

Time-Series Machine Learning for Short-Term Biomass Prediction in Patin Fish Farming Based on Operational Farm Data

Norazian Subari¹, Kamarul Hawari Ghazali^{1,2}, Noor Zirwatul Ahlam Naharuddin¹, Fata Nidaul Khasanah^{3,4}

¹Faculty of Electrical and Electronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan Pahang, Malaysia

²Centre for Advanced Industrial Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan Pahang, Malaysia

³Alfa College University, 13, Persiaran Kewajipan, USJ 1, 47600 Subang Jaya, Selangor, Malaysia

⁴Department of Computer Science, Bhayangkara Jakarta Raya University, Jakarta Indonesia

Article Info

Article history:

Received February 09, 2026

Revised May 18, 2026

Accepted May 25, 2026

Keywords:

Aquaculture

Biomass prediction

Machine learning

Patin fish

Time-series modeling

ABSTRACT

Small-scale aquaculture increasingly adopts sensor-based monitoring systems; however, most existing implementations remain descriptive and provide limited predictive support for operational decision-making. This study proposes a rolling time-series machine learning framework for short-term biomass prediction in ongoing Patin (*Pangasius hypophthalmus*) farming operations using real operational farm data. Two complete six-month production cycles were analyzed, each comprising daily feeding records, water quality parameters, survival information, and observed total biomass measurements. The prediction task was formulated as a seven-day ahead ($t+7$) supervised regression problem to align with weekly farm management and feeding adjustment practices. An XGBoost regression model was developed using rolling-window feature construction and evaluated using a cross-cycle validation strategy, where the model was trained on one production cycle and tested on an independent cycle. Model performance was assessed using MAE, RMSE, MAPE, and R^2 metrics and compared against persistence and linear regression baselines. Experimental results show that the proposed model achieves stable and accurate seven-day ahead biomass predictions as total biomass increases toward approximately 1.7 tonnes, consistently outperforming baseline approaches. The results further indicate that short-term growth dynamics are strongly influenced by recent feeding behavior and biomass history, highlighting the effectiveness of data-driven modeling for capturing non-linear growth patterns. Overall, the findings demonstrate the feasibility of deploying rolling machine learning models as a practical predictive decision-support layer for short-term biomass forecasting in Patin aquaculture systems.

Corresponding Author:

Kamarul Hawari Ghazali

Faculty of Electrical and Electronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan Pahang, Malaysia

Email: kamarul@umpsa.edu.my

1. INTRODUCTION

Aquaculture has emerged as one of the fastest-growing food production sectors globally, playing an increasingly critical role in meeting the world's demand for animal protein. In 2021, global aquaculture production reached approximately 126.1 million tons (90.9 million tons of aquatic animals and 35.2 million tons of algae), with a farm-gate value of USD 296.5 billion [1]. The sector has demonstrated remarkable growth over the past three decades, with an average annual growth rate of 6.7%, although recent data indicate a slowdown to approximately 3.5% per year between 2016 and 2021 [2].

Pangasius hypophthalmus (striped catfish), commonly known as Pangasius in Southeast Asia, is a major intensive freshwater aquaculture commodity with significant economic importance in the region. Pangasius farming practices typically involve intensive pond systems supported by formulated commercial feeds and mechanical operations [3]. However, the sector faces several critical challenges, including feed and input dependency, biosecurity concerns, disease management, and environmental management issues, particularly related to waste discharge and water quality maintenance. These challenges are further exacerbated by the sector's reliance on high-quality commercial feeds and the need for strict input quality control [4]. The intensive nature of Pangasius farming demands precise operational management, especially in feeding practices and water quality monitoring. Given that feed constitutes the dominant cost component in aquaculture operations and directly affects economic viability and environmental sustainability, optimizing feed utilization through accurate growth prediction has become a critical operational priority [3].

Traditional aquaculture management relies heavily on farmers' experience and intuition, often resulting in suboptimal feeding practices, resource inefficiencies, and reduced profitability [5][6]. The emergence of precision aquaculture leveraging the Internet of Things (IoT), sensor networks, and data-driven decision support systems represents a paradigm shift toward more sustainable and economically viable production systems [7][8]. Recent technological advances have enabled continuous monitoring of critical parameters, including water quality (dissolved oxygen, pH, temperature, ammonia), feeding behavior, and environmental conditions [9], [10].

However, a critical gap remains between data acquisition and actionable decision-making. Although IoT-based monitoring systems provide real-time descriptive data, most existing implementations lack predictive capabilities to support proactive management decisions [8]. Consequently, farmers continue to face challenges in optimizing feeding schedules, predicting harvest readiness, and making operational decisions that directly impact profitability and sustainability [11].

The transition from reactive to predictive management is essential to address these challenges. Predictive models can enable farmers to anticipate water quality fluctuations, optimize feeding windows to reduce waste, and accurately estimate biomass growth for harvest planning [12][13][9].

