Journal of Agricultural Sciences (Tarim Bilimleri Dergisi) 2025, 31 (1) : 71 – 79 DOI: 10.15832/ankutbd.1470111

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Application of the Machine Learning Methods to Assess the Impact of Physico-chemical Characteristics of Water on Feed Consumption in Fish Farms

Nedim Özdemir[a], Mustafa Çakir[b], Mesut Yilmaz[c], Hava Şimşek[a], Mükerrem Atalay Oral[d], Okan Oral[e*]

[a]*Faculty of Fisheries, Muğla Sıtkı Koçman University, Muğla, TURKEY*
[b]*Iskenderun Vocational School of Higher Education, İskenderun Technical University, İskenderun, TURKEY*
[c]*Faculty of Fisheries, Akdeniz University, Antalya, TURKEY*
[d]*Faculty of Maritime, Akdeniz University, Antalya, TURKEY*
[e]*Faculty of Engineering, Akdeniz University, Antalya, TURKEY*

ARTICLE INFO
Research Article
Corresponding Author: Okan Oral, E-mail: okan@akdeniz.edu.tr
Received: 17 April 2023 / Revised: 25 July 2024 / Accepted: 31 July 2024 / Online: 14 January 2025

Cite this article
Özdemir N, Çakir M, Yilmaz M, Şimşek H, Oral M A, Oral O (2025). Application of the Machine Learning Methods to Assess the Impact of Physico-chemical Characteristics of Water on Feed Consumption in Fish Farms. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):71-79. DOI: 10.15832/ankutbd.1470111

ABSTRACT
Machine learning methods, which are one of the subfields of artificial intelligence and have gained popularity in applications in recent years, play an important role in solving many challenges in aquaculture. In this study, the relationship between changes in the physico-chemical characteristics of water and feed consumption was evaluated using machine learning methods. Eleven physico-chemical characteristics (temperature, pH, dissolved oxygen, electrical conductivity, salinity, nitrite nitrogen, nitrate nitrogen, ammonium nitrogen, total phosphorus, total suspended solids, and biological oxygen demand) of water were assessed. Among all the measured physico-chemical characteristics of water, temperature was determined to be the most important parameter to be evaluated in fish feeding. Moreover, pH2, EC2, TP2, TSS2, S2 and NO_2-N parameters detected in the outlet water are more important than those detected in the inlet water in terms of feed consumption. Through regression analysis carried out using machine learning methods, the models developed with Random Forest, Gradient Boosting Machine and eXtreme Gradient Boosting algorithms exhibited higher success rates in predicting feed consumption compared to the other models. The present study highlights the pivotal role of machine learning methods in enhancing our understanding of fish feeding dynamics based on physico-chemical characteristics of water, thus contributing significantly to aquaculture management practices.

Keywords: Aquaculture, Feed intake, Artificial intelligence, Rainbow trout, Sustainability

1. Introduction

Fish farms release various amounts of waste into the aquatic environment, which increases the necessity of examining and analysing the impact of aquaculture on the environment (Ahmad et al. 2022). The polluted environment primarily harms biodiversity, disrupts ecological balances and prevents sustainability by affecting production (Leaf & Weber 1998; Sharma & Birman 2024).

The total trout production of Turkey was 145649 tons in 2022, and Muğla province contributed to this production with 18.2% of the total amount (Çöteli 2023). The Eşen River is at the heart of intensive trout production in Muğla province (Sezgin et al. 2023; Koçer et al. 2010; Pulatsü & Yıldırım 2011).

For sustainability, it is essential to ensure that natural resources are used effectively and efficiently, in a balanced way, and in harmony with nature (Qin et al. 2024; Moldan et al. 2012). It is necessary to know the potential and structure of natural resources well and to observe the changes that occur in these resources. Research and monitoring activities in aquaculture are necessary to manage resources (Subasinghe et al. 2009; Mandal & Ghosh 2024).

New application areas and algorithms are constantly being developed with artificial intelligence (Kaya et al. 2023; Akgül et al. 2023; Kaya 2023). Many successful activities have been reported in fish farms, such as planning production with computer support, monitoring environmental conditions and fish health (Yilmaz et al. 2022; Yilmaz et al. 2023; Cakir et al. 2023), and aeration tools, growth statistics, intensive data analysis, production of feed (Dikel & Öz 2022). Feed expenses constitute the largest part of the production cost in aquaculture farms (Li et al. 2020). Of course, the balance established between the amount of harvested product and the amount of feed consumed represents successful production (Pahlow et al. 2015). However, the concept of production solely for commercial concerns negatively affects sustainability (Folke & Kautsky 1992). Changes in water quality are the critical factor that directly affect fish production (Muir 2005). Evaluating the sensitive relationship between the physico-chemical characteristics of water and feed consumption with the help of computer-aided methods can help both optimize the amount of feed used and predict the amount of feed to be consumed on the farm (Zhao et al. 2021). Thus, while reducing production costs, predictable feed consumption will allow businesses to make more accurate future.

Machine learning (ML) methods, which are one of the subfields of artificial intelligence (AI) and have gained popularity in applications in recent years, play an important role in solving many challenges in aquaculture (Zhao et al. 2021).

