Acta Ichthyologica et Piscatoria 55, 2025, 19–29 | DOI 10.3897/aiep.55.140634

PENSOFT.

ACTA ICHTHYOLOGICA ET PISCATORIA

Assessment of the nursery environment and distribution of Ayu, *Plecoglossus altivelis* (Actinopterygii, Osmeriformes, Plecoglossidae), in Vietnam

Huy Xuan CHU[1], Toan Quang LE[1], Ngoc Minh NGUYEN[1], Ai Huyen Thi TONG[1], Huy Quang BUI[1], Hai HOANG[1], Hai Phuc NGUYEN[1], Hieu Duc NGUYEN[1], Anh Ngoc Thi DO[2,3]

1 *Space Technology Institute, Vietnam Academy of Science and Technology, Ha Noi, Vietnam*

2 *Laboratory of Ecology and Environmental Management, Science and Technology Advanced Institute, Van Lang University, Ho Chi Minh City, Vietnam*

3 *Faculty of Applied Technology, School of Technology, Van Lang University, Ho Chi Minh City, Vietnam*

https://zoobank.org/EEC00D48-9B8A-49CE-85AC-7EFE25E79639

Corresponding author: Anh Ngoc Thi Do (dothingocanh@vlu.edu.vn)

Academic editor: Wojciech Piasecki ♦ **Received** 30 October 2024 ♦ **Accepted** 21 January 2025 ♦ **Published** 19 March 2025

Citation: Chu HX, Le TQ, Nguyen NM, Tong AHT, Bui HQ, Hoang H, Nguyen HP, Nguyen HD, Do ANT (2025) Assessment of the nursery environment and distribution of Ayu, *Plecoglossus altivelis* (Actinopterygii, Osmeriformes, Plecoglossidae), in Vietnam. Acta Ichthyologica et Piscatoria 55: 19–29. https://doi.org/10.3897/aiep.55.140634

Abstract

The distribution of larvae and juveniles of Ayu, *Plecoglossus altivelis* (Temminck et Schlegel, 1846), is shaped by intricate interplays involving environmental variables and anthropogenic influences. The intricate interplay and equilibrium among these factors will govern the distribution and abundance of Ayu fish larvae and juveniles in estuarine settings. In this investigation, a hybrid Ant Colony Optimization-Adaptive Neuro-Fuzzy Inference System (ACO-ANFIS) model was utilized to enhance the precision of Ayu fish larvae and juvenile's occurrence estimation for the period from 2021 to 2022. The outcomes evinced that the hybrid model displays strong predictive capabilities, with $R^2_{test} > 0.75$, and AUC > 0.79. Among the environmental parameters, temperature, salinity, and turbidity exhibit the highest correlations with Ayu fish occurrence, with R values of 0.47, 0.54, and 0.40 for the Ka Long estuary, and 0.49, 0.50, and 0.42 for the Ba Lat estuary, respectively. The presence of Ayu species is limited to northern Vietnam, albeit with a declining pattern from the Ka Long estuary to the Ba Lat estuary. The study's outcomes propose that the identification of suitable habitats and the cartography of fish distribution are invaluable for scrutinizing the ramifications of natural and anthropogenic influences on species distribution.

Keywords

Ayu fish, distribution, environment condition, sun-glint, Sentinel-2

Introduction

The distribution of species arises from intricate interactions involving environmental parameters and biological factors (Do et al. 2022a; Do and Tran 2023a; Do et al. 2024a; Trieu et al. 2025). Within estuarine settings, fish species exhibit rapid responses to various environmental pressures, such as climate change and pollution, resulting in disruptions to species composition, ecological function, and diversity (Murawski 1993). Species distribution modeling stands as a crucial and swiftly advancing field in ecological research (Moore et al. 2009; Murase et al. 2009), offering not only enhanced comprehension of the relationships between species and their environment but also enabling the prediction of species distribution in unsampled locations.

Coastal and estuarine regions commonly feature waves and wind, leading to irregular water surfaces that can significantly distort the identification of characteristics in shallow water environments (Do et al. 2022b). Remote sensing data plays a pivotal role in estuarine research applications due to its capability to frequently capture extensive areas, facilitating cost savings, reduced field survey efforts, and enhanced accessibility to remote or challenging-to-reach areas (Do and Tran 2023b; Pham et al. 2023a; Pham et al. 2024). The fusion of remote sensing data and predictive models can aid in estimating and forecasting the distribution range of fish species across large areas (Do et al. 2022a; Do and Tran 2023a). One of the methodologies under consideration is the Generalized Additive Model (GAM). GAM serves as a non-parametric regression technique capable of handling non-linear relationships and diverse statistical data distributions, thus finding extensive application in investigations concerning the correlation between environmental factors and the spatial distribution of fish (Murase et al. 2009; Zhao et al. 2014). Nonetheless, the construction and estimation of GAM model components can become intricate, particularly in the presence of numerous predictor variables. Moreover, GAM necessitates a substantial and comprehensive dataset of observational data for estimation, a situation that may lead to over-fitting and diminished predictive efficacy (Do et al. 2022a).

Machine learning models have been widely adopted in ecological and environmental modeling due to their capacity to manage nonlinear relationships and offer precise simulation outcomes (Do et al. 2022a; Do and Tran 2023c). The utilization of the hybrid Ant colony optimization-Adaptive Neuro-fuzzy Inference System (ACO-ANFIS) model is increasingly prevalent among scientists and policymakers to bolster biodiversity management strategies (Do et al. 2022a; Aghelpour et al. 2023). Essential aspects of conservation and management efforts to maintain or rejuvenate estuarine ecosystems to an optimal ecological condition involve predicting distribution and comprehending the primary environmental factors linked to species' distribution (Do et al. 2022a; Do and Tran 2023c, 2023b). Various studies have leveraged machine learning to clarify the significance of key environmental variables in forecasting fish occurrences. The majority of these studies have concentrated on the impact of purely environmental factors on species distribution and the associations between them (Zhao et al. 2014; Do and Tran 2023b). Nevertheless, the integration of data from field surveys and remote sensing, along with cross-comparisons across diverse estuarine regions, remains limited (Do et al. 2022a).

Ayu, *Plecoglossus altivelis* (Temminck et Schlegel, 1846), reproduces in rivers, and its larvae drift passively downstream into coastal waters upon hatching, where they spend the winter before migrating back upstream to mature for the subsequent spawning season (Azuma 1981; Tanaka et al. 2011). Traditionally, Ayu uses sandy shorelines and estuaries as their primary breeding locations, which are currently undergoing alterations due to human activities and natural occurrences. The Ka Long estuary in Quang Ninh province and the Dinh estuary in Ninh Thuan province, Vietnam, exhibit high salinity intrusion, substantial tidal ranges, and the development of vast tidal flats (Do et al. 2022a). These three estuaries possess distinctive features that make them promising nursery grounds for various fish species during their larval and juvenile phases. Nonetheless, like other regions globally, Vietnamese estuaries face significant impacts from climate change and human interventions. Moreover, there is a scarcity of studies that employ a comprehensive approach integrating field data and remote sensing to ascertain suitable habitats and the distribution of Ayu larvae in estuarine settings (Do and Tran 2023b).

Therefore, the primary objective of this current research endeavor is to employ the ACO-ANFIS model to:

- identify the pivotal environmental variables influencing the presence of Ayu larvae; and
- utilize the hybrid ACO-ANFIS model to assess the distribution of Ayu larvae within estuaries.

Materials and methods

Study site and data sources

Study area. The Ka Long River originates from the northern edge of Quang Duc Commune, Hai Ha District, located at coordinates 21°36′36″N and 107°41′2″E, situated approximately 5 km west of the Bac Phong Sinh border gate. It flows in an east–northeast direction (Fig. 1A) and spans about 16 km in length. This river is situated in a region characterized by a tropical climate featuring a hot and humid summer as well as a cold and dry winter. The average temperature during winter remains below 20°C, whereas in summer, it rises above 25°C. The yearly average humidity stands at 84%, while the annual rainfall varies from 1700 to 2400 mm.

