European Journal of Prosthodontics and Restorative Dentistry (2026) 34 (5s), 178–191

AI-Driven Cost-Benefit Analysis of Smart Aquaculture Technologies

Abstract
Smart aquaculture technologies are gaining traction as solutions to better manage water quality, boost productivity, and enhance the efficiency of investments, but are still hard to quantify in economic value under farm-level uncertainty. This study created an AI based scenario-based cost benefit framework for the assessment of smart aquaculture technologies with the public pond level aquaculture data from Andhra Pradesh, India. The smart aquaculture functions were first categorized as monitoring, management with equipment, advisory decision support, follow-up monitoring, corrective-action implementation, and self-initiated management. A pond-level analytical dataset was then created by merging water-quality, stocking, harvest, equipment and sensor-campaign data records. Multiple machine-learning models were developed to predict yield, and the best model was integrated into a counterfactual cost-benefit analysis between a 'business-as-usual' (BAU) and a 'smart-aquaculture' (SA) scenario. The present net value, benefit-cost ratio, return on investment, payback period, sensitivity analysis and Monte Carlo simulation were used to assess economic feasibility. Extra Trees was the best performing model with an R2 value of 40.5%. The scenario-based framework estimated a median NPV of USD 46,718.91, median BCR of 7.42 and 65.70% economically feasible ponds based on the specified baseline economic assumptions. Analysis of the adoption gap revealed that there was a higher potential for marginal gains where the existing level of smart-technology intensity was lower, which informed targeted investment decisions. The framework provides a decision-support pathway for the assessment of smart aquaculture investment options that can be replicated in other situations with limited data and assumptions.

Keywords
Smart aquaculture, artificial intelligence, cost-benefit analysis, machine learning, net present value, water-quality monitoring

Authors
Bhagyashree Shendkar[1]
Department of Computer Science and Engineering, MIT School of Computing, MIT Art, Design and Technology University, Loni Kalbhor, Pune, India
Email: bhagyashree.d.shendkar@gmail.com
ORCID ID: 0000-0003-4808-2642

Gurunath Waghale[2]
School of Commerce and Management
Sri Balaji University, Pune
Email: guruwaghale111@gmail.com
ORCID: 0009-0007-4217-6249

Ashish Bhasme[3]
Department of Computer Science and Engineering, MIT School of Computing, MIT Art, Design and Technology University, Loni Kalbhor, Pune, India,
Email: ashish.bhasme@gmail.com
ORCID: 0009-0004-6488-3794

Kiran Shinde[4]
Department of Computer Science and Engineering, MIT School of Computing, MIT Art, Design and Technology University, Loni Kalbhor, Pune, India,
Email: kiran.shinde@mituniversity.edu.in
ORCID: 0000-0001-7388-9312

Dipak Umbarkar[5]
School of Management
Pimpri Chinchwad University, Pune, India
Email: dsumbarkar@gmail.com
ORCID: 0009-0006-9854-3987

Kiran Sakat[6]
Department of Economics
Padmabhushan Vasantraodada Patil Mahavidyalaya, Kavathe Mahankal, Sangli, India
Email: sakatkiran@gmail.com
ORCID: 0009-000-9181-7579

Corresponding Author:
Bhagyashree Shendkar, Department of Computer Science and Engineering, MIT School of Computing, MIT Art, Design and Technology University, Loni Kalbhor, Pune, India, Email: bhagyashree.d.shendkar@gmail.com

Received: 19-06-2026
Revised: 22-07-2026
Accepted: 24-07-2026
Doi: 10.1922/ejprd.v34i5s.1543

INTRODUCTION

Aquaculture has emerged as a key element in global aquatic food production, in the rural livelihood formation and in food-security planning, but the sector is now expected to increase production while at the same time decreasing production risk, input inefficiency, mortality and environmental pressure (FAO, 2024). While aquaculture has made a significant contribution to seafood availability, it is still very diverse in terms of species, culture systems, feeds, farm size, water-management and governance (Naylor et al., 2021). This heterogeneity means that generalized statements of technology adoption are not very useful and there is a need for analytical frameworks that can help to connect farm level production evidence with investment appraisal.

Water quality, stocking density, feeding and timely operational response are crucial components to efficiently managing aquaculture and are essential to sustainable intensification (Boyd et al., 2020). The use of smart aquaculture technologies, such as sensor-based monitoring, automated data collection, advisory platforms, machine-learning models and decision-support systems, have been suggested as tools to enable more anticipatory and evidence-based farm management, rather than reactive management (Føre et al., 2018). These technologies, in theory, can help detect poor water quality conditions, aid in making the right management decisions, minimize management delay, and stabilize production. But their use can't be rationalized from a technical point of view only.

However, the adoption of smart aquaculture is still limited by connectivity needs, maintenance, calibration, data quality issues, and the farmers' ability, as well as the uncertainty of the returns from the operation (Rastegari et al., 2023). The main issue, therefore, is not only can smart aquaculture technologies be used to monitor ponds effectively, but can their benefits outweigh the capital and operational expenses of implementing them? Although the precision aquaculture architecture is low cost and shows the potential for real-time monitoring and data-driven farm control, the economic value needs to be considered in the context of capital expenditure, recurring operating expenditure, expected yield improvement, labour savings, mortality reduction, and market-price uncertainty (Teixeira et al., 2021).

Although there have been significant developments in AI and smart aquaculture systems, there is still a methodological gap between technical prediction and economic decision making. Most of the existing studies focus on the accuracy of monitoring, the integration of sensors, or predictive performance, with fewer studies reporting production gains from the use of AI in terms of investment indicators such as net present value, benefit–cost ratio, return on investment, payback period, or feasibility probability. Similar AI-based approaches in other applied sectors, including healthcare big-data analytics and machine-learning-enabled early detection, show that predictive models can support decision making when heterogeneous data are converted into actionable indicators (Kulkarni and Shendkar, 2022; Damre et al., 2024). This gap is crucial in small and medium-scale aquaculture systems where technology adoption is sensitive to production variability, market access and cost.

In this regard, a scenario-based cost–benefit analysis framework is created to assess smart aquaculture technologies based on pond-level production evidence and explicit economic assumptions and driven by AI. The study aims to categorize the smart aquaculture functions; use machine learning to predict the production outcomes and translate the AI-predicted production changes into indicators of economic feasibility. The objectives of the research are:

1. to identify and categorize key smart aquaculture technologies and their functional roles in aquaculture production;
2. to construct an AI-driven cost-benefit analysis framework for measuring the economic costs, operational benefits, and productivity impacts of these technologies under stated baseline assumptions;
3. and to evaluate the scenario-based financial feasibility, adoption potential, and decision-support value of smart aquaculture technologies for sustainable aquaculture systems.

2. LITERATURE REVIEW

Smart aquaculture research has evolved from the initial idea of monitoring and image analysis to integrated systems for decision support. Computer vision is a technology that has been studied as a potentially useful but operationally challenging method for measuring fish in underwater conditions where lighting, turbidity, occlusion and stress sensitivity are issues (Zion, 2012). Since then, machine vision has been applied to biomass, product quality and fish-management tasks, but practical application still depends on sensor quality, calibration and farm conditions (Saberioon et al., 2017). The literature on machine-learning applications in intelligent fish aquaculture reveals that machine-learning models are currently applied for biomass estimation, species identification, behaviour analysis and water-quality prediction (Zhao et al., 2021). AI-based image interpretation has also been applied in related aquatic and maritime contexts, including marine vessel detection from satellite imagery for surveillance applications, demonstrating the wider relevance of machine-learning-based visual detection in water-associated environments (Bhosale et al., 2025).

Recent smart-water and IoT research highlight the need to make aquaculture decisions based on near real-time data for water-quality parameters such as pH, dissolved oxygen, temperature, turbidity and related indicators. The use of sensors and machine-learning prediction for biofloc studies can be integrated to provide alerts to farmers and improve water-management responses (Rashid et al., 2021). A smart aquaculture system is an extension of this logic, as it integrates sensing, AI and cloud-computing infrastructure for real-time environmental regulation and production monitoring (Wu et al., 2023). Comparable sensor-based and AI-enabled early-detection frameworks have also been proposed in healthcare, where wearable devices and machine-learning models are used to identify disease risk at an early stage (Damre et al., 2024), while big-data analytics has been used to support healthcare decision-making from complex and heterogeneous datasets (Kulkarni and Shendkar, 2022). Such studies strengthen the broader methodological argument for sensor-driven, AI-assisted decision support, although aquaculture-specific validation remains necessary to measure benefits against investment and running costs.