In the current study, the relationship between changes in the physico-chemical characteristics of water and feed consumption was evaluated using ML methods. At the same time, it was aimed to determine the most successful ML method for predicting the effect of environmental conditions on rainbow trout production.

2. Material and Methods

2.1. Study area

The Eşen River (Figure 1) originates at approximately 2000 meters altitude in the southwest of Turkey. This river is used for different purposes such as drinking water supply, electricity generation, agricultural irrigation, and Rainbow Trout breeding (*Oncorhynchus mykiss* Walbaum 1792).

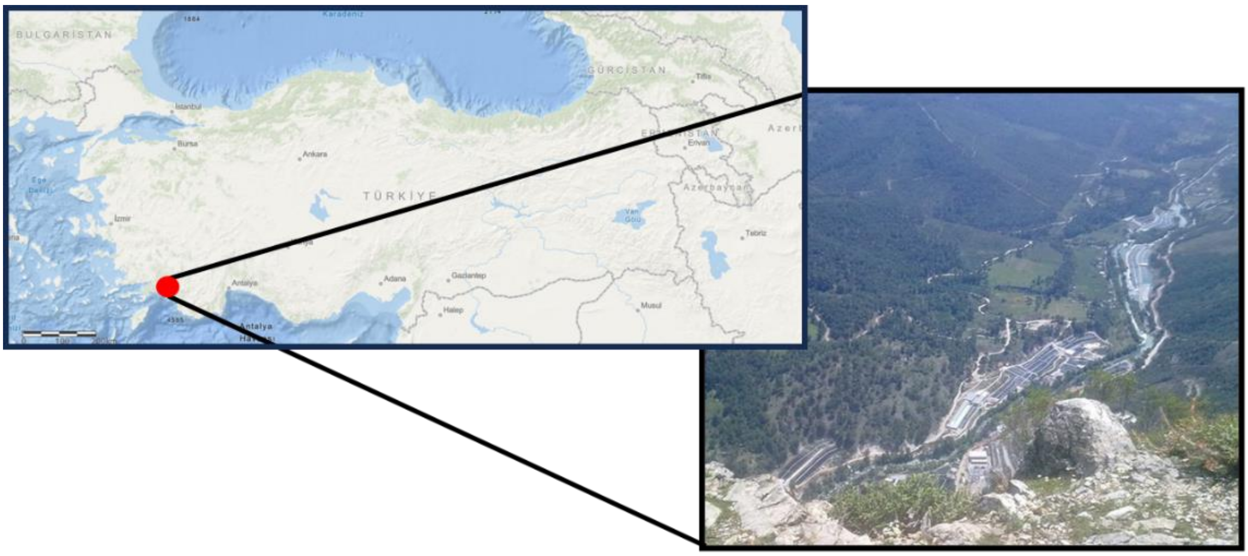

Figure 1- A view from Eşen River (Original)

Seven stations, including inlet and outlet waters of 3 farms and reference point, were monitored for 1 year. Moreover, feed consumption rates of farms were recorded. While determining the stations, care was taken to ensure that they were representative of the fish farms.

The Reference point is located 115 m upstream of the farms, which are located on Eşen River. This station, which is not exposed to pollutants, also serves as a reference for the rest of the river.

2.2. Water sampling and analysis

Water samples obtained from a total of 7 stations were filled into 2-liter polyethylene bottles and transported to Muğla Sıtkı Koçman University Laboratories in an ice-cooled box.

From the water samples taken from the stations, water temperature (T), pH, dissolved oxygen (DO), electrical conductivity (EC), and salinity (S) parameters were determined on-site in the field with a YSI multiparameter (MPS 556). Nitrite nitrogen (NO_2-N), nitrate nitrogen (NO_3-N), ammonium nitrogen (NH_4), total phosphorus (TP), total suspended solids (TSS) and biological oxygen demand (BOD) analyses were carried out in Muğla Sıtkı Koçman University Research Laboratories and Central Environmental Laboratory according to APHA (2012) methods.

2.3. ML methods

Regression methods are statistical techniques used to predict one variable based on one or more other variables. These methods model the relationship between data and make predictions using these models. There are many algorithms used in literature to obtain these predictions. However, some algorithms are more preferred in academic studies because they produce more successful predictions compared to others or because the models can be easily interpreted. These popular algorithms include Artificial Neural Network (ANN), Decision Tree (DT), Generalized Linear Model (GLM), Gradient Boosting Machine (GBM), K-Nearest Neighbour (KNN), Random Forest (RF), Support Vector Machine (SVM) and eXtreme Gradient Boosting (XGBoost). The success of these algorithms may vary according to the structure and complexity of the problem.

The regression models aim to estimate the relationship between at least one input variable and one output variable. Input variables can be defined as independent variables and output variables as dependent variables. The input variables used in this study are the physico-chemical characteristics of water, while the output variable is Monthly Feed Consumption per kg (MFC). Input variables are inlet and outlet water temperatures (T1, T2), pH values (pH1, pH2), dissolved oxygen amounts (DO1, DO2), electrical conductivities (EC1, EC2), salinity values (S1, S2), NH_4, NO_2 and NO_3 values ($NH_4$1, $NH_4$2, $NO_2$1, $NO_2$2, $NO_3$1, $NO_3$2), total suspended solids amount (TSS1, TSS2), Total phosphorus values (TP1, TP2) and biological oxygen needs (BOD1, BOD2). Since the stocking density for portion fish was the same in all farms included in the study (25 kg/m^3), stocking density was not included in the study as a variable.