Ba Lat Estuary is a coastal inlet located in northern Vietnam. It is positioned between Giao Thuy District, Nam Dinh Province (Fig. 1B). The tropical climate of this region experiences two distinct seasons: a dry period extending from November to April, and a rainy season occurring between May and October. The average temperature hovers around 23°C. Annual rainfall levels range from 1650 to 1800 mm, with the rainy season spanning from May to October and the dry season from November to February. The yearly sunshine duration is between 1600 and 1700 hours, while the average relative humidity falls within the range of 80%–90%.

The Dinh estuary is a water body that flows into the East Sea (South China Sea). It stretches across a length of 135 km and covers a drainage basin area of 3109 km^2. Originating from the elevated E Lam Thong Mountain range bordering Lam Dong Province in the Phan Rang region (Fig. 1C), this river basin experiences a typical tropical monsoon climate characterized by hot and dry conditions, strong winds, and high evaporation rates. The average annual temperature ranges from 26 to 27°C, with

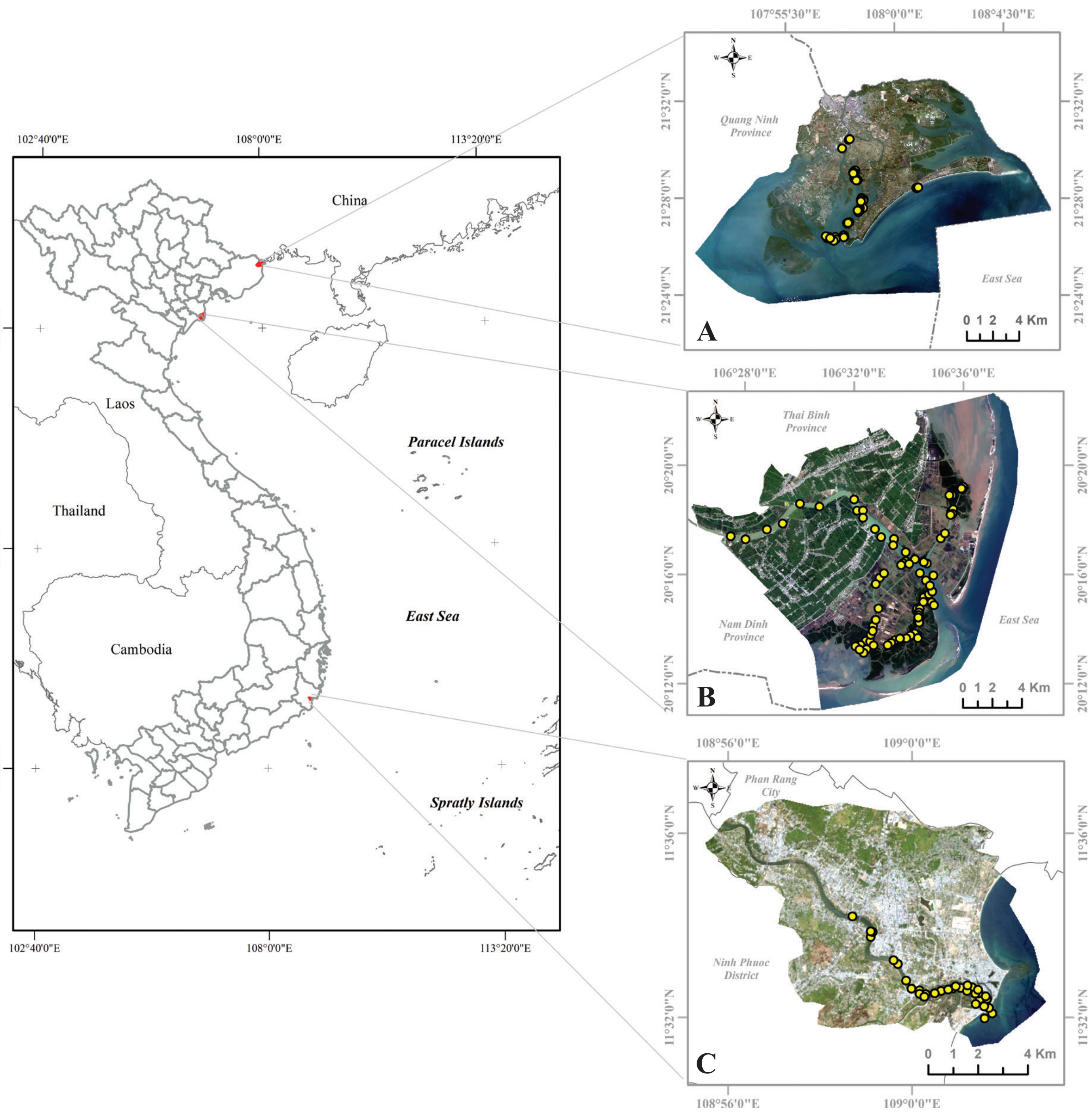

Figure 1. The study area in Ka Long (**A**); Ba Lat (**B**); and Dinh estuaries (**C**) in Vietnam. The spatial distribution of the 63; 77; and 54 sampling sites of Ayu, *Plecoglossus altivelis*, respectively at the three estuaries. Note: the East Sea is otherwise known as the South China Sea.

rainfall averaging between 700 to 800 mm in Phan Rang and exceeding 1100 mm in mountainous terrains. Air humidity levels fall between 75% and 77%. The climate exhibits two primary seasons: a rainy period from September to November, and a dry season spanning December to August. The coastal zones are abundant in lagoons and bays, offering opportunities for tourism development and aquaculture, particularly shrimp farming, which stands as a pillar of the local fisheries industry.

Data use

A total of 117 survey sites were sampled across the two estuaries, comprising 63 samples from the Ka Long estuary collected from September 2021 to December 2021, 77 samples from the Ba Lat estuary collected from March to April 2022, and the remaining 54 samples from the Dinh estuary (Fig. 1) collected from May to June 2022. Sampling was conducted using a small seine net (1 × 4 m, mesh aperture = 1 mm). Specimens were preserved in 5% formalin and then transferred to 80% ethanol. Fish larvae and juveniles were separated under a 10 × 40 magnification dissecting microscope and identified as Ayu based on external morphology following Do et al. (2022a). Water temperature (°C), salinity (ppt), turbidity (NTU), and Chlorophyll-a (mg/m^3) were measured at each site using a Water Quality Checker (WQC-22A, TOA DDK). Furthermore, additional data were obtained from Sentinel-2 satellite imagery, with details on preprocessing and sun-glint correction provided in Hedley method sun-glint correction section below.

Elevation. Elevation is a significant factor influencing fish distribution, with higher altitudes generally associated with reduced distribution and vice versa (Ley 2005).

Elevation data in the study area (Table 1) ranged from 0–22 m for the Ka Long estuary, 0–42 m for the Ba Lat estuary, and 0–38 m for the Dinh estuary, derived from a Digital Elevation Model (DEM) with a cell size of 10 × 10 m.

Slope. Slope directly impacts surface runoff, with higher slope angles increasing water velocity and affecting flow direction. Slope angle calculations were performed using the DEM spatial analysis tool in ArcGIS 10.8. The slopes of the Ka Long, Ba Lat, and Dinh estuaries exhibited a wide range, of 0–22.6, 0–17.2, and 0–15.24, respectively (Table 1).

Distance from source. The distance from the source is a crucial factor influencing fish distribution and salinity levels in the study area. Areas closer to the sea typically have higher salinity values, leading to increased abundance of fish inland. Details regarding the distance from the source are presented in Table 1.

Mean width. Similar to the distance from the source, the mean width of the estuary is an abiotic factor associated with the presence of fish larvae and juveniles, as estuary width impacts current speed and salinity levels. Previous research in the study area has indicated higher salinity in the bank water region compared to the center of the current (Do and Tran 2023c; Table 1).

