fishes

MDPI

Article

CART Rule-Guided MaxEnt Model Construction and Its Application in Fishing Ground Prediction of Chub Mackerel in the Northwestern Pacific Ocean

Zuli Wu †, Fenghua Tang †, Yumei Wu, Shengmao Zhang, Fei Wang and Xuesen Cui *

Key and Open Laboratory of Remote Sensing Information Technology in Fishing Resource, East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai 200090, China; wuzuli@hotmail.com (Z.W.); tangfh@ecsf.ac.cn (F.T.); wym_07@163.com (Y.W.); zhangsm@ecsf.ac.cn (S.Z.); essiwf@163.com (F.W.)

* Correspondence: cuixuesen@eastfishery.ac.cn

† These authors contributed equally to this work.

Abstract

Chub mackerel (*Scomber japonicus*) is a commercially important pelagic species in the northwest Pacific Ocean. Accurate identification of its fishing grounds can provide a more robust and targeted scientific basis for fishery management and ecological research. Based on fishing effort and five environmental factors (i.e., sea surface temperature [SST], chlorophyll-a concentration [CHL], SST gradient [GSST], sea surface height [SSH], and current speed), this study developed a Classification and Regression Tree (CART) rule-guided MaxEnt model. Specifically, rules generated by the CART model were first extracted and then incorporated as constrained feature functions into MaxEnt for model training. To select the optimal model scheme, four combinations of rule compositions and feature function outputs were designed, and model performance on the validation dataset was evaluated using ROC curves. Finally, the model was further verified with in situ environmental and fisheries data from April to November 2024. Results showed that the predicted fishing grounds were highly aligned with the actual monthly fishing grounds in 2024, and the predicted migration routes matched the movement trajectory of fishing vessels. The model also exhibited satisfactory performance, achieving an average AUC of 0.722 ± 0.033, a sensitivity of 0.604, a specificity of 0.834, and a negative predictive value (NPV) of 0.978. In conclusion, the CART rule-guided MaxEnt model, integrating the interpretability of CART and the predictive power of MaxEnt, effectively predicts the spatial distribution of chub mackerel fishing grounds in the northwest Pacific Ocean, providing technical support for fishery management and ecological research.

Keywords: chub mackerel; CART; MaxEnt; fishing ground; Northwestern Pacific Ocean

Key Contribution: This study developed a novel CART rule-guided MaxEnt model integrating fishing effort and five marine environmental factors to predict the fishing grounds of chub mackerel in the northwest Pacific; validated with 2024 in situ data, the hybrid model achieved reliable performance with high consistency between predicted and actual fishing grounds and migration routes, combining interpretability and predictive power to support local fishery management and ecological research.

Check for updates

Academic Editor: Patrice Brehmer

Received: 16 April 2026
Revised: 31 May 2026
Accepted: 1 June 2026
Published: 4 June 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Chub mackerel (*Scomber japonicus*), an economically valuable pelagic migratory fish in the Northwestern Pacific Ocean, is not only one of the main target species for distant-water fisheries of coastal countries bordering the Northwestern Pacific but also a key species maintaining the balance of the marine food web [1,2]. Among these coastal nations, China, Japan, and Russia are the primary fishing nations in the Northwest Pacific. Since 2017, the combined annual catch of chub mackerel by these three countries has shown a gradual declining trend, dropping from 530,000 tons to 130,000 tons [3]. Notably, China launched the high-seas trawling and purse seine fishery program in the Northwestern Pacific in 2014, with light purse seine as the dominant fishing method and chub mackerel as the primary target species [4]. However, affected by global climate change, such as sea surface temperature anomalies and altered ocean current patterns, as well as overfishing, the migration routes, aggregation patterns, and fishing ground distribution of chub mackerel exhibit significant spatiotemporal heterogeneity [5,6]. Given this complex spatiotemporal variability, traditional fishing ground prediction methods relying on empirical judgment or a single model are no longer sufficient to meet the demands of precise fishing and sustainable management. Therefore, there is an urgent need to construct an efficient and stable prediction model system to achieve accurate forecasting of chub mackerel fishing grounds.

The core of fishing ground prediction lies in revealing the coupling relationship between fish aggregation and marine environmental factors, whose accuracy directly determines the efficiency of fishery production and the rationality of resource utilization [7]. In recent years, machine learning algorithms have been widely applied in the fields of fishery resource assessment and fishing ground prediction due to their powerful capabilities for nonlinear fitting and feature mining, thereby gradually replacing traditional statistical regression methods [8–12]. Furthermore, with the resurgence of neural network technology, its applications in the fishery sector have gradually expanded. Building on this progress, deep learning approaches have increasingly been adopted for fishing ground prediction [13–16].

Although deep learning models have demonstrated high-precision prediction potential in fishery research, traditional machine learning models still outperform them in small-dataset scenarios due to their advantages, such as simple structure, strong data adaptability, good result interpretability, and low deployment costs [14,17]. Among various traditional machine-learning models, different types exhibit complementary performance characteristics across diverse application scenarios owing to their unique learning mechanisms. Specifically, tree-based models (e.g., decision trees, Boosting Regression Trees) have gained considerable popularity in fishery research for their excellent interpretability. For instance, in intelligent feeding control systems for aquaculture, decision tree models can accurately identify fish feeding states, and their intuitive, transparent decision paths clearly illustrate the model's operational logic, facilitating the understanding of intelligent feeding mechanisms [18]. In a study on walleye pollock (*Gadus chalcogrammus*) fishing ground prediction in the western Bering Sea, Boosting Regression Trees (BRT) successfully revealed the importance ranking of environmental factors such as chlorophyll concentration and pH value, providing reliable ecological references for resource management [19]. By simulating human decision-making logic, such models translate complex fishery phenomena into interpretable rule-based systems, thereby significantly enhancing the operability and credibility of model results in practical applications.

On the other hand, ecological niche models represented by MaxEnt (Maximum Entropy) have demonstrated significant advantages in predictive performance, particularly for biogeographic distribution [20]. By utilizing only occurrence data and environmental

variables, the MaxEnt model can accurately predict potential occurrence with high precision. For example, Yan et al. (2020) confirmed that when predicting the epizootic risk of Meriones unguiculatus plague, the MaxEnt model achieved an AUC of 0.987, delivering more reliable predictions [21]. In the prediction of potential habitats for rare bird species such as *Pteroptochos tarnii* and *Eugralla paradoxa* in Chilean tropical forests, the AUC values of the model on the test set achieved 0.868 and 0.994, respectively, demonstrating strong predictive accuracy and stability [22]. Such robustness and high prediction precision in handling small sample sizes and complex environmental gradients have enabled MaxEnt and its core principles to exhibit enormous application potential in fisheries science, especially in the assessment of habitat suitability for economically important fish species and fishery forecasting. Existing studies have attempted to transfer its principles to the prediction of the spatial distribution of fishery resources. By integrating marine remote sensing environmental data (e.g., sea surface temperature, chlorophyll-a concentration), high-precision fishery forecasting models have been constructed [4,23,24].

Different types of models each possess their own advantages and limitations. How to effectively integrate the strengths of diverse models to develop a hybrid model framework with both strong interpretability and high prediction accuracy has become a key scientific issue in enhancing fishery data analysis capabilities. Building on this, the present study focuses on chub mackerel in the northwest Pacific Ocean, integrating fishery logbook data from 2018 to 2023 with multi-source remote sensing environmental data (e.g., sea surface temperature [SST], chlorophyll-a concentration [CHL], sea surface height anomaly [SSHA]). Utilizing Bootstrap to enhance rule diversity, a CART rule-guided MaxEnt model was constructed for fishing ground forecasting. By comparing and analyzing the forecasting performance of the model under different methodological schemes, the superiority of the model was verified. This study not only contributes to the innovation and improvement of fishing ground-forecasting models in fisheries science but also serves as a practical reference for the sustainable utilization and management of chub mackerel resources.

2. Data and Method

2.1. Sources and Processing of Fishery Data and Environmental Data

Fishery data were obtained from the fishing records of China's light purse seine operations in the high seas of the Northwest Pacific from 2014 to 2024. The dataset includes information such as operation date, operation location (longitude and latitude, °), fishing effort (haul), and catch yield (kg) of chub mackerel. The study area is primarily located in the sea area east of the Exclusive Economic Zones (EEZ) of Japan and Russia, covering the range of 35° N–48° N and 144° E–170° E (Figure 1).

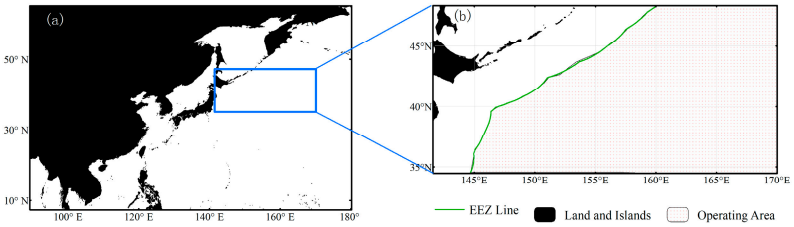

Figure 1. Schematic map of fishing ground for chub mackerel in the Northwest Pacific Ocean: (**a**) geographic location of the study area, (**b**) scope of fishing ground demarcated by the EEZ boundary.

Numerous studies have shown that the distribution of chub mackerel resources varies with changes in environmental factors, including sea surface temperature (SST, °C) [25–28], chlorophyll concentration (CHL, mg/m^3) [1,26,27,29], sea level anomaly (SLA, m) [30,31], SST gradient (GSST, °C/km) [25,32], and ocean currents [1,25,33]. Accordingly, this study

incorporated the above-mentioned environmental factors and time variable (month) into the model. Among these, data for SST, CHL, SLA, and ocean currents were retrieved from the Copernicus Marine Data Store (https://data.marine.copernicus.eu/, accessed on 22 October 2025), while the calculation method for the GSST followed the protocol described in the literature [34].