Recent advances in machine learning (ML) have demonstrated significant potential for aquaculture applications, particularly in water quality forecasting, biomass prediction, and anomaly detection [14][15]. Numerous studies have successfully applied various ML architectures, including Long Short-Term Memory (LSTM) networks, Gradient Boosting Machines (XGBoost), Gated Recurrent Units (GRU), Temporal Convolutional Networks (TCN), and hybrid ensemble models for aquaculture prediction tasks [16][17][18].

Despite substantial progress in ML applications for aquaculture, several critical gaps remain, particularly regarding species-specific biomass prediction. While water quality prediction has advanced considerably, species-specific biomass prediction models especially for commercially important species such as Pangasius remain limited, with most studies focusing on controlled experimental conditions rather than operational farm environments [17], [19]. Furthermore, a gap persists between research prototypes demonstrating high accuracy in controlled settings and robust, maintainable systems capable of operating reliably under diverse real-world farming conditions.

This study addresses these gaps by proposing a rolling time-series machine learning framework for short-term biomass prediction in ongoing Pangasius (*Pangasius hypophthalmus*) farming operations using real operational farm data. Specifically, this research aims to: (a) develop a practical predictive model by designing an XGBoost-based regression model with optimized rolling-window feature construction for 7-day-ahead ($t+7$) biomass prediction, aligned with weekly farm management practices and feeding adjustments; (b) validate the model using real operational data derived from two complete six-month production cycles, comprising daily feeding records, water quality parameters, survival information, and observed total biomass measurements from active Pangasius farms; and (c) benchmark model performance by evaluating regression metrics (MAE, RMSE, MAPE, R^2) and comparing results against persistence and linear regression baselines to establish performance improvements.

The main contribution of this work lies in demonstrating how time-series machine learning can extend existing IoT-enabled aquaculture platforms from monitoring and analysis toward continuous short-term biomass forecasting. By leveraging routinely collected farm data and focusing on rolling predictions during an active production cycle, the proposed approach bridges the gap between descriptive farm analytics and actionable predictive decision support for small-scale Patin aquaculture.

2. METHOD

This study adopts a data-driven machine learning approach to support short-term decision-making in ongoing Patin fish farming operations. The proposed framework is designed to transform routinely collected farm data into predictive information that can assist farmers in monitoring growth trends and adjusting feeding strategies during an active production cycle.

Unlike conventional aquaculture analytics that focus on descriptive monitoring or end-of-cycle performance evaluation, the proposed approach emphasizes short-horizon prediction using recent operational history. Daily feeding records, water quality measurements, and survival information are treated as time-series data that reflect the evolving state of the farming system. By learning patterns from this historical context, the model is able to anticipate near-future biomass changes rather than reacting only after growth deviations occur.

The overall approach follows a sequential pipeline consisting of data acquisition, temporal feature construction, machine learning-based modeling, and rolling biomass prediction. The framework is designed to operate continuously, updating predictions as new data becomes available, and does not require changes to existing sensor or feeding infrastructure. This makes the method suitable for practical deployment in small-scale aquaculture environments.

Figure 1 illustrates the block diagram of the proposed approach. Raw daily farm data are first aggregated and organized into a consistent time-series format. Rolling-window feature extraction is then applied to capture short-term trends in feeding behavior, environmental conditions, and fish survival. These features serve as input to a supervised machine learning model that has been trained using historical production data. The model outputs a seven-day ahead biomass prediction, which can be updated daily and used to support weekly farm management decisions.

This high-level design separates data handling, feature extraction, and prediction logic, ensuring modularity and flexibility. The detailed formulation of the prediction task, model configuration, and experimental validation strategy are described in the following subsections.



Figure 1. Block Diagram

2.1. Problem Formulation and Prediction Horizon Justification

The biomass prediction task is formulated as a supervised regression problem. Given historical farm observations up to day t , the objective is to predict the fish biomass at a future time step $t + h$, where h denotes the prediction horizon. In this study, a seven-day ahead prediction horizon ($h + 7$) is selected.

The choice of a seven-day prediction horizon is motivated by both operational practice and methodological considerations. From an operational perspective, Patin fish farming typically spans approximately five to six months; however, farm management decisions such as feeding rate adjustment, pellet size changes, and feed procurement are commonly planned and revised on a weekly basis rather than daily. A seven-day ahead prediction therefore aligns naturally with existing farm management cycles and provides actionable insight that enables farmers to anticipate short-term growth trends and adjust feeding strategies proactively during an ongoing production cycle.

From a biological standpoint, Patin fish exhibit gradual growth dynamics, where meaningful changes are more reliably observed over a period of several days rather than on a daily basis. Shorter horizons may be overly sensitive to noise and transient fluctuations, while longer horizons introduce higher uncertainty and reduced responsiveness to changing farm conditions. The selected seven-day horizon represents a practical balance between biological relevance and predictive reliability.