In the regression analysis, ML techniques were used to model the relationship between the above-mentioned input variables and the output variable. The dataset contains 22 input variables and one output variable for a total of 180 observations. A significant portion of the dataset (80%) was randomly divided to be used for model training and the remaining portion (20%) for testing. The 10-fold cross-validation method was applied for the validation of the model to be created during the training phase. In addition, hyperparameter optimization of the regression algorithms was achieved to obtain the best model in a limited solution space.

To measure the performance of regression models, three different evaluation metrics are commonly used in studies, which are Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and R^2 respectively.

MAE is the average of the absolute values of the differences between the predicted values (ŷ) and the actual values (y) and is expressed as in Equation 1. N represents the number of observations (James et al. 2013).

$$MAE = (1/\mathrm{N}) * \sum|\mathrm{y} - \hat{\mathrm{y}}| \quad (1)$$

RMSE is another evaluation metric used to measure the performance of a regression model, and it is the square root of the mean square of the differences between the predicted values and the actual values (Géron 2019).

Mathematically, for a data set with N observations, the RMSE is calculated as shown in Equation 2.

$$RMSE = \sqrt{(1/\mathrm{N}) * \sum|\mathrm{y} - \hat{\mathrm{y}}|} \quad (2)$$

Since RMSE is calculated using the squares of the error quantities, large errors have a more significant impact, while small errors have a smaller impact. This makes RMSE a more sensitive evaluation metric than MAE, making it more suitable for datasets with outliers. In cases where outliers lead to large errors, the RMSE will be higher. However, RMSE is preferred because it is more sensitive than MAE, because of the use of the squares of the errors.

R^2 or coefficient of determination is an evaluation metric used to measure how well a regression model is fitted. This metric expresses how much of the variance of the actual data is explained by the model (Draper & Smith, 1998).

Mathematically, R^2 is calculated by the formula in Equation 3:

$$R^2 = 1 - (SSres / SStot) \quad (3)$$

Where; SSres represents the sum of the error squares, which measures the deviation of the model's predictions from the actual data, and SStot represents the total variance of the actual data.

The R^2 value ranges from 0 to 1, with a higher value indicating a better fit to the data. An R^2 value of 1 means that the model explains the actual data perfectly while a value of 0 means that the model does not explain the data at all. In a successful regression model, the MAE and RMSE error metrics are expected to be as small as possible, approaching zero, while the R^2 metric is expected to be close to 1.

3. Results and Discussions

3.1. Evaluation metrics of regression models

In this section, we present the results of the regression models where 4/5 of the dataset consisting of 180 observations was used for training the model, and 1/5 of the dataset was used for testing. Three different metrics were used to evaluate the ML regression models. Among these metrics, MAE and RMSE are error metrics, and these metric values can take values ranging from zero to infinity. In models with good performance, MAE and RMSE values are expected to be close to zero. These metrics are obtained both in the training phase and in the testing phase. Since 10-fold cross-validation is applied in the training phase, the error metrics consist of ten different values instead of a single value.

In the fish farms, where the study was conducted, a certain proportion of biomass is given depending on the water temperature and the amount of dissolved oxygen in line with the recommendations of the feed manufacturer. However, fish behavior during feeding is still evaluated by technical personnel. Feeding is stopped when a situation affecting feed intake (such as a decrease in the fish's feed intake or reluctance to feed) is observed. Therefore, the amounts of feed consumption in farms are affected by instantaneous changes. Possible reasons that may reduce the fish's demand for food cannot be monitored instantly. The obtained results can explain the possible reasons for the parameters affecting feed consumption. The order of importance of physico-chemical characteristics of water affecting feed consumption was evaluated with BORUTA analysis (Figure 2).