To achieve spatiotemporal matching between fishery data and environmental data, both datasets were aggregated into 0.5° × 0.5° gridded data on a monthly scale. We then linked and matched the two datasets using dual keys: time (month) and space (longitude, latitude). Ultimately, a spatiotemporally integrated comprehensive analytical dataset was developed to support subsequent model training.

2.2. The Process of Dataset Generation

The dataset spanning from 2014 to 2023 was selected as the fundamental dataset for model training and validation, while the 2024 dataset was independently designated as the test set to objectively evaluate the generalization ability of the models. For the 2014–2023 dataset, a 7:3 partitioning ratio was adopted to split the data into a training set and a validation set.

In fishery-related research, catch per unit effort (CPUE) is commonly employed to characterize fishery resource density [35–37]. However, CPUE data suffers from systematic biases caused by fishermen's adaptive strategies and market-driven fishing behaviors, which may affect its application in reflecting actual resource abundance [38]. Referring to existing methodological practices [30,39,40], this study adopted fishing effort rather than CPUE for fishing ground delineation as a practical methodological choice. Therefore, in the present study, monthly fishing effort (haul) was selected as the indicator for distinguishing fishing grounds from non-fishing grounds, and the median value of monthly fishing effort was set as the threshold to demarcate fishing grounds and non-fishing grounds for each month.

2.3. The Construction of CART Rule-Guided MaxEnt Model

An operational forecasting model for the chub mackerel fishing ground was established via the integration of the Classification and Regression Tree (CART) and Maximum Entropy (MaxEnt) algorithms (Figure 2). Specifically, the dataset was partitioned into training and validation subsets following the method detailed in Section 2.2. The CART model was then trained on the training subset, from which if–then decision rules were extracted. These decision rules were further converted into feature functions and fed into the MaxEnt model as constraints. During the validation phase, the optimal modeling framework was identified through performance evaluation, and this optimized scheme was subsequently implemented to generate the 2024 chub mackerel fishing ground prediction.

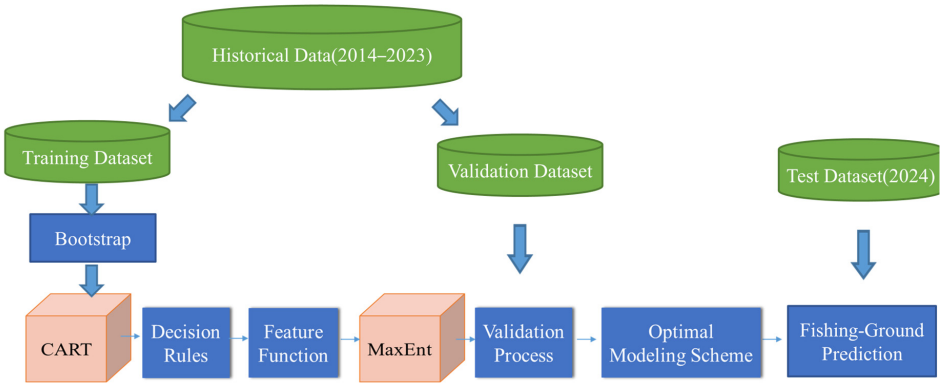

Figure 2. Flow chart of CART rule-guided MaxEnt model construction.

2.3.1. CART Rule Generation

CART, introduced by Breiman et al. [41], is a widely used decision tree algorithm celebrated for its interpretability and efficiency in extracting explicit decision rules from complex datasets [42–44]. In the present study, we built a CART model to extract two types of if–then decision rules, namely, Single-Variable Decision Rules (SVDRs) and Multivariate Decision Rules (MVDRs) for chub mackerel fishing grounds:

SVDRs: Rules were derived from individual environmental feature variables independently. These rules are concise, highly interpretable, and directly quantify the independent effect of a single environmental factor on the response variable.

MVDRs: MVDRs were jointly generated from multiple environmental variables and can capture interactions between features. To enhance the generalization ability of the rules, this study adopted the core mechanism of random forest and utilized the Bootstrap resampling method to improve the diversity of CART-derived decision rules [45,46]. Notably, the stochastic nature of random forest stems from the combination of Bootstrap sampling and random feature selection, where the former (i.e., Bootstrap sampling) acts as the key to improving the diversity of base learners. Given the original training set D (with a sample size of N), the Bootstrap strategy was applied to generate T subsets, ensuring that each sampling process was independent of the previous ones and the sample size of each subset was identical to that of the original training set. An independent unpruned CART was constructed based on each subset. During node splitting, the Breiman criterion was followed, and $m = \text{int}(\log_2 M + 1)$ candidate features were randomly selected from the total of M-dimensional features [45].

2.3.2. MaxEnt Feature Function Design

MaxEnt is a statistical model based on information theory, whose core idea is to select the probability distribution with the maximum entropy under the premise of satisfying known constraint conditions [47,48]. One of its main advantages is that it requires only a small amount of species occurrence data and environmental variables to generate high-precision distribution prediction maps, and this characteristic has been widely verified in relevant studies [22,49]. The formula for the class-conditional probability ultimately output is as follows:

$$p(y \mid x) = 1/Z(x)\exp(\sum_{i=1}^{n} \lambda_i f_i(x,y)) \quad (1)$$

where x denotes the input features, y represents the label value, $f_i(x,y)$ stands for the i-th feature function, λ_i is the i-th weight value, and $Z(x)$ serves as the normalization factor, i.e., $Z(x) = \sum_y \exp(\sum_{i=1}^{n} \lambda_i f_i(x,y)$. In this study, the feature function $f(x,y)$ is the carrier linking CART rules and MaxEnt. It was used to formalize the associative relationships in the rules into learnable constraints for the model. In the present study, two feature function schemes, i.e., Binary Feature Function (BFF) and Continuous Feature Function (CFF) for MaxEnt, were designed:

BFF: If the class probability value output by the CART rule was greater than 0.5, the sample was determined to satisfy the corresponding rule, and the value of $f(x,y)$ was set to 1; otherwise, it was set to 0. This scheme quantified the condition of "satisfied-unsatisfied" fishing grounds, i.e., "fishing grounds or non-fishing grounds". If the class probability output by the CART rule exceeded 0.5, the sample was considered to satisfy the corresponding rule, and $f(x,y)$ was assigned a value of 1; otherwise, it was assigned 0. This approach quantified the rule-satisfied/rule-unsatisfied condition as a binary classification of fishing ground (1) or non-fishing ground (0).

CFF: It directly took the class probability by the CART decision tree as the value of $f(x,y)$. This retained the continuous gradient features in the fishing ground probability

information, thereby more finely depicting the degree to which the sample conformed to the rules.

Taking a set of typical MVDRs extracted from CART as an example:

3) SST < 21.14371 2444 1213 TRUE (0.49631751 0.50368249)

6) Month < 5.5 699 276 FALSE (0.60515021 0.39484979)

12) CHL < 0.1290527 183 35 FALSE (0.80874317 0.19125683)

24) CHL $\geq$ 0.101188 58 3 FALSE (0.94827586 0.05172414) *.

These threshold-based decision rules are encoded as a CFF for the MaxEnt model, defined as:

$$f(Month, SST, CHL) = \begin{cases} 0.052, if\ SST < 21.144\ and\ Month < 5.5\ and\ CHL < 0.129\ and\ CHL \geq 0.101 \\ 0.948, if\ SST < 21.144\ and\ Month < 5.5\ and\ CHL < 0.129\ and\ CHL < 0.101 \end{cases}$$

By combining two types of CART rules (i.e., SVDRs and MVDRs) with two types of feature functions (i.e., BFF and CFF), a total of four rule-guided MaxEnt model schemes were constructed in this study: SVDR + BFF, MVDR + BFF, SVDR + CFF, and MVDR + CFF. These schemes were developed to evaluate and compare the predictive performance of different rule-feature function combinations.

2.3.3. Maximum Entropy Model Training

We adopted the Batch Gradient Descent (BGD) algorithm [50,51] to train the Maximum Entropy Model and solve for the weight parameters λ_i corresponding to each feature function in Equation (1). The specific process is as follows:

Parameter Initialization: Set the stopping criterion (stopEPS = 0.001). To guarantee an unbiased starting state for the model, the weight parameters λ_i corresponding to all feature functions were initialized to 0.

Gradient Iterative Calculation: Traverse all training samples in each iteration. Firstly, calculate the feature values of the samples via the feature functions $f_i(x, y)$ designed in Section 2.2. Subsequently, compute the linear combination dotProduct $= \sum_{i=1}^{n} \lambda_i f_i(x_k, y_k)$(where y_k denotes the actual label of the $k-$th sample), and convert it into the predicted probability $p_k = 1/(1 + \exp(-1/(1 + \text{dotProduct})))$. Finally, based on the error between the actual label (0 or 1) and the predicted probability, accumulate the gradient of each feature function as $\text{gradient}_i = \text{gradient}_i + (\text{y}_k - p_k) f_i(x_k, y_k)$.

Weight Update: Adjust the weight parameters according to the learning rate (lRate) with the update formula $\lambda_i = \lambda_i + lRate \cdot \text{gradient}_i$. By adjusting the model parameters (i.e., the weights λ_i of feature functions) along the direction of the fastest increase in the function value indicated by the gradient, the maximization of the model's log-likelihood function is ultimately achieved.

Convergence Criterion: The gradient norm (normGrad) was calculated. This study adopted the L2 norm (i.e., Euclidean norm) for the calculation, which is expressed as normGrad(gradient) = $\sqrt{\sum_{i=1}^{n} gradient^2}$. When normGrad < stopEPS, the model was determined to have converged, and the iteration was terminated.