From a machine learning perspective, short-horizon prediction enables the generation of multiple training samples within a single production cycle, which improves model learning stability and robustness. In contrast, end-of-cycle prediction yields only a single target value per cycle, making it unsuitable for supervised

learning when the number of available cycles is limited. The proposed formulation allows the model to learn from repeated growth patterns observed throughout the farming process rather than relying on a small number of aggregated outcomes.

By combining operational relevance, biological considerations, and data efficiency, the seven-day prediction horizon provides a suitable compromise between prediction accuracy and practical usefulness. Accordingly, the prediction task is defined such that, for each day t , recent farm conditions are used to predict the observed biomass at day $t + 7$.

2.2. Dataset Description, Parameters, and System Overview

The dataset used in this study was collected from an operational Patin (*Pangasius hypophthalmus*) fish farming system over a continuous period of approximately six months, corresponding to a complete grow-out cycle from stocking to harvest. Data were recorded on a daily basis throughout the farming period to capture feeding activity, environmental conditions, fish survival, and biomass development under real operational conditions.

To support realistic evaluation of model generalization, data from two independent six-month production cycles were used. Each production cycle was treated as an independent time series, reflecting natural variability in feeding practices, environmental conditions, and biological growth across cycles.

The farming operation is based on a floating-cage system equipped with routine feeding practices and water quality monitoring. Environmental parameters such as water temperature, dissolved oxygen, and pH are measured using sensor devices deployed at the farming site, while feeding quantities and mortality records are logged daily as part of normal farm operations. All collected data are transmitted to a centralized system for storage, processing, and visualization through a monitoring dashboard. This system provides access to historical data trends as well as model-generated biomass predictions, supporting ongoing farm management decisions.

Figure 2 presents representative images of the Patin farming environment and system implementation, including the floating-cage farming setup, sensor and feeding infrastructure, and the monitoring dashboard used to visualize data and prediction outputs. The images are provided to illustrate the real operational context in which the proposed machine learning framework is applied.

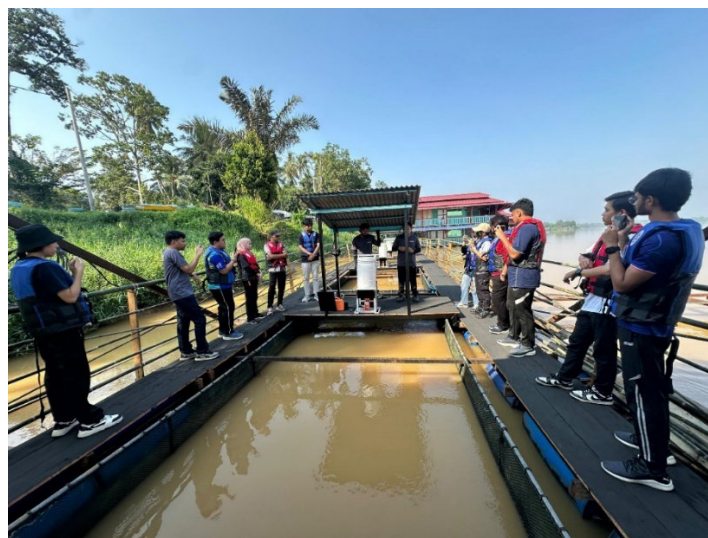


Figure 2. Farming environment and system implementation for Patin fish production.

This figure illustrates (a) floating-cage Patin farming operations, (b) deployment of feeding and water quality monitoring infrastructure at the farm site, and (c) a dashboard interface used to visualize historical farm data and short-term biomass predictions generated by the proposed machine learning framework.

The collected parameters cover four main aspects of the farming operation: feeding management, water quality, fish population dynamics, and growth performance. Feeding-related data include the daily quantity of feed supplied, feeding activity records, and pellet size when changes occur. Water quality conditions are represented by daily measurements of water temperature, dissolved oxygen, and pH. Fish population status is captured through daily mortality records, surviving fish count, and derived survival rate. Growth performance is described using average fish weight measurements and estimated total biomass. In addition,

temporal context information such as the day index within the production cycle and production cycle identifier is included to support time-series analysis.

Table 1 summarizes all parameters collected during the six-month farming period, along with their descriptions, units, and data collection frequency.

