Journal of
Marine Science and Engineering

MDPI

Article

Relationship between Resource Distribution and Vertical Structure of Water Temperature of Purpleback Flying Squid (*Sthenoteuthis oualaniensis*) in the Northwest Indian Ocean Based on GAM and GBT Models

Chen Shang [1,2], Haibin Han [1,2], Junlin Chen [2,3], Fenghua Tang [2], Wei Fan [2], Heng Zhang [2] and Xuesen Cui [2,*]

[1] College of Marine Sciences, Shanghai Ocean University, Shanghai 201306, China; shangchen1931996@163.com (C.S.); hhbfishery@163.com (H.H.)
[2] Key Laboratory of Fisheries Remote Sensing, Ministry of Agriculture and Rural Affairs, East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai 200090, China; chenjunlin383@gmail.com (J.C.); f-h-tang@163.com (F.T.); fanw@ecsf.ac.cn (W.F.); zhangziqian0601@163.com (H.Z.)
[3] School of Navigation and Naval Architecture, Dalian University of Ocean, Dalian 116023, China
* Correspondence: cuixuesen@eastfishery.ac.cn

Abstract: The Northwest Indian Ocean is a key fishing ground for China's pelagic fisheries, with the purpleback flying squid being a significant target. This study uses commercial fishing logs of the Indian Ocean between 2015 and 2021, alongside pelagic seawater temperature and its vertical temperature difference within the 0–200 m depth range, to construct generalized additive models (GAMs) and gradient boosting tree models (GBTs). These two models are evaluated using cross-validation to assess their ability to predict the distribution of purpleback flying squid. The findings show that factors like year, latitude, longitude, and month significantly influence the distribution of purpleback flying squid, while surface water temperature, 200 m water temperature, and the 150–200 m water layer temperature difference also play a role in the GBT model. Similar factors also take effects in the GAM. Comparing the two models, both GAM and GBT align with reality in predicting purpleback flying squid resource distribution, but the precision indices of GBT model outperform those of the GAM. The predicted distribution for 2021 by GBT also has a higher overlap with the actual fishing ground than that by GAM, indicating GBT's superior forecasting ability for the purpleback flying squid fishing ground in the Northwest Indian Ocean.

Keywords: generalized additive model; Northwest Indian Ocean; gradient boosted trees; *Sthenoteuthis oualaniensis*; fishing ground prediction

check for updates

Citation: Shang, C.; Han, H.; Chen, J.; Tang, F.; Fan, W.; Zhang, H.; Cui, X. Relationship between Resource Distribution and Vertical Structure of Water Temperature of Purpleback Flying Squid (*Sthenoteuthis oualaniensis*) in the Northwest Indian Ocean Based on GAM and GBT Models. *J. Mar. Sci. Eng.* **2023**, *11*, 1800. https://doi.org/10.3390/jmse11091800

Academic Editor: Francesco Tiralongo

Received: 2 August 2023
Revised: 26 August 2023
Accepted: 8 September 2023
Published: 15 September 2023

1. Introduction

The purpleback flying squid (*Sthenoteuthis oualaniensis*) belongs to the class Cephalopoda, order Teuthida, and family Ommastrephidae [1], a warm-water oceanic species that predominantly resides in the tropical and subtropical waters of the Indian and Pacific Oceans [2]. This species is notable for its broad distribution and significant economic value. Characterized by a strong swimming ability, rapid generational turnover, and a brief life cycle, the marine environment significantly influences the purpleback flying squid's resource distribution [3]. As an active predator, it plays a crucial role in tropical marine ecosystems, often preying on small fish and shellfish, with cannibalistic behavior also commonly observed [2,4]. Furthermore, this squid serves as prey for predatory fish and tropical seabirds [5,6], indicating its potential role in sustaining marine ecosystem levels [7].

With the ongoing growth of global fisheries and the increasing demands for commercial and food security purposes, the number of cephalopod species entering commercial fisheries has continued to grow [8]. The purpleback flying squid, once disregarded in commercial fishing, has now emerged as a pivotal cephalopod resource in the South China Sea [9], where China has been engaging in fishing activities targeting squid using light falling gear since 2010 [10]. Previous surveys indicate a substantial resource reserve of the purpleback flying squid, with 1–2 million tons in the South China Sea [11,12], 3–4 million tons in the Indian Ocean [13] and 5–7 million tons in the Pacific Ocean. China initiated fishing activities in the high seas of the Indian Ocean using light falling gear in 2014, establishing a fleet of six vessels by 2016 [14]. Current research on light falling gear fishing centers primarily on the South China Sea. In the Indian Ocean, however, there is a lack of reports on this species because of its short fishing history. The Northwest Indian Ocean has extensive upwelling because of the influence of countercurrent and monsoon currents [15] and is identified as a high-density distribution area for the purpleback flying squid [16,17].

The temperature of seawater and its spatial structure are important indicators of the upper-level fisheries resources and environmental changes in the ocean, which can significantly impact the distribution of marine biological resources [18–20]. Similarly, seawater temperature is an influencing factor for the activity and distribution of squid [21]. It is known that purpleback flying squid in the Northwest Indian Ocean show significant diel vertical migration that is influenced by the vertical structure of the water temperature, probably changing the catch efficiency. Therefore, clarifying the relationship between the squid's distribution and the water temperature will contribute to the sustainable development and management of the fishery.

