Journal of Fisheries
Volume 13 Issue 2 August 2025
Peer Reviewed | Open Access

eISSN 2311-3111
pISSN 2311-729X

DOI: https://doi.org/10.17017/j.fish.722

Original Article

# Application of machine learning to growth model in fisheries

**Semra Benzer[1] • Recep Benzer[2] • Ali Gül[3]**

[1] *Gazi Faculty of Education, Gazi University, Teknikokullar, Ankara 06500, Turkey*
[2] *School of Administrative and Social Sciences, Department of Management Information System, Ankara Medipol University, Ankara 06050, Turkey*
[3] *Education Faculty, Biology Education, Gazi University, Ankara 06050, Turkey*

**Correspondence**
Semra Benzer; Gazi Faculty of Education, Gazi University, Teknikokullar, Ankara 06500, Turkey
sbenzer@gazi.edu.tr and sbenzer@gmail.com

**Manuscript history**
Received 12 July 2024 | Accepted 17 March 2025 | Published online 10 May 2025

**Citation**
Benzer S, Benzer R, Gül A (2025) Application of machine learning to growth model in fisheries. Journal of Fisheries 13(2): 132202. DOI: 10.17017/j.fish.722

**Abstract**
Traditional growth models, such as length-weight relationships (LWRs) and the von Bertalanffy (VB) growth function, have been widely used in fishery science. Their limitations in capturing nonlinear patterns necessitate alternative approaches. Machine learning (ML) techniques have recently gained attention as a powerful tool for enhancing predictive accuracy in biological studies. In this study, the growth parameter of Eastern mosquitofish, *Gambusia holbrooki* (135 females: 21–58.78 mm and 0.152–3.424 g; 59 males: 19.25–43.20 mm; 0.108–1.075 g), was determined with traditional LWRs, VB, and machine learning algorithms. The LWRs growth equations of female and male individuals were $W=0.00002102\ L^{2.8849}$ and $W=0.00003064\ L^{2.8212}$, respectively. The VB equations were determined $L_t=80.990\ [1–e^{−0.990(t+0.208)}]$ for female and $L_t=64.172\ [1-e^{−0.610(t+0.271)}]$ for male. In general, the performance of both methods (VB model and ML algorithms) in predicting lengths, as measured by Mean Absolute Percentage Error (MAPE), was satisfactory, with the VB model demonstrating slightly superior performance (2.734). In addition, the ML algorithm gives better results in length data prediction with multilayer perceptron and in weight data prediction with Sequential Minimum Optimization (SMO) algorithm when ML algorithms are examined. The diverse ML algorithms positively impacted the investigations addressing growth-related issues in fisheries.

**Keywords:** artificial neural networks; Eastern mosquitofish; growth parameters; length-weight relationships; von Bertalanffy

## 1 | INTRODUCTION

Growth is an important factor influencing the timing of sexual maturation, reproduction, production, and fishing arrangements (Beaudouin *et al.* 2008). Growth is unique to each fish in this process (Xiong *et al.* 2015). The growth of a fish population is related to stock abundance, food source, area, temperature, growing season, and other environmental factors in natural waters (Tesch 1968). Length-weight data is an important data bank in population calculations in the backward calculation of length or weights from the methods used in freshwater growth studies (direct observation, tagging and marking, and back-calculation of lengths or weights). Length-weight relationships (LWR) emerge as an important tool in the science and management of fisheries (Xiong *et al.* 2015).

Although traditional growth models have been widely used to estimate fish length-weight relationships, their limitations in capturing nonlinear growth patterns remain challenging. Despite the growing interest in machine learning (ML) techniques, a systematic comparison between traditional and ML-based growth models in fisheries remains scarce. Traditional growth models, such as the von Bertalanffy function, assume fixed growth parameters that may not fully capture environmental variability (Schnute and Richards 1990). Moreover, LWR models often oversimplify the LWR by assuming a single exponential function, which may not be valid for all fish species and growth stages (Froese 2006). This study aims to address this gap by evaluating the predictive performance of different ML models against the traditional LWR and von Bertalanffy models.

Eastern mosquitofish, *Gambusia holbrooki* Girard, 1859, one of the world's 100 worst invasive alien species (ISSG 2000), disperses well after entering a new area. The survivability of mosquitofish in stagnant water has been used for biological control of mosquitoes (Leyse *et al.* 2004). The combination of very rapid reproduction (Zane *et al.* 1999) and aggressive predation (Goodsell and Kats 1999) makes the mosquitofish an important invader. *Gambusia holbrooki* has a very large distribution in the freshwater ecosystems. There are a few studies on LWRs (e.g. Eagderi and Radkhah 2015; Xiong *et al.* 2018, Kurtul and Sari 2020; Sellaoui and Bounaceur 2020), population and biology (e.g. Pyke 2005; Patimar *et al.* 2011; Erguden 2013) of *Gambusia holbrooki*.

Traditional models rely on predefined mathematical functions that may not adapt well to dynamic environmental conditions (Beverton and Holt 2012). In contrast, machine learning techniques, such as artificial neural networks (ANNs) and support vector regression (SVR), can learn patterns directly from data without requiring explicit assumptions about growth functions. Studies suggest that these models can outperform traditional approaches by capturing nonlinear dependencies and incorporating multiple environmental factors (Huang *et al.* 2019; Zhang *et al.* 2021; Rahman *et al.* 2021). However, a direct performance comparison between ML and traditional models in fisheries applications remains largely unexplored, highlighting the need for further investigation.