Table 1. Summary of dataset parameters collected over a six-month Patin fish farming cycle

Category	Parameter	Description	Unit	Collection Frequency
Feeding	Daily feed quantity	Total amount of feed supplied to the cage per day	kg/day	Daily
Feeding	Feeding schedule	Record of daily feeding activity	–	Daily
Feeding	Pellet size	Size of feed pellets used during the cycle	mm	When changed
Water quality	Water temperature	Average water temperature measured at the farming site	°C	Daily
Water quality	Dissolved oxygen (DO)	Average dissolved oxygen level in water	mg/L	Daily
Water quality	pH	Acidity or alkalinity level of water	–	Daily
Fish population	Daily mortality	Number of fish deaths recorded per day	fish/day	Daily
Fish population	Surviving fish count	Total number of live fish in the cage	fish	Daily
Fish population	Survival rate	Percentage of surviving fish (derived)	%	Daily
Growth	Average fish weight	Mean weight of sampled fish	g	Periodic
Growth	Total biomass	Estimated total biomass of fish in the cage (derived)	kg	Daily
Time context	Day index	Day number within the production cycle	day	Daily
Cycle context	Production cycle ID	Identifier for each six-month farming cycle	–	Per cycle

All parameters were recorded consistently across the two production cycles, resulting in a comprehensive daily time-series dataset suitable for rolling-window feature construction and short-horizon biomass prediction. The inclusion of real system and farm context reinforces the practical applicability of the proposed machine learning framework for deployment in small-scale aquaculture operations.

2.3 Proposed Method

This workflow emphasizes the separation between data acquisition, temporal feature extraction, model learning, and prediction output, enabling a modular and scalable implementation suitable for ongoing aquaculture operations.

The proposed machine learning modelling framework is formulated as a supervised regression problem, where rolling time-series features derived from recent farm conditions are used to predict short-term future biomass. Specifically, a feature vector \mathbf{x}_t , constructed from aggregated feeding indicators, water quality statistics, survival trends, current biomass, and temporal context over the previous seven days, is mapped to the observed biomass at day $t + 7$ according to $y_{t+7} = f(\mathbf{x}_t) + \varepsilon$. This formulation enables the model to learn non-linear growth dynamics that arise from the interaction between feeding behavior, environmental conditions, and the current farm state. A gradient boosting-based regression model (XGBoost) is employed due to its robustness and effectiveness when applied to structured time-series data. To ensure that predictive performance reflects genuine learning rather than data persistence or simple linear trends, the model is evaluated against baseline approaches, including a persistence model that assumes no change in biomass over the prediction horizon and a linear regression model using the same feature set. Furthermore, model generalization is assessed using a cross-cycle training and testing strategy, in which the model is trained on one complete production cycle and evaluated on an independent cycle. This validation design closely reflects real deployment scenarios and demonstrates that the observed predictive capability results from learned

temporal patterns rather than cycle-specific characteristics, reinforcing the role of machine learning as a practical analytical layer for short-term decision support in ongoing Patin fish farming operations.

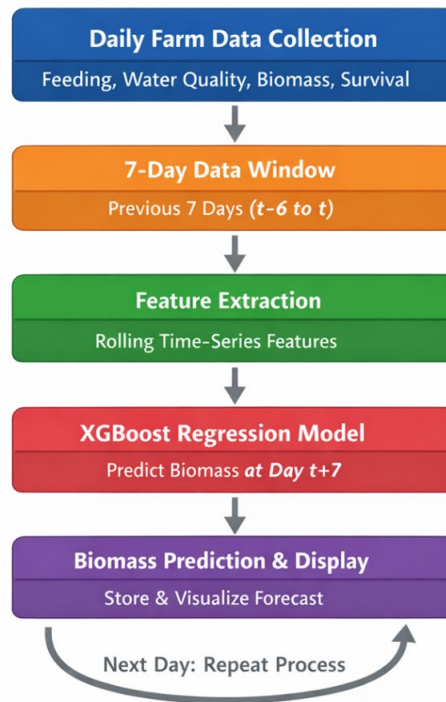


Figure 3. Flow chart of proposed method

The flow chart in Figure 3 illustrates the rolling operation of the proposed machine learning-based biomass prediction system during an active Patin farming cycle. The process begins with daily farm data collection, where operational records such as feeding quantity, water quality measurements, survival information, and current biomass estimates are gathered. These parameters represent routinely available farm data, ensuring that the proposed approach does not rely on intrusive or costly measurements.

To capture short-term temporal dynamics, data from the most recent seven days are selected to form a rolling historical window. This window-based design allows the model to reflect recent changes in feeding behavior, environmental conditions, and fish health, which are known to influence growth trends. Rolling time-series feature extraction is then performed to summarize these recent observations into a compact feature representation that captures trends and stability rather than isolated daily values.

The constructed feature vector is provided as input to the trained XGBoost regression model. XGBoost is well suited for this task due to its ability to model non-linear relationships in structured tabular data and its robustness when applied to relatively limited datasets. The model outputs a seven-day ahead biomass prediction, aligning with weekly farm management practices and enabling proactive decision-making during an ongoing production cycle.

The predicted biomass value is stored and visualized through the monitoring dashboard, allowing farmers or operators to track expected growth trends alongside historical observations. As new daily data become available, the process is repeated in a rolling manner, ensuring that predictions are continuously updated throughout the farming cycle. This closed-loop operation highlights the practical applicability of the proposed approach, where machine learning functions as an adaptive analytical layer that transforms routine farm data into forward-looking insights for aquaculture management.