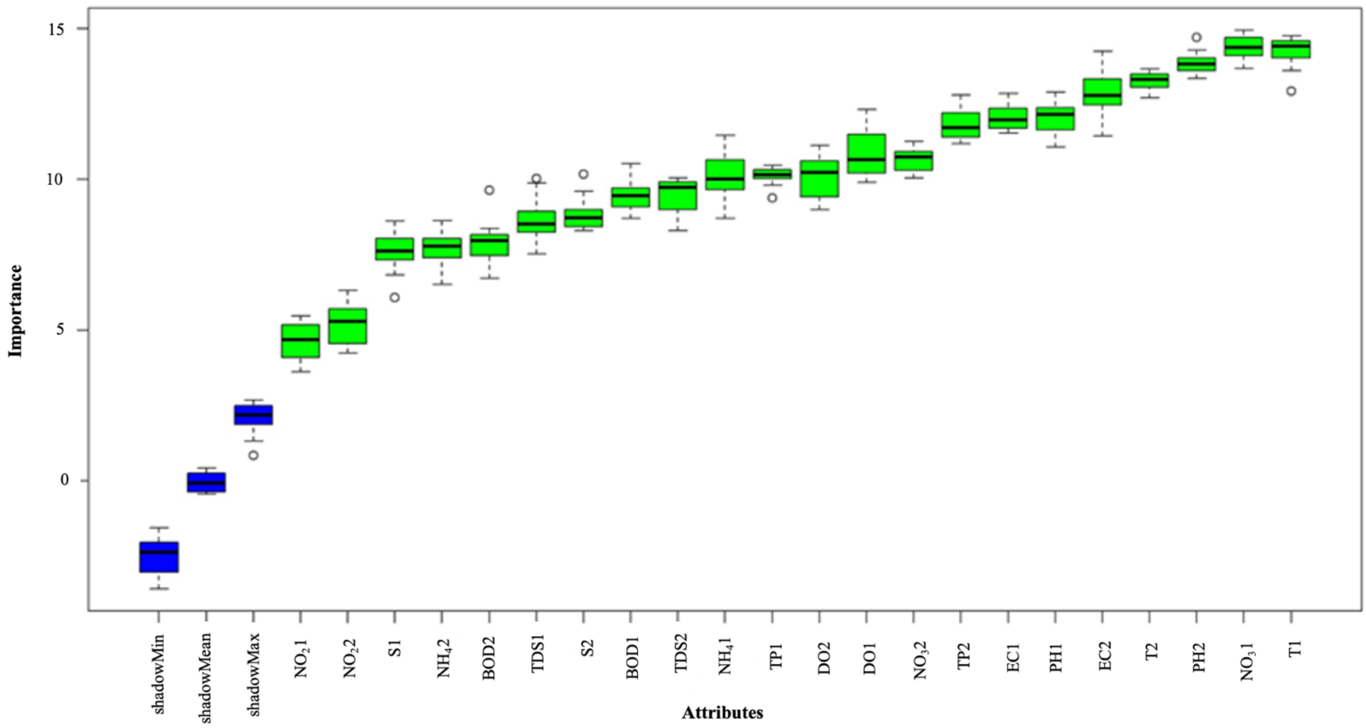

Figure 2- The order of importance of physico-chemical characteristics of water affecting feed consumption

These parameters of the wastes (caused by feed, fish and metabolic wastes) released into the water from the farm environment can be seen from the analysis results of the farm inlet and outlet waters that are directly affected by the physico-chemical characteristics of water (pH, TP, NH_4, BOD, NO_2-N, NO_3-N, EC, TSS, S, DO and T) evaluated in the current study in the farm environment (Table 1).

The pH2, EC2, TP2, TSS2, S2 and NO_2 parameters detected in the outlet water are more important than those detected in the inlet water in terms of feed consumption (Figure 2). Surprisingly, NO_3, pH, EC, TP2 are ranked as the most important and more important than DO, which is evaluated first (along with water temperature) in fish farms. In trout farms, the DO rate is desired to be above 6 mg/L for feed consumption (Pedersen 1987; McDaniel et al. 2005). In the farms where the study was carried out, DO mean values were measured in the range of 6.43-8.40 mg/L.

Among all the measured physico-chemical characteristics of water, temperature was determined to be the most important parameter to be evaluated in fish feeding (Figure 2). In trout farming, temperatures in the range of 1-25 °C (with an optimal 16 °C) are needed to grow fish (Woynarovich et al. 2011). In the current study, water temperature was measured between 11.30 and 14.88 °C. This value is within the desired range for trout farming and is close to the optimum value (16 °C).

Table 1- Measured physico-chemical characteristics of water (min-max / mean)

Physico-chemical Characteristics	*Reference*	*Farm A*		*Farm B*		*Farm C*	
		Inlet	*Outlet*	*Inlet*	*Outlet*	*Inlet*	*Outlet*
T (ºC)	11.32-13.75	11.37-13.77	11.31-13.78	12.50-13.68	12.30-14.29	11.30-14.75	11.30-14.88
	13.04	13.09	13.05	13.08	13.27	13.12	13.12
pH	7.53-9.31	7.70-8.73	7.44-8.70	7.67-9.12	7.46-8.75	7.50-9.12	7.66-8.87
	8.15	8.14	8.01	8.15	7.97	8.08	8.05
DO (mgL^{-1})	7.43-10.51	7.3-10.43	5.49-9.00	5.08-10.3	4.97-9.21	5.42-10.26	5.1-9.00
	8.57	8.40	6.85	7.09	6.43	7.14	6.67
EC (μScm^{-1})	254.80-400.00	253.80-364.00	257.30-401.00	254.20-404.00	258.60-405.00	254.40-390.00	253.30-390.00
	336.46	337.70	344.55	337.45	348.40	344.18	338.97
S (‰)	0.160-0.220	0.160-0.230	0.160-0.230	0.160-0.220	0.160-0.230	0.160-0.220	0.16-0.220
	0.185	0.190	0.190	0.186	0.190	0.187	0.188
NH_4 (mgL^{-1})	BDL-0.10	BDL-0.13	BDL-0.54	BDL-0.39	0.23-2.92	BDL-0.53	BDL-0.82
	0.01	0.02	0.31	0.27	0.78	0.34	0.57
NO_2-N (mgL^{-1})	BDL-0.05	BDL-0.02	BDL-BDL	BDL-0.25	BDL-0.20	BDL-0.50	BDL-0.10
	0.01	0.01	BDL	0.07	0.06	0.07	0.03
NO_3-N (mgL^{-1})	BDL-10.37	0.91-9.59	1.19-6.56	1.30-17.51	1.20-28.39	1.16-20.56	1.04-22.59
	2.84	2.89	2.29	3.76	4.92	4.33	4.18
TP (mg L^{-1})	BDL-0.285	BDL-0.005	BDL-0.046	BDL-0.043	0.013- 0.108	BDL-0.084	0.011- 0.914
	0.026	0.001	0.024	0.030	0.073	0.042	0.137
BOD_5 (mgL^{-1})	0.07-4.26	0.21-3.14	1.16-4.28	0.41-4.85	2.33-5.70	0.76-4.62	1.02-5.37
	1.94	2.16	2.7	2.71	3.97	2.77	3.21
TSS (mgL^{-1})	BDL-1.00	BDL-0.40	0.10-10.90	0.10-6.50	0.30-3.90	0.40-4.30	0.50-8.20
	0.23	0.11	1.91	1.65	2.23	1.58	2.53
Total Feed Consumption (ton)		649.5		1156.2		748.0	