The mathematical models used in traditional fisheries science are simplified representations of the processes that govern the growth and reproduction of populations. Like all models, they only appear as special cases of certain scientific theories. At any given point in time, it reveals the situation that can be achieved at best (Flinn and Midway 2021). Therefore, models must be able to change, and new dimensions of representation must be considered as the knowledge evolves or a different problem is addressed (Bimba *et al.* 2016). It is a fact that many aspects of the models presented in the last 30 years are insufficient, but still provide with information about the issues of interest. However, it is also time to ask whether the model has been applied to a situation corresponding to the specific situation in which it was developed. In this context, the results obtained in the LWRs (Ricker 1973) and von Bertalanffy (Sparre and Venema 1998) used in traditional methods should be compared with the results obtained by machine learning. Machine learning algorithms (linear regression model, multilayer perceptron, RBF network, RBF regressor, SMO regression) use parameters based on training data representing a large cluster. As the training data expands to represent the world more realistically, the algorithm calculates more accurate results. Machine learning-based studies on fish product (e.g. Rahman *et al.* 2021), fish age classification (Benzer *et al.* 2022), and intelligent fish aquaculture (Zhao *et al.* 2021) are available. A similar research was conducted with artificial neural networks methodology on *Pseudorasbora parva* in Hirfanlı Dam Lake, Türkiye using the LWRs regarding the superiority of ANNs estimation model over Von Bertalanffy model (Benzer and Benzer 2020). This study has three main objectives: (1) to assess the accuracy of traditional fish growth models (LWR and von Bertalanffy) by applying them to real world fish datasets; (2) to systematically compare these models with machine learning algorithms, including linear regression, multilayer perceptron, RBF network, RBF regression, and SMO regression, in terms of their predictive performance; and (3) to provide empirical evidence on whether ML techniques offer a statistically significant improvement over traditional models, thereby guiding future applications in fisheries management. By addressing these objectives, this study aims to contribute to the on-going discourse on integrating machine learning into fisheries science, potentially enhancing predictive accuracy and improving management strategies for sustainable fish populations.

## 2 | METHODOLOGY

### 2.1 Preparation of data

Fish specimens ($n$ = 194) were collected from the Gediz River and their tributaries and creeks (38°35′18″N, 26°48′57″E), within the borders of İzmir under the İzmir Province Biodiversity Inventory and Monitoring Project in July 2017. In the study, a 12 volt DC, 5 amper Samus brand 725MP, and PWM 2 model back type electrofisher device and hand net (500 µm mesh size) were used to collect fishes. The fish specimens were placed in nylon bags with 4% formal and moved to Gazi University Faculty of Education Science II Laboratory. They were weighed with a precision scale of ± 0.1 g, and their lengths were measured with an electronic caliper with an interval of 0.01 mm. Sex determinations were made by observing the morphology and gonad structure of the anal fin, which takes the form of the gonidium (Heinen-Kay *et al.* 2015). Age determination was done according to Lagler (1966) for *Gambusia holbrooki* individuals. The distribution of the data used in both traditional and ML approaches can be seen in Figure 1.

### 2.2 Traditional methods

The general regression equation

$$W = a\ L^{b} \qquad (1)$$

was used (Ricker 1973) in the calculation of the LWR. In the formulation, W is the weight (g) of fish, L represents length (mm) of fish and *a* and *b* are used as constants. Growth was estimated using the VB growth equations (Sparre and Venema 1998):

$$L_t = L_\infty \; [1 - e^{-k(t-t_0)}]^b \qquad (2)$$

$L_t$ is the total length (TL, cm) at age t; $L_\infty$ is asymptotic theoretical maximum TL, *k* is the growth coefficient, *t* is the age, $t_0$ is the age at zero length.