Because of the intricate interplay between marine organisms and the environment, many factors can influence species distribution [22]. Linear regression, owing to its ease of use and interpretability, is a common approach for investigating the impact of environmental factors on species patterns. However, the relationship between the environment and species can sometimes be challenging to ascertain as purely linear, necessitating models that allow for nonlinear effects [23]. Hence, models accommodating nonlinear effects might be better suited for exploring the intricate relationship between marine organisms and their environment [24]. The generalized additive model (GAM) has been widely applied in studying the relationship between species distribution and environmental factors [25,26]. This model allows for the use of nonparametric smoothing functions to simulate the nonlinear relationship between response variables and environmental factors. It has also been used to explore the impact of surface marine environmental elements on the distribution of purpleback flying squid resources. Zhang et al. [27] and Yan et al. [28] utilized the GAM to analyze the purpleback flying squid resources in the Northwest Indian Ocean and the South China Sea, respectively, both concluding that longitude, latitude, and temperature significantly influence the distribution of purpleback flying squid.

The GAM has the characteristic of being sensitive to extreme values [29]. However, when the model is used for extrapolation, unrealistic inferences can be generated [30]. gradient boosted trees (GBTs) is a model that sequentially fits multiple individual decision trees and aggregates the predicted results. Each additional tree adapts to the residuals of the previous tree [31]. The GBT model is insensitive to outliers, extreme values, and missing values in the data [32]. Some scholars have applied the GBT model in fisheries research [33,34], but there have not been any reports on its application in the distribution study of cephalopod resources.

Considering the scant previous research on the distribution of purpleback flying squid resources in the Northwest Indian Ocean, where most studies have primarily focused on surface environmental factors [18,35,36], while less attention has been given to subsurface seawater temperature influence, this study aims to utilize temperature values from pelagic water layers and vertical temperature differences in the Northwest Indian Ocean. By integrating Chinese commercial fishing logs data of purpleback flying squid, GBT and GAM can be constructed. The training accuracy of the two models are compared to explore their potential in analyzing the spatiotemporal variations of the fishing grounds and predicting their habitats. At the same time, the influence of the surface and subsurface seawater temperatures on the fishing grounds are investigated to provide a theoretical basis and more methodological choices for the resource management of purpleback flying squid in the Indian Ocean.

2. Materials and Methods

2.1. Data Sources

2.1.1. Fishing Ground Areas and Fishery Data

The study area of this research covers the Northwest Indian Ocean, with a time span from November 2015 to November 2021, and a spatial range of 10°–22° N and 55°–70° E. The distribution of fish catches of purpleback flying squid in the Northwest Indian Ocean during 2015–2021 was illustrated in Figure 1.

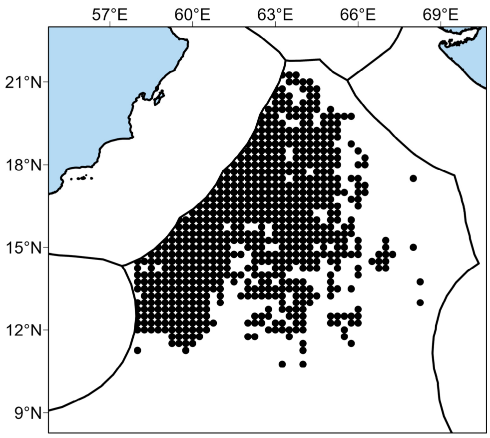

Figure 1. Distribution of fish catches of purpleback flying squid in the Northwest Indian Ocean during 2015–2021.

The fisheries data of light falling gear [14] in the Northwest Indian Ocean for this study were sourced from the fishing log of China high seas commercial fishing vessels. Because of the strong Southwest Monsoon in the summer, there is a fishing moratorium from June to August, and therefore, there is no production data for squid during these 3 months. The dataset consists of 20,258 operation records, providing information such as operation time, latitude and longitude coordinates of each operation's starting and ending points, and the catch quantity. The fishing areas were statistically analyzed using a $0.25° \times 0.25°$ latitude–longitude grid. The catch data within each fishing grid were aggregated into three-day intervals to calculate the catch per unit effort (CPUE, t/net), as shown in Equation (1):

$$\mathrm{CPUE} = \mathrm{Catch}/\mathrm{N} \quad (1)$$

Catch represents the total catch(t) within a fishing grid over a three-day interval, and N represents the total number(net) of fishing operations in that grid over the same period.

2.1.2. Selection and Processing of Environmental Data

Purpleback flying squid primarily inhabit waters above the 200 m layer [5,37], and the yield is closely associated with surface water temperature, 50 m water temperature, and 200 m water temperature [38]. Therefore, this study limits the vertical depth to within 200 m. In light of this, the study downloaded environmental data from 2015 to 2021 from the Copernicus Marine Service (https://resources.marine.copernicus.eu/products (accessed on 25 April 2023)), which are reanalysis values based on the NEMO model. The data have a daily time resolution and a spatial resolution of 0.25°. The temperatures at 0 m (T_0), 50 m (T_{50}), 100 m (T_{100}), 150 m (T_{150}), and 200 m (T_{200}) were extracted from this dataset. Moreover, because the temperature difference between different water layers may affect the distribution of purpleback flying squid [39], this study also calculated the vertical temperature differences between each adjacent 50 m layer from the top to 200 m depth as independent factors, denoted as $\Delta T_{0–50}$, $\Delta T_{50–100}$, $\Delta T_{100–150}$, and $\Delta T_{150–200}$.