2.4. Determination of the Optimal Probability Threshold

Given that this study aimed to objectively delineate suitable habitats rather than conduct cost-oriented fishery risk assessment, the Youden index was retained as a uniform and unbiased threshold criterion. This scheme ensures the objectivity and reproducibility of habitat zoning, and it has been widely validated in ecological suitability modeling. In binary classification problems in machine learning, Youden's J index is used to select the optimal probability threshold, with its core lying in balancing the sensitivity and specificity of the classification model [52]. This index is defined as the sum of sensitivity and specificity

minus 1, i.e., J = sensitivity + specificity — 1, thereby identifying a cutoff threshold that can optimally distinguish between the two classes of samples [53]. In this study, we input the validation data into the constructed model and calculated the predicted fishing ground probability for each sample. By combining these predicted probabilities with the true labels, ROC curves were plotted and Youden's J index was computed. Once Youden's J index reached its maximum value, the corresponding predicted probability was designated as the cutoff threshold for fishing ground prediction. Specifically, if the predicted probability of a sample was larger than this cutoff threshold, the fishing zone was classified as a "fishing ground"; otherwise, it was classified as a "non-fishing ground."

2.5. Overlap Analysis of Predicted Probability and Fishing Effort

Schoener's D was used to quantify the spatial overlap between different species or scenario simulations [54]. This index ranges from 0 to 1, with 0 indicating no spatial overlap and 1 indicating perfect spatial matching between the two distributions.

The calculation formula is as follows:

$$D = 1 - \frac{1}{2}\sum_{i=1}^{n}|P_i - F_i| \quad (2)$$

where P_i represents the normalized habitat prediction probability of grid cell i, F_i is the normalized fishing effort density of the corresponding grid cell, and n is the total number of grid cells in the study area.

3. Result

3.1. Monthly Spatiotemporal Distribution of Historical Operating Positions

Based on the long-term fishery data from 2014 to 2023, Figure 3 presents the spatiotemporal distribution pattern of fishing operation positions of chub mackerel in the northwestern Pacific Ocean from April to November, covering the sea area ranging from 145° E to 165° E and 35° N to 45° N. In terms of the temporal dimension, the fishing operation positions from April to June (green dots) are concentrated in the sea area from 145° E to 155° E and 35° N to 40° N. Those from July to August (yellow dots) expand northeastward to the region from 150° E to 160° E and 38° N to 43° N. From September to November (orange-red dots), they gradually migrate southward, distributing in the sea area from 150° E to 165° E and 35° N to 43° N. Overall, it shows a spatiotemporal dynamic characterized by an expansion from southwest to northeast from April to August and a southward migration after September. Moreover, there is a significant spatial correlation between the distribution of fishing operation positions and the Exclusive Economic Zone (EEZ) line, reflecting the stable migration habits and operation distribution characteristics of this species in the study area during 2014–2023.

3.2. Relationships Between Seasonal and Environmental Factors and Fishing Effort

The distribution of fishing effort varies significantly with the values of environmental factors. Specifically, for SST, the peak in April (652 hauls) occurs at 12.8 °C, while relatively higher peaks occur at 18.2–19.2 °C in July and August (3781 and 2570 hauls) (Figure 4a). In contrast, the peak position of CHL gradually decreases from 0.68 mg/m^3 (808 hauls) in April to 0.36 mg/m^3 (5578 hauls) in August (Figure 4b). Regarding current speed, the fishing effort peak shifts to higher speeds from 0.125 m/s (712 hauls) in April to 0.275 m/s (3674 hauls) in June, before dropping to the lowest speed of 0.025 m/s (3928 hauls) in November (Figure 4c). Similarly, the peak position of fishing effort in SST gradient ranges from 0.014 °C/km (379 hauls) in April to 0.05 °C/km (1617 hauls) in November (Figure 4d). Finally, the peak position of fishing effort in SSHA changes from −0.125 m to 0.075 m

(Figure 4e). Overall, it can be seen that the peak positions of fishing effort continuously change in the value ranges of each environmental factor with seasonal progression.

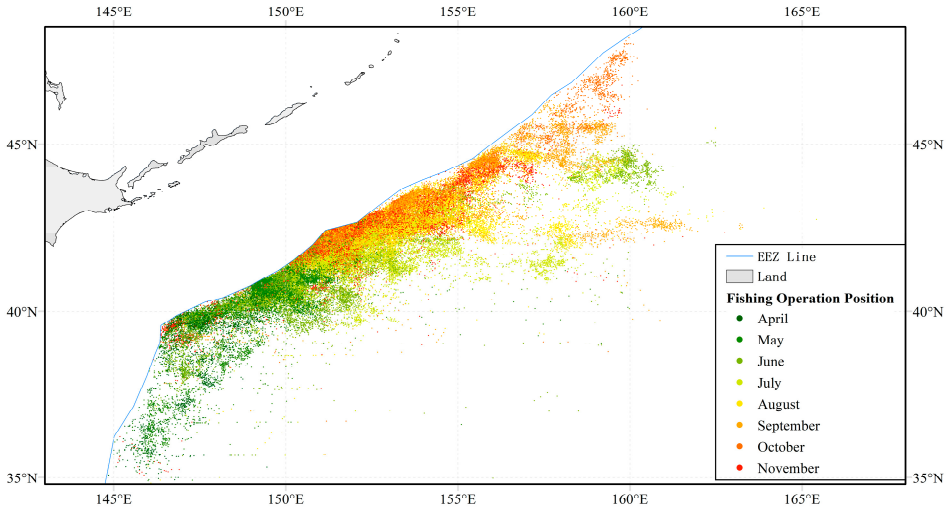

Figure 3. Monthly fishing positions of chub mackerel in the Northwest Pacific Ocean, 2014–2023.

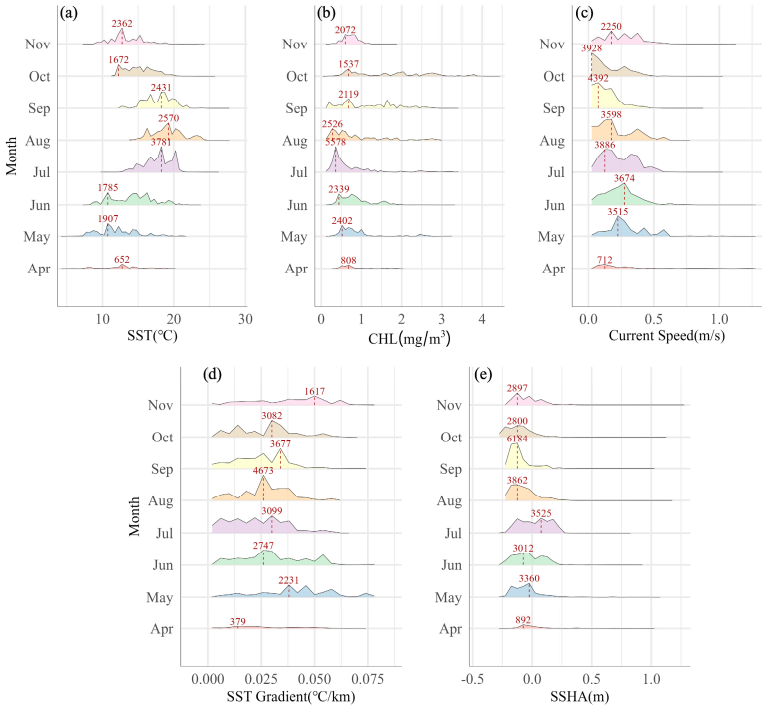

Figure 4. Variation in fishing effort in response to (**a**) SST, (**b**) CHL, (**c**) current speed, (**d**) SST gradient, and (**e**) SSHA across different seasons. Note: Red numbers in the figure indicate purse seine hauls at the peak.

3.3. Selection of Optimal Combination Scheme of Rules and Feature Functions

The monthly AUC distributions in the training and validation phases revealed significant differences among the five models (one CART and four rule-guided MaxEnt models). In the training phase (Figure 5a), CART (AUC: 0.684–0.843) and SVDR + BFF (0.695–0.819) performed the worst, followed by MVDR + BFF (0.762–0.875) and SVDR + CFF (0.750–0.863). The MVDR + CFF MaxEnt model was optimal, with AUC values of 0.840–0.941. In the validation phase (Figure 5b), SVDR + BFF remained the worst (AUC: 0.608–0.772), followed by SVDR + CFF (0.646–0.782), MVDR + BFF (0.676–0.798), and CART (0.696–0.830). The MVDR + CFF model still performed best (AUC: 0.727–0.883).

In summary, the MVDR + CFF model exhibited the best training fitting and validation generalization performance.

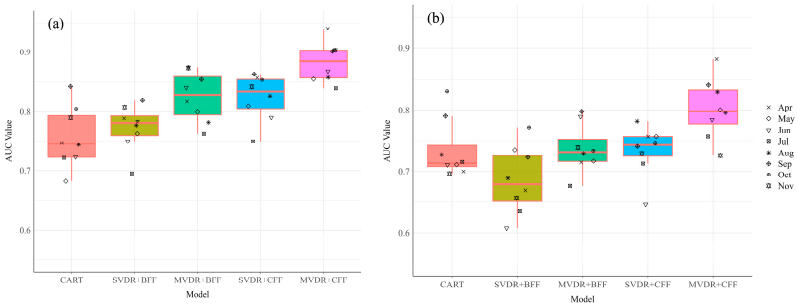

Figure 5. Monthly AUC distributions of five models (CART and four rule-guided MaxEnt) in the (**a**) training phase and (**b**) validation phase.