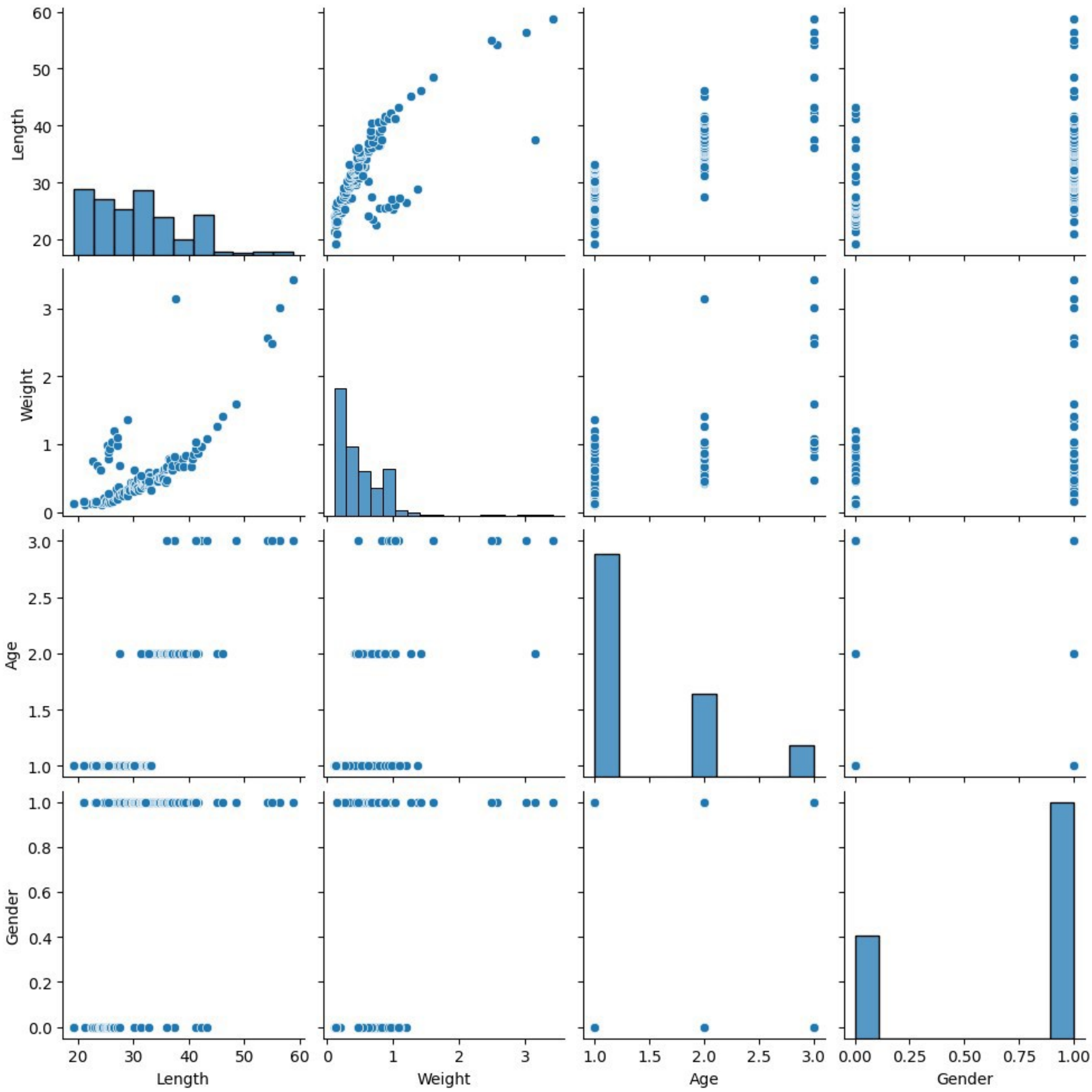


**FIGURE 1** The distribution of the data used in both traditional and machine learning approaches.

**2.3 Artificial intelligent methods**

Derived from artificial intelligence concept, ML has become a trend among science communities to facilitate and automate complicated tasks through experience (Jordan and Mitchell 2015). The selection of machine learning models was based on their ability to handle non-linear relationships and provide high predictive accuracy. However, although ML models offer flexibility, they may require large datasets and careful parameter tuning. This study addresses such concerns by systematically evaluating multiple models and fine-tuning their parameters using cross-validation to optimize predictive performance with the available dataset. In this study, the performance of five machine learning algorithms for the estimation of length and weight data of *G. holbrooki* were investigated. Available algorithms from the library in Weka tool (version 3.8.5, Hall *et al.* 2009) were used in this study. Linear regression (Rencher and Schaalje 2008) is used to model a continuous variable Y as a mathematical function of one or more X variables. This regression model can be used to predict the Y when only the X is known. The equation can be generalized as follows:

$$Y = a + bX + \epsilon \quad (3)$$

where, $a$ is the intercept, $b$ is the slope and $\epsilon$ is the error term.

Linear regression was selected for this study because it models the relationship between fish length and weight, which is a fundamental aspect of growth studies in fisheries. Traditional growth models, such as the LWR and von Bertalanffy growth function (VB), often rely on linear regression-based approaches for parameter estimation (Ricker 1973; Sparre and Venema 1998). Linear regression provides a baseline model to compare the performance of machine learning algorithms against conventional fisheries growth models (Suryanarayana *et al.* 2008). Additionally, many studies have shown that linear regression has been successfully used to predict fish growth patterns (Benzer and Benzer 2022; Ozcan 2024), further supporting its inclusion in this analysis.

The Multilayer Perceptron (MLP) mimics the human learning system, like a nervous system that processes complex data and makes logical decisions based on learning. The MLP consists of three layers: the input layer, the hidden layer, and the output layer, which are connected to each other through nodes representing neurons in biological nervous systems and learn from the input information and eventually work together to make a logical decision. The considered data enters the neural network and is processed in the hidden layer called the computational layer, and the value of the weights is regulated by the back propagation algorithm until it reaches the actual output values (Hecht-Nielsen 1992). The mathematical definitions of the algorithm can be generalized as follows:

$$U_k = \sum_{j=1}^{P} W_{kj} X_j + b_k \quad (4)$$

$$y_k = \emptyset(U_k) \quad (5)$$

where, $X_j$ is the input nod, $W_{kj}$ is the weight from $j^{th}$ to $k^{th}$ nod (neuron), $b_k$ is the bias of the $k^{th}$ neuron and $y_k$ is the output of $k^{th}$ neuron.

The MLP is a type of artificial neural network that is effective in learning multivariable relationships and nonlinear patterns in fish growth (Hecht-Nielsen 1992; Zhao *et al.* 2021). While traditional growth models are often based on certain assumptions, MLP can consider different environmental variables, allowing for more flexible and accurate predictions (Benzer and Benzer 2020). Since fish growth is influenced by environmental factors such as temperature, oxygen levels, and feeding rates, MLP has been shown to be superior in modeling the complex relationships between these variables (Zhao *et al.* 2021). Previous studies have reported that MLP has yielded successful results in modeling fish growth and biological processes (Rahman *et al.* 2021).

Radial Basis Function (RBF) network is a three-layer feedforward neural network. The first layer corresponds to the inputs of the network, the second is a hidden layer consisting of a number of RBF non-linear activation units, and the last one corresponds to the final output of the network. The network input/output algorithm can be mathematically written as follows (Sun *et al.* 2016):

$$h_j = exp\left(\frac{\|x-c_j\|^2}{2b_j^2}\right) \quad (6)$$

$$\delta = W^{*T} h(x) + \epsilon \quad (7)$$

where, $x$ is the network input, $i$ is the $i^{th}$ input of the network input layer, $j$ is the $j^{th}$ network input of the hidden layer, $h = [h_j]^T$ is the output of the Gaussian function, W* is the ideal weight of the network, and $\epsilon$ is the error of the ideal neural network approximating $\delta$, $\varepsilon \le \varepsilon_{max}$.

RBF Regression is an approach used in high-dimensional problems (Poggio and Girosi 1990). In an RBF Regression model (given x), the expectation function can be written mathematically as follows:

$$F(x) = \sum_{j=1}^{m} k(\|x - x_j\|) a_j \quad (8)$$

where $\{(\|x-x_j\|) \; j = 1, 2, …., m\}$ is a set of m RBFs, which are fixed and non linear on x, and $\|.\|$ denotes the Euclidean norm.

RBF Network and RBF Regression are particularly effective in modeling nonlinear relationships in fish growth curves and handling high-dimensional datasets (Poggio and Girosi 1990; Sun *et al.* 2016). Traditional growth models, such as the von Bertalanffy model, rely on specific assumptions and often fail to capture certain biological nuances. These methods provide more flexible and data-driven alternatives, allowing for more accurate modeling of complex growth processes (Suryanarayana *et al.* 2008; Benzer *et al.* 2022). Various studies have demonstrated that RBF-based approaches outperform traditional regression models in predicting fish growth patterns, as these methods efficiently capture complex and nonlinear biological processes (Rahman *et al.* 2021; Zhao *et al.* 2021). These characteristics make RBF Network and RBF Regression suitable choices for fish growth modeling, enabling more accurate and adaptive predictions under changing environmental conditions.

Support Vector Machine (SVM) analysis is a widely used machine learning tool for classification and regression, first introduced by Vladimir Vapnik and his colleagues in 1992 (Vapnik 1995). SMO Regression (Support Vector Machine for Regression - SVR) is particularly effective in achieving high accuracy with small datasets and excels in capturing complex relationships within the data (Osuna *et al.* 1997).

A small dataset is typically defined as a dataset with limited observations, making it challenging for traditional machine-learning models to achieve reliable generalization (James *et al.* 2013). In biological studies, such as fish growth modeling, data collection may be constrained due to environmental, logistical, or ethical limitations, leading to small datasets that require specialized algorithms for effective modeling (Friedrichs and Igel 2005). Due to its ability to generalize well even with limited training data, SMO Regression is a suitable choice for modeling fish growth, as it enables precise predictions despite the inherent constraints of biological data collection (Rahman *et al.* 2021; Benzer *et al.* 2022).