This study's AI component is based on proven machine-learning techniques. Random forests are effective in nonlinear prediction by combining many different random decision trees and by ensemble averaging to decrease the overfitting (Breiman, 2001). Extra Trees introduce more randomization in the splitting process and are generally very powerful for small to medium sized datasets where the predictors are mixed (Geurts et al., 2006). Gradient boosting is another additive-tree framework that is built sequentially to minimize loss (Friedman, 2001). Black-box prediction is not enough for applied decision support as one need model-agnostic tools to determine which variables impact predictions (Molnar, 2022).

The economic element is based on conventional cost-benefit reasoning. Cost-benefit analysis is a formal way of comparing streams of benefits and costs that are discounted over time under different policies or investment options (Boardman et al., 2018). Net present value, benefit cost ratio and sensitivity testing have been central to agricultural project appraisal as decision criteria for farm investments for a long time (Gittinger, 1982). Sensitivity analysis is particularly relevant if there are uncertain prices, yields and costs that might alter the ranking of the decisions (Pannell, 1997). Global uncertainty is another indicator that can help to distinguish the dominant drivers from less influential assumptions (Saltelli et al., 2008).

3. MATERIALS AND METHODS

3.1 Study design and analytical framework

The techno-economic design was done quantitatively using secondary data and AI to determine if a measurable production and financial gain can be achieved by smart aquaculture technologies. The analytical framework was designed to link operational aquaculture data with the prediction using machine learning and formal investment appraisal. The study did not view the adoption of smart aquaculture as simply a technology, but as a farm-management transition process consisting of monitoring, technology-based control, advisory decision support, corrective actions, and adaptive response.

The analysis was broken down into three phases. Smart aquaculture technology proxies were first generated from the observable monitoring, advisory, equipment, follow-up and management-response variables. Second, smart technology indicators, water-quality summaries, pond characteristics, stocking-related variables, feeding variables, and fish-health and stress indicators were used to train machine-learning models for predicting pond-level yield. Third, the best performing AI model was put inside a counterfactual cost–benefit analysis by simulating the enhanced smart-aquaculture scenario and comparing the yield predictions with the baseline scenario.

Pond was used as the unit of analysis. The major production output measured was yield (kg/acre). Net present value, benefit–cost ratio, return on investment, payback period and economic feasibility status were used to determine the economic performance. This design enabled the prediction of production and evaluation of investments to be linked in one decision support system. It also enabled the identification of current smart technology intensity from remaining smart technology adoption potential as ponds already using higher smart technology intensity might have less potential for smart technology adoption than those with lower current technology intensity.

3.2 Dataset and preprocessing

The data analyzed is from the Public ARA Farm Data released by Fish Welfare Initiative that includes anonymized aquaculture data from Andhra Pradesh, India. The data set comprises water quality monitoring data, enrolled pond data, stocking and harvest data, dropout data and continuous sensor campaign sheets (Fish Welfare Initiative, 2026). The water-quality records were for the period 19 June 2021 to 27 January 2026, and the stocking and harvest records were for the period 1 January 2020 to 27 January 2026. These components were combined to create a pond level data set appropriate for AI modelling and cost–benefit estimation.

Before analysis all variables were standardized. The column names were standardized to machine-readable format, date variables were parsed and numeric fields were converted to values that can be analyzed. The data for water quality parameters such as dissolved oxygen, pH, temperature, turbidity, ammonia, alkalinity, hardness and other water quality parameters were pooled at pond level and presented as mean, median, standard deviation, minimum and maximum. Stocking and harvest records were combined for each pond for biomass and yield estimates of the harvested ponds. Yield was determined by the weight of the biomass harvested per acre of pond.

Records in which the yield was not provided or were negative or invalid were not included in the AI modeling dataset. The target variable (raw yield) was winsorized at the 1st and 99th percentiles to avoid extreme values that would impact the model training and evaluation. This process minimized the effect of extreme observations but kept valid pond level observations. Winsorized yield was then log transformed and used for training the machine-learning models. A final data set of 172 pounds having valid yield observations was used for modeling.

3.3 Smart technology proxy construction

Since the public data did not include direct labels of commercial technologies, functional proxies were used to operate smart aquaculture technologies. These proxies were chosen to reflect the real-life roles played by smart aquaculture systems to impact pond management. The resulting domains were water-quality monitoring, equipment enabled monitoring, advisory decision support, follow up monitoring, corrective-action implementation, self-initiated management, water-quality risk identification, aerator or pump support, and continuous sensor campaign participation. The proxy variables were aggregated at pond level. The intensity of pond level water-quality assessment was represented through monitoring records. Advisory decision-support records are those cases where corrective action was requested following monitoring. Farm-level response to recommendations was considered corrective-action implementation. Repeated assessment and adaptive management were done during follow-up monitoring. Equipment-enabled variables were mechanized infrastructure support such as aerators and pumps. The greater the number of sensor campaigns in which a sensor participated, the greater the exposure to monitoring.

A smart technology intensity index was created by scaling the proxy components available and averaging them at the pond level. The index was developed to capture the extent to which each pond was exposed to smart monitoring, advisory, equipment support and adaptive management practices. Responses that were corrected with a corrected implementation parser were considered positive corrective-action implementation responses, e.g., “Yes, all of them” and “Only some”. This measure ensured that the activity of the smart management was not underestimated due to implementation.

3.4 AI model development and interpretation

The problem of machine learning was stated as predicting the yield of a pond. The response variable was log winsorized yield (kg/acre). Smart technology proxies, water-quality summaries, pond area, pond depth, equipment indicators, stocking intensity variables, feeding variables, fish-health or stress indicators, and continuous sensor variables were used as predictor variables. Prior to training, variables that could directly cause outcome leakage such as total harvest, yield variables, date of events, and pond identifiers were excluded from the model.

The data set was split into training and testing sets in the ratio of 75:25. The following four models were used for comparison: Ridge Regression, Random Forest, Extra Trees, Gradient Boosting. These models were chosen to represent linear regularized prediction and nonlinear ensemble learning. The root mean squared error, mean absolute error and coefficient of determination were used to evaluate the model performance. The best model was chosen to perform prediction diagnostics and for counterfactual cost–benefit simulation.

Permutation feature importance was used for model interpretation. This method assessed how much the prediction error was raised when each feature was randomly permuted, and determined which variables made the greatest contribution to the model performance. Other predictor domains such as smart technology/management, water quality, stocking/feeding/production, farm/pond characteristics, fish health/stress and others were also used to summarize feature importance. This interpretation step was required to identify if smart aquaculture-related variables substantially contributed to yield prediction.

3.5 Cost–benefit and sensitivity analysis

The most effective AI model was incorporated into a counterfactual cost–benefit framework. Baseline yield was predicted for each pond using the trained model. Selected smart technology and management variables were then stepped up to higher levels of adoption to simulate a smart-adoption scenario. The difference between the predicted smart-scenario yield and the predicted baseline yield was considered as the incremental yield gain estimated by the AI. The incremental yield gains for a pond were set to zero if they were negative as calculated by the model, to be conservative and not assign a false benefit to a pond that the model predicted would not be improved.

An economic analysis was performed with farmgate price, smart technology capital cost, annual operating cost, feed-cost saving, labour-cost saving, mortality-risk reduction value, discount rate and project lifetime being editable variables. Based on these assumptions, annual incremental revenue, annual operational savings, annual technology operating cost, initial investment, present value of benefits, present value of costs, net present value, benefit–cost ratio, return on investment and payback period were calculated. Positive net present value and benefit–cost ratio of greater than one was used to determine economic feasibility. The baseline scenario was parameterized with pre-specified economic assumptions to allow for discounted financial indicators to be calculated from the yield gains predicted by the AI. The assumed farmgate prices were USD 2.20 kg^{-1}, smart technology capital expenditure USD 950 acre^{-1}, annual smart technology operating expenditure USD 180 acre^{-1}, feed-cost saving USD 120 acre^{-1} year^{-1}, labour-cost saving USD 90 acre-1 year-1 and mortality-risk reduction value USD 75 acre^{-1} year^{-1}. The calculations for baseline NPV, BCR, ROI and payback period were made using a discount rate of 10% and a project life of five years. Thus, the cost–benefit estimates should be viewed as conditional scenario outcomes, and not necessarily as actual farm accounting outcomes.