3.4. Determination of Monthly Optimal Probability Thresholds

Figure 6 presents the detailed ROC curves of the training and validation phases under the optimal model scheme (e.g., MVDR + CFF). Based on the analysis of these ROC curves and their corresponding AUC values over the 8-month period from April to November, the AUC values of the training phase ranged from 0.840 in July (Figure 6d) to 0.941 in April (Figure 6a), with an average of approximately 0.884. Conversely, those of the validation phase ranged from 0.727 in November (Figure 6h) to 0.883 in April (Figure 6a), with an average of approximately 0.802. As clearly illustrated in the figure, the training set curves for all months lay above the validation set curves, and moreover, all resided above the diagonal line (i.e., the random guess line). Although the validation phase curves fluctuated, they still exhibited effective ability to distinguish between "fishing grounds" and "non-fishing grounds" overall.

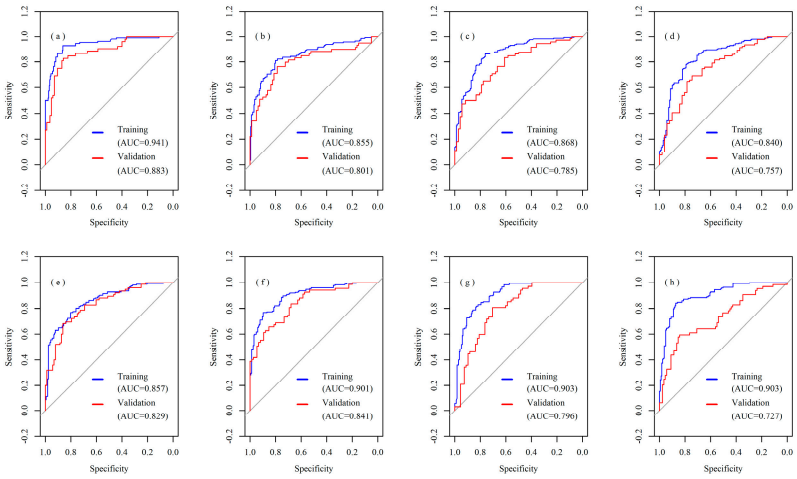

Figure 6. ROC Curves of training and validation datasets for the optimal model scheme (MVDR + CFF) . (**a**) April; (**b**) May; (**c**) June; (**d**) July; (**e**) August; (**f**) September; (**g**) October; (**h**) November. Note: The grey diagonal line represents the ROC baseline, corresponding to an AUC value of 0.5.

Subsequently, based on the validation phase ROC curves (red curves in Figure 6), a curve illustrating the relationship between Youden's J index and probability thresholds was plotted to determine the optimal values (Figure 7). Specifically, the thresholds for April–July and September were all less than 0.5, whereas those for August and October–November were all greater than 0.6.

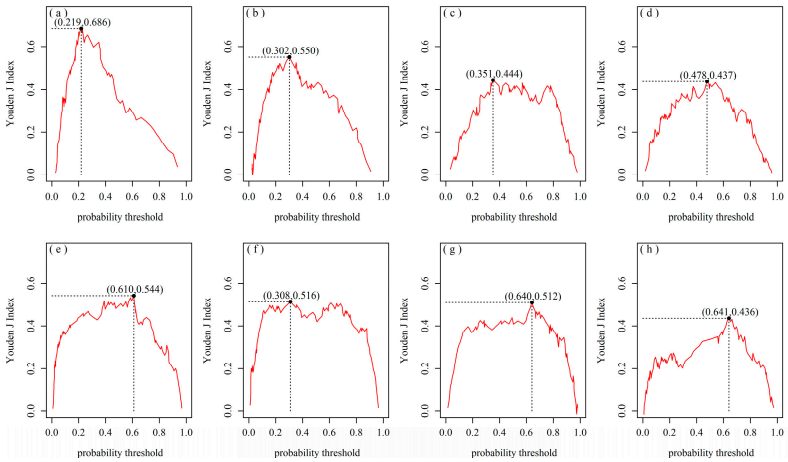

Figure 7. Relationship between Youden's J index and probability thresholds for the optimal model scheme (MVDR + CFF) . (**a**) April; (**b**) May; (**c**) June; (**d**) July; (**e**) August; (**f**) September; (**g**) October; (**h**) November. Note: The positions of black dots on the *X*-axis represent optimal probability thresholds.

3.5. Fishing Ground Forecasting Results and Performance Verification in 2024

From April to November 2024, the chub mackerel fishing season in the Northwest Pacific Ocean exhibited significant spatial dynamic changes, with good overall consistency between the predicted fishing grounds and actual fishing areas (Figure 8). Initially, in April, the fishing season was in its initial stage, with relatively few active vessels scattered in the southwestern waters of the Northwest Pacific fishing ground. High-probability areas predicted by the model appeared in the southern region around 149° E, 37° N, corresponding to a small number of actual fishing operations recorded there (Figure 8a). Subsequently, in May, the fishing region expanded northeastward, and the predicted fishing ground probability near the EEZ (149° E–155° E) increased significantly, which was highly consistent with the large number of fishing operations in this area. However, the southern waters (147° E, 37° N) showed high predicted probability but no actual fishing activities (Figure 8b). From June to August, the fishing positions further extended northeastward, crossing 165° E and 47° N. During this period, the areas with high predicted probability presented a northeast-southwest wavy strip distribution, showing high consistency with the actual fishing areas (Figure 8c–e). By September, the fishing grounds shifted westward and showed high spatial aggregation (Figure 8f). Notably, the high-probability predicted areas from August to September were all located north of 42° N with a wide coverage, and the actual fishing operations were basically concentrated in these waters. Finally, from October to November, the fishing area continued to shrink southwestward and was distributed in a strip along the EEZ line, corresponding to the narrowed spatial scope of the high-probability predicted areas. Specifically, considerable fishing activities occurred west of 160° E, corresponding to the area of high predicted probabilities, while no fishing activities were observed east of 160° E (Figure 8g,h).

To evaluate the performance of the 2024 fishing ground prediction model, we generated monthly ROC curves from April to November based on the predicted results (Figure 9). In each subplot, sensitivity is plotted against specificity with the diagonal line representing random guessing. The results showed that all ROC curves lay above this diagonal line, which indicates the model's consistent capability to distinguish between fishing grounds and non-fishing grounds. Monthly AUC values ranged from 0.657 (November) to 0.757 (September), with an average of 0.722 ± 0.033, which reflects reasonable to good overall classification performance for chub mackerel fishing ground prediction in 2024.

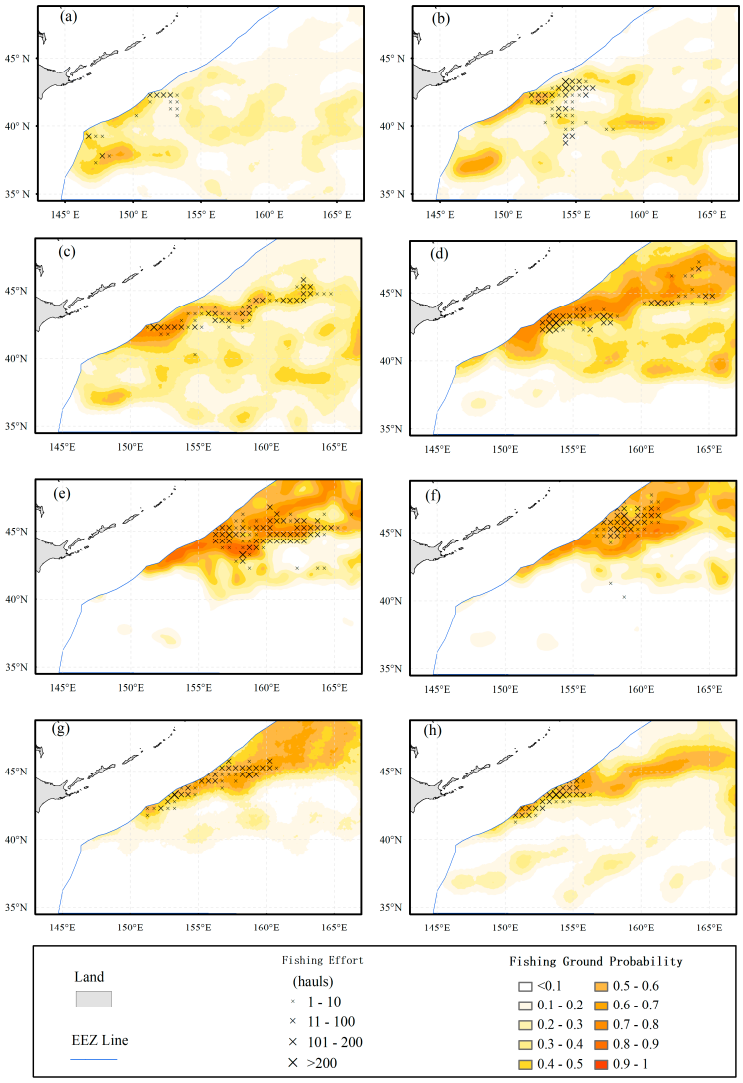

Figure 8. Overlay map of predicted fishing ground probability distribution and actual fishing positions: (**a**) April, (**b**) May, (**c**) June, (**d**) July, (**e**) August, (**f**) September, (**g**) October, and (**h**) November 2024.

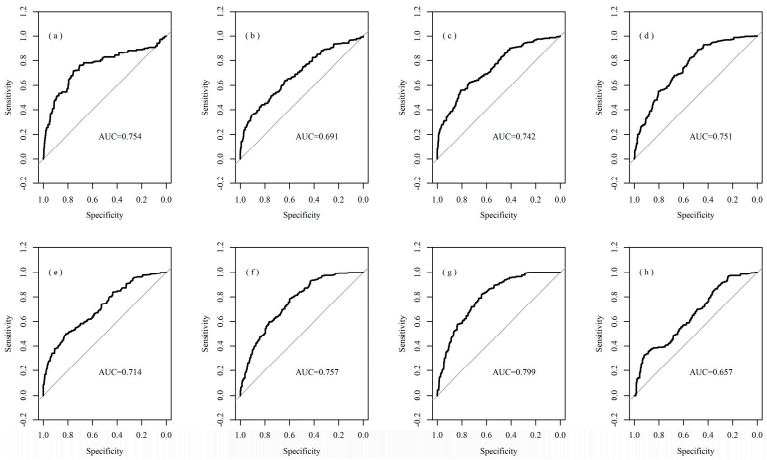

Figure 9. Monthly ROC curves and AUC evaluation of the 2024 fishing ground prediction model: (**a**) April, (**b**) May, (**c**) June, (**d**) July, (**e**) August, (**f**) September, (**g**) October, and (**h**) November.

