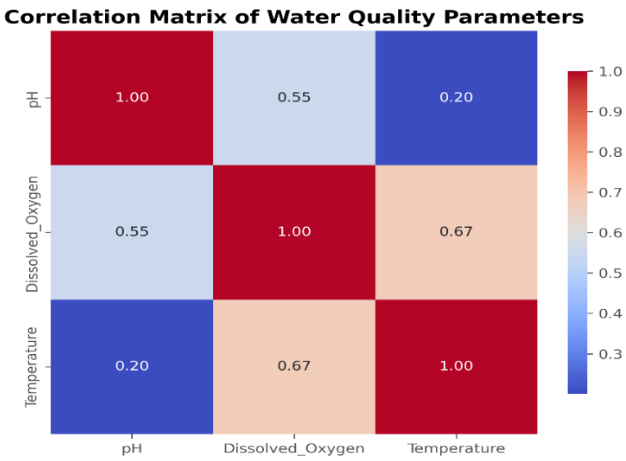

Figure 4. Pearson correlation matrix of water quality parameters (pH, dissolved oxygen, and temperature)

To statistically validate these findings, p-values were computed for each pairwise correlation. As shown in Table 2, all relationships were highly significant ($p < 0.001$), confirming that the observed patterns are unlikely to be attributed to random variability in the dataset.

Table 2. P-values of pairwise correlations among water quality parameters

	pH	Dissolved Oxygen	Temperature
pH	1.000000e+00	2.180403e-143	4.061121e-18
Dissolved Oxygen	2.180403e-143	1.000000e+00	1.731474e-240
Temperature	4.061121e-18	1.731474e-240	1.000000e+00

To further illustrate these associations, Figure 5 depicts the scatterplot and regression line between pH and dissolved oxygen. While the regression line suggests a moderate positive correlation ($r = 0.55$), the wide dispersion confirms that the relationship is not strictly linear. A positive association was expected because photosynthetic activity typically increases both pH and DO, whereas respiration decreases them simultaneously. However, in this system, the expected trend was attenuated by regulatory interventions such as pH dosing and aeration, as well as by natural fluctuations in fish metabolism and microbial processes. These combined factors explain why the relationship appears weaker and more scattered than theoretically anticipated.

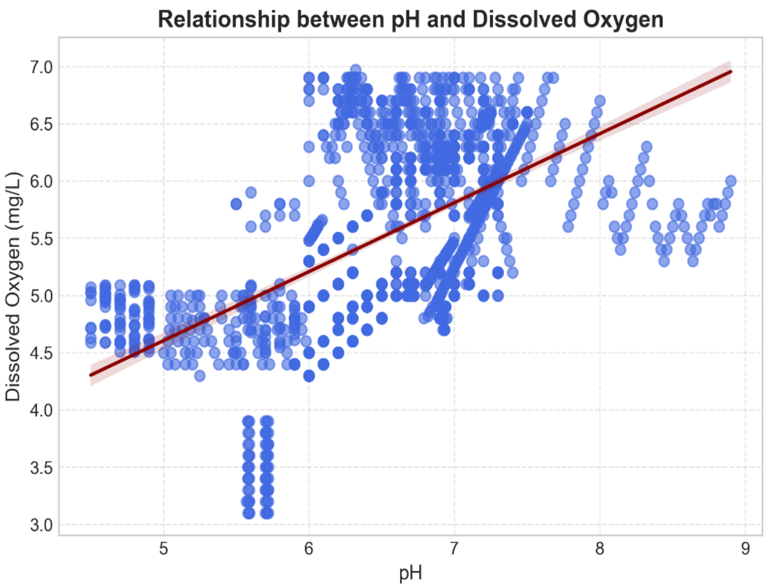

Figure 5. Scatter plot of pH versus dissolved oxygen with linear regression fit, illustrating the positive association between the two parameters

3.3. Principal component analysis (PCA)

To explore patterns and relationships among water quality parameters, a Principal Component Analysis (PCA) was conducted. Before analysis, variables (pH, dissolved oxygen, and temperature) were standardized using z-score normalization (mean = 0, standard deviation = 1) through the StandardScaler function in scikit-learn. This step ensured comparability across variables with different units and scales. The first two principal components captured most of the variance, with PC1 accounting for 65.8% and PC2 for 26.8% of the total variability (Figure 6).

PC1 showed negative loadings for temperature (–0.67), pH (–0.56), and dissolved oxygen (–0.49). PC2 was mainly influenced by dissolved oxygen (+0.78) and pH (–0.62), while temperature contributed minimally (–0.05). The scatterplot (Figure 6), colored by fish condition, revealed three distinct clusters corresponding to the dataset labels. Condition 1 formed a cluster on the left side of the plot, Condition 2 clustered in the upper-right region, and Condition 3 clustered in the lower-right region. Overall, PC1 separated the samples according to combined water quality gradients, while PC2 primarily differentiated observations based on variation in dissolved oxygen and pH.