The economic part was not included as a farm accounting audit, but as a transparent scenario-based cost–benefit module due to the lack of direct private financial information in the public data set. The assumption-based structure enabled the production effect predicted by the AI to be transformed into economic indicators, as well as the parameters that need to be locally validated prior to investment decisions. The default values should be substituted by actual data from farm surveys, market prices, vendor quotes and cost estimates of the species for field deployment.

3.6 Statistical and uncertainty analysis

To test the robustness of the economic results, sensitivity and uncertainty analyses were performed. One-way sensitivity analysis was performed to assess the impacts of farmgate price, technology capital cost, annual operating cost, feed-cost saving, labour-cost saving, mortality-risk reduction value and discount rate on the median NPV and economic feasibility share. Farmgate price and technology capital cost were selected as two-way sensitivity variables since they are two key market and investment parameters.

Uncertainty in the price, capital cost, operating cost, operational savings, discount rate, and the yield gain predicted by the AI was propagated using Monte Carlo simulation with 2000 iterations. The simulation results were median net present value, mean net present value, median benefit–cost ratio, median return on investment and feasibility share. Spearman rank correlation was performed to determine which uncertainty drivers are most important for median net present value.

Further statistical analysis was conducted to understand the relationship between the use of smart technology and economic outcomes. Using the Spearman correlation analysis, the association between smart technology intensity, adoption gap, yield and cost–benefit indicators were evaluated. Differences in economic outcomes between groups of adopters with different adoption gaps were analyzed using Kruskal–Wallis tests. The "adoption-gap index" was calculated as the difference between the observed smart technology intensity of a pond and the highest observed smart technology intensity. This index helped to differentiate ponds with high current technological intensity from ponds with greater remaining improvement potential.

4. RESULTS

4.1 Smart aquaculture technology proxy coverage

The smart aquaculture technology proxies showed meaningful variation in monitoring, advisory, equipment-supported, and adaptive management activities. Monitoring records were available for all monitored ponds, confirming that water-quality assessment was the core smart-management function represented in the data. Equipment-enabled monitoring was also highly represented, while water-quality risk identification, advisory decision support, aerator/pump support, follow-up monitoring, and self-initiated management were observed across substantial proportions of ponds. These findings support the operational classification of smart aquaculture technologies into monitoring, advisory, corrective, equipment-supported, and farmer-response domains. **Table 1** reports the pond-level distribution of the smart aquaculture technology proxies used to characterize monitoring, advisory, equipment-supported, and adaptive management functions.

Table 1. Coverage of smart aquaculture technology proxies.

Smart technology proxy	Ponds with proxy	Total ponds	Coverage (%)
Monitoring records	298	298	100.00
Equipment-enabled monitoring	295	298	98.99
Water-quality risk identification	220	298	73.83
Advisory decision support	199	298	66.78
Aerator/pump support	197	298	66.11
Follow-up monitoring	194	298	65.10
Self-initiated management	164	298	55.03
Corrective-action implementation	Corrected proxy retained	298	See corrected implementation output

The coverage profile indicates that monitoring records were available for all 298 monitored ponds, while equipment-enabled monitoring was recorded for 295 ponds, corresponding to 98.99% coverage. Water-quality risk identification was observed in 220 ponds, advisory decision support in 199 ponds, aerator/pump support in 197 ponds, follow-up monitoring in 194 ponds, and self-initiated management in 164 ponds. Corrective-action implementation was retained as a corrected proxy after response categories were reclassified to capture partial and full implementation responses. **Figure 1** compares the relative coverage of each smart aquaculture technology proxy across the monitored pond sample.

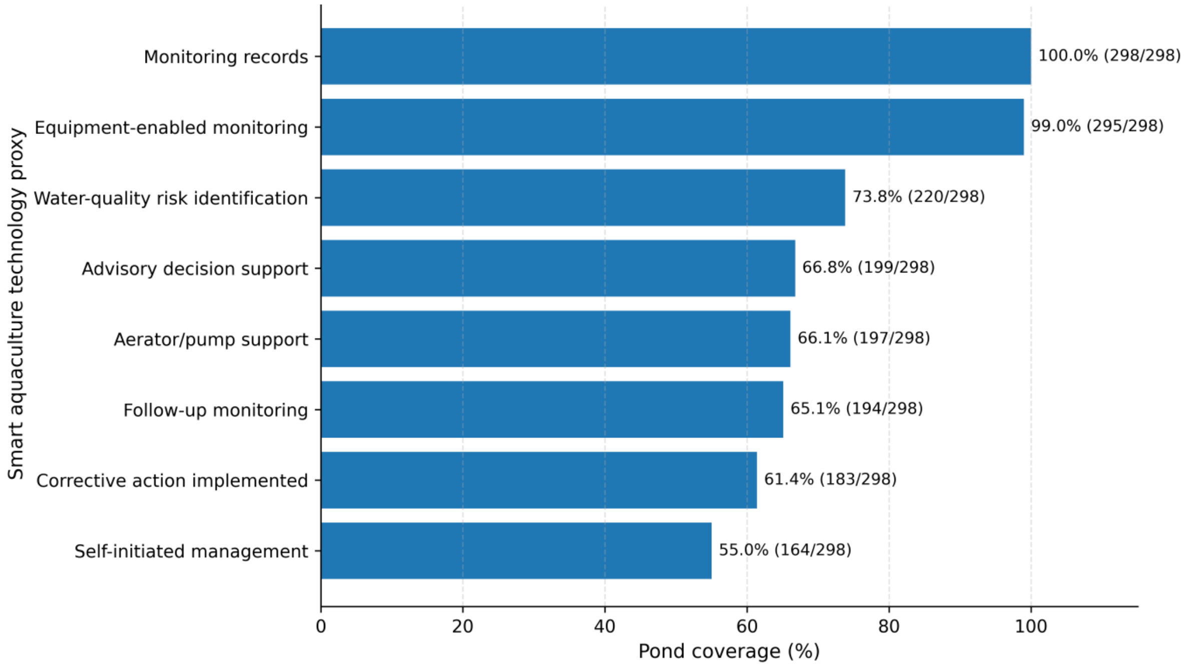

Figure 1. Coverage of smart aquaculture technology proxies across monitored ponds.

The horizontal bars show the relative proportions of monitoring records, monitoring using equipment, water-quality risk identification, advisory decision support, aerator/pump support, follow-up monitoring, corrective-action implementation, and self-initiated management. The distribution shows that the smart aquaculture activity was not within a single technology category but was distributed among various functional domains. The percentage labels show the percentage of ponds that are covered by the proxy as well as the number of ponds represented by the proxy.

4.2 Yield treatment and AI model performance

The raw yield variable showed substantial skewness because a small number of extreme observations inflated the mean. Raw mean yield was 32,442 kg per acre, while the median was 5,490 kg per acre. After winsorization, the mean yield decreased to 7,029 kg per acre, while the median remained 5,490 kg per acre. This indicated that winsorization reduced the influence of extreme observations without materially changing the central tendency of yield. The winsorized yield variable was therefore used for model training and prediction. **Figure 2** depicts the distributional effect of outlier treatment on the pond-level yield variable.