*BDL: Below Detection Limit

According to the European Union directives, 2006/44/EC, (EU, 2006) on the quality of fresh waters that need to be protected or improved to support fish life, the pH, TP, NH_4 and BOD parameters of the water for salmonids should be 6-9, 0.2, ≤1 and ≤3 mg/L, respectively. During the study, pH was found to be 7.44-9.12, TP BDL was found to be 0.914, NH_4 BDL-2.92 and BOD were found to be between 0.21 and 5.70 in the farms. Although the measured TP, NH_4 and BOD values exceeded the values required for the growth of salmonids, pH2 measured within the recommended value range is the 3rd most important parameter affecting feed consumption among all the parameters. The fact that temperature and pH have an impact on all chemical reactions, metabolism, and toxicity (Jana & Sarkar 2005) may explain why they are at the top of the list of importance.

According to BORUTA analysis, although NO_2-N is more toxic in terms of feed consumption, it is ranked lower than NO_3-N in terms of importance. Inlet water NO_3-N value is listed as the most important parameter, right after temperature. NO_3 is the final product of the two-stage oxidation of ammonia. The intermediate product, NO_2, is oxidized with the help of bacteria to produce nitrate (Hargreaves 1998). Therefore, it can be said that the NO_2 concentration in the environment is kept at levels that will not affect feed consumption through nitrification.

High values of total dissolved solids, which describe inorganic salts and dissolved materials in water (Devi et al. 2017; Firooz et al. 2012), mean that they are unsuitable for fish health (Ahmed et al. 2019). It has been reported that total dissolved solids are among the top physico-chemical characteristics of water that cause fish disease outbreaks (Yılmaz et al. 2022). However, TSS, which was revealed in the current study to be an important parameter in terms of feed consumption, comes after other physico-chemical characteristics of water in importance. Similarly, S is one of the parameters that affects the fish's feed consumption less.

EC has greater importance on feed intake than TSS and S, comparable to pH and less important than temperature. In fact, it is affected by these 4 physico-chemical characteristics of water and the presence of inorganic dissolved solids such as ions carrying a negative charge (nitrate and phosphate anions) and ions carrying a positive charge (sodium, calcium, iron, etc.). With this feature, EC can be considered a critical indicator in evaluating feed consumption in fish farms where ML methods are used.

It should be noted that the data in this study are limited to the current measured values of physico-chemical characteristics of water measured in the farms where the study was carried out. In other words, it can be said that the order of importance of physico-chemical characteristics of water in terms of feed consumption may vary in each farm's own dynamics. It was revealed in the current study that a critical parameter such as DO may be placed behind other physico-chemical characteristics of water in terms of importance in farm environments where it remains at values that will not negatively affect fish feed consumption throughout the season (Figure 2). This shows that the farm environment should be evaluated with all its variables for effective production.

Table 2 shows the minimum, 1st quantile, median, mean, 3rd quantile, and maximum values of the MAE-based comparative error values for each regression model during the training process. As shown in the table, the DT algorithm has the highest average MAE error among the regression models. The regression models with the lowest mean MAE error values are XGBoost, GBM and RF. It can be observed that the minimum and maximum error ranges are narrower for the GLM, RF and GBM models, but wider for the ANN and DT models in terms of MAE error values.

Table 2- MAE-based error table for training data

	MAE					
	Min.	***1st Qu.***	***Median***	***Mean***	***3rd Qu.***	***Max.***
DT	0.0162611160	0.0218714429	0.024323865	0.025098375	0.028672426	0.034461268
ANN	0.0112890995	0.0123856034	0.013982100	0.016852160	0.017949030	0.037744668
SVM	0.0109737158	0.0118979884	0.012412511	0.014065447	0.016024698	0.019771007
GLM	0.0111900914	0.0121773948	0.013181225	0.013256358	0.014288051	0.015159422
KNN	0.0062285714	0.0086330769	0.012843077	0.011462479	0.014234856	0.015253333
RF	0.0039281359	0.0046053656	0.004942887	0.005486680	0.006135318	0.008105229
GBM	0.0032916726	0.0043113957	0.005101034	0.005223984	0.005706219	0.008745502
XGBoost	0.0002086571	0.0009246325	0.003543157	0.003952663	0.006798657	0.008891349

Table 3 shows the minimum, 1st quantile, median, mean, 3rd quantile and maximum values of the RMSE-based comparative error values of each regression model during the training process. As can be seen from the table, the regression model with the highest mean RMSE error is the DT algorithm. The regression models with the lowest RMSE mean error values are GBM, XGBoost and RF. Regarding RMSE error values, it is understood that the minimum and maximum error ranges are narrower for GLM, RF and GBM models but wider for ANN, DT and XGBoost.