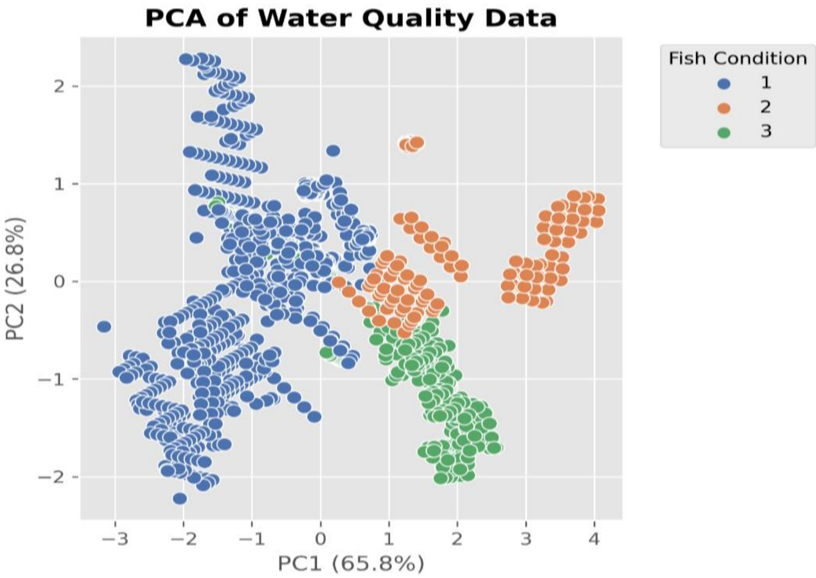

Figure 6. PCA scatterplot of water quality parameters. Points are colored according to fish condition. The axes indicate the proportion of variance explained by each principal component

3.4. Temporal stability of water quality parameters

The long-term evolution of water quality parameters in the aquaponic system was evaluated through time series of pH, dissolved oxygen (DO), and temperature. Statistical indicators (Table 3) confirmed that all three variables exhibited significant decreasing trends according to the Mann–Kendall test ($p < 0.001$). The linear trend analysis showed negative slopes for pH (–0.00045 units per measurement, $R^2 = 0.10$), DO (–0.00089 mg/L per measurement, $R^2 = 0.33$), and temperature (–0.00088 °C per measurement, $R^2 = 0.21$), although the magnitude of decline was more pronounced for DO. Variability was highest for dissolved oxygen (CV = 14.7%), followed by pH (CV = 11.3%), while temperature was the most stable parameter (CV = 4.0%), consistently close to the mean operational value of 24.7 °C.

Table 3. Statistical summary and trend analysis of water quality parameters

Variable	Mean	CV (%)	Linear trend slope	R^2	ADF p-value	MK trend	MK p-value
pH	6.60	11.30	–0.00045	0.10	0.088	Decreasing	<0.001
Dissolved oxygen	5.57	14.74	–0.00089	0.33	0.069	Decreasing	<0.001
Temperature (°C)	24.70	4.04	–0.00088	0.21	0.308	Decreasing	<0.001

To visualize these dynamics, Figure 7 illustrates the trajectories of each parameter across the measurement period. The pH series (top panel, blue) shows an initial range of 7.6–7.8, followed by a sharp decline to approximately 6.0 at the fourth measurement. A partial recovery is observed thereafter, but the series ends with a further drop to about 5.6. Dissolved oxygen (middle panel, green) remained relatively stable between 5.5 and 6.0 mg/L until the fifth measurement, after which it exhibited a marked decline, reaching ~3.5 mg/L at the final observation. By contrast, temperature (bottom panel, red) remained within a narrow range around the mean, confirming its stability relative to the other parameters.

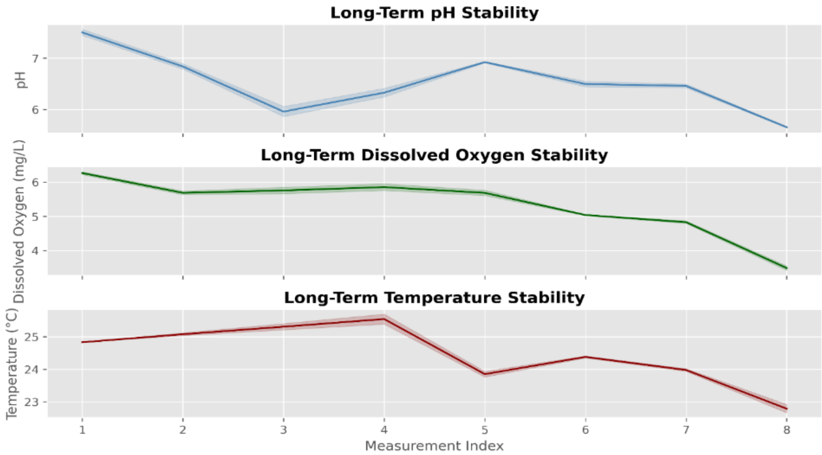

Figure 7. Long-term time series of water quality parameters. Panels show pH (top), dissolved oxygen (middle), and temperature (bottom) over the measurement period, highlighting temporal trends and critical decreases in pH and oxygen