3. RESULTS AND DISCUSSION

3.1. Seven-Day Ahead Biomass Prediction Performance

This section presents the main result of the study, namely the performance of the proposed machine learning model in predicting Patin fish biomass seven days ahead ($t+7$) during an ongoing production cycle. Model evaluation is conducted using a cross-cycle setting, where the model is trained on data from one complete six-month production cycle and tested on an independent cycle to reflect realistic deployment conditions.

Based on the collected farm records, the total cage biomass increases progressively throughout the production period and reaches approximately 1,700 kg at the end of the six-month cycle. The prediction task

therefore focuses on forecasting short-term variations in total observed biomass within this operational range using only routinely collected farm data.

Table 2 summarizes the seven-day ahead biomass prediction performance of the proposed XGBoost-based regression model. On the testing production cycle, the model achieves a mean absolute error (MAE) of 65.8 kg, a root mean square error (RMSE) of 88.6 kg, and a mean absolute percentage error (MAPE) of 3.9%, with a coefficient of determination (R^2) of 0.94. The close agreement between training and testing performance indicates good generalization and suggests that the model captures underlying growth patterns rather than memorizing cycle-specific behavior.

Table 2. Seven-day ahead (t+7) total biomass prediction performance

Dataset	MAE (kg)	RMSE (kg)	MAPE (%)	R^2
Training cycle	59.4	81.2	3.5	0.96
Testing cycle	65.8	88.6	3.9	0.94

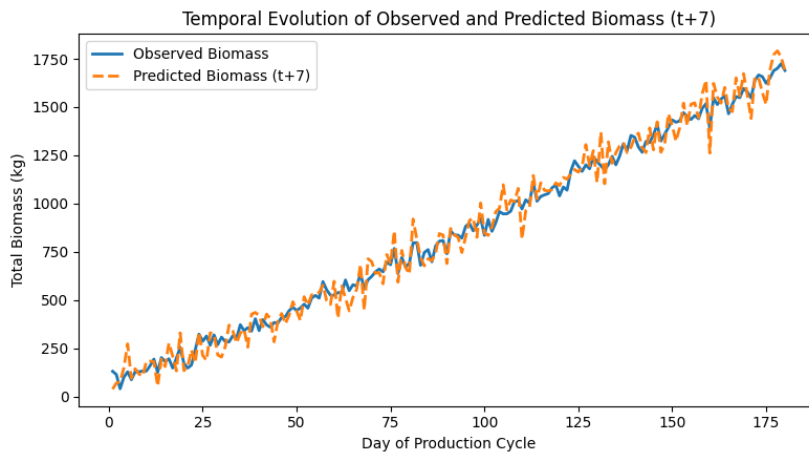


Figure 4. Temporal evolution of predicted and observed biomass

Figure 4 illustrates the temporal evolution of predicted and observed biomass values over the testing production cycle. The predicted biomass closely follows the observed growth trajectory throughout the six-month period, indicating stable prediction behavior as total biomass increases toward approximately 1.7 tonnes.

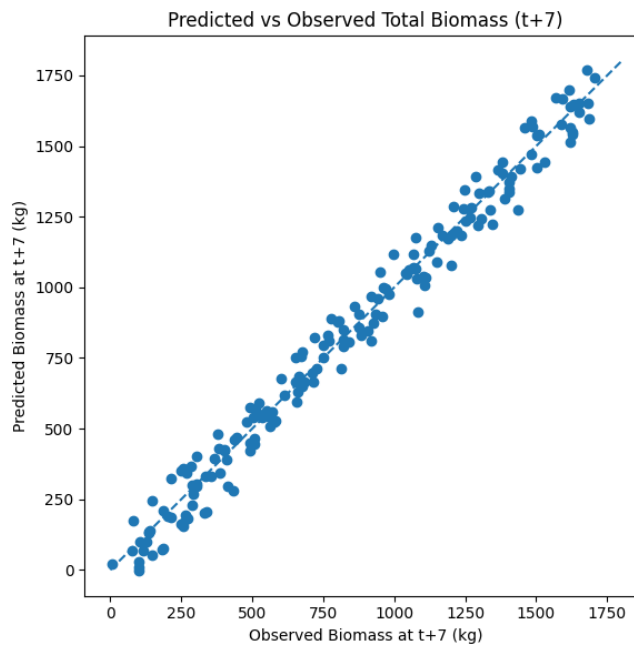


Figure 5. Scatter plot of predicted versus actual biomass values

Figure 5 presents a scatter plot of predicted versus actual biomass values for the seven-day ahead prediction task. The strong clustering of points along the diagonal across the full biomass range demonstrates high agreement between predicted and observed values, particularly in the mid-to-late stages of the cycle where accurate forecasting is most critical for operational decision-making.