**2.4 Metrics and artificial intelligence tool**

This study applied a 70% training - 30% test data splitting method to evaluate the predictive performance of machine learning models. The dataset was randomly shuffled before being divided, ensuring that the training set (70%) was used for model learning, while the test set (30%) was reserved for performance evaluation on unseen data. This approach allows for a consistent and fair comparison between traditional and machine learning-based models, providing a reliable assessment of their predictive accuracy.

Three widely used error metrics were employed to assess model performance: Mean Absolute Error (MAE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE). These metrics quantify prediction accuracy from different perspectives and help identify the strengths and weaknesses of various models (Wang and Xu 2004). Lower values for these metrics indicate better model performance, as they represent the deviation between predicted and actual values.

MAE measures the average absolute difference between predicted and actual values, making it an intuitive and easily interpretable metric (Willmott and Matsuura 2005). Since MAE retains the same unit as the original data, it is particularly useful for evaluating fish length and weight predictions in this study. MSE squares the error terms, penalizing more significant deviations more heavily (Chai and Draxler 2014). This characteristic makes MSE particularly suitable for identifying models that produce significant errors, as it amplifies their impact. In this study, MSE was utilized to distinguish models with higher prediction variability. MAPE expresses the prediction error as a percentage of actual values, making it a scale-independent metric (Armstrong and Collopy 1992). This is especially useful when comparing errors across different magnitude ranges, such as length and weight measurements in fish populations. By normalizing the errors, MAPE ensures a balanced assessment across different growth parameters. The mathematical formulations of these error metrics are as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|e_t| \tag{9}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e_t^2 \tag{10}$$

$$MAPE = \frac{100}{n}\sum_{j}^{n}\frac{|e_j|}{|A_j|} \tag{11}$$

The estimation process was evaluated using MAPE to determine the most accurate approach (Figure 2). MAE and MAPE are commonly used to measure prediction errors, but they serve different purposes depending on the dataset and application (Moon and Yao 2011). The Weka Machine Learning Workbench was used for data processing and model evaluation to enhance model performance analysis. Weka provides various machine-learning techniques, feature selection tools, and graphical analysis options, making it a suitable platform for this study (Frank *et al.* 2014).

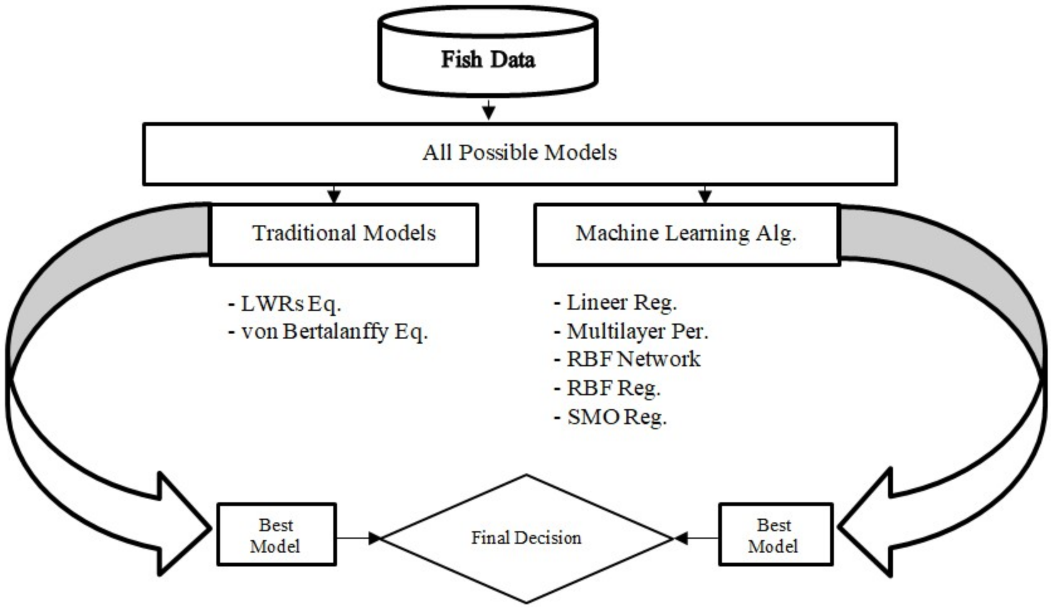


**FIGURE 2** Schematic workflow of the best method.

**3 | RESULTS AND DISCUSSION**

Descriptive statistics of *Gambusia holbrooki* specimens [135 females, total length (TL) and weight (W) of 21 – 58.78 mm; 0.152 – 3.424 g; 59 males, TL and W of 19.25 – 43.20 mm; 0.108 – 1.075 g] are presented in Table 1. The LWRs growth equations of female and male individuals were W = 0.00002102 L 2.8849 and W = 0.00003064 L 2.8212, respectively (Figure 3; Table 2).

The LWRs growth equations were calculated using the data of *G. holbrooki* individuals (Table 2, Figure 3). The *b* value for females was higher than that of males. The relationship between the population's age and length is shown in Figure 4 as a von Bertalanffy growth.

**TABLE 1** Mean Total length (TL, mm) mean weight (W, g) of *Gambusia Holbrooki*, analyzed in this study.

| | Female | | | Male | | |
|---|---|---|---|---|---|---|
| **Age** | ***n*** | **TL (mm; mean ± SD and range)** | **W (g; mean ± SD and range)** | ***n*** | **TL (mm; mean ± SD and range)** | **W (g; mean ± SD and range)** |
| 1 | 77 | 27.95 ± 4.20<br>21.00 – 33.18 | 0.3392 ± 0.18<br>0.152 – 1.370 | 51 | 22.70 ± 3.11<br>19.25 – 30.20 | 0.29 ± 0.31<br>0.108 – 1.205 |
| 2 | 45 | 37.01 ± 3.34<br>32.2 – 46.12 | 0.7955 ± 0.43<br>0.417 – 3.140 | 3 | 32.53 ± 2.74<br>27.49 – 33.81 | 0.56 ± 0.10<br>0.464 – 0.680 |
| 3 | 13 | 46.32 ± 7.10<br>41.20 – 58.78 | 1.6384 ± 0.89<br>1.025 – 3.424 | 5 | 40.03 ± 3.09<br>36. 50 – 43.20 | 0.85 ± 0.22<br>0.479 – 1.075 |