2.2. GAM and GBT Models

The GAM is a combination of generalized linear models and additive models, where the components are smooth functions. Nonlinear relationships can be established by adding smooth terms [40,41]. The GAM can directly handle the nonlinear relationship between the response variable and multiple explanatory variables. It utilizes nonparametric methods to detect the underlying data structure and identify patterns to obtain predictive results. The formula for the GAM constructed in this study is as follows:

$$\log(\text{CPUE}+1) \sim s(\text{year}) + s(\text{month}) + s(\text{lon}) + s(\text{lat}) + \sum_{t\in T} s(\text{T}) + \sum_{\Delta t\in \Delta T} s(\Delta\text{T}) \quad (2)$$

Because of the skewed distribution of CPUE, a logarithmic function is used to correct the bias. Additionally, to avoid zero values, a constant of 1 is added to the CPUE in this study. In the equation, s() represents the spline smoothing function, which captures nonlinear effects and improves the model's predictive accuracy. Lon and lat represent longitude and latitude, respectively. T denotes the collection of water temperature variables for different water layers (T_0, T_{50}, T_{100}, T_{150}, and T_{200}), and ΔT represents the collection of temperature differences at 50 m intervals ($\Delta T_{0–50}$, $\Delta T_{50–100}$, $\Delta T_{100–150}$ and $\Delta T_{150–200}$). The GAM was fitted using the mgcv package (version 1.8.41) in R software version 4.2.2 [42].

In this study, the GAM was constructed by progressively incorporating environmental factors. Akaike's information criterion (AIC) was used to test the model's goodness after the new addition of explanatory variables. The smaller the value of the AIC, the better the fit of the model was proved. The F-test was used to determine whether the effect of the explanatory variables on the response variables was significant.

The GBT model is an ensemble learning algorithm based on decision trees. It combines multiple iteratively trained weak classifiers to form a robust classifier [43]. This algorithm does not assume an additive relationship between explanatory variables and the response factor, allowing for high-order interactions among factors [44]. Based on the previous round's results, the GBT model adjusts the weights of samples during the iterations. It emphasizes the prediction of misclassified samples in the next round of iteration, leading to higher accuracy and greater robustness.

The GBT model in this study was constructed using the sklearn package (version 1.2.1) of Python 3.8. The optimal combination of hyperparameters of the model can be obtained automatically based on different datasets and problems, reducing the time and cost of manual trial. Moreover, it can also decrease errors and improve efficiency and accuracy during model training. In this study, the GBT model utilized nontransformed CPUE values. This study configured four essential hyperparameters: (1) Number of decision trees determines how many decision trees are combined in the model. (2) Maximum depth of the decision tree refers to how detailed the tree can be. (3) Learning rate controls the step size in adjusting the model. Lower rates make convergence gradual for better generalization, but need more iterations. (4) L1-regularization penalty coefficient adds a penalty for using fewer features, helping prevent overfitting and improve model simplicity. The parameter configuration of the model is shown in Table 1.

Table 1. Hyperparameters configuration for GBT models.

Hyperparameter	Values
Number of decision trees	200, 300, 400
Maximum depth of the decision tree	7, 9, 11, 13
Learning rate	0.01, 0.05, 0.1
L1-Regularization penalty coefficient	0, 0.1, 0.2

Finally, this study used Python GridSearchCV function to train, evaluate, and automatically optimize the GBT model. In the process of machine learning, the configuration of hyperparameters significantly influences the model's performance [45]. GridSearchCV is a method used to determine the optimal values for model hyperparameters and set the number of cross-validation iterations for each set of hyperparameters. All combinations are tested to determine the optimal result [46,47]. During the training process, 5-fold cross-validation was employed, with mean squared error (MSE) as the evaluation metric for each combination [48].

2.3. Validation of the Two Models and Analysis of Factor Contributions

Cross-validation was employed in this study to assess the forecasting performance of the two models. The GAM and GBT model were trained using 70% of the samples for the models' construction through random selection. The remaining 30% of the data were reserved for evaluating the performance of the models as a test set [49]. This process was repeated 100 times, and the average results were calculated. Regression analysis [38,50], mean squared error (MSE), and coefficient of determination (R^2) were employed as 3 evaluation indicators for the models.

The explained deviance represented the factor contribution of the GAM. It is determined by the difference in explained deviance between a particular independent factor added and removed [23,51]. The feature importance in the GBT model was calculated, and the average value was obtained as the final factor contribution after performing 100 iterations of 5-fold cross-validation.

2.4. Prediction of the Spatial Distribution of Fishing Grounds

This study utilized well-trained GAM and GBT models to predict monthly spatial distribution with spatiotemporal and environmental data in 2021. The predicted CPUE values were presented at $0.25° \times 0.25°$ grids at different spatial locations. A comparison was conducted between the two models to evaluate their respective abilities and drawbacks in predicting the distribution of catch quantities by their prediction results.