Proximity of urban areas. Estuarine areas near urban centers are typically under high pressure from human activities, such as pollution, exploitation, and destruction of fish habitats. Consequently, fish tend to avoid areas in close proximity to urban centers, leading to a more abundant and diverse fish community further away from urban areas.

Proximity of mangrove areas. Mangrove forests offer plentiful food sources, shelters, and ideal spawning conditions for fish. Estuarine areas adjacent to mangrove forests exhibit superior fish distribution and diversity, whereas regions distant from mangroves generally show reduced fish populations and diversity.

Chlorophyll-a. Chlorophyll-a plays a crucial role in determining the distribution of juvenile Ayu in aquatic environments. As an indicator of the presence of phytoplankton, Chlorophyll-a not only provides a rich food source for juvenile fish but also reflects the water quality in their habitat. Areas with high concentrations of Chlorophyll-a typically have good dissolved oxygen levels, which facilitate the growth of fish. Therefore, the abundance of Chlorophyll-a directly impacts the distribution and survival of juvenile Ayu within the ecosystem.

Fish CPUE (catch per unit effort). Fish were captured at wadeable depths (approximately 20 to 100 cm) in the two estuaries employing a seine net (1 × 4 m, 1 mm mesh-aperture) (Kinoshita et al. 1988). Two individuals held the net extended, wading backward in the water from ankle to neck depth along the shoreline for about 50 m within a timeframe of roughly 2 min. Typically, one to two hauls were conducted at each sample station per day. The catch per unit effort (CPUE) was computed as the number of individuals in each haul (about 2 min or 50 m). (Kinoshita et al. 1988).

Hedley method sun-glint correction. Hedley method establishes a linear relationship based on a sample of image pixels rather than solely two image pixels (Kay et al. 2009, 2013). All chosen pixels are integrated in a linear regression of NIR brightness with visible channel brightness.

If the slope of the regression line for channel *i* is b_i, then all pixels in image channel *i* can be de-glinted by applying the subsequent equation

$$R_i^l = R_1 - b_i\,(R_{NIR} - \mathrm{Min}_{\mathrm{NIR}}) \qquad [1]$$

This approach relies on a sample of image pixels, eliminating the necessity to mask out non-water pixels like land, clouds, and boats before de-glinting (Busch et al. 2013). The key steps involve:

- Correcting atmospheric effects;
- Opting for one or more representative regions in the image showing a variety of solar flares, along with a homogeneous region free of solar flares. Identify the $\mathrm{Min}_{\mathrm{NIR}}$, which signifies the minimum brightness on the NIR channel within the selected sample;
- Conducting a linear regression analysis on the brightness values of the NIR channel (*x*-axis) against the brightness of the single channel (*y*-axis) using the chosen pixels. The slope of the regression line, denoted as b_i, is the key output for channel *i*.
- Eliminating sun-glint interference from all pixels [1]

The outcomes of the sun-glint adjustment process are demonstrated in Fig. 2 and will subsequently be utilized to extract data for urban and mangrove regions as input variables for the ANFIS model.

Table 1. Environment descriptors used as input variables and CPUE (catch per unit effort) of Ayu, *Plecoglossus altivelis*, in Vietnam is used as output.

Variable	Min			Max			Mean			Median		
	Ka Long	Ba Lat	Dinh	Ka Long	Ba Lat	Dinh	Ka Long	Ba Lat	Dinh	Ka Long	Ba Lat	Dinh
Temperature [°C]	16.40	19.30	22.23	23.60	22.00	25.27	18.90	18.00	20.09	17.40	17.30	19.62
Turbidity [NTU]	0.00	3.00	0.00	23.00	37.00	43.90	13.30	15.70	35.27	14.00	16.00	34.03
Distance from source [km]	0.00	0.00	0.00	8.00	14.10	12.69	1.70	3.90	4.60	1.63	4.10	4.21
Slope [°]	0.00	0.00	0.00	22.60	17.20	15.24	4.40	5.00	4.37	3.60	4.10	4.41
Mean width [km]	0.11	0.30	0.14	2.30	1.50	0.69	0.80	0.40	0.16	0.90	0.40	0.16
Salinity [‰]	0.00	0.00	0.00	27.90	30.20	34.43	12.30	10.20	16.34	13.10	7.60	17.03
Elevation [m]	0.00	0.00	0.00	22.00	42.00	38.00	4.90	10.30	9.34	3.50	9.00	9.50
To urban [km]	0.00	0.00	0.00	4682.33	8199.20	352.56	848.19	1781.24	86.35	890.65	2543.00	94.56
To mangrove [km]	0.00	0.00	0.00	12 753.69	11 169.60	0.00	4120.22	2514.40	0.00	3572.26	2409.91	0.00
Chlorophyll-a [mg/m³]	0.01	0.027	0.00	3.88	7.17	0.71	0.58	0.92	0.161	0.48	0.65	0.17
CPUE [Individuals/haul]	0.35	0.60	0.00	756.82	268.00	0.00	98.40	60.80	0.00	21.30	13.40	0.0

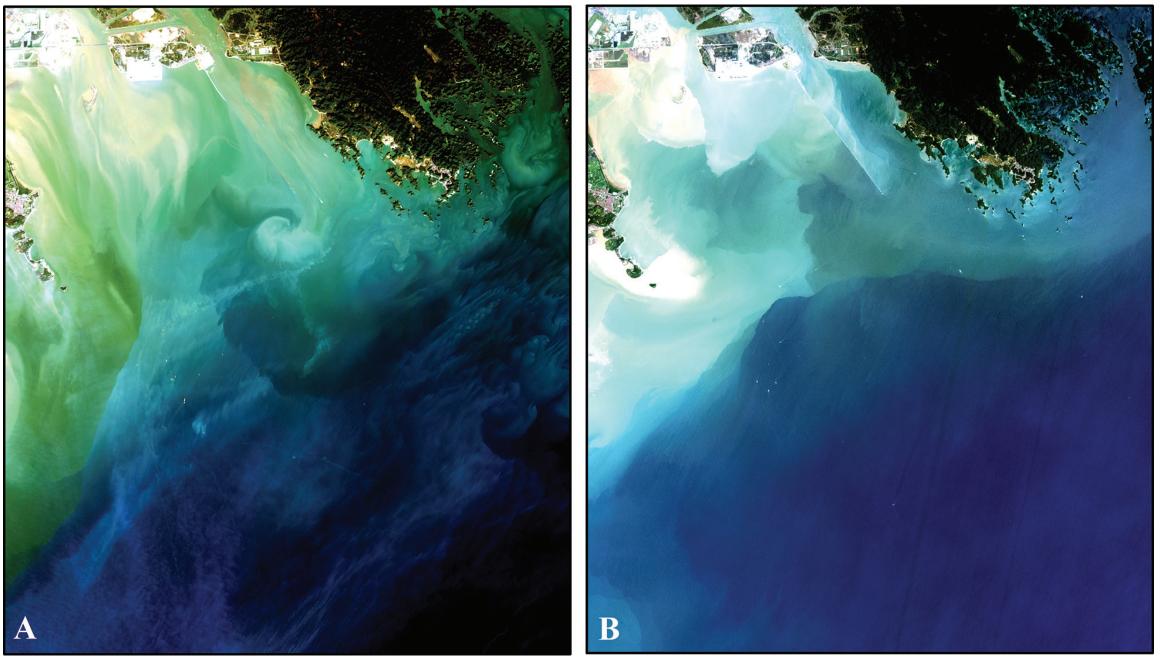

Figure 2. Satellite images of the study area before (**A**) and after (**B**) the pre-processing and sun-glint correction process.

Adaptive neuro fuzzy inference system (ANFIS) in predicting fish distribution

Adaptive neuro fuzzy inference system (ANFIS) is grounded on a fuzzy inference system, trained using an algorithm rooted in neural network principles (Azeem et al. 2000). The ANFIS architecture comprises five layers, illustrated in Fig. 3.