The monthly Schoener's D values between observed fishing catch distribution and predicted fishing ground probability from April to November ranged from 0.484 to 0.595,

with values of 0.513, 0.595, 0.533, 0.484, 0.580, 0.509, 0.527 and 0.589, respectively. Overall, the overlap index remained at a moderate spatial agreement level. The model could generally capture the spatial pattern of actual fishing distribution, while partial spatial mismatches still existed in some months.

To further validate the predictive performance of the model, a confusion matrix was constructed by matching the model outputs with the actual fishing area data of 2024 in both spatial and temporal dimensions. Using the optimal probability thresholds derived from Figure 7 as the criteria for distinguishing fishing grounds from non-fishing grounds, the model performance metrics were calculated based on the 2024 in situ fishing data (Table 1). Specifically, the model successfully identified 218 true fishing grounds (true positives, TP), while missing 143 actual fishing grounds (false negatives, FN). Additionally, it correctly classified 6384 non-fishing areas (true negatives, TN) but misclassified 1275 non-fishing areas as fishing grounds (false positives, FP).

Table 1. Confusion matrix of the model prediction of fishing grounds in 2024.

		Actual Labels	
		Non-Fishing Zone	**Fishing Zone**
Predicted Labels	Non-Fishing Zone	6384	143
	Fishing Zone	1275	218

In terms of specific classification performance, the model's sensitivity (True Positive Rate, TPR) was 0.60388, which meant that the model could correctly identify 60% of actual fishing ground (positive class) instances. Its specificity (True Negative Rate, TNR) was 0.83353, indicating that the model could accurately recognize 83% of actual non-fishing ground (negative class) instances. The Negative Predictive Value (NPV) was 0.97809, meaning that about 98% of the instances predicted as negative class by the model were truly negative class. These results indicated that the model was far more reliable in predicting non-fishing grounds than in predicting actual fishing grounds.

4. Discussion

4.1. Roles of Seasonal Factors and Marine Environmental Factors in the Model

In this study, a total of six predictive variables were used for modeling, including one temporal variable (month) and five environmental variables: sea surface temperature, chlorophyll-a concentration, ocean current, SST gradient, and sea surface height anomaly. Validated by independent test data, the model achieved an average AUC value of 0.8 (Figure 5b), indicating strong discriminative ability. Consistent with this performance, the model also exhibited reliable predictive power for the fishing ground distribution of chub mackerel in 2024 (Figure 8). While the model's predictive efficacy has been sufficiently verified, the underlying mechanisms driving its performance lie in the critical concurrent effect between temporal and environmental variables, which can be elaborated from two aspects:

First, as a temporal variable, month reflects the extensive regulatory effect of seasons on the marine environment and is closely linked to the biological characteristics of chub mackerel. The Pacific stock of chub mackerel is widely distributed from the southern Pacific coast of Japan to the offshore waters of the Kuril Islands. Spawning occurs in the coastal waters of central and southwestern Japan from February to June; larvae are transported by ocean currents to the feeding grounds in the Kuroshio-Oyashio Transition Zone (KOTZ) and reach the northernmost feeding areas in September, before gradually migrating back to coastal regions after late summer [55,56]. Additionally, chub mackerel's

adaptability to environmental factors varies across seasons [57]. This season-dependent adaptive strategy further highlights that temporal variables exert a significant impact on chub mackerel's distribution, even greater than spatial variables [58]. In this study, the CART model effectively captured interactive relationships between the month variable and other environmental indicators. The temporal factor month not only helped reveal drivers of the species' seasonal behaviors but also played a pivotal role in the proposed model.

Second, the five environmental variables each perform distinct ecological functions. Among these, SST is regarded as a key regulatory factor for fish survival and migration, with fish being highly sensitive to temperature changes and their spatial distribution often constrained by suitable temperature ranges [29,57]. Similarly, Chl-a concentration, as a critical indicator of primary productivity, directly reflects the abundance of food resources on which fish depend and thus exerts a significant influence on the spatial distribution of chub mackerel [1,27,59]. Furthermore, the dynamic characteristics of ocean currents in the Northwest Pacific further shape the distribution of nutrients and phytoplankton [60], whereas SST gradients, which characterize frontal zone ecology, form high-productivity areas that serve as habitat hotspots for fish [6,32]. In addition to these factors, SSHA directly influences habitat localization through mesoscale dynamic processes (e.g., warm eddies and cold eddies). Fan et al. (2020)'s chub mackerel HSI model, with environmental weights derived from catch and fishing effort, showed SSHA weights (0.41, 0.33) ranking second to SST [30]. Previous studies have also indicated that the fishing grounds of chub mackerel are consistently located at the junction of high and low sea level anomaly (SLA) values, with high-yield areas always on the high-value side [30,31]. Collectively, these findings provide solid quantitative support for the development of fishing ground prediction models.

Therefore, the bidirectional migration of Pacific chub mackerel is an integrated outcome of intrinsic life-history rhythms and varying marine environments. Such seasonal biological habits determine the species' periodic northeast feeding migration and southwest spawning migration. Correspondingly, critical marine variables, including temperature, food availability, and mesoscale oceanographic features, further modulate its spatial distribution. Owing to the effective capture of seasonal-environmental interactions, the CART-MaxEnt model can reliably reflect the dynamic distribution of chub mackerel fishing grounds.

4.2. Analysis of the Model's Prediction Performance

The fishery forecasting results of 2024 indicated that the model exhibited characteristics of moderate sensitivity (60.39%), high specificity (83.35%), and high NPV (97.81%), reflecting its application value in the forecasting of chub mackerel fishing grounds in the Northwest Pacific Ocean.

From the perspective of fishery application scenarios, high specificity and high NPV represent the core advantages of this model. The NPV implies that nearly 98% of the areas classified as non-fishing grounds by the model were true non-fishing grounds, a feature with direct guiding significance for fishery production. Given the vast fishing area in the Northwest Pacific and the high costs of fuel and time for fishing vessels, the model can accurately screen out regions with low resource density, helping fishermen avoid ineffective operations and significantly improving fishing efficiency and economic benefits. This aligns with the goals of reducing operational risks and optimizing resource utilization in fishery production. The high specificity (83.35%) further demonstrates the model's strong stability in identifying non-fishing ground environments, and the low false positive rate of only 16.65% can effectively reduce resource waste caused by false reporting of fishing grounds.

In contrast, the moderate sensitivity (60.39%) indicates that approximately 39.61% of the true fishing grounds were not identified by the model. This phenomenon requires a

comprehensive analysis combining the ecological characteristics of chub mackerel and the constraints of model construction. As a typical seasonal migratory fish [55], the formation of chub mackerel fishing grounds is regulated by the coupling of multiple environmental factors [61,62]. Some fishing grounds may be associated with transient environmental processes such as mesoscale eddies and water temperature gradients [5,28]. However, such dynamic environmental variables were not incorporated into the input features of the model in this study. Additionally, the spatiotemporal resolution of the environmental data employed herein may have been insufficient to capture the formation mechanisms of local fishing grounds, thereby masking the environmental signals associated with true fishing grounds. These two limitations collectively impaired the model's ability to identify positive class samples (i.e., fishing grounds). Furthermore, within the entire fishing ground domain of the Northwest Pacific, the spatial coverage of true fishing ground samples was relatively limited, which resulted in the statistically derived predictive accuracy of the model being biased toward the non-fishing ground class (Table 1 and Figure 8).

It is further worth noting that the MaxEnt model requires only presence-only data for its implementation [63]. In the present study, fishing effort was used merely as a practical methodological basis, and records with fishing effort exceeding the median were selected as the threshold for defining fishing grounds. This setup thus represents a relatively stringent screening criterion for the application of the MaxEnt model. Similarly, such rigorous screening strategies have been widely adopted in previous studies on fishing ground forecasting models [40,64,65]. This is because fisheries production demands reliable predictions. Strict delineation of fishing grounds ensures the model's high-probability predictions are valid, better meeting the practical need to avoid unproductive voyages. We used this criterion for operational purposes only, without evaluating its advantages over CPUE from an ecological perspective.

4.3. Construction Logic and Performance Optimization of the Rule-Guided MaxEnt

In the process of integrating the CART and MaxEnt models in this study, the model construction was not achieved through a simple superposition of the two models but rather via the deep integration of if-then rules and feature functions. Based on the principle of locally optimal splitting, CART can generate a binary tree structure, which facilitates the interpretation of the causes underlying fishing ground formation in fishing ground forecasting [42,44]. However, the prediction function of the CART model is a step function, which is discontinuous and non-smooth [41]. Its extrapolation capability outside the range of training data is almost zero, leading to frequent misclassification of edge samples [66,67]. Compared with the four schemes defined in Section 2.3.2, the CART model exhibited the lowest training accuracy and the second-lowest validation accuracy (Figure 5a,b), confirming that the standalone CART model has no significant advantages in terms of predictive accuracy for chub mackerel fishing ground forecasting.