3.5. Autocorrelation and rolling variability

To further investigate temporal dependencies in water quality parameters, autocorrelation functions (ACF) were computed alongside rolling mean and standard deviation bands for pH, dissolved oxygen (DO), and temperature (Figure 8). The pH series exhibits an initial autocorrelation near 1 that gradually decreases but remains positive and significant up to lag ~40–50, indicating long-term memory: if the water has a certain pH today, it is highly likely to maintain similar values in subsequent measurements. Dissolved oxygen shows a similar pattern, remaining significant until approximately lag 35–40, reflecting the continuity of oxygen dynamics influenced by environmental conditions and sustained biological consumption. Temperature displays a nearly identical pattern to pH, with high persistence and a slow decay of autocorrelation, consistent with gradual environmental changes. The rolling mean and standard deviation bands within the same figure highlight periods of higher variability. While temperature remains stable, pH and DO show intermittent fluctuations, corroborating the trends observed in the time series.

Overall, all three water quality parameters demonstrate strong positive autocorrelation at short and medium lags, confirming that their evolution is strongly influenced by prior values. In practical terms, the physicochemical processes governing pH, DO, and temperature exhibit temporal inertia, changing gradually rather than abruptly between consecutive measurements.

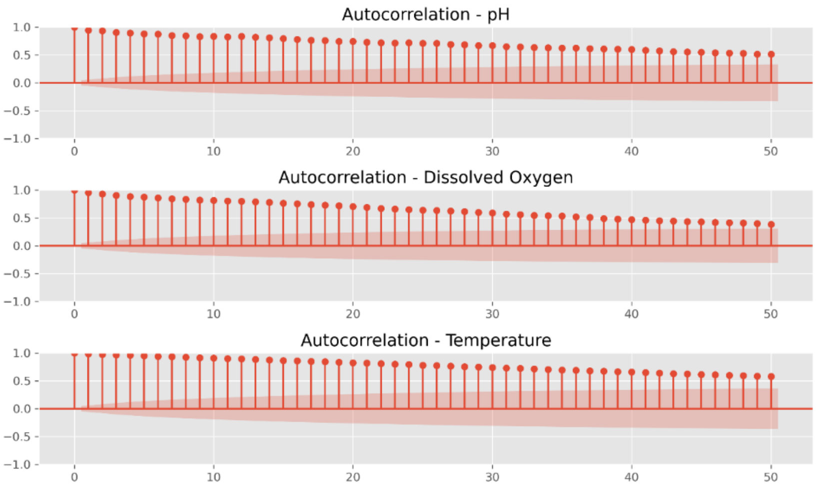

Figure 8. Combined visualization of autocorrelation functions (ACF) and rolling mean ±1 standard deviation bands for pH (top), dissolved oxygen (middle), and temperature (bottom), illustrating both temporal dependence and variability patterns

3.6. Trend analysis with rolling mean and variability

A long-term trend analysis with variability was conducted for pH, dissolved oxygen (DO), and temperature using a 30-point rolling mean and ±1 standard deviation bands (Figure 9). The pH series (top panel, blue) exhibits a general downward trend, decreasing from values around 7.5–8.0 to approximately 6.0–6.2 in the final measurements. The variability, indicated by the shaded band, remains relatively stable for most of the period but shows increased dispersion towards the end. This pattern indicates a progressive acidification of the water, with occasional episodes of short-term fluctuation.

Dissolved oxygen (middle panel, green) begins near 6.2–6.5 mg/L and shows a slight peak around the third to fourth measurement (~6.7 mg/L). Subsequently, a marked decline occurs, reaching approximately 3.5–4.0 mg/L in the last observations. The expansion of the variability band in the final points reflects greater instability in oxygen levels, indicating a sustained loss of dissolved oxygen over time.

Temperature (bottom panel, red) starts around 25 °C, rises to ~27 °C at the fourth measurement, and then gradually decreases to near 23 °C by the end. Variability is more pronounced at intermediate points, reflecting temporary thermal oscillations. Overall, temperature displays a slight downward trend with minor fluctuations throughout the period.