Table 3- RMSE-based error table for training data

	RMSE					
	Min.	***1st Qu.***	***Median***	***Mean***	***3rd Qu.***	***Max.***
DT	0.021817266	0.025538575	0.029346811	0.030055109	0.032247176	0.04257985
ANN	0.014761440	0.015181156	0.018514751	0.020932702	0.021602019	0.04369090
SVM	0.014338495	0.015829987	0.017048988	0.018649991	0.021480378	0.02544829
GLM	0.013903197	0.015474243	0.016332362	0.016662840	0.017832495	0.01985353
KNN	0.008120521	0.011424263	0.016634153	0.015426489	0.018643893	0.02231042
RF	0.005425699	0.006055001	0.007568518	0.007929979	0.008630886	0.01313262
XGBoost	0.000285502	0.001776021	0.007803664	0.007383174	0.012750269	0.01427064
GBM	0.004772173	0.005297180	0.007186077	0.007276941	0.007980441	0.01348817

The minimum, 1st quantile, median, mean, 3rd quantile and maximum values of the R^2 based comparative model explanatory power values of each regression model during the training process are shown in Table 4. As can be seen from the Table 4, the regression models with the highest mean R^2 values are RF, GLM and XGBoost algorithms. The regression model with the lowest mean R^2 model explanatory power is DT. It is understood that the XGBoost model produced the maximum R^2 model explanatory power values.

Table 4- R^2-based performance table for training data

	R^2					
	Min.	*1st Qu.*	*Median*	*Mean*	*3rd Qu.*	*Max.*
RF	0.8925463	0.9705421	0.9745838	**0.9701575**	0.9862443	0.9956343
GLM	0.9231790	0.9671374	0.9761152	**0.9698487**	0.9793847	0.9914711
XGBoost	0.8680894	0.9167459	0.9710129	**0.9536456**	0.9981806	0.9999758
KNN	0.7948448	0.8324811	0.9086410	0.8809891	0.9142515	0.9495288
GLM	0.8147133	0.8313405	0.8613456	0.8624089	0.8809775	0.9267126
SVM	0.7267628	0.7996369	0.8334013	0.8262501	0.8627225	0.8870258
ANN	0.2662195	0.7661713	0.8500663	0.7779701	0.8693918	0.8865680
DT	0.1951602	0.4227606	0.5267219	0.5051544	0.6157219	0.7078576

The highest mean R^2 values were indicated in bold

After analysing the model evaluation metrics obtained for the training data, it is necessary to measure the performance of the regression models with test data that the regression models have never seen in the training phase. The results of the evaluation metrics of the performance test with the test data are given in Table 5.

Table 5- Model evaluation values with test data

		Evaluation Metrics		
		MAE	*RMSE*	*R^2*
Regression Models	**DT**	0.022643467	0.027030900	0.598961001
	ANN	0.016318823	0.019655126	0.787960261
	SVM	0.014680428	0.019378929	0.793877621
	GLM	0.014104027	0.017393552	0.833948736
	KNN	0.011044444	0.014378379	0.886528852
	RF	0.004423242	0.006448258	0.977178175
	XGBoost	**0.003074641**	0.006178471	0.979047897
	GBM	0.004402947	**0.005872574**	**0.981071218**

The lowest MAE, RMSE and the highest R^2 values were indicated in bold

Based on the evaluation metrics in Table 5, XGBoost has the lowest MAE error value, followed closely by GBM and RF. Therefore, XGBoost, GBM and RF are the most successful models according to the MAE metric.

According to the RMSE metric (Table 5), the lowest error value belongs to the GBM model with 0.005872574. Similarly, XGBoost and RF models are also close to the GBM model with the lowest error value.

When the results are analysed according to the R^2 model explanatory power metric (Table 5), the GBM model can explain the test data in the best way. Close R^2 values are also observed for XGBoost and RF models.

In the regression analysis carried out using ML techniques, the models developed with RF, GBM and XGBoost algorithms yielded better results for both the training and test datasets compared to models developed with other algorithms. Based on the MAE, RMSE and R^2 metrics used to evaluate the models, it was observed that RF, GBM and XGBoost algorithms generated similar results, although the order of performance varied, indicating that all three models are effective. Hence, the analysis of the eight different models highlighted that RF, GBM and XGBoost are the top three most effective algorithms. On the other hand, the DT algorithm produced the least effective results both in terms of model training and test performance.