**TABLE 2** Length-weight relationships (LWRs) and von Bertalanffy parameters of *Gambusia holbrooki*.

| **Sex** | **LWRs growth equations** | $R^2$ | $L_\infty$ | $t_o$ | $k$ |
|---|---|---|---|---|---|
| Female<br>(*n* = 135) | W = 0.00002102 L 2.8849<br>Log W= –4.6773 + 2.8849 Log L | 0.848 | 80.990 | –0.208 | 0.99 |
| Male<br>(*n* = 59) | W = 0.00003064 L2.8212<br>Log W= –4.5137 + 2.8212 Log L | 0.811 | 64.172 | –0.271 | 0.61 |
| Pooled<br>(*n* = 194) | W = 0.00003946 L 2.7144<br>Log W= –4.4039 + 2.7144 Log L | 0.839 | 71.264 | –0.266 | 0.70 |

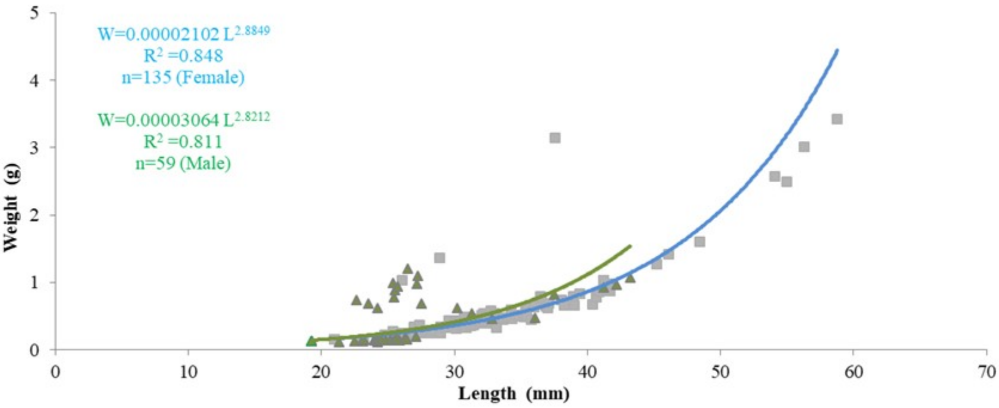


**FIGURE 3** Length weight relationship of *Gambusia holbrooki* in the Gediz River, Turkey.

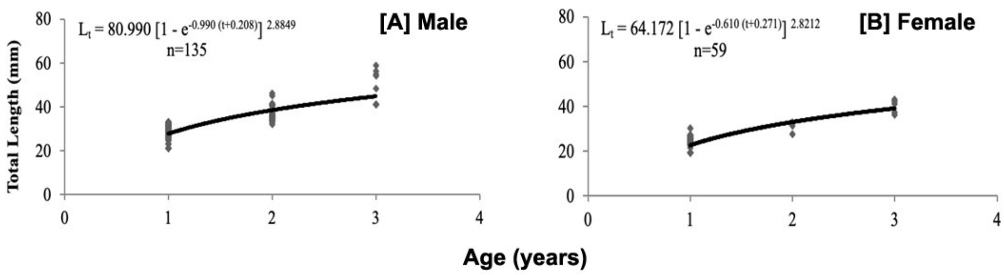


**FIGURE 4** von Bertalanffy growth of *Gambusia* holbrooki in the Gediz River, Turkey.

Predict models were created with Weka Platform for ML algorithms. Length and weight predict results obtained through the Weka Platform (Supplementary material A and B). LWR and von Bertalanffy evaluations were made, and the results of the traditional methods were compared with the statistical methods used to measure the prediction accuracy (prediction performance) of the models (Table 3). In general, the performance of both methods in predicting lengths, as measured by MAPE, was satisfactory, with von Bertalanffy's model demonstrating slightly superior performance (2.734). Considering the LWRs and von Bertalanffy studies of *G. holbrooki* individuals in the literature (e.g. Beaudouin *et al.* 2008; Erguden 2013; Xiong *et al.* 2018; Sellaoui and Bounaceur 2020), it has been determined that these results are similar to the results we obtained.

The length, weight, and age data of all *G. holbrooki* individuals in the Gediz River was modeled by using the ML algorithms. Evaluations of linear regression model, multilayer perceptron, RBF network, RBF regression, and SMO regression model were made, and the results of the traditional methods were compared with the statistical methods used to measure the prediction accuracy (prediction performance) of the models (Table 4).

**TABLE 3** Observed data and calculated values for traditional methods (LWR and VB).

| Age | Observed Data | LWR | MAPE (%) | von Bertalanffy (VB) | MAPE (%) |
|---|---|---|---|---|---|
| | TL; W | TL; W | TL; W | TL | TL |
| 1 | 25.86; 0.321 | 27.60; 0.270 | 6.729; 15.888 | 25.860 | 0.000 |
| 2 | 36.45; 0.781 | 38.62; 0.737 | 5.953; 5.634 | 37.451 | 2.746 |
| 3 | 44.57; 1.421 | 47.71; 1.182 | 7.045; 16.819 | 47.001 | 5.454 |
| Total | 30.47; 0.537 | – | 6.576; 12.780 | | 2.734 |

MAPE: Mean Absolute Percentage Error

**TABLE 4** Observed data and calculated values for artificial intelligent methods.

| Parameters | Linear Regression Model | Multilayer Perceptron | RBF Network | RBF Regressor | SMO Regression |
|---|---|---|---|---|---|
| **Length** | | | | | |
| Actual | 30.47 | 30.47 | 30.47 | 30.47 | 30.47 |
| Predict | 29.943 | 29.893 | 31.195 | 29.943 | 29.977 |
| MAE | 2.777 | 1.990 | 4.768 | 2.685 | 2.661 |
| MSE | 10.143 | 6.123 | 36.086 | 9.476 | 10.077 |
| MAPE | 0.099 | 0.073 | 0.173 | 0.097 | 0.097 |
| **Weight** | | | | | |
| Actual | 0.537 | 0.537 | 0.537 | 0.537 | 0.537 |
| Predict | 0.517 | 0.513 | 0.580 | 0.487 | 0.450 |
| MAE | 0.188 | 0.131 | 0.284 | 0.128 | 0.118 |
| MSE | 0.057 | 0.038 | 0.106 | 0.041 | 0.040 |
| MAPE | 0.558 | 0.477 | 0.952 | 0.353 | 0.274 |