This combined analysis of rolling mean and variability confirms the temporal patterns observed in the raw time series, highlighting progressive acidification, decreasing oxygen availability, and moderate temperature fluctuations, which are critical for maintaining optimal aquaponic system performance. Among these parameters, dissolved oxygen and pH emerged as the most limiting factors, suggesting that oxygen depletion and progressive acidification were the main stressors contributing to reduced fish survival during the study.

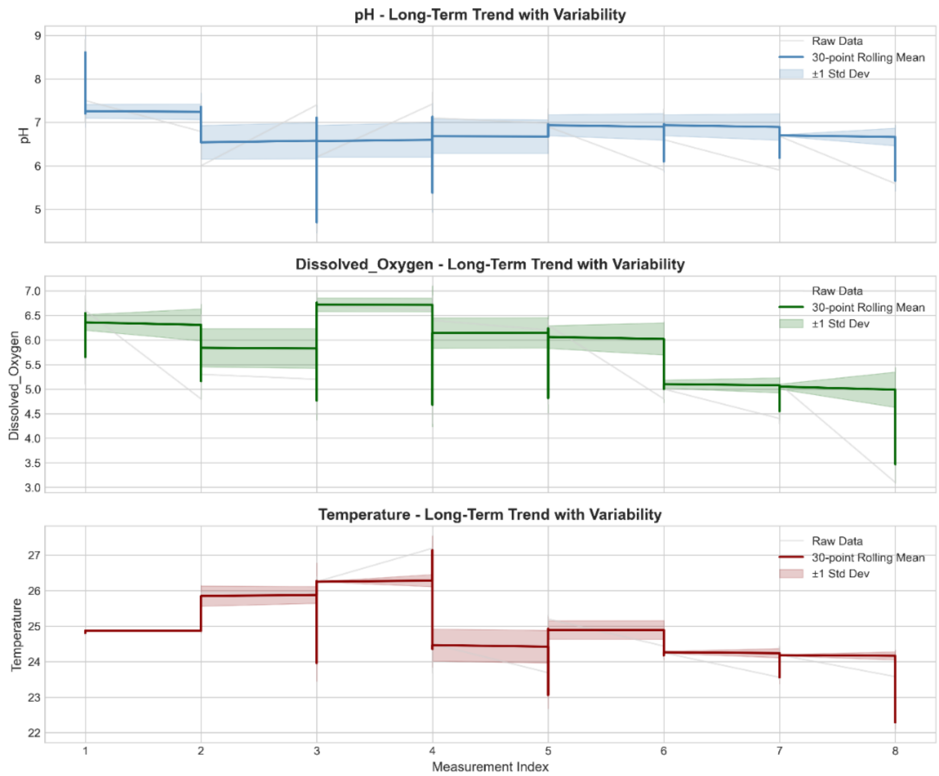

Figure 9. Long-term trends of water quality parameters with 30-point rolling mean and ±1 standard deviation bands for pH (top), dissolved oxygen (middle), and temperature (bottom)

An advanced analysis of water quality parameters was conducted using seasonal-trend decomposition with the seasonal_decompose method (Figure 10) to separate long-term trends from potential cyclical patterns. The decomposition confirmed that the previously observed trends—progressive acidification of pH, declining dissolved oxygen, and moderate cooling of temperature—were consistent across the series, while the seasonal components remained negligible for all parameters, indicating no marked cyclic behavior. These results reinforce that the deterioration of water quality in the aquaponic system followed a continuous, non-cyclic trend, in agreement with the rolling mean and variability analysis (Figure 9).