In contrast, the MaxEnt model is renowned for its ability to handle complex environmental data and provide continuous suitability predictions, and comparative studies with other models have further shown that it yields more accurate predictions of fish habitat locations and distributions [68,69]. However, MaxEnt has been criticized for its limited interpretability [20,70]. In the present study, the MaxEnt model directly converted the intermediate splitting rules of CART into feature functions. This approach avoids the blindness associated with traditional MaxEnt feature design while retaining core interpretive information such as variable thresholds and category constraints, thereby addressing to a certain extent the issue of insufficient interpretability in traditional MaxEnt models [71,72]. Furthermore, through smoothing via the maximum entropy probability distribution (Equation (1)), MaxEnt transformed the rigid classification boundaries of CART into continuous

suitability probabilities (ranging from 0 to 1), effectively reducing the risk of misclassifying edge samples.

4.4. Analysis of Performance Differences Among Models with Different Schemes

From the perspective of model rule composition, the simple rule construction method (e.g., SVDRs) processes individual environmental variables independently. By contrast, complex rules (e.g., MVDRs) are formed by combining multiple types of environmental variables. Such complex rules are more consistent with the ecological conditions of practical chub mackerel fishing ground forecasting (Figure 5). In the performance evaluation on the validation set, the model adopting the "MVDR + CFF" complex rule scheme exhibited particularly outstanding performance, with the median AUC value stably around 0.80 (Figure 5b). It demonstrated the robustness and generalization ability of the complex rule scheme. The formation of habitats and aggregation of fishing grounds for pelagic commercial fishes such as chub mackerel is inherently an ecological process driven by the synergistic effects of multiple factors, including SST, salinity, and prey density [57,61,62], and the independent effects of a single environmental variable cannot fully characterize this complex regularity. More importantly, the feature functions defined based on CART rules are not merely data-driven results but also possess clear ecological interpretability [73,74]: the splitting thresholds and combination logic of each environmental variable in the rules can be directly correlated with the ecological habits of fish. For instance, elevated catch yields of chub mackerel were correlated with SST ranging from 15 to 21 °C and low chlorophyll-a (Chl-a) concentrations (<2 mg m^{-3}) along the southern and southwestern coasts of Portugal [75]. Such a combination deepens the model's interpretive dimension from the traditional ranking of single-variable importance to the analysis of the contribution of multi-variable combinations.

In terms of the return values of feature functions, MaxEnt models utilizing continuous feature functions (CFF) outperformed those utilizing binary feature functions (BFFs), the construction of which involves a discretization or binning process (Figure 5). While data discretization can significantly reduce model complexity and improve the learning accuracy and speed of certain classifiers [76,77], the process of converting CART rules into MaxEnt feature functions via binning inevitably entails information loss [78,79]. For the present study, binning results in the loss of probabilistic gradient information inherent in CART rules, failing to distinguish differences in probabilities within the same interval. Although a probability threshold of 0.5 is commonly used as the binary classification criterion in machine learning models [80,81], the continuous class probabilities derived from CART can more fully retain original information. This enables the model to learn more precise relationships and patterns in the data. Thus, the MaxEnt model's overall performance is typically enhanced when its feature functions are of the continuous type.

4.5. Limitations and Future Prospects of the Model

In the present study, the rules generated by the CART model were adopted as the feature functions of MaxEnt. Comparative experimental results demonstrated that the rule features derived from combinations of multiple environmental variables could effectively improve the predictive accuracy of the model (see Figure 5). However, in research fields such as fishery science and ecology, it is equally crucial to deduce ecological mechanisms through model interpretability. Although the CART rules have strong interpretability, their inherent complexity seriously undermines the scientific interpretability of the model. This is mainly because the number of rules generated by the CART model increases exponentially with greater tree depth and more splitting nodes [41], which tends to produce a large number of redundant if-then combination rules. More critically, the weights (λ) of feature

functions are distributed across the combinatorial relationships of different environmental factors. As a result, the complex interactions among environmental factors exacerbate the challenges of model interpretation. To address this contradiction, pruning operations and tree depth parameter adjustments can be performed on the CART model [41,43]. Such measures can reduce the number of node splits, simplify the complexity of rule combinations, and thus improve ecological interpretability.

Despite the acceptable overall performance, the model exhibited a moderate sensitivity of 0.604, indicating that a portion of actual fishing grounds remained undetected. This limitation can be attributed to several factors. Chub mackerel displays patchy aggregation patterns driven by mesoscale oceanic fluctuations [5], resulting in scattered fishing spots that are difficult to capture accurately. The lack of relevant mesoscale dynamic variables in model inputs inevitably weakens the model's spatial precision and restricts its overall predictive performance. Meanwhile, this limitation also requires cautious interpretation of model outputs. For future work, we plan to integrate higher-frequency oceanographic observations and dedicated mesoscale indicators into the modeling framework to further improve prediction accuracy. Additionally, fishery datasets are commonly biased by preferential sampling, whereby fishers tend to target historically productive sea areas [82], thereby introducing inherent data noise and hindering the model's ability to identify sporadic fishing grounds. Notably, the relatively high specificity and favorable NPV (0.978) ensured low false-positive predictions and reliable identification of valid fishing areas. Future work will optimize classification thresholds and further refine the hybrid model structure to improve the detection capacity for discrete and marginal fishing grounds.

Additionally, this study ignored the inherent spatial autocorrelation of gridded oceanographic and fishery data. Adjacent grid cells often possess similar environmental conditions and fishing records, which violates the independent sample assumption of conventional machine learning evaluation. This issue may lead to mild overestimation of model performance metrics and reduce the spatial generalization ability of the model. Furthermore, the absence of spatial cross-validation made the model unable to completely eliminate spatial clustering bias. Nevertheless, monthly averaging of raw datasets and CART-based sample filtering moderately alleviated strong spatial dependence among sampling points. Future studies will introduce block spatial cross-validation to reduce spatial autocorrelation bias and further improve the reliability of model evaluation.

5. Conclusions

In this study, a CART rule-guided MaxEnt feature function construction method was proposed. By leveraging Bootstrap to enhance the robustness and diversity of feature functions, the complementary advantages of CART's interpretability and MaxEnt's predictive accuracy were achieved. Statistical and experimental comparisons confirmed that multivariate complex rules and continuous feature functions exhibited higher predictive accuracy in chub mackerel fishing ground forecasting in the northwest Pacific Ocean. This technical pathway integrates high accuracy and strong interpretability, serving as a modeling approach with theoretical rationality and practical value. Furthermore, this study objectively analyzed the limitations of rule interpretation in experimental validation and practical application and proposed directions for improvement focused on rule simplification. In the future, through iterative optimization of the method system and expansion of application scenarios, this technical pathway is expected to provide strong support for research on the balance between predictive accuracy and model interpretability in the field of fishing ground forecasting. Additionally, we will incorporate high-frequency datasets and mesoscale dynamic variables to further improve the model's spatial precision and predictive performance.

Author Contributions: Methodology, Z.W. and F.T.; conceptualization, X.C.; software, S.Z.; validation, X.C.; formal analysis, Y.W.; data curation, F.W. and Y.W.; writing—original draft preparation, Z.W. and F.T.; writing—review and editing, Z.W., F.T., and X.C.; visualization, S.Z.; supervision, X.C.; investigation F.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Ministry of Science and Technology of the People's Republic of China, grant number 2023YFD2401305.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The environmental data used in this study are available in the Copernicus Marine Data Store (https://data.marine.copernicus.eu/, accessed on 22 October 2025).

Acknowledgments: The authors would like to thank the Technical Group for High Seas Purse Seine and Trawl Fisheries, China Oceanic Fisheries Association for providing the catch data, and the Copernicus Marine Service for supplying the environmental data used in this study. This research was supported by the National Key R&D Program of China (Grant No. 2023YFD2401305).

Conflicts of Interest: The authors declare that they have no competing interests.