In recent years, machine learning methods have been tried to be adapted to aquaculture (Zhao et al. 2021). Since feed cost is the main determinant of production cost (Li et al. 2020), many studies including computer-aided feeding systems (Hu et al. 2022) and the development of smart systems that enable optimization of feeding amount by monitoring fish behaviour (Ubina et al. 2021; Du et al. 2023; Zhou et al. 2019) have been reported. In the current study, RF, GBM, and XGBoost were determined to be the most effective algorithms that enable the estimation of the feeding amount by monitoring physico-chemical characteristics of water in fish farms. Comparable to the current study, it has been reported that methods such as convolution neural network (Ubina et al. 2021) and deep learning techniques using the image processing method (Hu et al.

2022; Zhou et al. 2019) can be successfully used to optimize feeding in aquaculture. In addition, studies investigating the effects of changes in water quality on the stock (Rana et al. 2021), fish length estimation (Li et al. 2020) and biomass detection (Yang et al. 2021) have been successfully carried out using ML methods.

4. Conclusions

In fish farming, feed expenses are the main factor determining the production cost. Feeds that cannot be consumed on farms are directly dumped into the water. These unused feeds deteriorate physico-chemical characteristics of water in the farm environment, reducing production performance and quality. In addition, these unconsumed feeds dissolve in water and provide an environment for the proliferation of undesirable pathogenic microorganisms. The emergence of diseases threatens fish health and causes great losses in production. Moreover, since unused feed are released to nature, they are loaded into the receiving environment as pollutants.

For aquaculture production to be sustainable, appropriate business policies must be developed for the use of natural resources and input costs must be kept at reasonable levels. The development of computer-aided smart production systems has the potential to prevent losses, protect the environment and support healthy and economical production. Sustainable aquaculture can be supported through effective risk management using smart systems.

References

Ahmad A L, Chin J Y, Harun M H Z M & Low S C (2022). Environmental impacts and imperative technologies towards sustainable treatment of aquaculture wastewater: A review. *Journal of Water Process Engineering* 46: 102553

Ahmed M, Rahaman M O, Rahman M & Kashem M A (2019). Analyzing the quality of water and predicting the suitability for fish farming based on iot in the context of bangladesh. In 2019 *International Conference on Sustainable Technologies for Industry* 4.0 (STI) (pp. 1–5). IEEE

Akgül İ, Kaya V & Zencir Tanır Ö (2023). A novel hybrid system for automatic detection of fish quality from eye and gill color characteristics using transfer learning technique. *Plos one* 18(4): e0284804

Baird R B, Eaton A D & Rice E W (2017) Standart Methods for The Examination of Water and Wastewater, 23st Edition, American Public Health Association, Washington.

Cakir M, Yilmaz M, Oral M A, Kazanci H Ö & Oral O (2023). Accuracy assessment of RFerns, NB, SVM, and kNN machine learning classifiers in aquaculture. *Journal of King Saud University-Science* 35(6): 102754

Çöteli F T (2023). Aquaculture Product Report (In Turkish). Tepge publication No: 373, Republic of Turkey Ministry of Agriculture and Forestry, Ankara.

Dikel S & Öz M (2022). Artificial intelligence (AI) application in aquaculture. In ISPEC 10th International Conference on Agriculture, Animal Sciences and Rural Development, July 18-19, Sivas, Turkey.

Devi P A, Padmavathy P, Aanand S & Aruljothi K (2017). Review on water quality parameters in freshwater cage fish culture. *International Journal of Applied Research* 3(5): 114–120

Draper N R & Smith H (1998). Applied regression analysis. John Wiley & Sons.

Du L, Lu Z & Li D (2023). A novel automatic detection method for breeding behavior of broodstock based on improved YOLOv5. *Computers and Electronics in Agriculture* 206: 107639

EU 2006. Directive 2006/44/EC of the European Parliament and of the Council of 6 September 2006 on the quality of fresh waters needing protection or improvement in order to support fish life

Firooz F, Mehdi R, Mostafa F, Alireza M & Gholamhossein N F (2012). Evaluation of physicochemical parameters of waste water from rainbow trout fish farms and their impacts on water quality of Koohrang stream –Iran. *International Journal of Fisheries and Aquaculture* 4(8): 170–177

Folke C & Kautsky N (1992). Aquaculture with its environment: prospects for sustainability. *Ocean & coastal management* 17(1): 5-24

Géron A (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc.

Hargreaves J A (1998). Nitrogen biogeochemistry of aquaculture ponds. *Aquaculture* 166(3-4): 181-212

Hu W C, Chen L B, Huang B K & Lin H M (2022). A computer vision-based intelligent fish feeding system using deep learning techniques for aquaculture. *IEEE Sensors Journal* 22(7): 7185-7194

James G, Witten D, Hastie T, Tibshirani R & Taylor J (2013). An Introduction to Statistical Learning: with Applications in R. Springer, Switzerland.