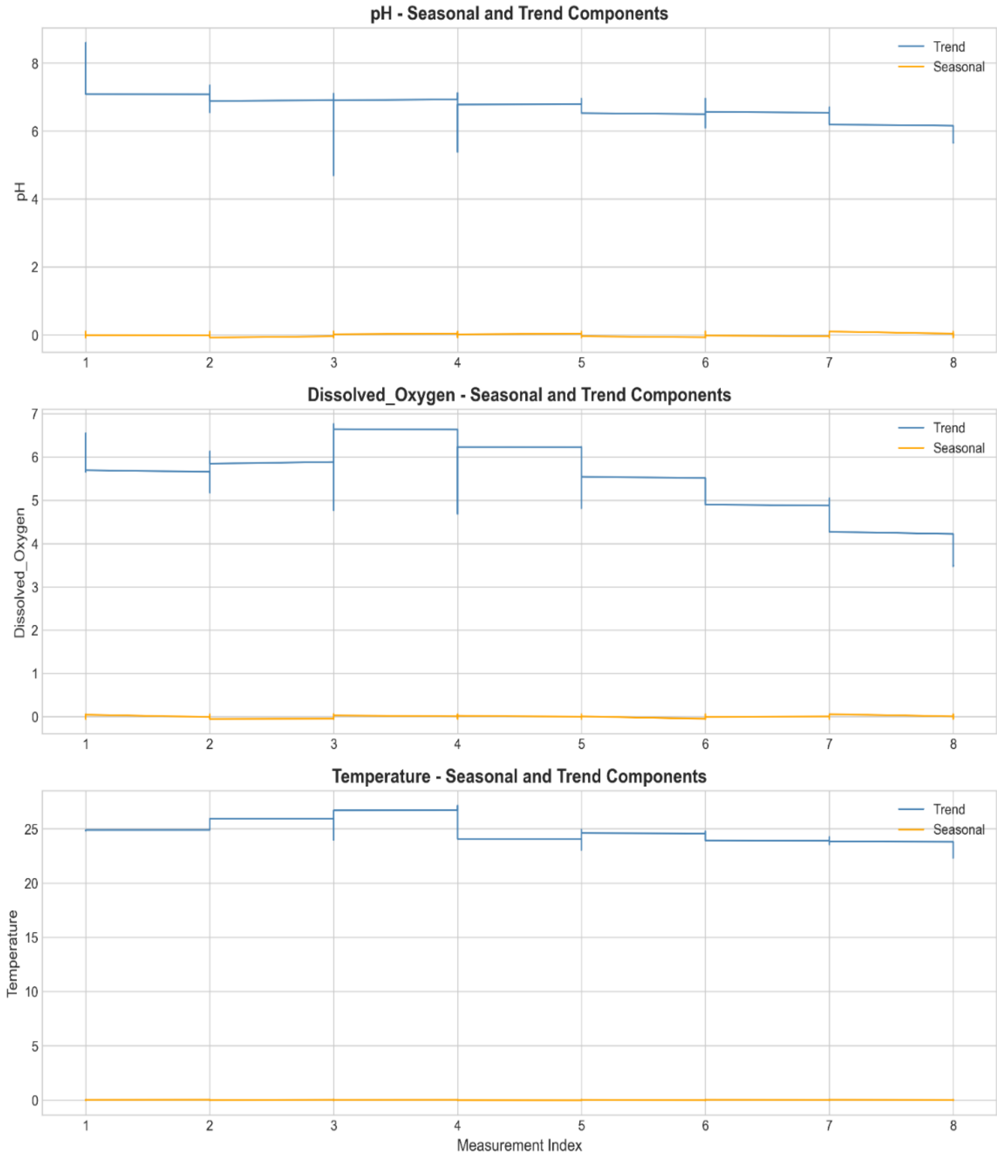

Figure 10. Seasonal-trend decomposition of pH (top), dissolved oxygen (middle), and temperature (bottom), illustrating the long-term downward trends and negligible seasonal components

3.7. Distribution of water quality parameters

The distribution of water quality parameters was analyzed using histograms overlaid with kernel density estimates (KDE) for pH, dissolved oxygen (DO), and temperature (Figure 11). The pH distribution (top panel, blue) is unimodal with a slight left skew. Most measurements occur between 6.8 and 7.2, corresponding to near-neutral water, while lower values between 5 and 6 appear less frequently, indicating occasional episodes of acidification. Overall, the majority of samples maintain a near-neutral pH, with rare deviations toward acidic conditions. Dissolved oxygen (middle panel, green) exhibits a bimodal distribution, with two main peaks around 5 mg/L and 6.2–6.5 mg/L. This pattern suggests the presence of two different conditions, potentially reflecting temporal variations, spatial heterogeneity, or temperature-related effects on oxygen solubility. Temperature (bottom panel, red) shows a multimodal distribution with peaks near 24, 25, and 26 °C, with the highest frequency around 25 °C. The tails of the distribution indicate occasional lower (~22–23 °C) and higher (~27 °C) temperatures, likely associated with environmental or operational variability. Taken together, these distributions indicate that pH remains mostly neutral with occasional acidification events, dissolved oxygen reflects two distinct water quality scenarios, and temperature exhibits broader variability, which could influence both pH and oxygen dynamics.