References

1. Dai, S.; Tang, F.; Fan, W.; Zhang, H.; Cui, X. Distribution of Resource and Environment Characteristics of Fishing Ground of *Scomber japonicas* in the North Pacific High Seas. *Mar. Fish.* **2017**, *39*, 372–382.
2. Zhou, H.; Yan, L.; Zhang, H.; Li, J. The Relationship between Age, Egg Size, and Fecundity of *Scomber japonicus* in the East China Sea. *J. Fish. Sci. China* **2022**, *29*, 1189–1197.
3. Yukami, R.; Nishijima, S.; Isu, S.; Furuichi, S.; Watanabe, R.; Higasikuchi, K.; Saito, R.; Ishikawa, K. *Stock Assessment and Evaluation for Chub Mackerel (Fiscal Year 2024). Marine Fisheries Stock Assessment and Evaluation for Japanese Waters*; Japan Fisheries Agency: Tokyo, Japan; Japan Fisheries Research and Education Agency: Kanagawa, Japan, 2025; pp. 1–97.
4. Xue, J.; Fan, W.; Tang, F.; Guo, G.; Tang, W.; Zhang, S. Analysis of Potential Habitat Distribution of *Scomber japonicus* in Northwest Pacific Ocean Using Maximum Entropy Model. *South China Fish. Sci.* **2018**, *14*, 92–98. [CrossRef]
5. Fan, X.; Cui, X.; Yang, S.; Tang, F. The Spatial Distribution Relationship between Mesoscale Eddies and Chub Mackerel and Its Preliminary Analysis of Causes in the Northwest Pacific Ocean. *Front. Mar. Sci.* **2025**, *12*, 1634527. [CrossRef]
6. Li, J.; Cui, X.; Tang, F.; Fan, W.; Han, Z.; Wu, Z. Spatiotemporal Analysis of Marine Environmental Influence on the Distribution of Chub Mackerel in the Northwest Pacific Ocean Based on Geographical and Temporal Weighted Regression. *J. Sea Res.* **2024**, *200*, 102514. [CrossRef]
7. Chen, X.; Gao, F.; Guan, W.; Lei, L.; Wang, J. Review of Fishery Forecasting Technology and Its Models. *J. Fish. China* **2013**, *37*, 1270–1279. [CrossRef]
8. Chen, X. An Approach to the Relationship between the Squid Fishing Ground and Water Temperature in the Northwestern Pacific. *J. Shanghai Fish. Univ.* **1995**, *4*, 181–185.
9. Cui, X.; Zhou, W.; Tang, F.; Dai, Y.; Zhang, S.; Jin, S. The Construction of Habitat Suitability Index Forecast Model of *Ommastrephes bartramii* Fishing Ground Based on Constrained Linear Regression. *Prog. Fish. Sci.* **2018**, *39*, 64–72. [CrossRef]
10. Fan, W.; Chen, X.; Shen, X. Tuna Fishing Grounds Prediction Model Based on Bayes Probability. *J. Fish. Sci. China* **2006**, *13*, 426–431.
11. Yang, X.; Dai, X.; Tian, S.; Xu, L. Forecasting Fishing Grounds for Tuna Purse Seine Fisheries in the Western and Central Pacific Ocean. *J. Fish. Sci. China* **2016**, *23*, 1417–1425.
12. Zhou, S.; Fan, W.; Wu, J. Prediction of Probable Tuna Fishing Grounds Based on Bayesian Theorem. *Artif. Intell. Comput. Intell. Int. Conf.* **2009**, *4*, 156–162. [CrossRef]
13. Mao, J.; Chen, X.; Yu, J. Forecasting Fishing Ground of *Thunnus alalunga* Based on BP Neural Network in the South Pacific Ocean. *Haiyang Xuebao* **2016**, *38*, 34–43.
14. Wang, P.; Fan, E.; Wang, P. Comparative Analysis of Image Classification Algorithms Based on Traditional Machine Learning and Deep Learning. *Pattern Recognit. Lett.* **2021**, *141*, 61–67. [CrossRef]
15. Xu, J.; Chen, X.; Yang, M. Forecasting on Fishing Ground of Red Flying Squid (*Ommastrephes bartramii*) in the North Pacific Ocean Based on Artificial Neural Net. *J. Shanghai Ocean Univ.* **2013**, *22*, 432–438.
16. Zhu, H.; Wu, Y.; Tang, F.; Jin, S.; Pei, K.; Cui, X. Construction of Fishing Ground Forecast Model of *Ommastrephes bartramii* Using Convolutional Neural Network in the Northwest Pacific. *Trans. Chin. Soc. Agric. Eng.* **2020**, *36*, 153–160. [CrossRef]

17. Eyinade, W.; Ezeilo, O.J.; Ogundeji, I.A. Deep Learning vs. Traditional Machine Learning in Financial Market Predictions. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.* **2024**, *10*, 379–406. [CrossRef]
18. Gao, Y.; Sha, Z.; Zhang, C.; Qiao, H.; Tang, R.; Li, D.; Wang, C. Intelligent feeding control of fish shoal using image texture features and decision tree. *Trans. Chin. Soc. Agric. Eng.* **2025**, *41*, 183–193. [CrossRef]
19. Zhao, Y.; Zou, X.; He, Y. A Boosted Regression Tree Based Forecast Model for Walleye Pollock (*Gadus chalcogrammus*) Fishing Grounds in the Western Bering Sea. *J. Dalian Fish. Univ.* **2025**, *40*, 472–480. [CrossRef]
20. Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum Entropy Modeling of Species Geographic Distributions. *Ecol. Model.* **2006**, *190*, 231–259. [CrossRef]
21. Yan, D.; Liu, G.; Hou, Z.; Kang, D.; Yang, S.; Lan, X. Using Two Ecological Niche Models to Predict the Potential Risk of Epizootic Situation in the Foci of *Meriones unguiculatus* Plague. *Chin. J. Vector Biol. Control* **2020**, *31*, 12–15. [CrossRef]
22. Moreno, R.; Zamora, R.; Molina, J.R.; Vasquez, A.; Herrera, M.Á. Predictive Modeling of Microhabitats for Endemic Birds in South Chilean Temperate Forests Using Maximum Entropy (Maxent). *Ecol. Inform.* **2011**, *6*, 364–370. [CrossRef]
23. Wang, J.; Tabeta, S. MaxEnt Modeling to Show Patterns of Coastal Habitats of Reef-Associated Fish in the South and East China Seas. *Front. Ecol. Evol.* **2023**, *11*, 1027614. [CrossRef]
24. Zhang, X.; Shi, Y.; Li, F.; Zhu, M.; Wei, Z. Prediction of Potential Fishing Ground for Pacific Saury (*Cololabis saira*) Based on MAXENT Model. *J. Shanghai Ocean. Univ.* **2020**, *29*, 280–286. [CrossRef]
25. Gao, F.; Chen, X.; Guan, W.; Li, G. A New Model to Forecast Fishing Ground of *Scomber japonicus* in the Yellow Sea and East China Sea. *Acta Oceanol. Sin.* **2016**, *35*, 74–81. [CrossRef]
26. Lee, D.; Son, S.; Kim, W.; Park, J.M.; Joo, H.; Lee, S.H. Spatio-Temporal Variability of the Habitat Suitability Index for Chub Mackerel (*Scomber japonicus*) in the East/Japan Sea and the South Sea of South Korea. *Remote Sens.* **2018**, *10*, 938. [CrossRef]
27. Li, X.; Pang, Z.; Zhu, J.; Ying, Y.; Sun, S. Spatial-Temporal Patterns in Fishing Ground of *Scomber japonicus* in Center Eastern Central Atlantic Ocean. *Fish. Sci.* **2018**, *37*, 31–37. [CrossRef]
28. Wang, L.; Li, Y.; Zhang, R.; Tian, Y.; Zhang, J.; Lin, L. Relationship between the Resource Distribution of *Scomber japonicus* and Seawater Temperature Vertical Structure of Northwestern Pacific Ocean. *Period. Ocean Univ. China* **2019**, *49*, 29–38. [CrossRef]
29. Guan, W.; Chen, X.; Gao, F.; Li, G. Environmental Effects on Fishing Efficiency of *Scomber japonicus* for Chinese Large Lighting Purse Seine Fishery in the Yellow and East China Seas. *J. Fish. Sci. China* **2009**, *16*, 949–958.
30. Fan, X.; Tang, F.; Cui, X.; Yang, S.; Zhu, W.; Huang, L. Habitat Suitability Index for Chub Mackerel (*Scomber japonicus*) in the Northwest Pacific Ocean. *Haiyang Xuebao* **2020**, *42*, 34–43.
31. Li, G.; Chen, X. Study on the Relationship between Catch of Mackerel and Environmental Factors in the East China Sea in Summer. *J. Mar. Sci.* **2009**, *27*, 1–8.
32. Miao, Z. Relation between Pneumatophorus and Carangidae Fishing Grounds in the Summer-Autumn and Ocean Hydrologic Environment in the Northern Part of the East China Sea. *J. Zhejiang Coll. Fish.* **1993**, *12*, 32–39.
33. Yatsu, A.; Watanabe, T.; Ishida, M.; Sugisaki, H.; Jacobson, L.D. Environmental Effects on Recruitment and Productivity of Japanese Sardine *Sardinops melanostictus* and Chub Mackerel *Scomber japonicus* with Recommendations for Management. *Fish. Oceanogr.* **2005**, *14*, 263–278. [CrossRef]
34. Pi, Q.; Hu, J. Analysis of Sea Surface Temperature Fronts in the Taiwan Strait and Its Adjacent Area Using an Advanced Edge Detection Method. *Sci. China Earth Sci.* **2010**, *53*, 1008–1016. [CrossRef]
35. Owiredu, S.A.; Onyango, S.O.; Song, E.-A.; Kim, K.-I.; Kim, B.-Y.; Lee, K.-H. Enhancing Chub Mackerel Catch Per Unit Effort (CPUE) Standardization through High-Resolution Analysis of Korean Large Purse Seine Catch and Effort Using AIS Data. *Sustainability* **2024**, *16*, 1307. [CrossRef]
36. Tao, Y.; Yi, M.; Li, B.; Feng, B.; Lu, H.; Yan, Y. Comparative Analysis of CPUE of Different Fishing Types in the South China Sea Based on the Fishing Port Sampling Survey. *Prog. Fish. Sci.* **2019**, *40*, 1–10.
37. Zhu, G.; Liu, Z.; Xu, G.; Zhang, J.; Meng, T.; Huang, H.; Xu, Y. Spatial-Temporal and Environmental Effects of Catch Rate on Antarctic Krill Fishery in the South Georgia Island in the Austral Winter Season Based on the Fine Scale Data. *Chin. J. Appl. Ecol.* **2014**, *25*, 2397–2404. [CrossRef]
38. Bordalo-Machado, P. Fishing Effort Analysis and Its Potential to Evaluate Stock Size. *Rev. Fish. Sci.* **2006**, *14*, 369–393. [CrossRef]
39. Tian, S.; Chen, X.; Chen, Y.; Xu, L.; Dai, X. Evaluating Habitat Suitability Indices Derived from CPUE and Fishing Effort Data for *Ommatrephes bratramii* in the Northwestern Pacific Ocean. *Fish. Res.* **2009**, *95*, 181–188. [CrossRef]
40. Tang, F.; Ba, Y.; Zhang, S.; Yang, S.; Zhao, G.; Wu, Z.; Li, J.; Cui, X. Prediction of Potential Fishing Grounds for Chub Mackerel in the Northwest Pacific Utilizing a Combination of Multivariable Gaussian Mixture Model and Bayesian Approach. *Reg. Stud. Mar. Sci.* **2025**, *85*, 104139. [CrossRef]
41. Breiman, L.; Friedman, J.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Chapman and Hall/CRC: New York, NY, USA, 1984; ISBN 978-0-412-04841-8.
42. Cui, X.-S.; Wu, Y.-M.; Zhang, J.; Zhou, S.-F.; Fan, W. Fishing Ground Forecasting of Chilean Jack Mackerel (*Trachurus murphyi*) in the Southeast Pacific Ocean Based on CART Decision Tree. *Period. Ocean Univ. China* **2012**, *42*, 53–59.