RBF: Radial Basis Function; SMO: Sequential Minimum Optimization; MAE: Mean Absolute Error; MSE: Mean Squared Error; MAPE: Mean Absolute Percentage Error

Table 4 shows the prediction results of the five ML approaches, according to the different statistics results. In the literature, no study was recorded where this method was employed for *G. holbrooki*. However, there are studies that examined the performance data with multilayer perceptron in other fish species (Benzer and Benzer 2018, 2020). It has been determined that MAPE values in all ML algorithms used in this study yielded better results when compared to traditional methods (e.g. LWRs and von Bertalanffy). Among the machine learning approaches, the one with the lowest MAPE value is considered to provide the best predictive performance (Table 3 and Table 4). This means that all of the results obtained through ML approaches result in maximum performance of their predictabilities (Table 4). In addition, it is seen that ML algorithm gave better results in length data prediction with Multilayer Perceptron and in weight data prediction with SMO algorithm (Table 4, Figures 5 – 6). To investigate the relationships between the predicted and actual values, we selected the ML algorithms, as their MAPE prediction performance was better than the other MAPE values. The predicted length and weight values has a higher similarity to what the linear model looks like, whereas the values are highly dispersed in the RBF approaches. One of the ML approaches discussed in our study, the multilayer perceptron algorithm, was studied for *Pseudorasbora parva* and *Atherina boyeri*, and results revealed better MAPE values than the traditional methods, similar to what we found in this study (Benzer and Benzer 2018, 2020).

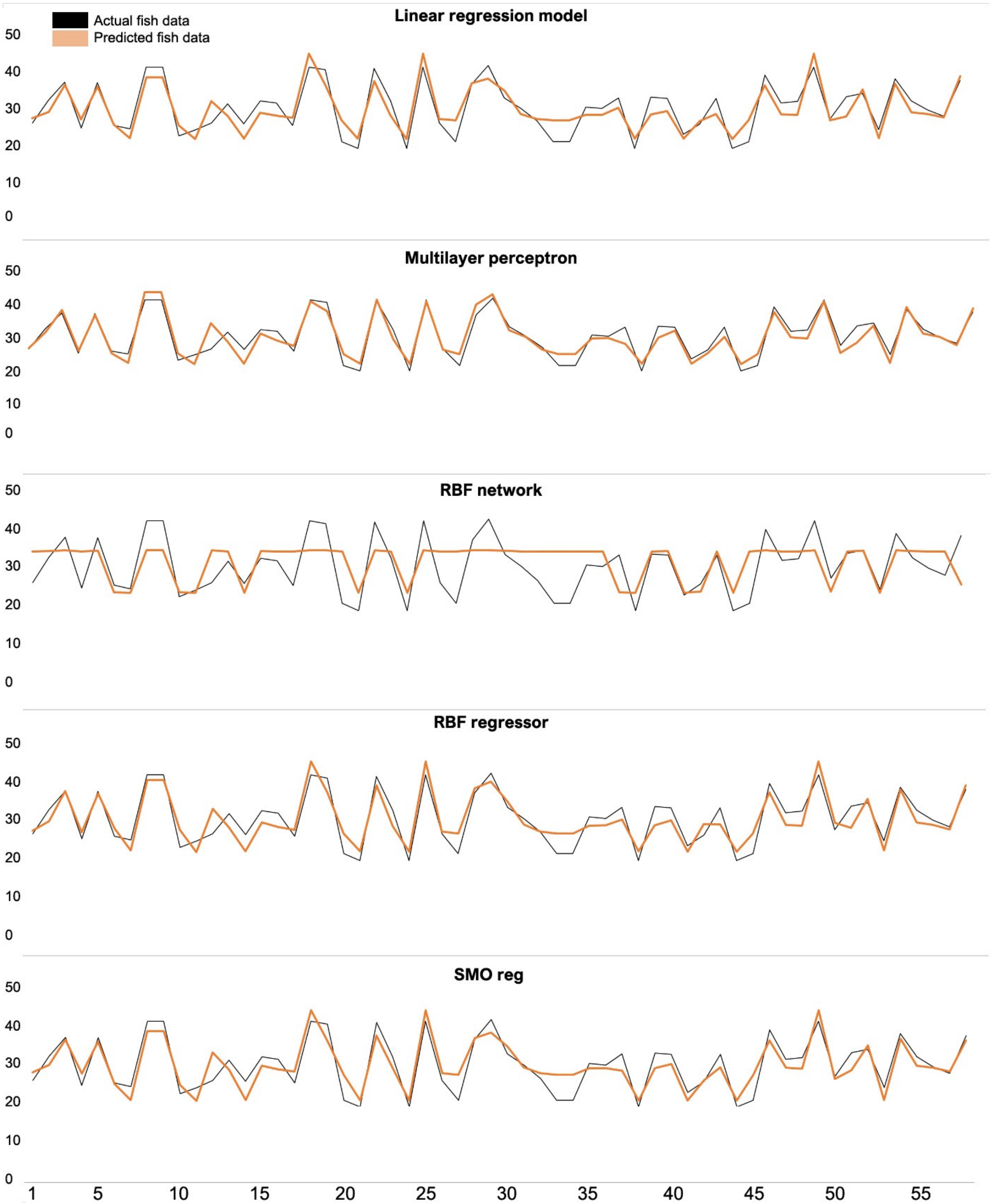


**FIGURE 5** Performance results obtained from different machine learning algorithms for the fish length. X-axis represents fish ID, Y-axis represents fish length (cm).