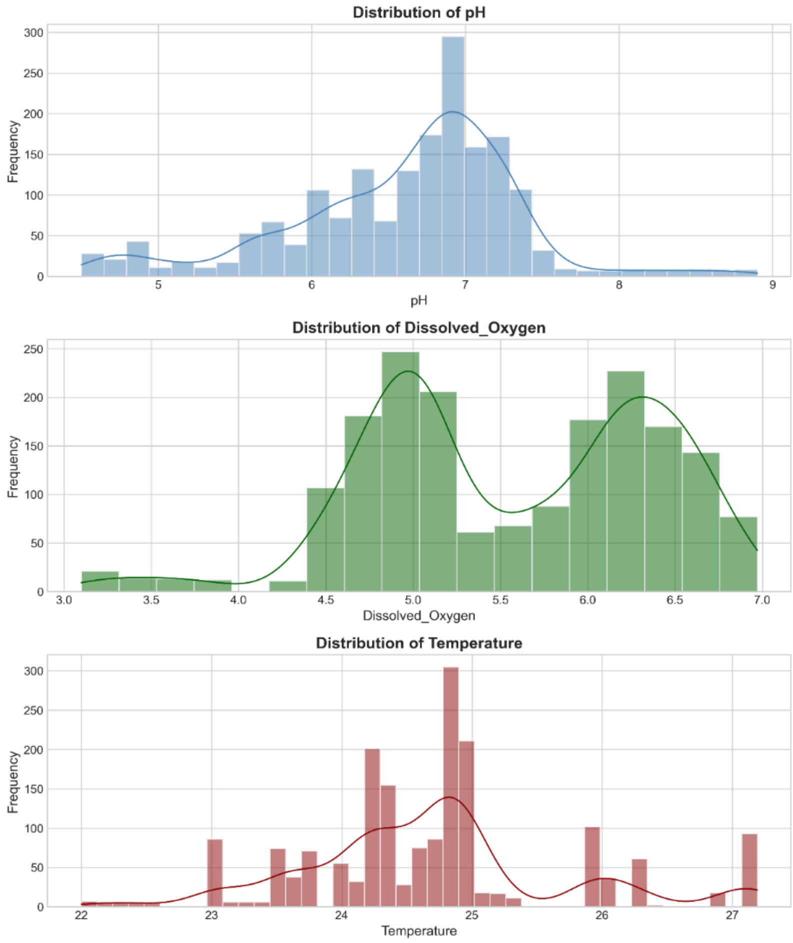

Figure 11. Histograms with KDE for pH (top), dissolved oxygen (middle), and temperature (bottom), illustrating the distribution patterns of water quality parameters in the aquaponic system

4. Conclusion and discussion

Unlike prior studies that mainly focused on parameter regulation through PID or fuzzy control [7], [12], the main contribution of this work lies in validating a long-term, biologically consistent dataset under PID regulation. This dataset captures both stable operating conditions and critical deviations, which are often overlooked but are essential for developing predictive tools. By ensuring completeness, consistency, and realistic variability, the dataset offers a reliable foundation for future AI-based models aimed at improving aquaponic management [13], [15].

The results demonstrate that, although the PID controller maintained average values within acceptable ranges for the species under study—typically a pH of 6.5–8.5 and dissolved oxygen (DO) above 5 mg/L— it failed to prevent critical excursions in both pH and DO. Among the three parameters, pH exhibited the highest proportion of out-of-range values, confirming its critical role in overall stability. Time series and rolling statistics revealed progressive acidification and sustained oxygen depletion, while autocorrelation confirmed that these deviations persisted across consecutive measurements, increasing biological stress. Seasonal-trend decomposition further indicated that these changes were progressive rather than cyclical, reflecting a genuine deterioration of water quality.

Correlation analyses provided additional insights into system dynamics. The strong negative relationship between temperature and DO is consistent with oxygen solubility laws, while the positive association between pH and DO reflects coupled biological and chemical processes. The weaker link between temperature and pH suggests indirect interactions, likely influenced by microbial activity and feeding. Together, these findings emphasize that aquaponic resilience depends on monitoring parameters in an integrated manner, rather than in isolation. Multivariate analysis via PCA revealed distinct clusters of water quality conditions that aligned with different fish performance states. These patterns suggest that simultaneous declines in pH and DO create stress scenarios, reinforcing the need for multi-parameter monitoring to safeguard fish welfare. Similar conclusions have been reported in broader aquaponics and fish welfare studies [5], [15], supporting the generalizability of these results. Overall, the study highlights both the utility and limitations of PID regulation. While effective at stabilizing averages, PID alone cannot anticipate or prevent critical events in dynamic environments. This limitation opens opportunities for advanced predictive and adaptive control strategies. The validated dataset presented here can serve as a benchmark for training machine learning models—including Random Forests, neural networks, and discriminant analysis—to detect early warning signs and optimize system stability [16], [21], [22].

This work was conducted on a domestic-scale aquaponic unit, with a limited number of fish and plants, and using low-cost sensors. While these conditions restrict direct extrapolation to industrial systems, they reflect realistic challenges of small-scale and emerging aquaponic operations. Future research should expand to larger and commercial systems, integrate additional biological indicators such as growth or mortality, and adopt higher-precision instrumentation. By addressing these aspects, predictive control systems can move from experimental validation toward scalable solutions for sustainable aquaponics.