Layer 1: This segment is referred to as the fuzzification layer, representing the training data and the initial seven predictor variables for fish, as detailed in Table 1. Adaptive nodes with node functions defined as:

$$Q_i^1 = \mu A_k^1(X_1) \qquad [2]$$

Parameters $\{a, b, c\}$ determine the shape of the membership function, termed as the premise parameters

$$\mu A_k\left(X_i\right) = \frac{1}{1 + \left|\frac{x - c_i}{a_i}\right|^{2\delta_i}} \qquad [3]$$

where $i = 1{:}n$, with n being the number of input variables; k representing the clusters and rules count; $X_1, X_2 \ldots X_n$ denoting the input nodes; $A, B \ldots C$ as the linguistic labels; and $\mu(X_n)$ indicating the membership value.

Layer 2: The second layer is called the rule layer; each node in this layer is a circular node labeled II, responsible for multiplying the incoming signals.

$$Q_i^2 = w_i = \mu A_k^1(X_1) \times \mu A_k^2(X_2) \times \ldots \times \mu A_k^n(X_n) \qquad [4]$$

Layer 3: Each circular node labeled N in this layer calculates the normalized firing strengths w_i as per the formula [5]:

$$Q_i^3 = \overline{w_i} = \frac{w_i}{\operatorname{sum}\left(w_i\right)} \qquad [5]$$

Layer 4: This layer is commonly referred to as the defuzzification layer, where each individual represents an adaptable node.

$$Q_i^4 = \overline{w_i} f_i \qquad [6]$$

where $f_i = z_i x + r_i y + t_i$ with f_i is the fuzzy if-then rule; z_i, r_i, and t_i represent the designated parameters, known as the consequence parameters.

Layer 5: The CPUE result of ANFIS is derived from the summation of outcomes acquired for each rule within the defuzzification layer.

$$Q_i^5 = \sum_i \overline{w_i} f_i = \sum_i \frac{w_i \times f_i}{\operatorname{sum}\left(w_i\right)} = y \qquad [7]$$

RMSE is utilized to operate the functional component in the quest for ACO, with m representing the size of the training set:

$$\mathrm{RMSE} = \left[\frac{\left(\sum_k^n f_i - Q_i\right)^2}{m}\right]^{1/2} \qquad [8]$$

Of which, 70% is allocated for training purposes and 30% for validation, resulting in 44, 54, and 38 training instances for Ka Long, Ba Lat, and Dinh estuaries, respectively.

ACO algorithm in search optimization

Ant colony optimization (ACO) is a methodology inspired by simulating the behavior of ant colonies in nature, aimed at addressing intricate optimization issues (Lorpunmanee et al. 2007, 2007). In this study, to enhance the ANFIS model's performance, the rapid hybrid learning method was utilized, incorporating the ACO algorithm to optimize

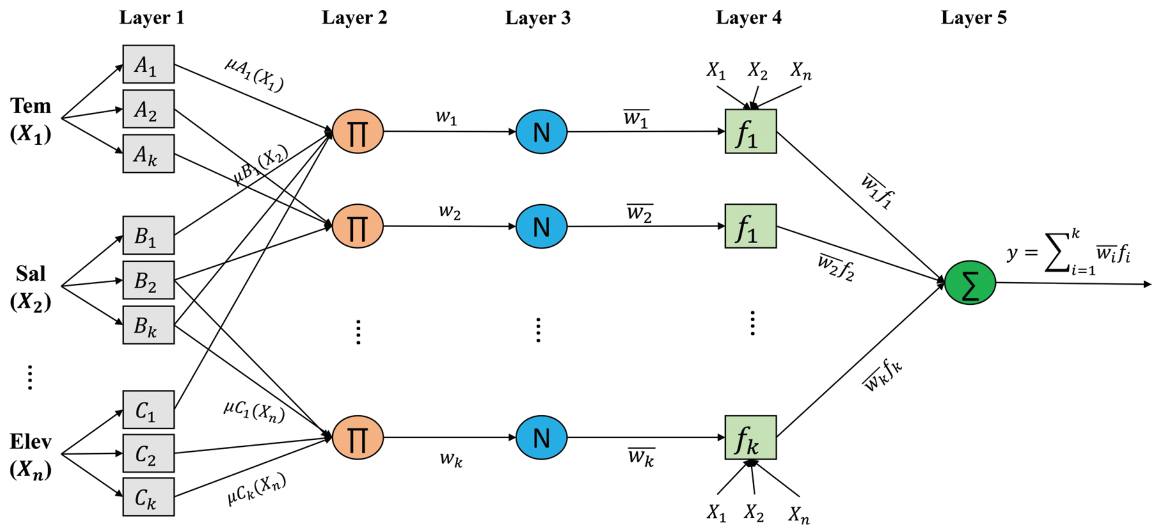

Figure 3. Fish CPUE (catch per unit effort) outputs of Ayu, *Plecoglossus altivelis*, in Vietnam of the Adaptive neuro fuzzy inference system (ANFIS).

parameters (Zhou et al. 2021). Essentially, ACO is employed to identify the optimal ANFIS parameters.

To replicate ant behavior, artificial ants have been created, each linked to an edge (i; j), the pheromone concentration τ_{ij}, and the heuristic parameter ϑ_{ij} on that specific edge. Routes with lower pheromone concentrations will eventually be phased out, leading all ants to converge on a path that typically becomes the shortest route from the ant nest to their food source. Post each iteration, the pheromone on each edge is adjusted based on the formula (Rabiner and Juang 1986):

$$\tau_{ij}(t+1) = (1-p) \times \tau_{ij}(t) + \sum_{k=1}^{m} \Delta\tau_{ij}^{k}(t)\ \forall(i,j) \quad [9]$$

where, $0 < p \leq 1$ represents the pheromone evaporation rate; and $\Delta\tau_{ij}^{k}(t)$ denotes the amount of pheromone that ant k deposits on edge ij, calculated as follows:

$$\Delta\tau_{ij}^{k} = \begin{cases} \frac{Q}{f(k)}, \text{If ant} k \text{passes through edge}(i,j) \\ 0, \text{Vice versa} \end{cases} \quad [10]$$

where, Q is a constant, and $f(k)$ stands for the objective value in each iteration.

Various values are chosen for optimization to select the suitable input parameter for CPUE estimation (Table 1). The search process concludes upon reaching the predefined maximum number of iterations (three objective functions in the ongoing study). By combining ACOs with ANFIS, the optimal parameters of ANFIS were ascertained to compute fish CPUE across the entire study region.

In this investigation, root mean square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), and receiver operating curve (ROC) were adopted to validate the GA-ANFIS model (Table 2). Results showcasing low RMSE and MAE, along with high R^2 are deemed as the most favorable outcomes (Do et al. 2022b; Do et al. 2022c; Pham et al. 2023b; Do et al. 2024b; Do 2024a, 2024b, 2024d; Pham and Do 2024; Van Pham et al. 2025).

Table 2. Statistical criteria for evaluating the ACO-ANFIS model.

ID	Function Name	Equation
1	Root mean square error (RMSE)	$\text{RMSE} = \left[\left(\sum_{i=1}^{k}(Y_i - X_i)^2\right)/k\right]^{1/2}$
2	Mean absolute error (MAE)	$\text{MAE} = \left[\left(\sum_{i=1}^{k}(Y_i - X_i)\right)/k\right]^{1/2}$
3	Coefficient of determination (R^2)	$R^2 = 1 - \sum_{i=1}^{k}(Y_i - X_i)^2 / \sum_{i=1}^{k}\left(X_i - \bar{X}_i\right)^2$

X_i = the measured dependent variable; Y_i = the calculated value, $\bar{X}_i$ = the mean value of the measured variable; and k = the number samples.

Results

Variable importance and suitable rearing environment

The investigation expedition gathered CPUE information on Ayu larvae at three river estuaries: Ka Long, Ba Lat, and Dinh, with respective survey tallies of 63, 77, and 54. Nonetheless, the findings revealed the absence of Ayu larvae at the Dinh estuary (Table 1). Consequently, the comprehension of crucial variables and an appropriate rearing habitat for Ayu larvae has become imperative.