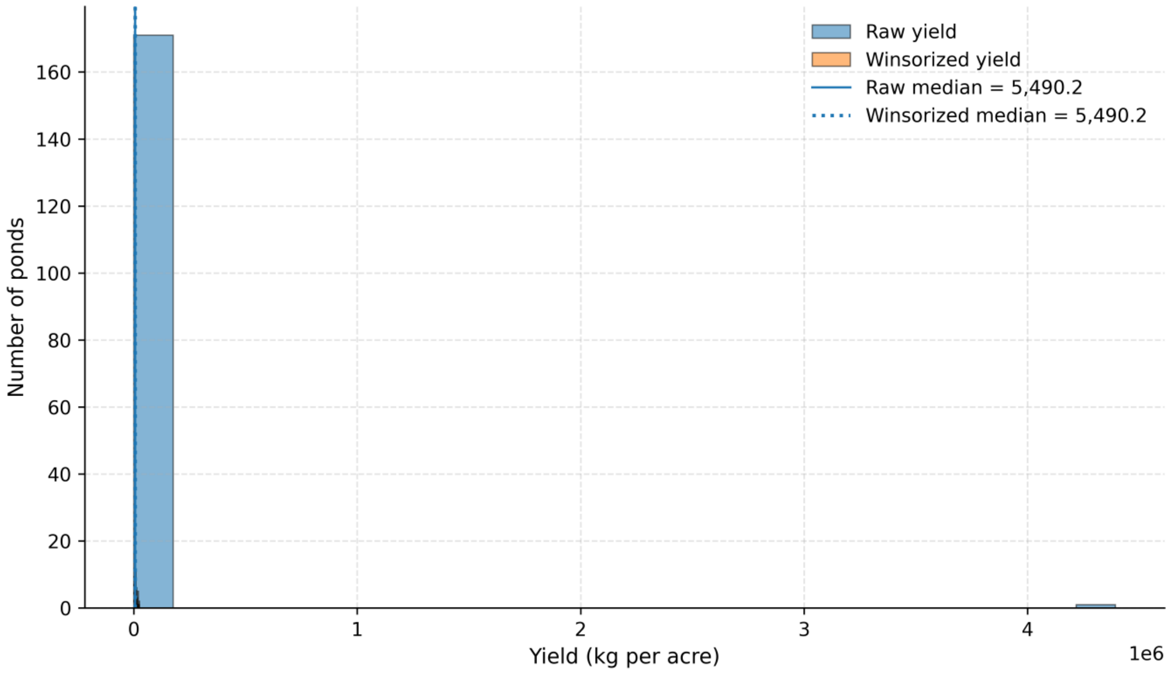

Figure 2. Yield distribution before and after outlier treatment.

The raw yield distribution was highly skewed because of extreme high-yield observations, which expanded the scale of the x-axis and compressed most observations near the lower range. The winsorized yield distribution retained the same median value of approximately 5,490 kg per acre but reduced the influence of extreme observations on the upper tail. The distribution therefore indicates how the target variable was stabilized before machine-learning model development.

Among the evaluated machine-learning models, Extra Trees produced the best overall predictive performance. The model achieved a test root mean squared error of 4,172.27 kg per acre and an original-scale R^2 of 0.405. The observed mean yield in the test set was 6,963.62 kg per acre, while the predicted mean yield was 6,092.45 kg per acre. The mean absolute error was 2,292.37 kg per acre. These results indicate moderate but meaningful predictive performance under heterogeneous pond-level production conditions. **Table 2** reports the diagnostic performance indicators of the best-performing AI model used for pond-level yield prediction.

Table 2. Best AI model performance for pond-level yield prediction.

Metric	Value
Modeling observations	172
Selected predictor features	127
Training observations	129
Testing observations	43
Best model	Extra Trees
Observed yield mean	6,963.62 kg/acre
Predicted yield mean	6,092.45 kg/acre
Mean absolute error	2,292.37 kg/acre
RMSE	4,172.27 kg/acre
R^2 original scale	0.405

The Extra Trees model was chosen from 172 modeling observations, and 129 observations were used for training and 43 observations were used for testing. The observed mean yield of the model in the test-set was 6,963.62 kg/acre while the predicted mean yield of the model was 6,092.45 kg/acre. The prediction error statistics were a mean absolute error of 2292.37 kg per acre, root mean squared error of 4172.27 kg per acre, and original-scale R2 of 0.405. **Figure 3** shows the comparative predictive performance of the machine-learning models evaluated based on test-set RMSE.

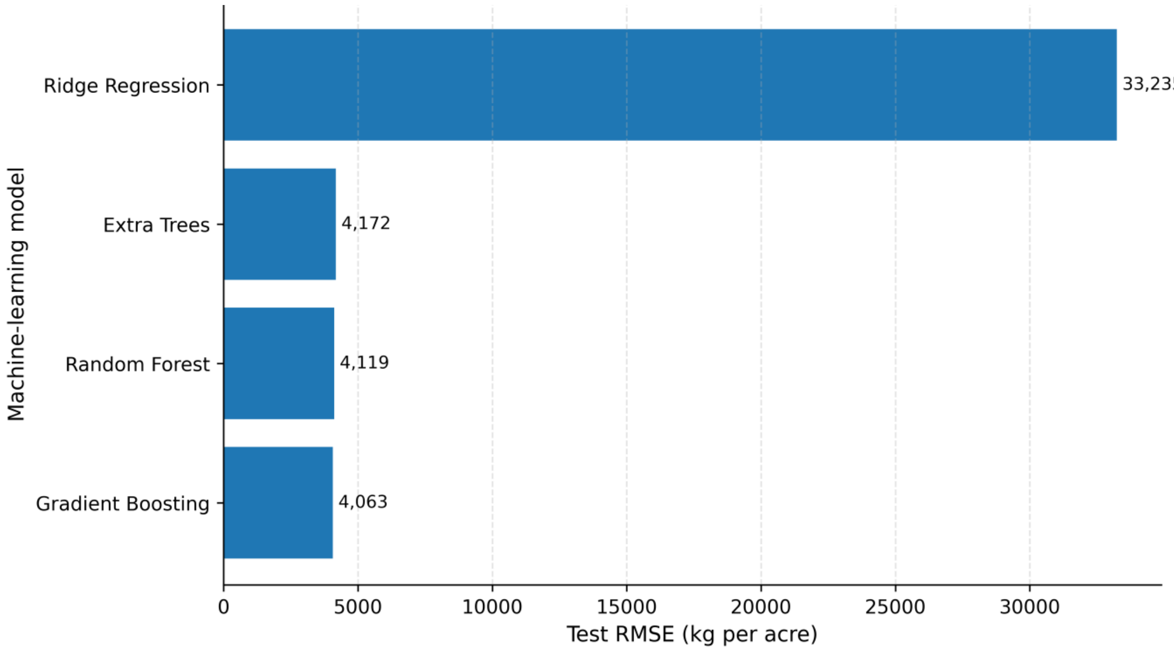

Figure 3. AI model comparison for yield prediction using test-set RMSE.

The comparison shows that the Ridge Regression had the largest prediction error with a significantly higher RMSE than the tree-based ensemble models. The RMSE values for Gradient Boosting, Random Forest and Extra Trees were much lower, suggesting better performance of the nonlinear models in terms of prediction. Extra Trees was kept as the best performing model due to its comparative predictive accuracy and was applied as yield-prediction engine in the cost–benefit analysis. The observed versus predicted diagnostic plot for the Extra Trees model is shown in **Figure 4**.

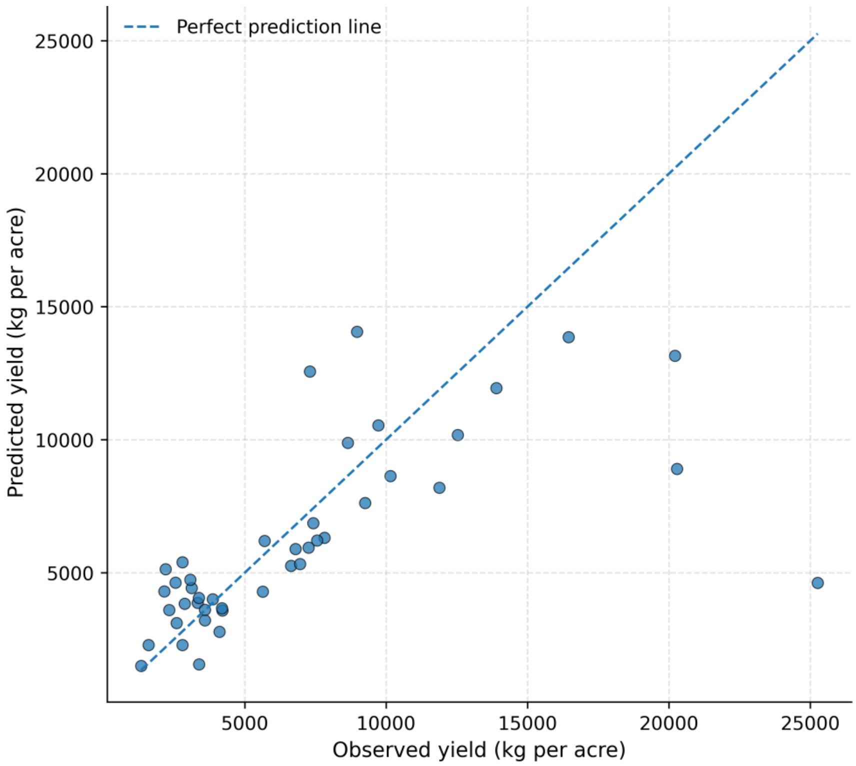

Figure 4. Observed versus predicted aquaculture yield for the Extra Trees model.