Declaration of competing interest

The authors declare that they have no known financial or non-financial competing interests in any material discussed in this paper.

Funding information

No funding was received from any financial organization to conduct this research.

Acknowledgements

The authors, Rocío Cazes Ortega and Jorge Saúl Fandiño Pelayo, gratefully acknowledge the Bicentennial Excellence Doctoral Scholarships from Minciencias (Becas de Excelencia Doctoral del Bicentenario) for their support during this research.

Author contribution

The contribution to the paper is as follows: Jorge Saúl Fandiño Pelayo and Rocío Cazes Ortega contributed to study conception and design, data collection, analysis and interpretation of results, and draft preparation; Luis Sebastian Mendoza Castellanos and O. Lengerke contributed to the review and critical revision of the manuscript; Luis Guillermo Hernández Rojas contributed to the review of methodology and adjustments in the Results and Discussion sections. All authors approved the final version of the manuscript.

Abbreviations and acronyms

DO: Dissolved Oxygen
PID: Proportional-Integral-Derivative (control)
PCA: Principal Component Analysis
KDE: Kernel Density Estimation
All abbreviations are defined at first mention in the text.

References

[1] S. Goddek, B. Delaide, U. Mankasingh, K. V. Ragnarsdottir, H. Jijakli, and R. Thorarinsdottir, “Challenges of sustainable and commercial aquaponics,” Sustainability, vol. 7, no. 4, pp. 4199–4224, Apr. 2015. [Online]. Available: https://doi.org/10.3390/su7044199

[2] S. Goddek and A. Körner, “A fully integrated simulation model of multi-loop aquaponics: A case study for system sizing in different environments,” Agricultural Systems, vol. 171, pp. 143–154, Mar. 2019. [Online]. Available: https://doi.org/10.1016/j.agsy.2019.01.010

[3] L. A. Ibrahim, H. Shaghaleh, G. M. El-Kassar, M. Abu-Hashim, E. A. Elsadek, and Y. Alhaj Hamoud, “Aquaponics: A sustainable path to food sovereignty and enhanced water use efficiency,” Water, vol. 15, no. 24, p. 4310, Dec. 2023. [Online]. Available: https://doi.org/10.3390/w15244310

[4] B. König, J. Janker, T. Reinhardt, M. Villarroel, and R. Junge, “Analysis of aquaponics as an emerging technological innovation system,” Journal of Cleaner Production, vol. 180, pp. 232–243, Jun. 2018. [Online]. Available: https://doi.org/10.1016/j.jclepro.2018.01.037

[5] H. Y. Yildiz, L. Robaina, J. Pirhonen, E. Mente, D. Domínguez, and G. Parisi, “Fish welfare in aquaponic systems: Its relation to water quality with an emphasis on feed and faeces—A review,” Water, vol. 9, no. 1, p. 13, Jan. 2017. [Online]. Available: https://doi.org/10.3390/w9010013

[6] S. Goddek, A. Joyce, B. Kotzen, and G. M. Burnell, “Aquaponics and global food challenges,” in Aquaponics Food Production Systems, S. Goddek, A. Joyce, B. Kotzen, and G. M. Burnell, Eds. Cham, Switzerland: Springer, 2019, pp. 3–17. [Online]. Available: https://doi.org/10.1007/978-3-030-15943-6_1

[7] O. Supriadi, A. Sunardi, H. A. Baskara, and A. Safei, “Controlling pH and temperature aquaponics use proportional control with Arduino and Raspberry,” IOP Conference Series: Materials Science and Engineering, vol. 550, no. 1, p. 012016, Jul. 2019. [Online]. Available: https://doi.org/10.1088/1757-899X/550/1/012016

[8] C. L. Kok, I. M. B. P. Kusuma, Y. Y. Koh, H. Tang, and A. B. Lim, “Smart aquaponics: An automated water quality management system for sustainable urban agriculture,” Electronics, vol. 13, no. 5, art. 820, Feb. 2024. [Online]. Available: https://doi.org/10.3390/electronics13050820

[9] H.-C. Li, K.-W. Yu, C.-H. Lien, C. Lin, C.-R. Yu, and S. Vaidyanathan, “Improving aquaculture water quality using dual-input fuzzy logic control for ammonia nitrogen management,” Journal of Marine Science and Engineering, vol. 11, no. 6, p. 1109, Jun. 2023. [Online]. Available: https://doi.org/10.3390/jmse11061109

[10] S. Graber and R. Junge, “Aquaponic systems: Nutrient recycling from fish wastewater by vegetable production,” Desalination, vol. 246, nos. 1–3, pp. 147–156, Feb. 2009. [Online]. Available: https://doi.org/10.1016/j.desal.2008.03.048