The correlation coefficient (R) of the ten chosen variables (Table 1) for scrutinizing larval dispersion is depicted in Fig. 4. Employing the ACO algorithm for search optimization, three variables—temperature, salinity, and turbidity exhibited the highest correlation with Ayu presence (Fig. 4).

The R coefficients for temperature, salinity, and turbidity were 0.47, 0.54, and 0.40 in the Ka Long estuary; and 0.49, 0.50, and 0.42 in the Ba Lat estuary (Fig. 4). Furthermore, other variables like mangrove cover, distance from source, and proximity to urban areas also displayed relatively strong correlation, with R values of 0.45, –0.38, and 0.3 for the Ka Long estuary, and 0.4, –0.49, and 0.34 for the Ba Lat estuary (Fig. 4). These outcomes suggest that landscape and human-induced factors significantly influence

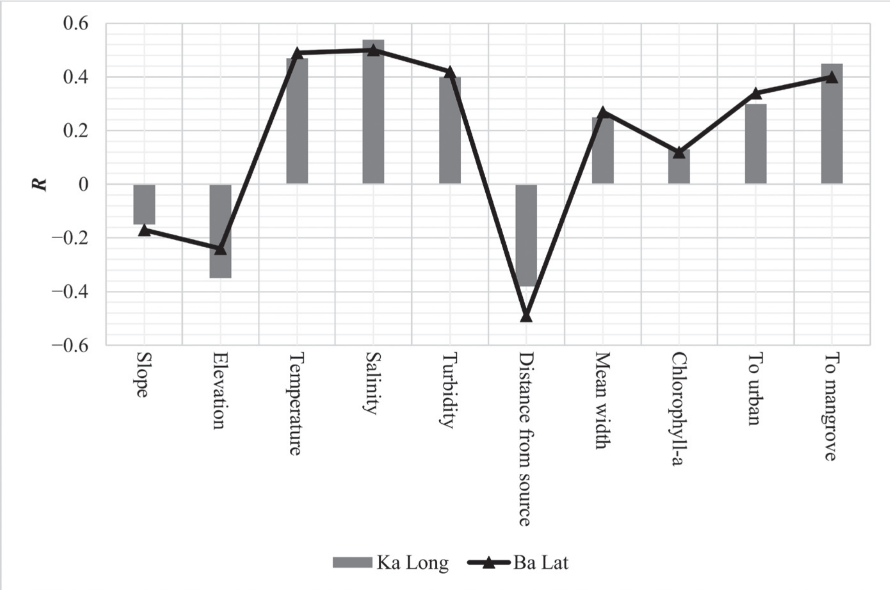

Figure 4. Correlation coefficient between environmental variables and CPUE (catch per unit effort) of Ayu, *Plecoglossus altivelis* in Vietnam of Ka Long and Ba Lat estuaries.

species presence, particularly in the context of larvae and juveniles. Conversely, some variables exhibited lesser significance for Ayu presence, such as Chlorophyll-a, and Slope, with R values of 0.13 and –0.15 (Ka Long estuary), and 0.12 and –0.17 (Ba Lat estuary). Hence, the ACO algorithm assisted in discarding redundant and less crucial variables for Ayu larvae distribution, facilitating the selection of the most pertinent variables for the predictive model.

Based on the survey outcomes and CPUE data, in the presently reported study, the values of the three most critical variables were categorized into five levels ranging from very low to very high (Table 3). The "very high" ranges for temperature, salinity, and turbidity are 21–25°C, <13‰, and <10 NTU, respectively.

For the temperature variable: When the temperature falls below 13°C and ranges between 13–15°C, the CPUE of Ayu fish larvae is observed at notably low levels as indicated in Table 3. In these temperature intervals, the activity and growth of Ayu fish larvae might face inhibition. As the temperature escalates to 15–19°C, the CPUE steadily increases to a moderate level. This particular temperature range proves to be relatively conducive for the requirements of Ayu larvae, enabling them to engage in active behaviors, forage effectively, and reproduce reasonably well. The most favorable temperature range identified for the Ayu species is situated between 19–25°C. It is evident that deviations towards excessively low or high temperatures in comparison to this range result in a decline in CPUE.

For the salinity variable: Among the array of environmental factors, salinity emerges as one of the most critical elements influencing the distribution patterns of the Ayu species. When the salinity surpasses 25‰, the CPUE is notably low, while salinity levels ranging from 20‰–25‰ correspond to a low CPUE level, as depicted in Table 3. Conversely, salinity values oscillating within 13‰–15‰ and below 13‰ are acknowledged as the optimal range for the thrivingness and well-being of Ayu larvae, facilitating their active engagement in foraging and successful reproduction.

Regarding the turbidity variable: The outcomes delineated in Table 3 underscore a negative correlation between turbidity levels and the likelihood of Ayu presence. The optimal turbidity threshold for this species is determined to be below 18 NTU, with an even more preferable level falling below 10 NTU, as illustrated in Table 3. This prescribed range of turbidity stands as the ideal milieu for Ayu larvae, ensuring adequate illumination, food availability, and fostering optimal conditions for their various activities.

In general, Table 3 effectively demonstrates how environmental parameters, such as temperature, salinity, and turbidity, exert a profound influence on the CPUE levels of Ayu larvae. Instances of temperatures deviating from the suitable range, salinity exceeding the permissible threshold, or turbidity levels registering at extremes will induce stress, curtail growth prospects, and elevate mortality rates among the larvae, consequently leading to a substantial alteration in their spatial distribution within the region.

Table 3. Fish CPUE (catch per unit effort) levels of Ayu, *Plecoglossus altivelis*, in Vietnam depending on temperature, salinity, and turbidity selected from ACO.

Variables	Very low	Low	Medium	High	Very high
Temperature [°C]	<13	(13, 15)	(15, 19)	(19, 21)	(21, 25)
Salinity [‰]	>25	(20, 25)	(15, 20)	(13, 15)	<13
Turbidity [NTU]	>25	(23, 25)	(18, 23)	(10, 18)	<10

Generation and analysis of the fish CPUE using ACO-ANFIS

The integration of ACO has enhanced the predictive capability of the ANFIS model in estimating the CPUE of fish in the Ka Long and Ba Lat estuaries. The delineation of Ayu fish larvae CPUE distribution consists of 5 categories, spanning from 0.35 to 732.28 individuals/haul at Ka Long and from 0.6 to 218.50 individuals/haul at Ba Lat (Fig. 5). At the Ka Long and Ba Lat estuaries, the mean CPUE values for fish are 105.10 and 52.86 individuals/haul, respectively (Fig. 5, Table 4).

The prognostic outcomes derived from the ACO-ANFIS model reveal that the CPUE of fish in the Ka Long estuary surpasses that in the Ba Lat estuary. Nevertheless, the CPUE distribution of Ayu larvae in the Ba Lat estuary displays a more uniform pattern in contrast to the distribution observed in the Ka Long estuary (Fig. 5). This discrepancy in distribution could potentially be attributed to the significant difference between the minimum and maximum CPUE values.

The efficacy of the ACO-ANFIS model in predicting the CPUE of Ayu species is illustrated in Fig. 6 and Table 5.

The computational findings demonstrate that the hybrid model attains commendable performance levels, with R^2_{train} values of 0.84 and 0.81 at the Ka Long and Ba Lat estuaries, respectively. Furthermore, for the subset of samples utilized for assessment, the model exhibits strong predictive accuracy, with R^2_{test} values of 0.79 and 0.75 at the Ka Long and Ba Lat estuaries, respectively.

Table 4. Predictions of fish CPUE (catch per unit effort) of Ayu, *Plecoglossus altivelis*, in Vietnam (in individuals/haul) in 2023.