3. Results

3.1. Performance Comparison of GAM and GBT Models

The 100 rounds of cross-validation results showed that the GBT model has a better fit than the GAM (Table 2). The MSE of the GBT model was 0.45, much smaller than that of the GAM, which was 1.09. The slope and the coefficient of determination of the GBT model were also closer to 1, and the intercept was closer to 0.

Table 2. Performance comparison of the 2 models.

Model	MSE	R^2	Intercept	Slope
GAM	1.09 ± 0.0476	0.81 ± 1.284	0.43 ± 0.005	0.74 ± 0.03
GBT	0.45 ± 0.0442	0.88 ± 0.002	0.30 ± 0.008	0.86 ± 0.02

The values following the $\pm$ symbol represent variance.

3.2. Factor Importance in GAM and GBT Models

3.2.1. GAM

From the Accumulation of deviance explained and the AIC value, the optimal GAM had a total explanation deviation rate of 80.4% for CPUE (Table 3). As shown in Table 4, among the spatiotemporal factors, the most significant contribution to the explanatory rate was the year (37.6%), followed by the month (22.9%), and the importance of latitude and longitude was 9% and 6.4%, respectively. Among the environmental factors, T_0 contributed the most (1.4%). T_{150} and $\Delta T_{0–50}$ contributed 0.6% and 0.5%, respectively. T_{50}, T_{100}, $\Delta T_{100–150}$, and $\Delta T_{150–200}$ all contributed 0.4%, while T_{200} and $\Delta T_{150–200}$ contributed 0.2%.

Table 3. GAM statistical results.

Formula	AIC	Accumulation of Deviance Explained/%	Determination Coefficient	*p*-Value
log(CPUE+1)~s(year)	6031.437	37.6	0.375	<0.0001
log(CPUE+1)~s(year)+s(month)	4225.134	60.5	0.604	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)	3245.926	69.5	0.692	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)+s(lat)	2348.830	75.9	0.755	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)+s(lat)+s(T_0)	2143.497	77.3	0.768	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)+s(lat)+s(T_0)+ s(T_{50})	2082.958	77.7	0.772	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)+s(lat)+s(T_0)+ s(T_{50})+s(T_{100})	2004.079	78.1	0.777	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)+s(lat)+s(T_0)+ s(T_{50})+s(T_{100})+s(T_{150})	1911.056	78.7	0.780	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)+s(lat)+s(T_0)+ s(T_{50})+s(T_{100})+s(T_{150})+s(T_{200})	1874.310	78.9	0.785	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)+s(lat)+s(T_0)+ s(T_{50})+s(T_{100})+s(T_{150})+s(T_{200})+s($\Delta T_{0\text{–}50}$)	1800.722	79.4	0.789	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)+s(lat)+s(T_0)+ s(T_{50})+s(T_{100})+s(T_{150})+s(T_{200})+s($\Delta T_{0\text{–}50}$)+s($\Delta T_{50\text{–}100}$)	1777.817	79.6	0.790	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)+s(lat)+s(T_0)+ s(T_{50})+s(T_{100})+s(T_{150})+s(T_{200})+s($\Delta T_{0\text{–}50}$)+s($\Delta T_{50\text{–}100}$)+ s($\Delta T_{100\text{–}150}$)	1716.990	80.0	0.794	<0.0001
log(CPUE+1)~s(year)+s(month)+s(lon)+s(lat)+s(T_0)+ s(T_{50})+s(T_{100})+s(T_{150})+s(T_{200})+s($\Delta T_{0\text{–}50}$)+s($\Delta T_{50\text{–}100}$)+ s($\Delta T_{100\text{–}150}$)+s($\Delta T_{150\text{–}200}$)	1668.537	80.4	0.797	<0.0001

Table 4. Deviance explained of the different factors in the optimal GAM.

Predictor Factor	Deviance Explained/%
Year	37.6%
Month	22.9%
Longitude	9.0%
Latitude	6.4%
T_0	1.4%
T_{150}	0.6%
$\Delta T_{0\text{–}50}$	0.5%
T_{50}	0.4%
T_{100}	0.4%
$\Delta T_{100\text{–}150}$	0.4%
$\Delta T_{150\text{–}200}$	0.4%
T_{200}	0.2%
$\Delta T_{50\text{–}100}$	0.2%

In terms of spatiotemporal factors in the GAM (Figure 2), it showed that the effect of the year factor on CPUE increased year by year from 2015 to 2019, and then began to decline until 2021. The highest effect of the month factor on CPUE was December and the second highest month was April; the optimal periods were February–April and September–December, and at these two stages the confidence interval of the factor was small, and its effect was significant. The effect of longitude on CPUE first showed a fluctuating upward trend and then gradually decreased after reaching the maximum value at 63° E. The effect of latitude on CPUE showed an upward trend and then fluctuated at a high level after 15.5° N, with a small confidence interval from 15.5° N to 16.25° N, and its effect was significant.