[11] D. Dhinakaran, S. Gopalakrishnan, M. D. Manigandan, and T. P. Anish, “IoT-based environmental control system for fish farms with sensor integration and machine learning decision support,” arXiv preprint arXiv:2311.04258, Nov. 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2311.04258

[12] X. Yang, S. Zhang, J. Liu, Q. Gao, S. Dong, and C. Zhou, “Deep learning for smart fish farming: Applications, opportunities and challenges,” Reviews in Aquaculture, vol. 13, no. 1, pp. 66–90, 2021. [Online]. Available: https://doi.org/10.1111/raq.12464

[13] R. Hossam, A. Heakl, and W. Gomaa, “Precision aquaculture: An integrated computer vision and IoT approach for optimized tilapia feeding (Preprint),” arXiv, Sep. 13, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2409.08695

[14] G. Ciarlo, E. Bonica, B. Bosio, and N. Bonavita, “Assessment and testing of sensor validation algorithms for environmental monitoring applications,” Chemical Engineering Transactions, vol. 57, pp. 331–336, 2017. [Online]. Available: https://doi.org/10.3303/CET1757056

[15] H. Segner, H. Sundh, K. Buchmann, J. Douxfils, K. S. Sundell, C. Mathieu, N. Ruane, F. Jutfelt, H. Toften, and L. Vaughan, “Health of farmed fish: Its relation to fish welfare and its utility as welfare indicator,” Fish Physiology and Biochemistry, vol. 38, no. 1, pp. 85–105, Feb. 2012. [Online]. Available: https://doi.org/10.1007/s10695-011-9517-9

[16] P. Chandramenon, A. Aggoun, and F. Tchuenbou-Magaia, “Smart approaches to aquaponics 4.0 with focus on water quality − Comprehensive review,” Computers and Electronics in Agriculture, vol. 216, p. 109256, 2024. [Online]. Available: https://doi.org/10.1016/j.compag.2024.109256

[17] A. A. Bracino, R. S. Concepcion II, D. G. D. Evangelista, R. R. P. Vicerra, and E. P. Dadios, “Fuzzy logic-based automated pH and temperature control in aquaponics systems,” Journal of Control and Intelligent Engineering Applications, vol. 5, no. 1, pp. 1–10, 2020. [Online]. Available: https://www.dlsu.edu.ph/wp-content/uploads/pdf/research/journals/jciea/vol-5-1/4bracino.pdf

[18] S. Wan, K. Zhao, Z. Lu, J. Li, T. Lu, and H. Wang, “A modularized IoT monitoring system with edge-computing for aquaponics,” Sensors, vol. 22, no. 23, p. 9260, 2022. [Online]. Available: https://doi.org/10.3390/s22239260

[19] M. G. A. Bautista, M. G. B. Palconit, M. A. Rosales, R. S. Concepcion II, A. A. Bandala, E. P. Dadios, and B. Duarte, “Fuzzy logic-based adaptive aquaculture water monitoring system based on instantaneous limnological parameters,” Journal of Advanced Computational Intelligence and Intelligent Informatics, vol. 26, no. 6, pp. 937–943, 2022. [Online]. Available: https://doi.org/10.20965/jaciii.2022.p0937

[20] A. Jayadi, S. Samsugi, E. K. Ardilles, and F. D. Adhinata, “Monitoring water quality for catfish ponds using fuzzy Mamdani method with Internet of Things,” in Proc. 2022 Int. Conf. Inf. Technol. Res. Innov. (ICITRI), 2022, pp. 77–82. [Online]. Available: https://doi.org/10.1109/ICITRI56423.2022.9970242

[21] T. A. Firdaus, E. D. Widianto, and D. Eridani, “Designing and implementing IoT-based water quality monitoring and control system of a pilot scale deep flow technique aquaponics for enhanced crop-fish production,” E3S Web Conf., vol. 448, p. 02031, 2023. [Online]. Available: https://doi.org/10.1051/e3sconf/202344802031

[22] M. Anila and O. Daramola, “Applications, technologies, and evaluation methods in smart aquaponics: A systematic literature review,” Artificial Intelligence Review, vol. 58, article 25, 2025. [Online]. Available: https://doi.org/10.1007/s10462-024-11003-x

[23] T. Khaoula, R. Ait Abdelouahid, I. Ezzahoui, and A. Marzak, “Architecture design of monitoring and controlling of IoT-based aquaponics system powered by solar energy,” Procedia Computer Science, vol. 191, pp. 493–498, 2021. [Online]. Available: https://doi.org/10.1016/j.procs.2021.07.063