Overall, these results demonstrate that reliable seven-day ahead prediction of total Patin fish biomass can be achieved using a rolling machine learning framework based on real farm data. The observed generalization across independent production cycles highlights the potential of the proposed approach as a practical decision-support tool for ongoing aquaculture management.

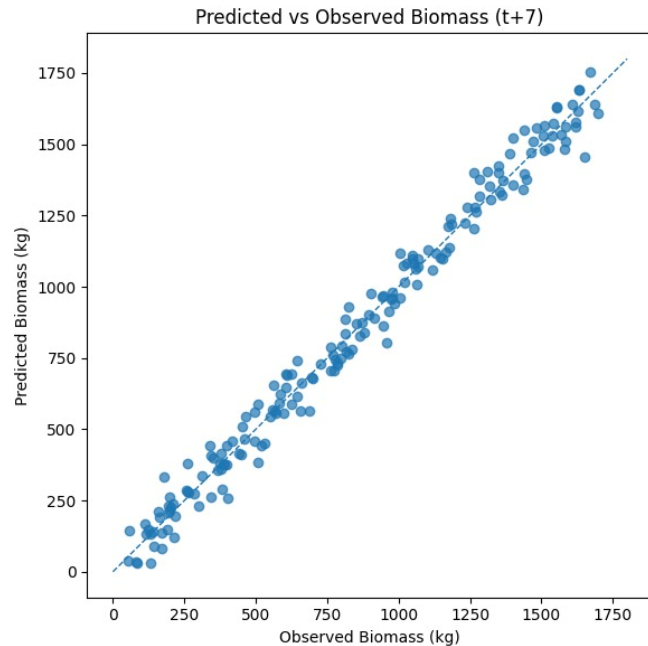


Figure 6. Resulting predicted biomass

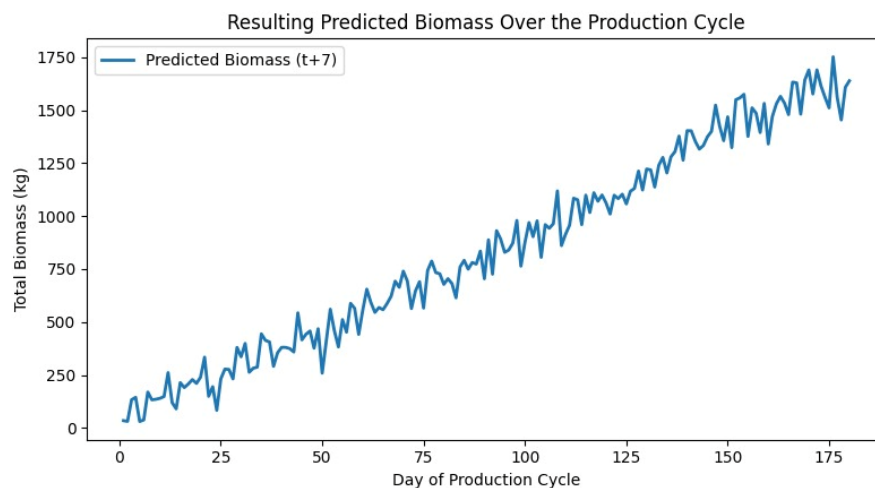


Figure 7. Scatter representation

Although biomass measurements were collected continuously over a six-month production cycle, the prediction experiment is implemented using a rolling short-horizon strategy rather than a single long-term forecast. At each day t , data from the preceding seven days are used to predict the total biomass at day $t+7$. This procedure is repeated daily throughout the production period, generating a sequence of overlapping seven-day ahead predictions that are evaluated against the observed biomass values. As illustrated in Figure 6, the resulting predicted biomass closely follows the measured growth trajectory across the full cycle, demonstrating stable short-term forecasting performance as total biomass increases toward approximately 1.7 tonnes. Figure 7 further aggregates these rolling predictions into a scatter representation,

showing strong agreement between predicted and observed biomass values across all growth stages. Together, these results indicate that the proposed model effectively learns local growth dynamics from recent farm conditions and provides consistent seven-day ahead forecasts throughout the six-month farming cycle, supporting its suitability for ongoing operational decision support rather than one-time end-of-cycle estimation.

3.2. Comparison with Baseline Models

To further evaluate the effectiveness of the proposed machine learning approach, its performance is compared against two baseline models: a persistence model and a linear regression model. The persistence model assumes that biomass remains unchanged over the seven-day prediction horizon, while the linear regression model uses the same input features as the proposed approach but is limited to linear relationships. These baselines provide reference performance levels to assess whether the observed prediction accuracy arises from genuine learning of growth dynamics rather than simple temporal continuity or linear trends.