Study area	Min	Mean	Max
Ka Long	0.35	105.10	732.28
Ba Lat	0.60	52.86	218.50

The predictive capability of the ACO-ANFIS model at the Ka Long estuary outperforms that at the Ba Lat estuary, potentially due to more concentrated sampling procedures, as indicated in Table 5 by RMSE values of 7.00 and 5.54 individuals/haul at the Ka Long and Ba Lat estuaries, respectively. Moreover, the receiver operating characteristic curve was employed to assess the ANFIS model integrated with the ACO algorithm, yielding AUC values of 0.804 and 0.793 at the Ka Long and Ba Lat estuaries, respectively (Fig. 6). These findings suggest that the ACO-ANFIS hybrid model is well-suited for forecasting the presence of Ayu fish larvae in the estuaries of Vietnam.

Table 5. Evaluation of accuracy of ACO-ANFIS model to estimate CPUE (catch per unit effort) of Ayu, *Plecoglossus altivelis*, in Vietnam.

ACO-ANFIS	RMSE	MAE	R^2_{train}	R^2_{test}
Ka Long	7.00	3.10	0.84	0.79
Ba Lat	5.54	2.41	0.81	0.75

ACO-ANFIS = Ant Colony Optimization-Adaptive Neuro-Fuzzy Inference System; RMSE: Root mean square error; MAE: Mean absolute error; and R^2_{train}: Coefficient of determination for the training dataset; R^2_{test}: Coefficient of determination for the testing dataset.

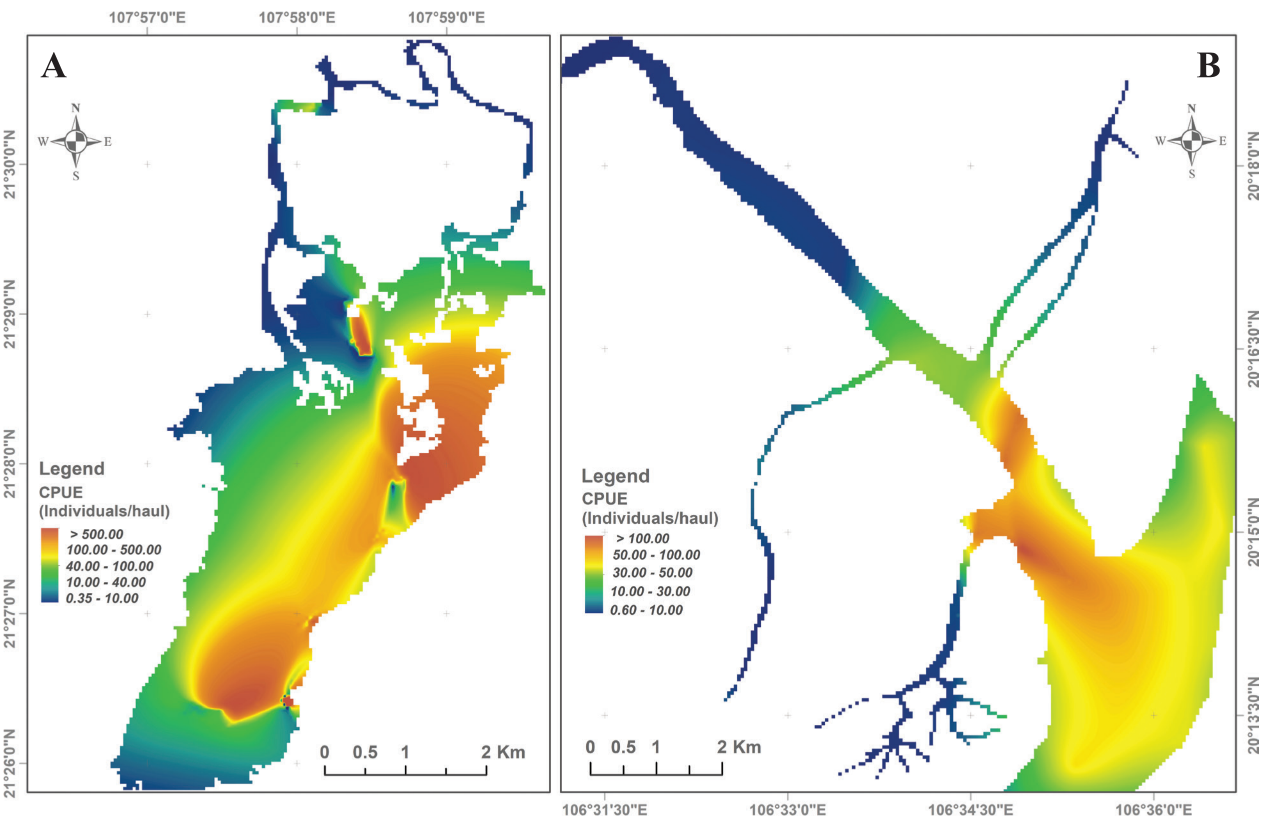

Figure 5. Map of predicted Ayu fish CPUE (catch per unit effort) of Ayu, *Plecoglossus altivelis*, in Vietnam distribution based on the hybrid ACO-ANFIS model. (**A**) Ka Long estuary, and (**B**) Ba Lat estuary.

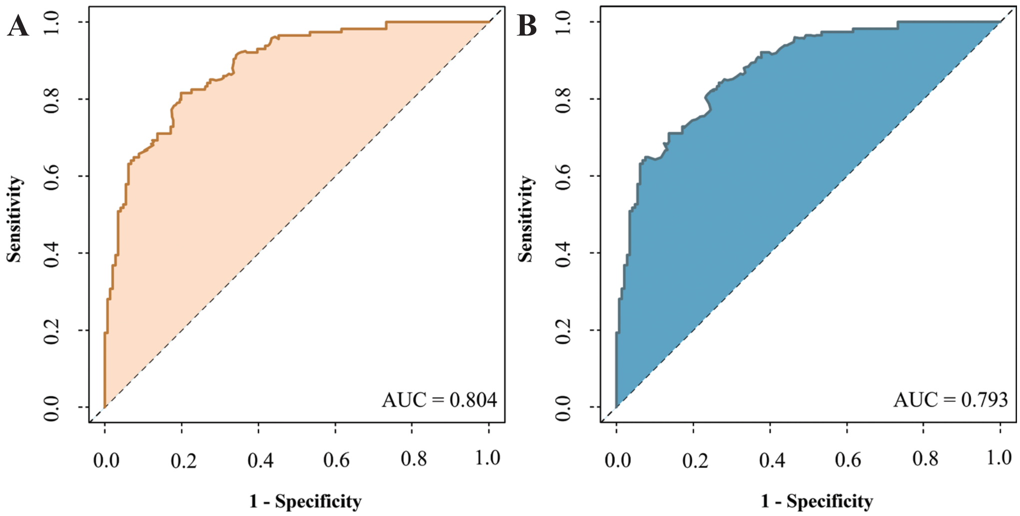

Figure 6. ROC curve for evaluating the performance of the ACO-ANFIS prediction model. (**A**) Ka Long estuary, and (**B**) Ba Lat estuary.

Discussion

Establishing robust methodologies for comprehending and ascertaining the spatial dispersion of fish, particularly in the early developmental stage, plays a pivotal role in the realm of fisheries management and the safeguarding of biodiversity. Essential insights into the distribution patterns of the Ayu species furnish fundamental information for the precise evaluation of population dynamics and resource productivity in aquatic environments, thereby facilitating the formulation of efficacious exploitation and governance strategies (Tenma et al. 2021). Moreover, knowledge concerning distribution patterns during the early life stages is essential for prognostic purposes and the effective regulation of the repercussions stemming from socio-economic development endeavors on fish habitat and distribution (Do and Tran 2023c).