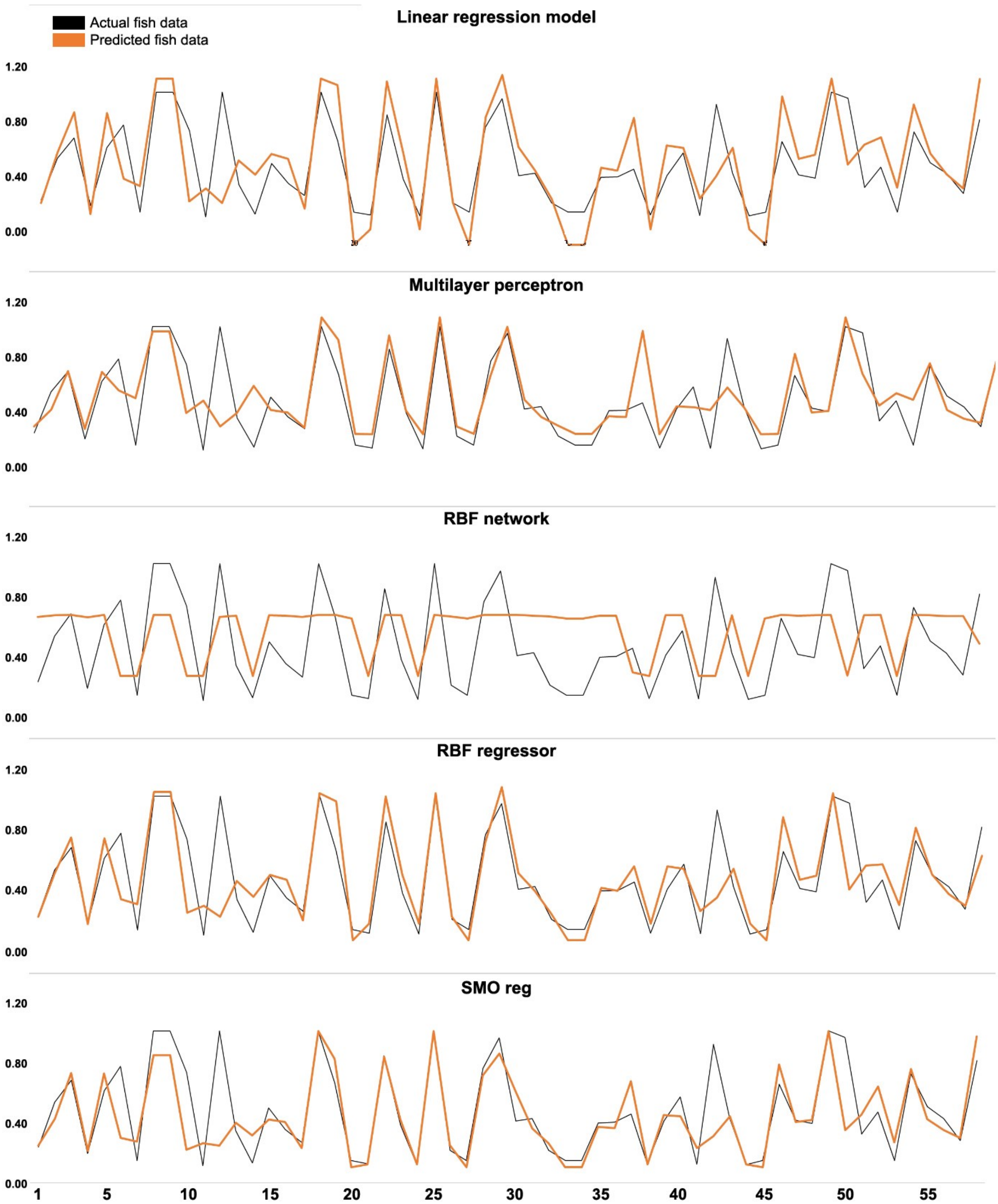


**FIGURE 6** Performance results obtained from machine learning algorithms for the fish weight. X-axis represents fish ID, Y-axis represents fish length (cm).

## 4 | CONCLUSIONS

This study investigates the best predictability by establishing prediction models with traditional approaches (LWRs and von Bertalanffy) and machine learning approaches (linear regression, multilayer perceptron, RBF network, RBF regression and SMO regression) based on *G. holbrooki* collected from the Gediz River. This study concludes that machine learning algorithms had a significant impact on addressing growth-related issues in fisheries. It is expected that the combined use of traditional and machine learning approaches for predictive models will facilitate future research in this field of fisheries.

**ACKNOWLEDGEMENTS**
The authors would like to express their gratitude to the editor and anonymous reviewers for their valuable comments and suggestions aimed at improving the quality of the paper. Similarly, the authors wish to extend their thanks for the proofreading of the article carried out by Gazi University Academic Writing Application and Research Center.

**CONFLICT OF INTEREST**
The author declares no conflict of interest.

**AUTHORS' CONTRIBUTION**
SB & AG: collected the material, analysed data and writing original draft. RB: designed the study concept, calculated ML and editing.

**DATA AVAILABILITY STATEMENT**
The data that support the findings of this study are available on a reasonable request from the corresponding author.

**REFERENCES**
Armstrong JS, Collopy F (1992) Error measures for generalizing about forecasting methods: empirical comparisons. International Journal of Forecasting 8(1): 69–80.

Beaudouin R, Ginot V, Monod G (2008) Growth characteristics of eastern mosquitofish *Gambusia holbrooki* in a northern habitat (Brittany, France). Journal of Fish Biology 73(10): 2468–2484.

Benzer R, Benzer S (2018) Alternative approaches for growth models: artificial neural networks (pp. 1–4). In: 2018 26th Signal Processing and Communications Applications Conference (SIU). IEEE, İzmir.

Benzer S, Benzer R (2020) Growth properties of *Pseudorasbora parva* in Süreyyabey Reservoir: traditional and artificial intelligent methods. Thalassas: An International Journal of Marine Sciences 36: 149–156.

Benzer S, Benzer R (2022) Morphometric analysis of Crayfish–traditional and artificial intelligent approach. Thalassas: An International Journal of Marine Sciences 38(2): 989–996.

Benzer S, Garabaghi FH, Benzer R, Mehr HD (2022) Investigation of some machine learning algorithms in fish age classification. Fisheries Research 245: 106151.

Beverton RJ, Holt SJ (2012) On the dynamics of exploited fish populations, Volume 11. Springer Science & Business Media, UK.

Bimba AT, Idris N, Al-Hunaiyyan A, Mahmud RB, Abdelaziz A, ... Chang V (2016) Towards knowledge modeling and manipulation technologies: a survey. International Journal of Information Management 36(6): 857–871.

Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geoscientific Model Development 7(3): 1247–1250.

Eagderi S, Radkhah A (2015) Length-weight relationship and condition factor of mosquitofish (*Gambusia holbrooki*) in three inland basins of Iran. Poeciliid Research 5(1): 39–43.

Erguden SA (2013) Age, growth, sex ratio and diet of eastern mosquitofish *Gambusia holbrooki* Girard, 1859 in Seyhan Dam Lake (Adana/Turkey). Iranian Journal of Fisheries Sciences 12(1): 204–218.

Flinn SA, Midway SR (2021) Trends in growth modeling in fisheries science. Fishes 6(1): 1.

Frank E, Hall M, Trigg L, Holmes G, Witten IH (2004) Data mining in bioinformatics using Weka. Bioinformatics 20(15): 2479–2481.