Each point corresponds to a test-set pond observation, while the dashed diagonal line denotes perfect agreement between observed and predicted values. Points clustered near the diagonal indicate better prediction alignment, whereas points farther from the line indicate residual error. The diagnostic plot reveals stronger prediction alignment in the lower-to-moderate yield range, whereas larger deviations were observed for several high-yield observations.

4.3 Predictor-domain importance

The results of permutation feature importance revealed that the smart technology and management variables were the best predictor domain for aquaculture yield. The total positive permutation importance for the smart technology/management domain (0.0939) was higher than the other domains: stocking/feeding/production variables (0.0697), water-quality variables (0.0388), farm/pond characteristics (0.0363), and fish-health/stress indicators (0.0297). This finding is in line with the main hypothesis of the study, which was that there is predictive, meaningful information for production performance in the smart aquaculture management variables. The relative contribution of predictor domains by positive permutation feature importance is reported in **Table 3**.

Table 3. Predictor-domain importance based on permutation feature importance.

Predictor domain	Positive feature count	Total importance	Mean importance
Smart technology / management	17	0.093882	0.005522
Other	6	0.057920	0.009653
Stocking / feeding / production	25	0.021991	0.000880
Water quality	18	0.020275	0.001126
Farm / pond characteristics	3	0.006907	0.002302
Fish health / stress	4	0.001072	0.000268

The domain with the highest total importance value was the smart technology/management domain with 17 positive-importance features. The "Other" domain had the second highest total importance of 0.057920, followed by stocking/feeding/production, water quality, farm/pond characteristics, and fish health/stress indicators. The domain-level results show that the variables of smart technology and management factors made the biggest contribution to the explanatory signal of the AI yield-prediction model. The results from Table 3 are displayed in a visual comparison of the predictor-domain importance results in **Figure 5**.

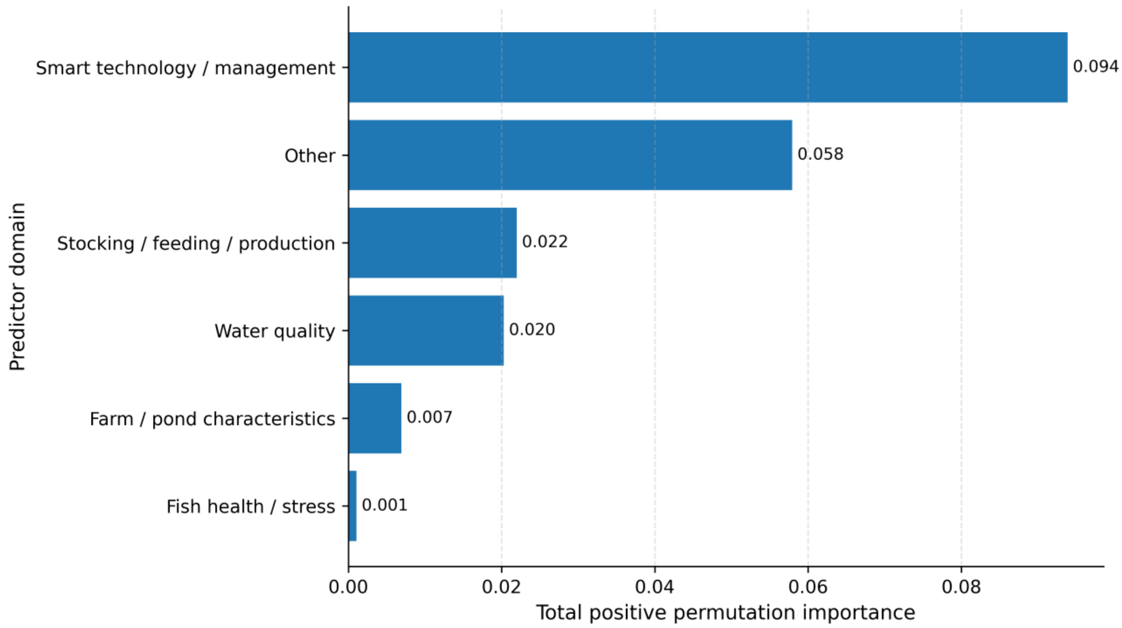

Figure 5. Relative contribution of predictor domains to AI-based yield prediction.

The ranked domain profile places smart technology/management as the largest contributor, followed by the "Other" domain, stocking/feeding/production, water quality, farm/pond characteristics, and fish health/stress indicators. The domain distribution indicates the relative predictive contribution of each grouped variable category. The pattern further indicates that the AI model relied most strongly on variables associated with monitoring, advisory support, equipment-enabled management, and adaptive management behaviour.

4.4 AI-driven cost-benefit outcomes

The AI-driven counterfactual cost–benefit analysis estimated positive economic indicators under the specified baseline assumptions. The median AI-predicted incremental yield was 1,323.56 kg per acre, while the mean incremental yield was 1,305.78 kg per acre. The median net present value was USD 46,718.91, and the mean net present value was USD 197,507.49. The median benefit-cost ratio was 7.42, and the mean benefit-cost ratio was 7.33. Under the baseline scenario, 113 out of 172 evaluated ponds were classified as economically feasible, representing 65.70% of the modeled ponds. **Table 4** reports the baseline AI-driven cost–benefit estimates generated under the specified economic assumptions.

Table 4. Baseline AI-driven cost-benefit analysis summary.

Indicator	Value
Median NPV (USD)	46,718.91
Mean NPV (USD)	197,507.49
Median BCR	7.42
Mean BCR	7.33
Economically feasible ponds	113
Economically feasible ponds (%)	65.70
Median AI-predicted incremental yield	1,323.56 kg/acre
Mean AI-predicted incremental yield	1,305.78 kg/acre

With the baseline economic assumptions, the median net present value was USD 46,718.91 and the mean net present value was USD 197,507.49, suggesting positive skewness in the distribution of the model-estimated net present value. The median benefit–cost ratio was 7.42 and the mean benefit–cost ratio was 7.33. In the baseline scenario, 113 of the 172 ponds assessed were deemed economically feasible (65.70% of the sample). The median and mean incremental yields predicted by AI were 1,323.56 kg/acre and 1,305.78 kg/acre respectively. The distribution of net present value (NPV) for the smart-aquaculture adoption scenario is shown in **Figure 6**.