Table 3 presents the seven-day ahead (t+7) biomass prediction performance of all models evaluated on the testing production cycle. The XGBoost-based model consistently outperforms both baseline approaches across all evaluation metrics. In particular, the persistence model exhibits the largest prediction errors, indicating that short-term biomass changes cannot be adequately captured by assuming static growth over a one-week period. The linear regression model achieves improved performance relative to the persistence baseline but remains less accurate than the proposed model, reflecting its limited capacity to model non-linear interactions among feeding behavior, environmental conditions, and biomass history.

Table 3. Comparison of seven-day ahead (t+7) biomass prediction performance on the testing cycle

Model	MAE (kg)	RMSE (kg)	MAPE (%)	R ²
Persistence model	112.6	145.3	6.8	0.82
Linear regression	86.9	118.4	5.1	0.88
XGBoost (proposed)	65.8	88.6	3.9	0.94

The reduction in prediction error achieved by the proposed model is substantial, with approximately 41% lower MAE compared to the persistence baseline and 24% lower MAE compared to linear regression. These improvements demonstrate that incorporating non-linear machine learning modeling provides meaningful gains over simpler predictive strategies. The higher coefficient of determination further indicates that the proposed approach explains a larger proportion of the variability in observed biomass growth.

Overall, the baseline comparison confirms that the predictive performance reported in Section 4.1 is attributable to the model's ability to learn complex, non-linear growth patterns from recent farm data rather than relying on static assumptions or linear approximations. This reinforces the value of machine learning as an analytical tool for short-term biomass forecasting in real-world Patin aquaculture operations.

4. CONCLUSION

This study investigated the feasibility of short-term biomass prediction for Patin fish farming using a rolling machine learning framework based on routinely collected farm data. Using daily measurements collected over six-month production cycles, the biomass prediction problem was formulated as a seven-day ahead (t+7) supervised regression task. The proposed approach focuses on continuous short-horizon forecasting rather than end-of-cycle estimation, aligning the prediction output with practical farm management needs.

The experimental results demonstrate that reliable seven-day ahead biomass prediction can be achieved using the proposed XGBoost-based model. Evaluation on an independent production cycle shows stable prediction accuracy as total biomass increases toward approximately 1.7 tonnes, with low prediction error and strong agreement between predicted and observed values. The close correspondence between training and testing performance indicates good generalization and suggests that the model captures underlying growth dynamics rather than cycle-specific behavior. Comparisons with persistence and linear regression baselines further confirm that the observed performance gains arise from the model's ability to learn non-linear relationships among feeding behavior, environmental conditions, and biomass history.

From an operational perspective, the rolling prediction strategy enables continuous forecasting throughout the farming cycle, allowing predictions to be updated daily as new data become available. This design supports proactive decision-making, such as early detection of growth deviations and timely adjustment of feeding strategies, rather than reactive responses based solely on historical observations. Importantly, the approach relies only on data that are typically available in commercial farming operations, enhancing its practical applicability.

Despite these promising results, several limitations should be acknowledged. The study focuses on a single prediction horizon (t+7) and a limited number of production cycles. Future work could explore multiple prediction horizons, incorporate additional environmental or behavioral indicators, and evaluate model

performance across a wider range of farming conditions. Integration of the prediction model into an automated decision-support system and validation in long-term field deployments would further strengthen its practical impact.

In conclusion, the results demonstrate that short-horizon, rolling machine learning models provide an effective and practical means of forecasting biomass growth in Patin aquaculture. By transforming routine farm data into forward-looking insights, the proposed approach offers a foundation for data-driven management and improved operational efficiency in modern aquaculture systems.