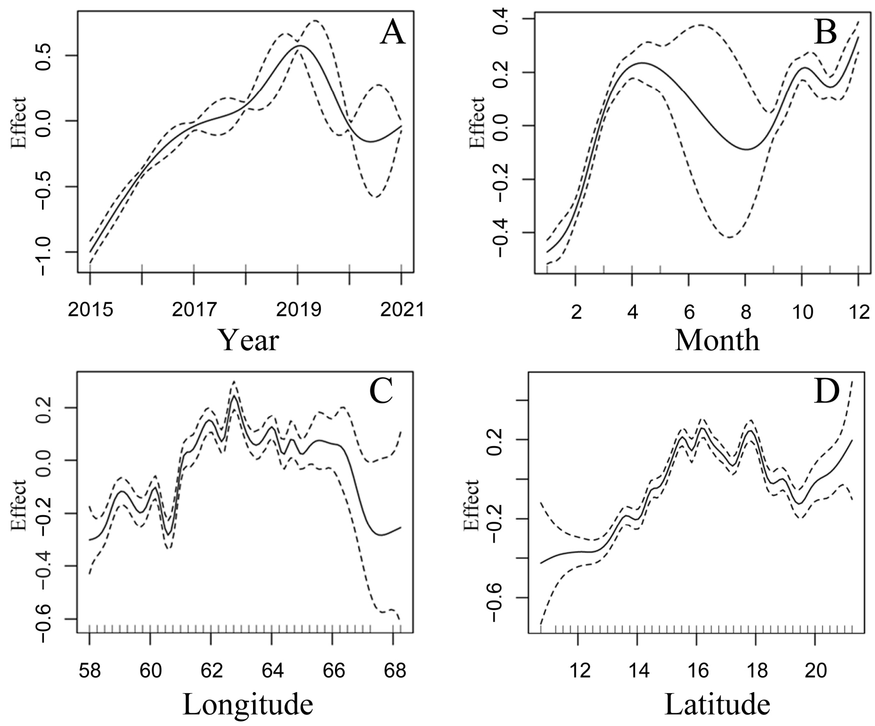

Figure 2. Analysis of the GAM results of the influence of spatiotemporal factors on CPUE: (**A**) year; (**B**) month; (**C**) longitude; (**D**) latitude. The solid line is the influence curve, and the 95% confidence interval is between the two dashed lines.

The results of the effect of water temperature on CPUE (Figure 3) showed that CPUE was roughly negatively correlated with T_0; CPUE reached the lowest value when T_0 was 31 °C; the suitable T_0 range was 24~28 °C; the confidence interval was mostly minor; and the confidence level was high at 26~28 °C, which indicates that this factor had a significant effect on CPUE. CPUE showed a slow decline with the increase of T_{50}, and it reached the lowest value when T_{50} was 26 °C, and then it increased with the increase in temperature. And then, with the increase in temperature, in the vicinity of a 26 °C, the confidence interval was small, and the appropriate T_{50} temperature was 20~23 °C and 26~28 °C. When the T_{100} was less than 22 °C, the CPUE increased with the water temperature, reaching the highest value when the T_{100} was 22 °C and showed a decreasing trend when it was more than 20 °C, and the optimal T_{100} range was 20~23 °C. In the 150 m water layer, when the water temperature was less than 20 °C, CPUE showed an increasing trend with the increase in the water temperature and reached the highest value when T_{150} was 20 °C. Then CPUE slowly decreased with the rise in temperature, and the suitable T_{150} range was 18~22 °C, and the confidence interval was the smallest when it is 19~20 °C, which is the closest temperature value. In the 200 m water layer, CPUE decreased with increasing temperature until the water temperature reached 18 °C, then showed an increasing trend.

In terms of the performance of the vertical temperature difference of the water layer (Figure 3), in the range of 0~50 m water layer, the CPUE increased with a greater temperature difference of 0–50 m (ΔT_{0-50}), but gradually decreased when greater than 5 °C of the temperature difference. In the 50~100 m pelagic range, CPUE generally increased with an increasing temperature difference, and the confidence interval was minimized when the temperature difference was 3.5 °C. In the 100~150 m aquatic range, CPUE increased and then decreased with an increasing temperature difference and reached the regional minimum at 2.5 °C, with a small confidence interval at 2~3.5 °C, which was closely affected. In the range of the 150~200 m water layer, CPUE firstly decreased and then increased with the increase in the temperature difference, and gradually decreased after reaching the regional maximum value at 2.8 °C, and then showed an increasing trend after reaching the regional minimum value at 3.4 °C, with a small and close influence in the confidence interval of 2.5~2.8 °C.