Numerous endeavors have been dedicated to prognosticating fish distribution, predominantly concentrating on evaluations within marine domains, while neglecting the influences of human-induced activities. Investigations conducted by Moore et al. (2009), Murase et al. (2009), and Zhao et al. (2014) have leveraged the GAM model for the prediction of fish distribution. Nevertheless, a primary constraint of the GAM model pertains to the necessity of a substantial volume of comprehensive and precise data encompassing environmental, biological, and fishing parameters, which are frequently arduous to procure (Do et al. 2024a). To tackle these challenges, diverse studies have been carried out to delineate fish distribution through the utilization of varied algorithms like ANN (Maravelias et al. 2003), SVM (Do and Tran 2023c), ANFIS (Do et al. 2022a). The operational mechanisms and outcomes of these studies have demonstrated that machine learning models generally exhibit superior capabilities in accommodating intricate nonlinear relationships compared to the GAM model (Brosse et al. 1999). Furthermore, machine learning models possess the capability to automatically select features and fine-tune hyperparameters, whereas the GAM model mandates manual intervention and comprehension (Do et al. 2022a).

The presently reported study evaluates the performance of the ANFIS model by assessing the correlation between fish population indicators and environmental factors. A model was specifically designed to examine and forecast the spatial dispersion of Ayu fish larvae within estuaries. Through the utilization of the ANFIS technique in conjunction with the ACO algorithm, the study aimed to enhance precision and reduce uncertainty arising from the strong interrelation among environmental factors in ANFIS (Table 5, Fig. 6). The outcomes derived from the ACO-ANFIS model exhibit a commendable performance, as evidenced by R^2_{test} values of 0.79, and 0.75; alongside AUC values of 0.804 and 0.793 for the Ka Long and Ba Lat estuaries, respectively. These findings substantiate the applicability of the ACO-ANFIS model within the research area and potentially extend its usage to other estuaries situated in tropical zones like Vietnam. Additionally, our discoveries suggest a consistent alignment between the presence of Ayu fish larvae and the environmental factors foreseen by the ACO-ANFIS model as compared to existing literature.

The examination pinpointed key variables in the estuaries utilizing the ACO algorithm, including temperature, salinity, and turbidity, as the most crucial elements influencing the existence and distribution of the Ayu species. Notably, water temperature plays a direct role in the fish's growth and reproductive cycles, while salinity and turbidity impact their foraging behaviors and ability to evade predators. This observation implies that the Dinh estuary may not be a suitable environment for Ayu larval nurturing. Data in Table 1 highlights maximum temperature, salinity, and turbidity levels at 25.27°C, 34.43‰, and 43.90 NTU, respectively, with corresponding mean values of 20.09°C, 16.34‰, and 35.27 NTU. Contrastingly, the set thresholds stand at 19–25°C for temperature, below 15‰ for salinity, and less than 18 NTU for turbidity. Furthermore, additional significant variables like proximity to urban zones and mangrove presence indicate an excessive level of urbanization and human interventions in the Dinh estuary. Past studies Do (2024a) and Do et al. (2022a) have demonstrated that landscape fragmentation, urban expansion, and rising temperatures have reshaped fish distributions, compelling them towards the seashore. Therefore, it is evident that tolerance boundaries and alterations in the environment may degrade water quality, impacting the species' progression and survival, possibly leading to their extinction.

This study has mapped the geographical distribution of Ayu fish larvae in these two northern Vietnamese estuaries, Ka Long and Ba Lat, since only these two have adequate living circumstances for Ayu fish larvae. The findings indicate that the Ka Long estuary exhibits greater suitability for the Ayu species compared to the Ba Lat estuary, displaying a maximum CPUE approximately 3.4 times higher than that of the Ba Lat estuary. Nonetheless, the distribution is non-uniform, particularly within the Ka Long estuary, suggesting that the early life phases of the fish are constrained to suitable habitats and not spread throughout the entire estuary. Consequently, this restriction contributes to the dwindling presence of Ayu fish larvae, which are presently confined to the northern regions of Vietnam (Tran et al. 2017). The present investigation has underscored the significance of pinpointing the environmental thresholds governing the nursery habitat of this species, notably in estuarine regions subjected to natural and human-induced pressures (Do, 2024a; Do et al. 2024b). Through distribution mapping, it becomes feasible to anticipate distribution alterations resulting from external influences, facilitating prompt mitigation strategies. This study encompassed not just one estuary, but three estuaries spanning from northern to south-central Vietnam. The outcomes distinctly reveal a diminishing trend in Ayu presence from north to south, with no sightings documented at the Dinh estuary (Ninh Thuan province). Moreover, the research highlights the efficacy of the ACO-ANFIS model. Consequently, the methodologies employed in this study hold promise for forecasting fish presence during the nursery phase in estuaries sharing similar environmental conditions with those in Vietnam.

Conclusions

By utilizing a new hybrid GA-ANFIS model combined remote sensing data, the present investigation demonstrates that water temperature, salinity, and turbidity are the primary influencing factors on the presence of Ayu larvae and juveniles in tropical estuaries. The research highlights that Ayu larvae exhibit a preference for utilizing the Ka Long estuary as a nursery habitat and show a decreasing trend from north to south, with the absence of the Ayu species observed in the Dinh estuary located in south-central Vietnam. These results hold significant implications for ecological research and the management of biodiversity, offering valuable insights for conservation efforts. Nevertheless, due to the limited validation of estuaries in Vietnam within this study, the dataset remains incomplete. Consequently, the subsequent research will aim to validate the model in various other prominent estuaries to enhance the depth of the investigation.

Acknowledgments

The authors thank anonymous reviewers for their helpful comments and suggestions, which improved this manuscript.

This research was financially supported by the VAST under the code: QTRU02.13/21-22; Project name: "Research and development of Chlorophyll-a algorithm for VNREDSat-1 and similar images."

References

Aghelpour P, Graf R, Tomaszewski E (2023) Coupling ANFIS with ant colony optimization (ACO) algorithm for 1-, 2-, and 3-days ahead forecasting of daily streamflow, a case study in Poland. Environmental Science and Pollution Research International 30(19): 56440–56463. https://doi.org/10.1007/s11356-023-26239-3

Azeem MF, Hanmandlu M, Ahmad N (2000) Generalization of adaptive neuro-fuzzy inference systems. IEEE Transactions on Neural Networks 11(6): 1332–1346. https://doi.org/10.1109/72.883438

Azuma M (1981) On the origin of Koayu, a landlocked form of amphidromous Ayu-fish, *Plecoglossus altivelis*: With 4 figures and 1 table in the text. SIL Proceedings, 1922–2010 21: 1291–1296. https://doi.org/10.1080/03680770.1980.11897175

Brosse Sé, Lek S, Dauba F (1999) Predicting fish distribution in a mesotrophic lake by hydroacoustic survey and artificial neural networks. Limnology and Oceanography 44(5):1293–1303. https://doi.org/10.4319/lo.1999.44.5.1293

Busch JA, Hedley JD, Zielinski O (2013) Correction of hyperspectral reflectance measurements for surface objects and direct sun reflection on surface waters. International Journal of Remote Sensing 34(19): 6651–6667. https://doi.org/10.1080/01431161.2013.804226

Do ANT (2024a) Assessing the impact of habitat fragmentation on the distribution of juvenile and larval *Sillago* species in the Ka Long estuary located in Northern Vietnam. Ocean Science Journal 59(2): 24. https://doi.org/10.1007/s12601-024-00149-y

Do ANT (2024b) Utilizing a fusion of remote sensing data and machine learning models to forecast flood risks to agriculture in Hanoi City, Vietnam. Letters in Spatial and Resource Sciences 17(1): 21. https://doi.org/10.1007/s12076-024-00382-y

Do ANT, Tran HD (2023a) Combining a deep learning model with an optimization algorithm to detect the dispersal of the early stages of spotted butterfish in northern Vietnam under global warming. Ecological Informatics 102380: 102380. https://doi.org/10.1016/j.ecoinf.2023.102380

Do ANT, Tran HD (2023b) Application of deep learning in assessing the impact of flooding on the endangered freshwater fish *Neolissochilus benasi* (Cyprinidae) in a northern province of Vietnam. Aquatic Ecology 57(4): 951–967. https://doi.org/10.1007/s10452-023-10056-4

Do ANT, Tran HD (2023c) Potential application of artificial neural networks for analyzing the occurrences of fish larvae and juveniles in an estuary in northern Vietnam. Aquatic Ecology 57(4): 813–831. https://doi.org/10.1007/s10452-022-09959-5