Friedrichs S, Igel C (2005) Evolutionary tuning of multiple SVM parameters. Neurocomputing 64: 107–117.

Froese R (2006) Cube law, condition factor and weight-length relationships: history, meta-analysis and recommendations. Journal of Applied Ichthyology 22: 241–253.

Goodsell JA, Kats LB (1999) Effect of introduced mosquitofish on Pacific treefrogs and the role of alternative prey. Conservation Biology 13: 921–924.

Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. ACM SIGKDD Explorations Newsletter, 11(1): 10–18.

Hecht-Nielsen R (1992) Theory of the backpropagation neural network. Neural networks for perception computation, learning, and architectures. Academic Press Inc, San Diego, CA. pp. 65–93.

Heinen-Kay JL, Morris KE, Ryan NA, Byerley SL, Venezia RE, ... Langerhans RB (2015) A trade-off between natural and sexual selection underlies diversification of a sexual signal. Behavioral Ecology 26(2): 533–542.

Huang L, Asteris PG, Koopialipoor M, Armaghani DJ, Tahir MM (2019) Invasive weed optimization technique-based ANN to the prediction of rock tensile strength. Applied Sciences 9(24): 5372.

ISSG (2000) 100 of the world's worst invasive alien species: a selection from the global invasive species database. Invasive Species Specialist Group, University of Auckland, Auckland.

James G, Witten D, Hastie T, Tibshirani R (2013) An Introduction to statistical learning with applications in R. Springer, New York.

Jordan MI, Mitchell TM (2015) Machine learning: trends, perspectives, and prospects. Science 349(6245): 255–260.

Kurtul I, Sari HM (2020) Length-weight relationships of invasive mosquitofish (*Gambusia holbrooki* Girard, 1859) in 23 river basins of Turkey. Turkish Journal of Zoology 44(4): 324–334.

Lagler KF (1966) Freshwater fishery biology, second edition. WM. C. Brown Company. Dubuque. Iowa. 421 pp.

Leyse KE, Lawler SP, Strange T (2004) Effects of an alien fish, *Gambusia affinis*, on an endemic California fairy shrimp, *Linderiella occidentalis*: implications for conservation of diversity in fishless waters. Biological Conservation 118: 57–65.

Moon Y, Yao T (2011) A robust mean absolute deviation model for portfolio optimization. Computers & Operations Research 38(9): 1251–1258.

Osuna E, Freund R, Girosi F (1997) An improved training algorithm for support vector machines (pp. 276–285). In: Neural networks for signal processing VII. Proceedings of the 1997 IEEE signal processing society workshop. IEEE, New Jersey.

Ozcan EI (2024) Performance of artificial neural networks and traditional methods in determining selected growth parameters of *Alburnus sellal* Heckel, 1843. Oceanological and Hydrobiological Studies 53(2): 153–163.

Patimar R, Ghorbani M, Gol-Mohammadi A, Azimi-Glugahi H (2011) Life history pattern of mosquitofish *Gambusia holbrooki* (Girard, 1859) in the Tajan River (southern Caspian Sea to Iran). Chinese Journal of Oceanology and Limnology 29(1): 167–173.

Poggio T, Girosi F (1990) Regularization algorithms for learning that are equivalent to multilayer networks. Science 247: 978–982.

Pyke GH (2005) A review of the biology of *Gambusia affinis* and *G. holbrooki*. Reviews in Fish Biology and Fisheries 15(4): 339–365.

Rahman LF, Marufuzzaman M, Alam L, Bari MA, Sumaila UR, Sidek LM (2021) Developing an ensembled machine learning prediction model for marine fish and aquaculture production. Sustainability 13(16): 9124.

Rencher AC, Schaalje GB (2008) Linear models in statistics. John Wiley & Sons, Hoboken, NJ.

Ricker WE (1973) Linear regression in fisheries research. Journal of the Fisheries Board of Canada 30: 409–434.

Schnute JT, Richards LJ (1990) A unified approach to the analysis of fish growth, maturity, and survivorship data. Canadian Journal of Fisheries and Aquatic Sciences 47(1): 24–40.

Sellaoui N, Bounaceur F (2020) Growth and length-weight relationships of *Gambusia affinis* (Baird et Girard, 1853) population in Algeria (Cyprinodontiformes Poeciliidae). Biodivers J 11(4): 951–959.

Sparre P, Venema SC (1998) Introduction to tropical fish stock assessment. Part 1 manual. FAO Fisheries Technical Paper (306.1, Rev. 2), Food and Agriculture Organization of the United Nations, Rome. 407 pp.

Sun Z, Deng Z, Zhang Z (2016) Intelligent control theory and technology, 2nd edition. Tsinghua University Press, Beijing, China.

Suryanarayana I, Braibanti A, Rao RS, Ramam VA, Sudarsan D, Rao GN (2008) Neural networks in fisheries research. Fisheries Research 92(2–3): 115–139.

Tesch FW (1968) Age and growth. In: Ricker WE (Ed) Methods for assessment of fish production in freshwaters. Blackwell Scientific Publications, Oxford, UK. pp. 93–123.

Vapnik V (1995) The nature of statistical learning theory. Springer, New York.

Wang W, Xu Z (2004) A heuristic training for support vector regression. Neurocomputing 61: 259–275.

Willmott CJ, Matsuura K (2005) Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research 30(1): 79–82.

Xiong W, Tao J, Zhang DC, Liu C, He DK, Chen YF (2015) Length-weight relationships for four small fish species caught in wetland of central Yangtze River, China. Journal of Applied Ichthyology 31: 219–220.

Xiong W, Zhu GP, Wang ZL, Ye N (2018) Length-weight relationships of four fish species from mangrove of Zhanjiang, China. Journal of Applied Ichthyology 34(1): 167–168.

Zane L, Nelson WS, Jones AG, Avise JC (1999) Microsatellite assessment of multiple paternity in natural populations of a live-bearing fish, *Gambusia holbrooki*. Journal of Evolutionary Biology 12: 61–69.

Zhang Y, Zhang F, Cheng J, Zhao H (2021) Classification and recognition of fish farming by extraction new features to control the economic aquatic product. Complexity 2021: 5530453.

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

***S Benzer*** http://orcid.org/0000-0002-8548-8994
***R Benzer*** https://orcid.org/0000-0002-5339-0554
***A Gül*** https://orcid.org/0000-0001-5751-4705