Jana B B & Sarkar D (2005). Water quality in aquaculture-Impact and management: A review. *The Indian Journal of Animal Sciences 75*(11)

Kaya V (2023). A Perspective on Transfer Learning in Computer Vision. In: Kılıc G B (Ed.). Advances in Engineering Sciences, Platanus Publishing, Ankara, Turkey

Kaya V, Akgül İ & Tanır Ö Z (2023). IsVoNet8: a proposed deep learning model for classification of some fish species. *Journal of Agricultural Sciences* 29(1): 298-307

Koçer M A, Muhammetoğlu A, Emre Y, Sevgili H, Türkgülü İ, Kanyılmaz M, Yılayaz A, Uysal R, Emre N, Mefut A, Uysal G, Yalım B & Topcuoğlu Ö A (2010). Determination of the Effects of Fish Farming and Basin-Related Pollution on the Eşen Stream Ecosystem by Using Mathematical Modeling Methods and Control of Nutrient Flux to the Mediterranean (in Turkish). (TÜBİTAK Project No: 107Y084).

Leaf A. & Weber P C (1998). Cardiovascular effects ofn- 3 fatty acids. *The New England Journal of Medicine* 318: 549-557

Li D, Wang Z, Wu S, Miao Z, Du L & Duan Y (2020). Automatic recognition methods of fish feeding behavior in aquaculture: A review. *Aquaculture* 528: 735508

Mandal A & Ghosh A R (2024). Role of artificial intelligence (AI) in fish growth and health status monitoring: A review on sustainable aquaculture. *Aquaculture International* 32(3): 2791-2820

McDaniel N K, Sugiura S H, Kehler T, Fletcher J W, Coloso R. M, Weis P & Ferraris R P (2005). Dissolved oxygen and dietary phosphorus modulate utilization and effluent partitioning of phosphorus in rainbow trout (Oncorhynchus mykiss) aquaculture. *Environmental Pollution* 138(2): 350-357

Moldan B, Janoušková S & Hák T (2012). How to understand and measure environmental sustainability: Indicators and targets. *Ecological indicators* 17: 4-13

Muir J (2005). Managing to harvest? Perspectives on the potential of aquaculture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360(1453): 191-218

Pahlow M, Van Oel P R, Mekonnen M M & Hoekstra A Y (2015). Increasing pressure on freshwater resources due to terrestrial feed ingredients for aquaculture production. *Science of the Total Environment* 536: 847-857

Pedersen C L (1987). Energy budgets for juvenile rainbow trout at various oxygen concentrations. *Aquaculture* 62(3-4): 289-298

Pulatsü S & Yıldırım H B (2011) Evaluation of Outlet Water Characteristics of Land-Based Trout Farms of Different Capacities (Mugla, Fethiye) (in Turkish). Ankara University Scientific Research Project Final Report, Project no: 09B4347009

Qin C, Xue Q, Zhang J, Lu L, Xiong S, Xiao Y & Wang J (2024). A beautiful China initiative towards the harmony between humanity and the nature. *Frontiers of Environmental Science & Engineering* 18(6): 1-9

Rana M, Rahman A, Dabrowski J, Arnold S, McCulloch J & Pais B (2021). Machine learning approach to investigate the influence of water quality on aquatic livestock in freshwater ponds. *Biosystems Engineering* 208: 164-175

Sezgin S S, Yilmaz M, Arslan T & Kubilay A (2023). Current antibiotic sensitivity of Lactococcus garvieae in rainbow trout (Oncorhynchus mykiss) farms from Southwestern Turkey. *Journal of Agricultural Sciences* 29(2): 630-642

Sharma I & Birman S (2024). Biodiversity Loss, Ecosystem Services, and Their Role in Promoting Sustainable Health. In The Climate-Health-Sustainability Nexus: Understanding the Interconnected Impact on Populations and the Environment (pp. 163-188). Cham: Springer Nature, Switzerland.

Subasinghe R, Soto D & Jia J (2009). Global aquaculture and its role in sustainable development. *Reviews in aquaculture* 1(1): 2-9

Ubina N, Cheng S C, Chang C C & Chen H Y (2021). Evaluating fish feeding intensity in aquaculture with convolutional neural networks. *Aquacultural Engineering* 94: 102178

Woynarovich A, Hoitsy G & Moth-Poulsen T (2011). Small-scale rainbow trout farming. FAO fisheries and aquaculture technical paper, (561). Food and agriculture organization of the united nations, Rome

Yang L, Liu Y, Yu H, Fang X, Song L, Li D & Chen Y (2021). Computer vision models in intelligent aquaculture with emphasis on fish detection and behavior analysis: A review. *Archives of Computational Methods in Engineering* 28(4): 2785–2816

Yilmaz M, Çakir M, Oral M A, Kazanci H Ö & Oral O (2023). Evaluation of disease outbreak in terms of physico-chemical characteristics and heavy metal load of water in a fish farm with machine learning techniques. *Saudi Journal of Biological Sciences* 30(4): 103625.

Yilmaz M, Çakir M, Oral O, Oral M A & Arslan T (2022). Using machine learning technique for disease outbreak prediction in rainbow trout (Oncorhynchus mykiss) farms. *Aquaculture Research* 53(18): 6721-6732

Zhao S, Zhang S, Liu J, Wang H, Zhu J, Li D & Zhao R (2021). Application of machine learning in intelligent fish aquaculture: A review. *Aquaculture* 540: 736724

Zhou C, Xu D, Chen L, Zhang S, Sun C, Yang X & Wang Y (2019). Evaluation of fish feeding intensity in aquaculture using a convolutional neural network and machine vision. *Aquaculture* 507: 457-465