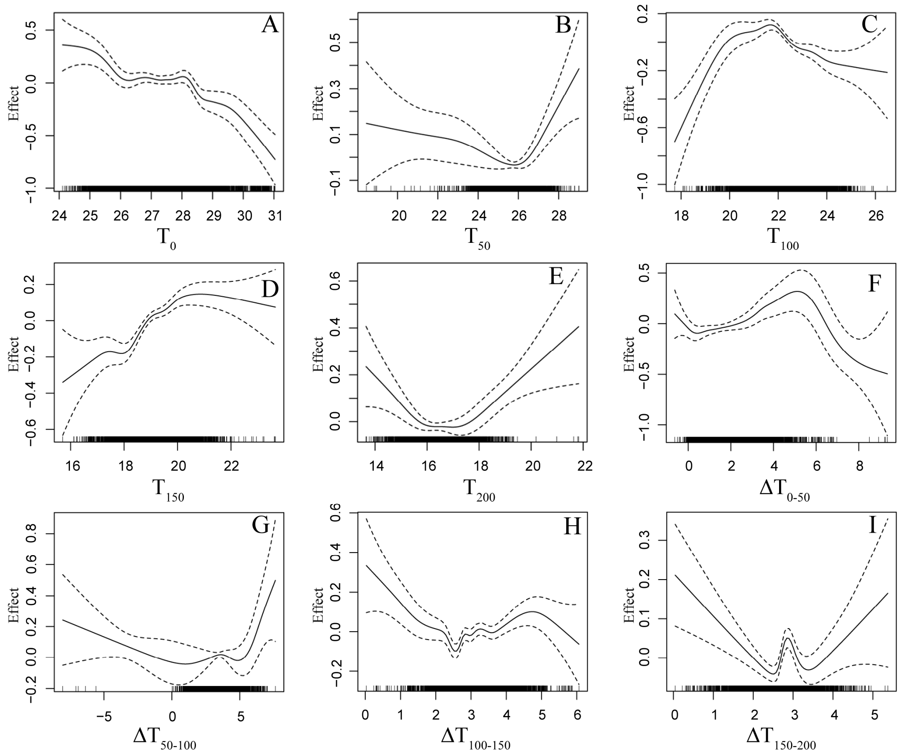

Figure 3. Analysis of the GAM results of the influence of environmental factors on CPUE: (**A**) T_0; (**B**) T_{50}; (**C**) T_{100}; (**D**) T_{150}; (**E**) T_{200}; (**F**) $\Delta T_{0–50}$; (**G**) $\Delta T_{50–100}$; (**H**) $\Delta T_{100–150}$; (**I**) $\Delta T_{150–200}$. The solid line is the influence curve, and the 95% confidence interval is between the two dashed lines.

3.2.2. GBT Model

The importance of the different factors in the GBT model is shown in Table 5, where the spatiotemporal factors are much more important than the environmental factors, and the most important of the spatiotemporal and environmental factors are year (22.68%) and T_0, respectively (5.62%).

Table 5. Mean feature importance of the GBT model.

Predictor Factor	Mean Value of Feature Importance	Standard Deviation of Feature Importance
Year	22.68%	0.008
Latitude	22.34%	0.143
Longitude	15.31%	0.126
Month	8.45%	0.004
T_0	5.62%	0.003
T_{200}	4.97%	0.004
$\Delta T_{150–200}$	4.32%	0.005
$\Delta T_{0–50}$	4.08%	0.003
$\Delta T_{100–150}$	3.98%	0.004
T_{50}	3.18%	0.002
T_{150}	2.51%	0.001
$\Delta T_{50–100}$	2.37%	0.001
T_{100}	2.20%	0.002

The partial dependency plot of the GBT model (Figure 4) shows that the effect of year on CPUE increased year by year and reached a peak in 2019, followed by a decreasing trend; the effect of month on CPUE reached a peak in November, and the optimal operating months are April and September to December.

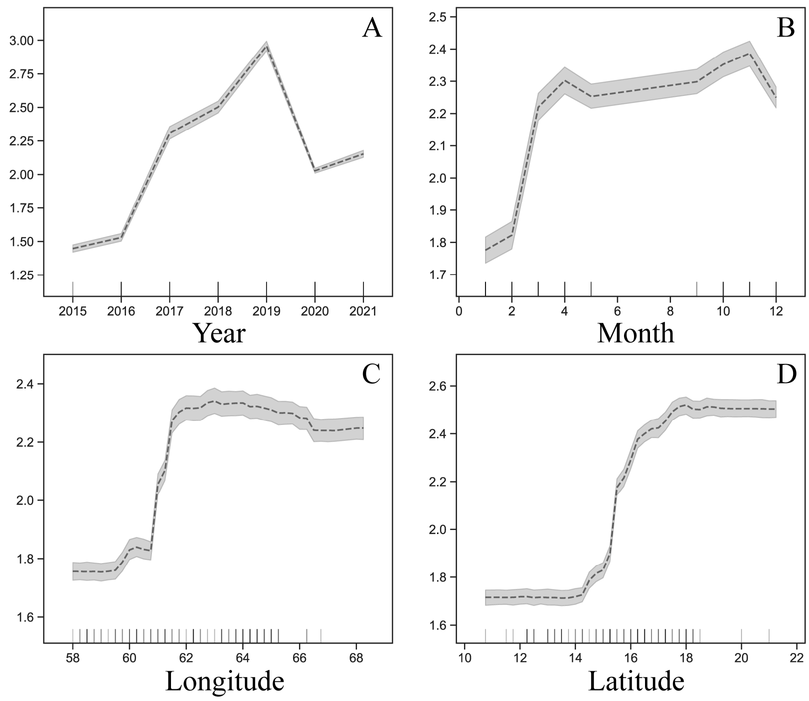

Figure 4. Analysis of the GBT model results of the influence of spatiotemporal factors on CPUE: (**A**) year; (**B**) month; (**C**) longitude; (**D**) latitude. The black curve is the influence curve of factor on CPUE, and the grey area is the 95% confidence interval.