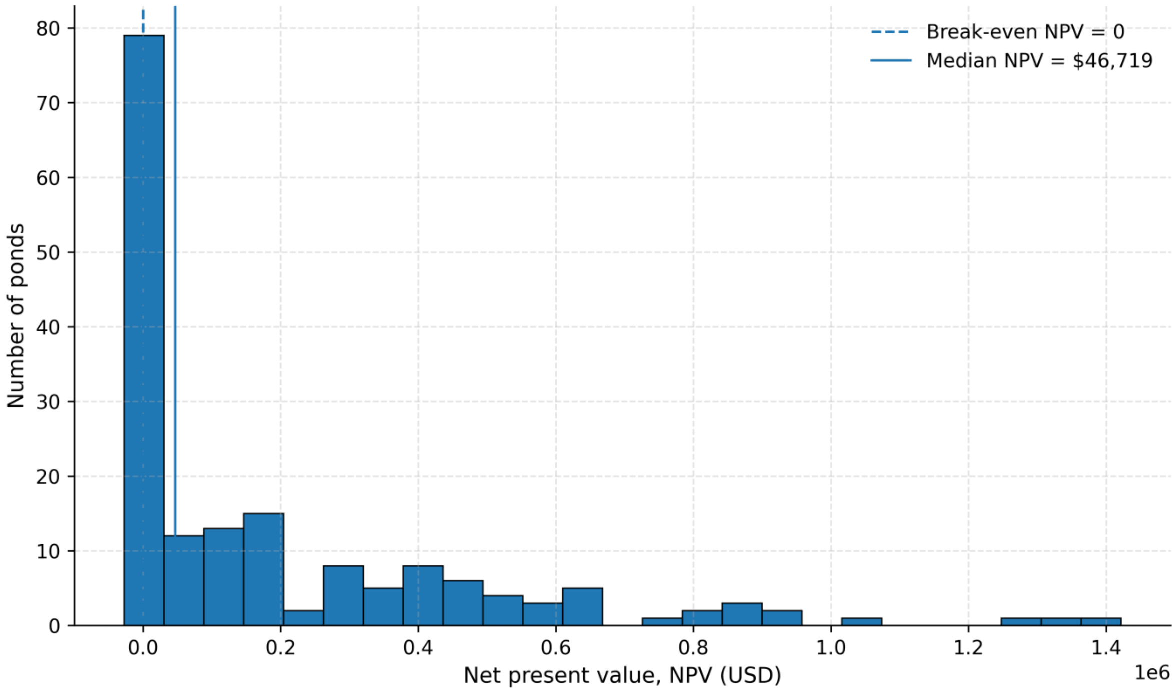

Figure 6. Distribution of net present value for smart aquaculture technology adoption.

The histogram indicates that many ponds were concentrated near lower positive NPV values, whereas a smaller number of ponds generated substantially higher NPV values, producing a right-skewed distribution. The break-even reference line marks NPV = 0, and the median reference line indicates the central economic estimate of approximately USD 46,719. The distribution represents the spread of economic outcomes beyond the average or median result.

4.5 Sensitivity and uncertainty analysis

The one-way sensitivity analysis revealed that the farmgate price had the greatest influence on median net present value in the scenario-based economic framework. The median net present value varied from USD 34590.71 to USD 58847.11 when the farmgate price varied between low and high. Discount rate was the next most important factor, with median net present value values of between USD 43,124.71 and USD 50,784.19. The difference between the technology capital cost within the range tested and the farmgate price was smaller, suggesting that the model-estimated feasibility outcomes were more sensitive to the assumptions regarding market prices than to the assumptions regarding technology capital cost.

The same trend was observed in the scenario-based economic framework, using Monte Carlo uncertainty analysis. The farmgate price was the most significant uncertainty factor for the median net present value with a Spearman correlation of 0.737. The second-best correlation was for the yield-gain multiplier predicted by AI with 0.547. The correlation between discount rate and median net present value was negative (-0.302). The key sensitivity and uncertainty-driver results are summarized in **Table 5**.

Table 5. Sensitivity and Monte Carlo uncertainty-driver summary.

Analysis component	Main result
One-way sensitivity: farmgate price	Median NPV range = USD 34,590.71 to USD 58,847.11
One-way sensitivity: discount rate	Median NPV range = USD 43,124.71 to USD 50,784.19
Monte Carlo driver: farmgate price	Spearman rho with median NPV = 0.737
Monte Carlo driver: AI yield-gain multiplier	Spearman rho with median NPV = 0.547
Monte Carlo driver: discount rate	Spearman rho with median NPV = -0.302

In the one-way sensitivity analysis, farmgate price produced the widest median NPV range, from USD 34,590.71 to USD 58,847.11. Discount rate produced the second reported sensitivity range, from USD 43,124.71 to USD 50,784.19. The Monte Carlo uncertainty-driver analysis showed positive Spearman correlations between median NPV and farmgate price (0.737) and between median NPV and the AI yield-gain multiplier (0.547), while discount rate showed a negative correlation of -0.302. **Figure 7** presents the one-way sensitivity ranges for median net present value across the tested economic assumptions.

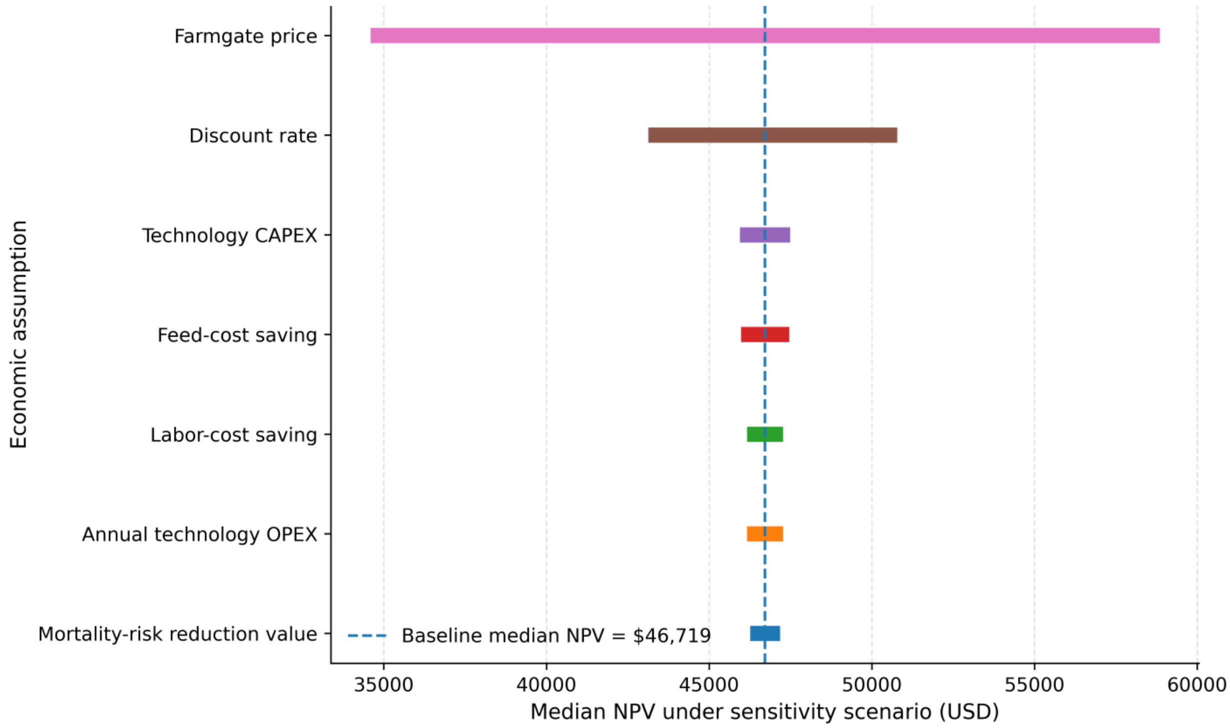

Figure 7. One-way sensitivity analysis of median net present value.

Each horizontal bar shows the difference in median NPV that results from changing only one of the assumptions, while keeping all other parameters the same. The sensitivity range is the farmgate price that results in the largest difference in median NPV between the low and high scenarios. Discount rate resulted in the next biggest variation, followed by technology CAPEX, feed-cost saving, labour-cost saving, annual technology OPEX, and mortality-risk reduction value, which resulted in smaller variations around the median NPV.

4.6 Smart technology adoption-gap analysis

The adoption-gap analysis clarified the relationship between current smart technology intensity and marginal economic benefit within the model-estimated scenario framework. Current smart intensity was positively associated with observed and baseline yield, but negatively associated with incremental yield, net present value, benefit-cost ratio, and return on investment. This indicates that ponds already operating at higher smart-management intensity had higher baseline productivity, but less remaining improvement potential. The adoption-gap index captured this remaining potential by measuring the distance between each pond's current smart intensity and the maximum observed smart intensity.

Kruskal-Wallis tests showed statistically significant differences across adoption-gap groups for all major economic outcomes. These included AI-predicted incremental yield, net present value, benefit-cost ratio, return on investment, payback period, and economic feasibility. This finding indicates that smart aquaculture investment may generate the greatest marginal value in ponds with larger adoption gaps. **Table 6** reports the Kruskal–Wallis group-comparison results for productivity and economic outcomes across smart technology adoption-gap groups.

Table 6. Kruskal-Wallis tests across smart technology adoption-gap groups.