REFERENCES

- [1] FAO, *The State of World Fisheries and Aquaculture*. Rome: FAO, 2022. doi: <https://doi.org/10.4060/cc0461en>.
- [2] J. H. Tidwell and G. L. Allan, "Fish as Food : Aquaculture ' s Contribution Ecological and Economic Impacts and Contributions of Fish Farming and Capture Fisheries," *EMBO Rep.*, vol. 2, no. 11, pp. 958–963, 2001, doi: 10.1093/embo-reports/kve236.
- [3] R. Anjar, A. Yustiati, and Y. Andriani, "Teknik Pembenihan Ikan Patin (Pangasius Hypopthalmus) Sistem Corong Fishing Engineering Catfish (Pangasius Hypopthalmus) Funnel System," *J. Akuatek*, vol. 3, no. 1, pp. 33–40, 2022.
- [4] J. Bostock *et al.*, "Aquaculture : Global Status and Trends," *Philos. Trans. R. Soc.*, vol. 365, no. 1554, pp. 2897–2912, 2010, doi: 10.1098/rstb.2010.0170.
- [5] S. R. L. Kruk, S. R. Bush, and M. Phillips, "Federating ' Aquaculture 4 . 0 ' for Data-Driven Social and Environmental Sustainability," *Mar. Policy*, vol. 169, no. 2024, p. 106355, 2024, doi: 10.1016/j.marpol.2024.106355.
- [6] O. V. Ocama *et al.*, "A Review on Advancing Technologies in Precision Agriculture : Applications , Challenges , and The Way Forward," *Procedia Comput. Sci.*, vol. 265, no. 2025, pp. 572–577, 2025, doi: 10.1016/j.procs.2025.07.221.
- [7] M. A. A. M. Hridoy, C. Bordin, A. Masood, and K. Masood, "Predictive Modelling of Aquaculture Water Quality Using IoT and Advanced Machine Learning Algorithms," *Results Chem.*, vol. 16, no. May, pp. 1–14, 2025, doi: 10.1016/j.rechem.2025.102456.
- [8] P. Erawati, D. Prasti, and T. B. Kriswinarso, "Internet of Things (IoT) -Based Water Quality Monitoring System Design for Tilapia Fish Farming Ponds," *Brilliance*, vol. 5, no. 2, pp. 860–869, 2025.
- [9] C. Chen, Y. Wu, J. Zhang, and Y. Chen, "IoT-Based Fish Farm Water Quality Monitoring System," *Sensors (Switzerland)*, vol. 22, no. 17, 2022.
- [10] K. H. Ghazali, N. Subari, R. Muda, and M. R. Yusof, "Data-Driven Insights in Patin Fish Farming Based on a Six- Month Analytical Study of Biomass Feeding Efficiency and Profitability," *Adv. Comput. Intell. Syst.*, vol. 1, no. 1, pp. 42–53, 2025.
- [11] B. Setiawan *et al.*, "Implementasi Sistem IoT Pada Akuakultur dan Hydroponik (Akuaponik) Modern Untuk Pertumbuhan Ikan Nila," *J. Inform. J. Pengemb.*, vol. 9, no. 1, pp. 47–53, 2024.
- [12] A. F. Abdullah, H. C. Man, A. Mohammed, M. M. A. Karim, S. U. Yunusa, and N. A. B. M. Jais, "Charting the aquaculture internet of things impact : Key applications , challenges , and future trend Charting the aquaculture internet of things impact : Key applications , challenges , and future trend," *Aquac. Reports*, vol. 39, no. December, p. 102358, 2024, doi: 10.1016/j.aqrep.2024.102358.
- [13] F. Ma, Z. Fan, A. Nikolaeva, and H. Bao, "Redefining Aquaculture Safety with Artificial Intelligence : Design Innovations , Trends , and Future Perspectives," *Fishes*, vol. 10, no. 88, pp. 1–27, 2025.
- [14] N. M. Abdikadir, A. S. Abdullah, H. O. Abdullahi, and A. A. Hassan, "Predicting Water Quality Parameters in Mahseer Fish Farming Using Machine Learning Techniques," *SSRG Int. J. Electron. Commun. Eng.*, vol. 11, no. 11, pp. 286–296, 2024.
- [15] G. Dharmarathne, A. M. S. R. Abekoon, M. Bogahawaththa, J. Alawatugoda, and D. P. P. Meddage, "A Review of Machine Learning and Internet-of-Things on The Water Quality Assessment : Methods, Applications and Future Trends," *Results Eng.*, vol. 26, no. May, p. 105182, 2025, doi: 10.1016/j.rineng.2025.105182.
- [16] P. W. Khan and Y. C. Byun, "Optimized Dissolved Oxygen Prediction Using Genetic Algorithm and Bagging Ensemble Learning for Smart Fish Farm," *IEEE Sens. J.*, vol. 23, no. 13, pp. 15153–15164, 2023, doi: 10.1109/JSEN.2023.3278719.
- [17] M. I. Maulana, F. Nugraha, and A. Setiawan, "Model Prediksi Kualitas Air Untuk Budidaya Ikan Lele Dengan Algoritma Extreme Gradient Boosting," *Build. Informatics, Technol. Sci.*, vol. 6, no. 3, pp. 1373–1382, 2024, doi: 10.47065/bits.v6i3.5998.
- [18] N. Özdemir, M. Çakır, M. Yılmaz, H. Şimşek, M. Oral, and O. Oral, "Application of the Machine Learning Methods to Assess the Impact of Physico- chemical Characteristics of Water on Feed Consumption in Fish Farms," *J. Agric. Sci.*, vol. 31, no. 1, pp. 71–79, 2025, doi:

- 10.15832/ankutbd.1470111.
- [19] R. A. Maula *et al.*, “Handling Missing Value dengan Pendekatan Regresi pada Dataset Akuakultur Berukuran Kecil,” *J. Rekayasa Elektr.*, vol. 18, no. 3, pp. 175–184, 2022, doi: 10.17529/jre.v18i3.25903.
- [20] P. Thaichon and S. Quach, *Introduction to artificial intelligence (AI)*. New York: Routledge, 2022.