Among the environmental factors (Figure 5), CPUE was the lowest at T_0 at 26 °C, and the optimum operating temperature was 27~28.5 °C. The T_{50} temperature was the optimum operating temperature at 24.2~25 °C, and then CPUE decreased with the increase of temperature, and after the temperature exceeded 26 °C, CPUE was generally at its lowest state. The optimum T_{100} range was 21.3~22 °C, and at a depth of 150 m water temperature, CPUE was the lowest at 18 °C, after which the CPUE response gradually increased with the increase of temperature, and after the temperature reached 19 °C, the CPUE and temperature curves flattened out, and the suitable operating temperature was 19~21 °C. In the depth of 200 m water temperature, CPUE started from 14 °C, gradually increased with the increase of temperature, and reached the highest value around 17.8 °C, and then CPUE fluctuated little with the change of temperature, and the optimal operating temperature was 17.8~18 °C. In the vertical gradient of water temperature, in the range of 0–50 m water layer, CPUE showed an increasing trend when greater temperature difference, reached the highest value at 3 °C, and then slowly decreased. In the 50~100 m range, CPUE showed a negative correlation with the temperature difference. In the 100–150 m pelagic range, the effect of temperature difference on CPUE showed a fluctuating trend, decreasing with the increase in the difference in the range of 2~2.6 °C and the lowest at 2.6 °C, then showing an increasing trend, reaching a high value at 3.2 °C, and then showing a decreasing trend in general. In the 150–200 m water layer range, CPUE showed a general decreasing trend with an increasing $\Delta T_{150–200}$, and a regional high value appeared at 2.7 °C.

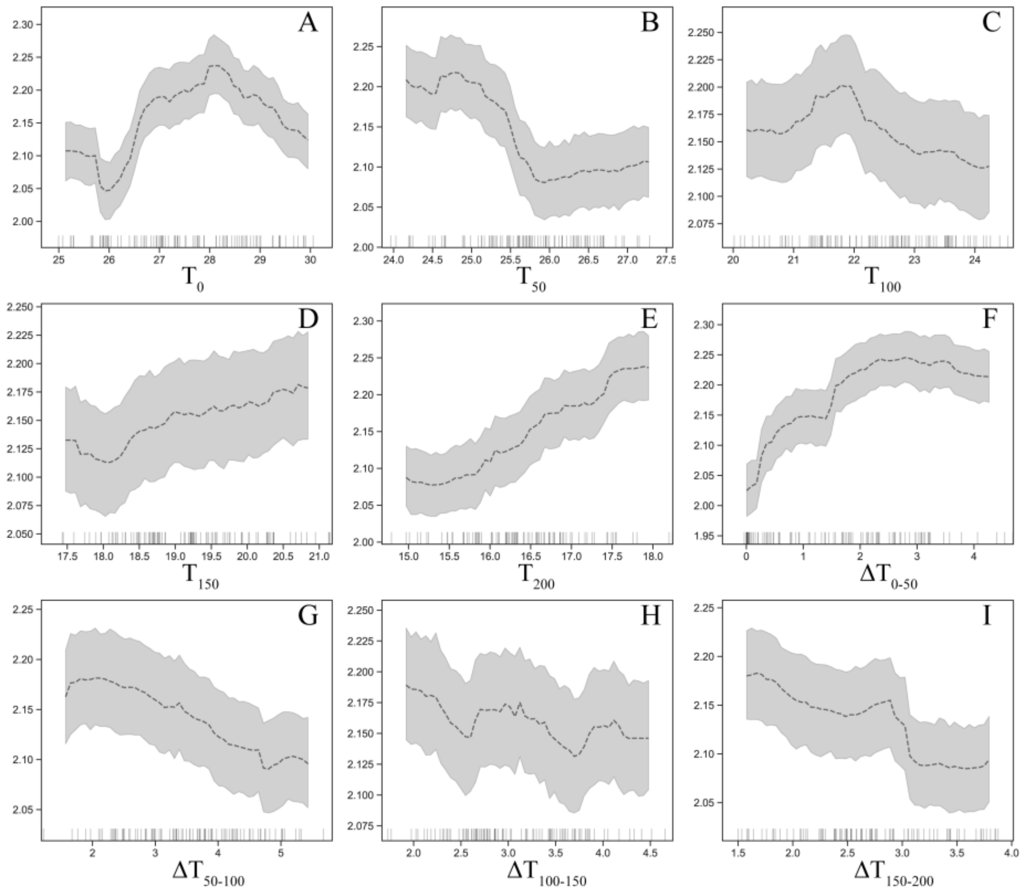

Figure 5. Analysis of the GBT model results of the influence of environmental factors on CPUE: (**A**) T_0; (**B**) T_{50}; (**C**) T_{100}; (**D**) T_{150}; (**E**) T_{200}; (**F**) $\Delta T_{0–50}$; (**G**) $\Delta T_{50–100}$; (**H**) $\Delta T_{100–150}$; (**I**) $\Delta T_{150–200}$. The black curve is the influence curve of factor on CPUE, and the grey area is the 95% confidence interval.