Outcome	Test statistic	p-value	Interpretation
AI-predicted incremental yield	90.962	1.769 × 10^-20	Significant group difference
NPV	78.308	9.901 × 10^-18	Significant group difference
BCR	88.057	7.562 × 10^-20	Significant group difference
ROI	85.777	2.364 × 10^-19	Significant group difference
Payback period	84.040	5.636 × 10^-19	Significant group difference
Economic feasibility	94.675	2.764 × 10^-21	Significant group difference

Statistically significant group differences were observed for AI-predicted incremental yield, net present value, benefit–cost ratio, return on investment, payback period, and economic feasibility. The test statistics ranged from 78.308 for NPV to 94.675 for economic feasibility, with all p-values far below 0.05. These results indicate that the adoption-gap grouping separated ponds into statistically distinct economic and productivity-response categories. **Figure 8** depicts the relationship between the smart technology adoption-gap index and AI-predicted improvement potential.

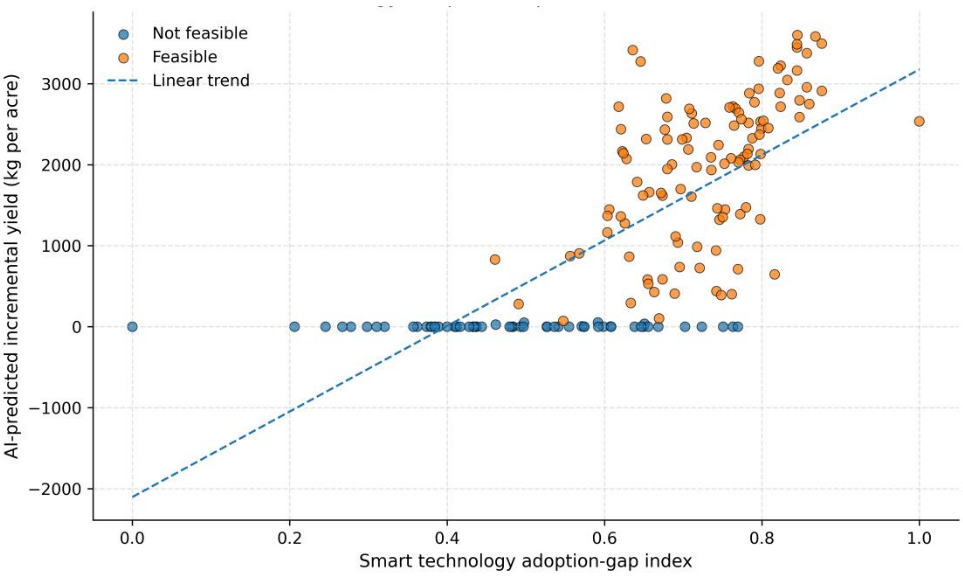

Figure 8. Smart technology adoption gap and AI-predicted economic improvement potential.

The x-axis represents the remaining smart technology adoption gap, while the y-axis represents the predicted yield gain in kilograms per acre. Points are separated by economic feasibility status, with feasible observations concentrated more heavily at higher adoption-gap values. The upward trend line indicates that larger adoption gaps were associated with greater predicted incremental yield, whereas ponds with lower adoption gaps tended to exhibit limited or zero marginal production gain.

5. DISCUSSION

Results suggest that smart aquaculture technologies could provide positive economic value, as model estimated economic value can be translated into AI-predicted production gains and discounted investment criteria can be used to monitor, provide advisory support, and manage adaptively. But the results also indicate that not all ponds are created equal when it comes to their economic value in the smart aquaculture system. This aligns with the general knowledge that development of aquaculture depends not only on the availability of technology, but also on the commercial incentives, farm capability, availability of inputs, production system design and the quality of farm management, especially in inland aquaculture systems in Asia (Belton and Little, 2011). The current analysis builds on this by showing how the heterogeneous pond-level data can be transformed to AI-based productivity estimates and cost–benefit indicators.

Based on the AI modeling result, it is proposed that the smart-management variables will have predictive information relevant to the productivity in aquaculture. The Extra Trees model explained moderate amounts of variation in the original yield scale, which is reasonable considering the variability in pond conditions, production histories, water-quality records, management intensity, and data completeness. The aim was not to create an ideal biological growth model, but to generate a defensible estimate of biological growth for use as a decision-support tool which could be incorporated into economic appraisal. The results of the feature-domains are thus significant: The smart technology and management variables had the greatest contribution of any other predictor domain. This implies that monitoring, advisory activity, follow-up assessment, equipment support and farmer response should be considered as economically important management signals and not as side lines descriptive variables.

The cost–benefit results indicate that smart aquaculture investment may be financially appealing when based on the assumed baseline conditions but should be viewed as an output of the scenarios and not as evidence from the farm accounting. Given the absence of private financial data in the public data set, the capital expenditure, operating expenditure, savings, farm-gate price and discount-rate assumptions were parameterized. This is significant as feed inputs are a key factor in aquaculture production cost and profitability (Tacon and Metian, 2015), and environmental management can influence system design, nutrient discharge and water use, which in turn can impact productivity and sustainability (Verdegem, 2013). Therefore, the economic estimates are provided as a template of an AI-CBA that is enhanced by local economic data.

The interpretation was informed by sensitivity and Monte Carlo analysis which identified which assumptions are most important. The farmgate price was the biggest contributor to the median net present value followed by the AI-predicted yield-gain multiplier and discount rate. This means that the cost of the device alone should not be the measure for smart aquaculture appraisal based on scenario. Whether a technically viable smart system will become financially viable may depend on access to the market, volatility of the price, and the response of the realized yield. More general sustainability evaluation also highlights the necessity of considering all three dimensions—economic, environmental, and social—when assessing sustainability and not just from the technology itself (Anderson et al., 2024).

The managerial implication that is most important is the adoption gap result. The higher the current smart technology intensity of a pond, the higher the yield, but the lower the potential gain from applying more smart technology. On the other hand, the bigger the adoption gap of ponds, the higher the incremental yield and economic returns under the simulated smart-upgrade scenario. Therefore, the concept is that smart aquaculture should not be thought of as a package of technologies for all ponds. The highest marginal value is likely to be in situations where baseline monitoring and adaptive management are not in full swing, but production scale and management responsiveness are adequate to achieve technology-enabled gains. Further research is needed to confirm this approach with actual financial data from the producers, vendor cost estimates, and species-specific market prices as well as with long-term adoption testing in the production areas.

6. CONCLUSION

A public pond dataset was used to create a scenario-based cost–benefit framework to estimate the production and financial value of smart aquaculture technologies. Smart aquaculture was operationalized by means of proxies such as monitoring, advisory decision support, corrective action implementation, follow-up monitoring, equipment support, self-initiated management and adoption-gap indicators. This allowed for the evaluation of smart aquaculture as a complete management system instead of a single device. Machine-learning models were used to predict pond-level yield, with Extra Trees being the most successful model. This model was embedded in a counterfactual approach to calculate incremental yield and investment feasibility in an improved smart-aquaculture scenario. The framework predicted positive economic parameters under baseline scenario, such as median net present value of USD 46,718.91, median benefit–cost ratio of 7.42 and 65.70% economically feasible ponds. Farmgate price, AI-predicted yield gain and discount rate were identified as feasibility determinants through sensitivity and Monte Carlo analyses. The results of the adoption-gap analysis indicated that adoption of smart management in ponds with lower smart-management intensity and larger potential for improvement led to higher marginal gains. Therefore, smart investment in aquaculture is not all-inclusive. Due to private financial information not being available, the results should be viewed as conditional scenario-based estimates and not actual farm profitability. The study provides a reproducible AI-CBA framework that needs to be validated with financial records, prices, quotations and field trials.

ACKNOWLEDGEMENT

[Insert acknowledgement of institutional, technical, or advisory support here. If no acknowledgement is required, state: The authors acknowledge the public availability of the dataset used in this study.]

Funding

[Insert funding information here. If no funding was received, state: This research received no external funding.]

Data Availability Statement

[Insert final dataset URL, repository name, access date, and any license information here. Generated analytical tables and figures may be made available by the corresponding author upon reasonable request.]

Ethics Statement

[Insert ethics approval details if applicable. If only anonymized public secondary data were used, state: This study used anonymized publicly available secondary data and did not involve direct experimentation on humans or animals by the authors.]

Author Contributions

[Insert author contribution statement using roles such as conceptualization, methodology, software, validation, formal analysis, writing-original draft, and writing-review and editing.]