43. De'ath, G.; Fabricius, K.E. Classification and Regression Trees: A Powerful Yet Simple Technique for Ecological Data Analysis. *Ecology* **2000**, *81*, 3178–3192. [CrossRef]
44. Song, L.-M.; Ren, S.-Y.; Hong, Y.-R.; Zhang, T.-J.; Sui, H.-S.; Li, B.; Zhang, M. Comparison on Fishing Ground Forecast Models of *Thunnus alalunga* in the Tropical Waters of Atlantic Ocean. *Oceanol. Limnol. Sin.* **2022**, *53*, 496–504. [CrossRef]
45. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
46. Breiman, L. Bagging Predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
47. Anderson, R.P. Maximum Entropy (Maxent) Modeling of Species Geographic Distributions Based on Presence-Only Occurrence Data. *Ann. N. Y. Acad. Sci.* **2012**, *1260*, 66–80. [CrossRef] [PubMed]
48. Jaynes, E.T. *Clearing Up Mysteries—The Original Goal*; Springer: Dordrecht, The Netherlands, 1989.
49. Nursan, M.; Yonvitner, Y.; Agus, S. Distribution of Skipjack (*Katsuwonus pelamis*) Fishing Areas Using Purse Seine Fishing Equipment in WPP 573. *J. Trop. Fish. Manag.* **2022**, *6*, 1–12. [CrossRef]
50. Robbins, H.; Monro, S. A Stochastic Approximation Method. *Ann. Math. Stat.* **1951**, *22*, 400–407. [CrossRef]
51. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
52. Youden, W.J. Index for Rating Diagnostic Tests. *Cancer* **1950**, *3*, 32–35. [CrossRef]
53. Yuan, M.; Li, P.; Wu, C. Semiparametric Inference of the Youden Index and the Optimal Cut-off Point under Density Ratio Models. *Can. J. Stat.* **2021**, *49*, 965–986. [CrossRef]
54. Schoener, T.W. The Anolis Lizards of Bimini: Resource Partitioning in a Complex Fauna. *Ecology* **1968**, *49*, 704–726. [CrossRef]
55. Watanabe, C.; Yatsu, A. Effects of Density-Dependence and Sea Surface Temperature on Interannual Variation in Length-at-Age of Chub Mackerel (*Scomber japonicus*) in the Kuroshio-Oyashio Area during 1970–1997. *Fish. Bull.* **2004**, *102*, 196–206.
56. Watanabe, C.; Yatsu, A. Long-Term Changes in Maturity at Age of Chub Mackerel (*Scomber japonicus*) in Relation to Population Declines in the Waters off Northeastern Japan. *Fish. Res.* **2006**, *78*, 323–332. [CrossRef]
57. Li, Y.; Chen, X.; Guo, A.; Zhou, W. Comparison of habitat suitability index model for Scomber japonicus in different spatial and temporal scales. *J. Fish. China* **2019**, *43*, 935–945. [CrossRef]
58. Li, G.; Chen, X.; Tian, S. CPUE Standardization of Chub Mackerel (*Scomber japonicus*) for Chinese Large Lighting-Purse Seine Fishery in the East China Sea and Yellow Sea. *J. Fish. China* **2009**, *33*, 1050–1059.
59. Smith, R.C.; Dustan, P.; Au, D.; Baker, K.S.; Dunlap, E.A. Distribution of Cetaceans and Sea-Surface Chlorophyll Concentrations in the California Current. *Mar. Biol.* **1986**, *91*, 385–402. [CrossRef]
60. Nakata, K.; Hada, A.; Matsukawa, Y. Variation in Food Abundance for Japanese Sardine Larvae Related to the Kuroshio Meander. *Fish. Oceanogr.* **2007**, *3*, 39–49. [CrossRef]
61. Ji, Z.; Xing, Q.; Yu, W. Analysis of Spatio-Temporal Distribution of *Scomber japonicus*and Its Relationship with Marine Environmental Factors in Northwest Pacific. *South China Fish. Sci.* **2025**, *21*, 22–30. [CrossRef]
62. Song, L.; Xu, H.; Chen, M.; Ebango Ngando, N. Relationship between Spatiotemporal Distribution of Chub Mackerel and Marine Environment Variables in the Waters near Mauritania. *J. Shanghai Ocean Univ.* **2020**, *29*, 868–877.
63. Elith, J.; Phillips, S.J.; Hastie, T.; Dudík, M.; Chee, Y.E.; Yates, C.J. A Statistical Explanation of MaxEnt for Ecologists. *Divers. Distrib.* **2011**, *17*, 43–57. [CrossRef]
64. Galparsoro, I.; Pouso, S.; García-Barón, I.; Mugerza, E.; Mateo, M.; Paradinas, I.; Louzao, M.; Borja, Á.; Mandiola, G.; Murillas, A. Predicting Important Fishing Grounds for the Small-Scale Fishery, Based on Automatic Identification System Records, Catches, and Environmental Data. *ICES J. Mar. Sci.* **2024**, *81*, 453–469. [CrossRef]
65. Su, S.; Mao, Q.; Li, Y.; Li, H.; Leng, J.; Lu, C. Deep Learning-Based Fishing Ground Prediction for Albacore and Yellowfin Tuna in the Western and Central Pacific Ocean. *Fish. Res.* **2024**, *278*, 107103. [CrossRef]
66. Diao, L.; Yi, G.Y. Classification Trees with Mismeasured Responses. *J. Classif.* **2023**, *40*, 168–191. [CrossRef]
67. Murphy, K.P. *Probabilistic Machine Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 2022.
68. Moore, C.; Drazen, J.C.; Radford, B.T.; Kelley, C.; Newman, S.J. Improving Essential Fish Habitat Designation to Support Sustainable Ecosystem-Based Fisheries Management. *Mar. Policy* **2016**, *69*, 32–41. [CrossRef]
69. Pittman, S.J.; Brown, K.A. Multi-Scale Approach for Predicting Fish Species Distributions across Coral Reef Seascapes. *PLoS ONE* **2011**, *6*, e20583. [CrossRef]
70. Dudík, M.; Phillips, S.J.; Schapire, R.E. Maximum Entropy Density Estimation with Generalized Regularization and an Application to Species Distribution Modeling. *J. Mach. Learn. Res.* **2007**, *8*, 1217–1260.
71. Mazzoni, S. *Distribution Modelling by MaxEnt: From Black Box to Flexible Toolbox*; University of Olso: Oslo, Norway, 2016.
72. Phillips, S.J.; Anderson, R.P.; Dudík, M.; Schapire, R.E.; Blair, M.E. Opening the Black Box: An Open-Source Release of Maxent. *Ecography* **2017**, *40*, 887–893. [CrossRef]
73. Steen, P.J.; Zorn, T.G.; Seelbach, P.W.; Schaeffer, J.S. Classification Tree Models for Predicting Distributions of Michigan Stream Fish from Landscape Variables. *Trans. Am. Fish. Soc.* **2008**, *137*, 976–996. [CrossRef]

74. Yu, W.; Dai, X.; Yang, Y.; Wan, R.; Pu, Y.; Yao, X. The Relationship between Water-Level Fluctuation Factors and the Distribution of Carex in Floodplain Grassland around Poyang Lake. *J. Lake Sci.* **2018**, *30*, 1672–1680. [CrossRef]
75. Lamas, L.; Oliveira, P.B.; Pinto, J.P.; Almeida, S.; Deus, R.; Silva, A.J.d.; Almeida, N. Fishing Areas Characterisation Using the SIMOcean Platform. *Aquat. Living Resour.* **2017**, *30*, 19. [CrossRef]
76. Elhilbawi, H.; Eldawlatly, S.; Mahdi, H. A Taxonomy of Discretization Techniques Based on Class Labels and Attributes' Relationship. In *Proceedings of the 2019 14th International Conference on Computer Engineering and Systems (ICCES)*; IEEE: New York, NY, USA, 2019; pp. 316–321.
77. Jin, R.; Breitbart, Y.; Muoh, C. Data Discretization Unification. In *Proceedings of the Seventh IEEE International Conference on Data Mining (ICDM 2007)*; IEEE: Piscataway, NJ, USA, 2007; pp. 183–192. [CrossRef]
78. Clarkson, E.; Kupinski, M. Quantifying the Loss of Information from Binning List-Mode Data. *J. Opt. Soc. Am. A* **2020**, *37*, 450–457. [CrossRef]
79. García, S.; Luengo, J.; Sáez, J.A.; López, V.; Herrera, F. A Survey of Discretization Techniques: Taxonomy and Empirical Analysis in Supervised Learning. *IEEE Trans. Knowl. Data Eng.* **2013**, *25*, 734–750. [CrossRef]
80. Hernández-Orallo, J.; Flach, P.; Ferri, C. A Unified View of Performance Metrics: Translating Threshold Choice into Expected Classification Loss. *J. Mach. Learn. Res.* **2012**, *13*, 2813–2869.
81. Rainio, O.; Teuho, J.; Klén, R. Evaluation Metrics and Statistical Tests for Machine Learning. *Sci. Rep.* **2024**, *14*, 6086. [CrossRef]
82. Karp, M.A.; Brodie, S.; Smith, J.A.; Richerson, K.; Selden, R.L.; Liu, O.R.; Muhling, B.A.; Samhouri, J.F.; Barnett, L.A.K.; Hazen, E.L.; et al. Projecting Species Distributions Using Fishery-Dependent Data. *Fish Fish.* **2023**, *24*, 71–92. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.