Do ANT, Tran HD, Ashley M (2022a) Employing a novel hybrid of GA-ANFIS model to predict distribution of whiting fish larvae and juveniles from tropical estuaries in the context of climate change. Ecological Informatics 71: 101780. https://doi.org/10.1016/j.ecoinf.2022.101780

Do ANT, Tran HD, Ashley M, Nguyen AT (2022b) Monitoring landscape fragmentation and aboveground biomass estimation in Can Gio Mangrove Biosphere Reserve over the past 20 years. Ecological Informatics 101743: 101743. https://doi.org/10.1016/j.ecoinf.2022.101743

Do TAT, Do ANT, Tran HD (2022c) Quantifying the spatial pattern of urban expansion trends in the period 1987–2022 and identifying areas at risk of flooding due to the impact of urbanization in Lao Cai city. Ecological Informatics 101912: 101912. https://doi.org/10.1016/j.ecoinf.2022.101912

Do ANT, Do TAT, Tran HD (2024a) Distribution of fish larvae and juveniles on salinity in an estuary predicted from remote sensing and fuzzy logic approach. Aquatic Ecology 58(3): 983–998. https://doi.org/10.1007/s10452-024-10119-0

Do ANT, Do TAT, Van Pham L, Tran HD (2024b) Assessment of the role of mangroves for *Periophthalmus modestus* applying machine learning and remote sensing: A case study in a large estuary from Vietnam. Aquatic Ecology 58(3): 879–893. https://doi.org/10.1007/s10452-024-10111-8

Do ANT, Tran HD, Do TAT (2024d) Impacts of urbanization on heat in Ho Chi Minh, southern Vietnam using U-Net model and remote sensing. International Journal of Environmental Science and Technology 21: 3005–3020. https://doi.org/10.1007/s13762-023-05118-x

Kay S, Hedley JD, Lavender S (2009) Sun glint correction of high and low spatial resolution images of aquatic scenes: A review of methods for visible and near-infrared wavelengths. Remote Sensing 1(4): 697–730. https://doi.org/10.3390/rs1040697

Kay S, Hedley J, Lavender S (2013) Sun glint estimation in marine satellite images: A comparison of results from calculation and radiative transfer modeling. Applied Optics 52(23): 5631–5639. https://doi.org/10.1364/AO.52.005631

Kinoshita I, Fujita S, Takahashi I, Azuma K (1988) Occurrence of larval and juvenile Japanese snook, *Lates japonicus*, in the Shimanto estuary, Japan. Ichthyological Research 34(4): 462–467. https://doi.org/10.1007/BF02905651

Ley JA (2005) Linking fish assemblages and attributes of mangrove estuaries in tropical Australia: Criteria for regional marine reserves. Marine Ecology Progress Series 305: 41–57. https://doi.org/10.3354/meps305041

Lorpunmanee S, Sap MN, Abdullah AH, Chompoo-inwai C (2007) An Ant Colony Optimization for Dynamic JobScheduling in Grid Environment. World Academy of Science, Engineering and Technology 29: 314–321.

Maravelias CD, Haralabous J, Papaconstantinou C (2003) Predicting demersal fish species distributions in the Mediterranean Sea using artificial neural networks. Marine Ecology Progress Series 255: 249–258. https://doi.org/10.3354/meps255249

Moore CH, Harvey ES, Van Niel KP (2009) Spatial prediction of demersal fish distributions: Enhancing our understanding of species–environment relationships. ICES Journal of Marine Science 66(9): 2068–2075. https://doi.org/10.1093/icesjms/fsp205

Murase H, Nagashima H, Yonezaki S, Matsukura R, Kitakado T (2009) Application of a generalized additive model (GAM) to reveal relationships between environmental factors and distributions of pelagic fish and krill: A case study in Sendai Bay, Japan. ICES Journal of Marine Science 66(6): 1417–1424. https://doi.org/10.1093/icesjms/fsp105

Murawski SA (1993) Climate change and marine fish distributions: Forecasting from historical analogy. Transactions of the American Fisheries Society 122(5): 647–658. https://doi.org/10.1577/1548-8659(1993)122<0647:CCAMFD>2.3.CO;2

Pham VT, Do ANT (2024) Quantifying rice dry biomass to determine the influence of straw burning on BC and NO2 emissions in the Hanoi metropolitan region. Environmental Monitoring and Assessment 197(1): 85. https://doi.org/10.1007/s10661-024-13493-2

Pham TV, Do TAT, Tran HD, Do ANT (2023a) Assessing the impact of ecological security and forest fire susceptibility on carbon stocks in Bo Trach district, Quang Binh province, Vietnam. Ecological Informatics 74: 101962. https://doi.org/10.1016/j.ecoinf.2022.101962

Pham VT, Do TAT, Tran HD, Do ANT (2023b) Classifying forest cover and mapping forest fire susceptibility in Dak Nong province, Vietnam utilizing remote sensing and machine learning. Ecological Informatics 102392. https://doi.org/10.1016/j.ecoinf.2023.102392

Pham TV, Thi Do TA, Tran HD, Thi Do AN (2024) Assessing groundwater potential for mitigating salinity issues in agricultural areas of southern Dong Nai province, Vietnam. Groundwater for Sustainable Development 25: 101177. https://doi.org/10.1016/j.gsd.2024.101177

Rabiner L, Juang B (1986) An introduction to hidden Markov models. IEEE ASSP Magazine 3: 4–16. https://doi.org/10.1109/MASSP.1986.1165342

Tanaka Y, Iguchi K, Yoshimura J, Nakagiri N, Tainaka K (2011) Historical effect in the territoriality of ayu fish. Journal of Theoretical Biology 268(1): 98–104. https://doi.org/10.1016/j.jtbi.2010.09.040

Tenma H, Tsunekawa K, Fujiyoshi R, Takai H, Hirose M, Masai N, Sumi K, Takihana Y, Yanagisawa S, Tsuchida K, Ohara K, Jo T, Takagi M, Ota A, Iwata H, Yaoi Y, Minamoto T (2021) Spatiotemporal distribution of *Flavobacterium psychrophilum* and ayu *Plecoglossus altivelis* in rivers revealed by environmental DNA analysis. Fisheries Science 87(3): 321–330. https://doi.org/10.1007/s12562-021-01510-z

Tran HD, Iida M, Maeda K (2017) Downstream migration of newly-hatched Ayu (*Plecoglossus altivelis*) in the Tien Yen River of northern Vietnam. Environmental Biology of Fishes 100(10): 1329–1341. https://doi.org/10.1007/s10641-017-0646-6

Trieu TA, Tran HD, Do ANT (2025) Integrating a novel algorithm in assessing the impact of floods on the genetic diversity of a high commercial value fish (Cyprinidae: *Spinibarbus* sp.) in Lang Son province of Vietnam. Zoology: Analysis of Complex Systems, ZACS 126240: 126240. https://doi.org/10.1016/j.zool.2025.126240

Van Pham T, Do ANT, Do TAT (2025) Evaluation of aboveground biomass in mangrove biosphere reserves from 1993 to 2023 under the influence of landscape fragmentation. Journal of the Indian Society of Remote Sensing 53: 451–468. https://doi.org/10.1007/s12524-024-02027-5

Zhao J, Cao J, Tian S, Chen Y, Zhang S, Wang Z, Zhou X (2014) A comparison between two GAM models in quantifying relationships of environmental variables with fish richness and diversity indices. Aquatic Ecology 48(3): 297–312. https://doi.org/10.1007/s10452-014-9484-1

Zhou J, Li C, Arslan CA, Hasanipanah M, Bakhshandeh Amnieh H (2021) Performance evaluation of hybrid FFA-ANFIS and GA-ANFIS models to predict particle size distribution of a muck-pile after blasting. Engineering with Computers 37(1): 265–274. https://doi.org/10.1007/s00366-019-00822-0