Conflict of Interest

[Insert conflict-of-interest statement. If none, state: The authors declare no conflict of interest.]

AI Use Disclosure

[Insert disclosure according to journal policy. Example if applicable: The authors used AI-assisted tools only for language editing and formatting support. All intellectual content, analysis, interpretation, and references were verified by the authors, who take full responsibility for the manuscript.]

REFERENCES

1. Belton, B. and Little, D.C. (2011) ‘Immanent and interventionist inland Asian aquaculture development and its outcomes’, *Development Policy Review*, 29(4), pp. 459–484. https://doi.org/10.1111/j.1467-7679.2011.00542.x.
2. Bhosale, R., Jadhav, A., Jagadale, S., Shendkar, B., Gadekar, A. and Chandre, P. (2025) ‘AI-driven marine vessel detection from satellite imagery for enhanced naval surveillance’, in *2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG)*. Indore, Madhya Pradesh, India, pp. 1–11. https://doi.org/10.1109/ICTBIG68706.2025.11323721.
3. Boardman, A.E., Greenberg, D.H., Vining, A.R. and Weimer, D.L. (2018) *Cost-Benefit Analysis: Concepts and Practice*. 5th edn. Cambridge: Cambridge University Press.
4. Boyd, C.E., D’Abramo, L.R., Glencross, B.D., Huyben, D.C., Juarez, L.M., Lockwood, G.S., McNevin, A.A., Tacon, A.G.J., Teletchea, F., Tomasso, J.R., Tucker, C.S. and Valenti, W.C. (2020) ‘Achieving sustainable aquaculture: Historical and current perspectives and future needs and challenges’, *Journal of the World Aquaculture Society*, 51(3), pp. 578–633. https://doi.org/10.1111/jwas.12714.
5. Breiman, L. (2001) ‘Random forests’, *Machine Learning*, 45(1), pp. 5–32. https://doi.org/10.1023/A:1010933404324.
6. Damre, S.S., Shendkar, B.D., Kulkarni, N., Chandre, P.R. and Deshmukh, S. (2024) ‘Smart healthcare wearable device for early disease detection using machine learning’, *International*

Journal of Intelligent Systems and Applications in Engineering, 12(4s), pp. 158–166.
7. FAO (2024) *The State of World Fisheries and Aquaculture 2024: Blue Transformation in Action*. Rome: Food and Agriculture Organization of the United Nations. https://doi.org/10.4060/cd0683en.
8. Fish Welfare Initiative (2026) *public_ara_data: Public dataset from the Alliance for Responsible Aquaculture (ARA) program — water quality, stocking/harvest, enrollment, and dropout data from fish farms in Andhra Pradesh, India*. GitHub repository. Available at: https://github.com/fish-welfare-initiative/public_ara_data (Accessed: 3 July 2026).
9. Føre, M., Frank, K., Norton, T., Svendsen, E., Alfredsen, J.A., Dempster, T., Eguiraun, H., Watson, W., Stahl, A., Sunde, L.M., Schellewald, C., Skøien, K.R., Alver, M.O. and Berckmans, D. (2018) 'Precision fish farming: A new framework to improve production in aquaculture', *Biosystems Engineering*, 173, pp. 176–193. https://doi.org/10.1016/j.biosystemseng.2017.10.014.
10. Friedman, J.H. (2001) 'Greedy function approximation: A gradient boosting machine', *The Annals of Statistics*, 29(5), pp. 1189–1232. https://doi.org/10.1214/aos/1013203451.
11. Garlock, T.M., Asche, F., Anderson, J.L., Eggert, H., Anderson, T.M., Che, B., Chávez, C.A., Chu, J., Chukwuone, N., Dey, M.M., Fitzsimmons, K., Flores, J., Guillen, J., Kumar, G., Liu, L., Llorente, I., Nguyen, L., Nielsen, R., Pincinato, R.B.M., Sudhakaran, P.O., Tibesigwa, B. and Tveteras, R. (2024) 'Environmental, economic, and social sustainability in aquaculture: the aquaculture performance indicators', *Nature Communications*, 15, Article 5274. https://doi.org/10.1038/s41467-024-49556-8.
12. Geurts, P., Ernst, D. and Wehenkel, L. (2006) 'Extremely randomized trees', *Machine Learning*, 63(1), pp. 3–42. https://doi.org/10.1007/s10994-006-6226-1.
13. Gittinger, J.P. (1982) *Economic Analysis of Agricultural Projects*. 2nd edn. Baltimore: Johns Hopkins University Press.
14. Kulkarni, N. and Shendkar, B. (2022) 'Big data analytics in healthcare using AI', *Harbin Gongye Daxue Xuebao/Journal of Harbin Institute of Technology*, 54(5), pp. 201–208.
15. Molnar, C. (2022) *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. 2nd edn. Munich: Christoph Molnar.
16. Naylor, R.L., Hardy, R.W., Buschmann, A.H., Bush, S.R., Cao, L., Klinger, D.H., Little, D.C., Lubchenco, J., Shumway, S.E. and Troell, M. (2021) 'A 20-year retrospective review of global aquaculture', *Nature*, 591(7851), pp. 551–563. https://doi.org/10.1038/s41586-021-03308-6.
17. Pannell, D.J. (1997) 'Sensitivity analysis of normative economic models: Theoretical framework and practical strategies', *Agricultural Economics*, 16(2), pp. 139–152. https://doi.org/10.1016/S0169-5150(96)01217-0.
18. Rashid, M.M., Nayan, A.-A., Simi, S.A., Saha, J., Rahman, M.O. and Kibria, M.G. (2021) 'IoT based smart water quality prediction for biofloc aquaculture', *International Journal of Advanced Computer Science and Applications*, 12(6), pp. 470–477. https://doi.org/10.14569/IJACSA.2021.0120608.
19. Rastegari, H., Nadi, F., Lam, S.S., Ikhwanuddin, M., Kasan, N.A., Rahmat, R.F. and Mahari, W.A.W. (2023) 'Internet of Things in aquaculture: A review of the challenges and potential solutions based on current and future trends', *Smart Agricultural Technology*, 4, 100187. https://doi.org/10.1016/j.atech.2023.100187.
20. Saberioon, M., Gholizadeh, A., Cisar, P., Pautsina, A. and Urban, J. (2017) 'Application of machine vision systems in aquaculture with emphasis on fish: State-of-the-art and key issues', *Reviews in Aquaculture*, 9(4), pp. 369–387. https://doi.org/10.1111/raq.12143.
21. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. (2008) *Global Sensitivity Analysis: The Primer*. Chichester: John Wiley & Sons.
22. Tacon, A.G.J. and Metian, M. (2015) 'Feed matters: Satisfying the feed demand of aquaculture', *Reviews in Fisheries Science & Aquaculture*, 23(1), pp. 1–10. https://doi.org/10.1080/23308249.2014.987209.
23. Teixeira, R.R., Puccinelli, J.B., Poersch, L., Pias, M.R., Oliveira, V.M., Janati, A. and Paris, M. (2021) 'Towards precision aquaculture: a high performance, cost-effective IoT approach', *arXiv* [preprint]. https://doi.org/10.48550/arXiv.2105.11493.
24. Verdegem, M.C.J. (2013) 'Nutrient discharge from aquaculture operations in function of system design and production environment', *Reviews in Aquaculture*, 5(3), pp. 158–171. https://doi.org/10.1111/raq.12011.
25. Wu, B., Wang, C. and Du, K.-L. (2023) 'A smart aquaculture system exploiting IoT, AI and cloud computing', in *2023 IEEE International Conference on Smart Internet of Things (SmartIoT)*. Xining, China, 25–27 August 2023. Piscataway, NJ: IEEE, pp. 251–256. https://doi.org/10.1109/SmartIoT58732.2023.00045.
26. Zhao, S., Zhang, S., Liu, J., Wang, H., Zhu, J., Li, D. and Zhao, R. (2021) 'Application of machine learning in intelligent fish aquaculture: A review', *Aquaculture*, 540, 736724. https://doi.org/10.1016/j.aquaculture.2021.736724.
27. Zion, B. (2012) 'The use of computer vision technologies in aquaculture: A review', *Computers and Electronics in Agriculture*, 88, pp. 125–132. https://doi.org/10.1016/j.compag.2012.07.010.