3.3. CPUE Prediction of the 2 Models

The GBT and GAMs were used to predict the spatial distribution of CPUE of the Northwest Indian purpleback flying squid, respectively, and the data used in the forecast were different pelagic temperature, pelagic gradient, and latitude/longitude data for the proposed forecast months in 2021.

To test the predictive effect of the models, the actual CPUE values from January to May and September to November 2021 were spatially superimposed and compared with the predicted values in this study.

The CPUE predictions of the two models are shown in Figures 6 and 7. It can be seen that the CPUE distribution of the purpleback flying squid is more concentrated, with more obvious seasonally variations in distribution. Regarding spatial distribution, the predicted CPUE generally showed high in the north and low in the south of Northwest Indian Ocean. The high production area of the purpleback flying squid predicted by the GBT model was roughly consistent with the actual fishing area. In the residual table comparing the predicted values against the observed values for different months using both models (Table 6), it is evident that the GBT model exhibits smaller residuals overall. Furthermore, the GBT model demonstrates a lower variance of residuals, substantiating its enhanced stability.

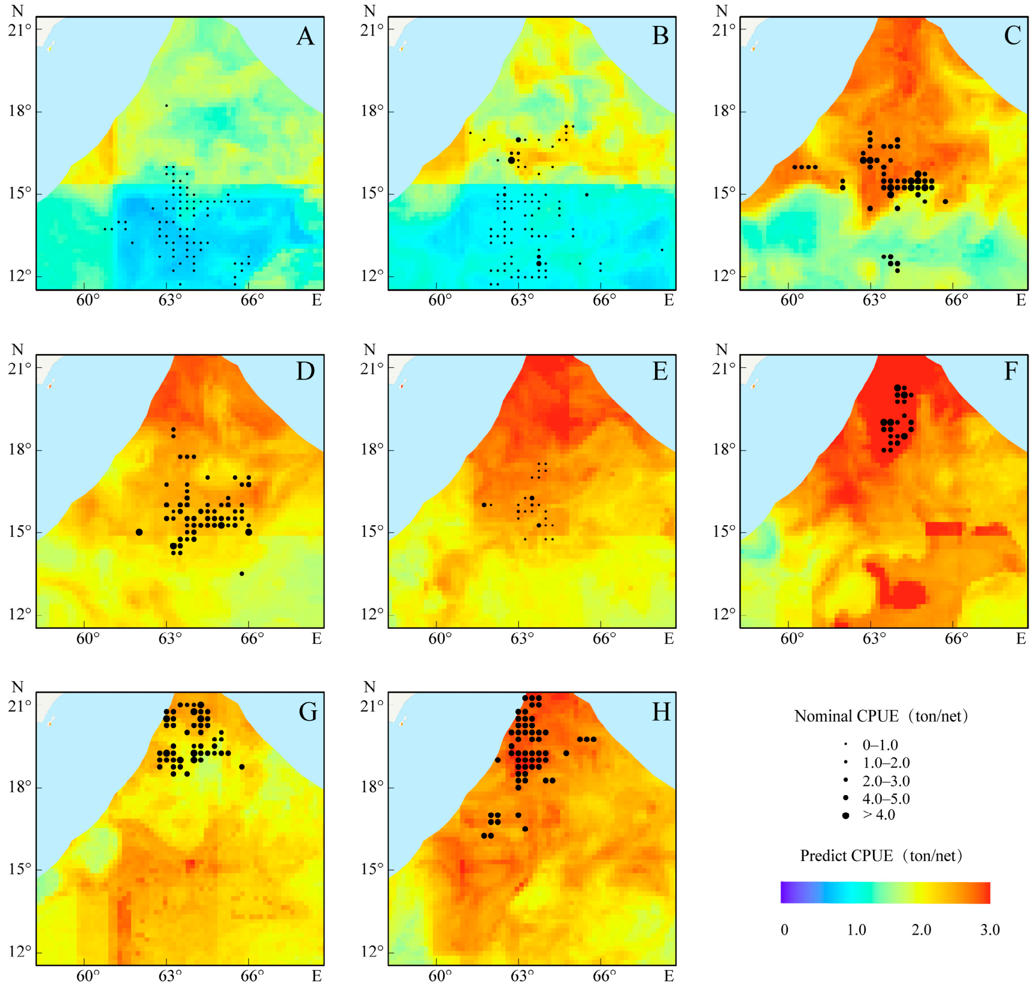

Figure 6. The monthly observed CPUE and forecast results distribution of purpleback flying squid based on GBT model in the Northwest Indian Ocean in 2021: (**A**) January; (**B**) February; (**C**) March; (**D**) April; (**E**) May; (**F**) September; (**G**) October; (**H**) November.

Table 6. Residuals of GBT and GAMs for Predicted CPUE and Nominal CPUE Across Different Months.

Month	GBT Model	GAM
1	0.82 ± 0.081	0.62 ± 0.199
2	0.65 ± 0.325	0.57 ± 0.502
3	−0.84 ± 0.310	−1.32 ± 0.551
4	0.41 ± 0.556	−0.74 ± 0.141
5	1.21 ± 0.161	1.58 ± 0.620
9	−0.38 ± 0.439	−0.73 ± 0.840
10	−0.65 ± 0.479	−0.59 ± 1.757
11	−0.69 ± 0.114	−1.31 ± 0.504

The values following the ± symbol represent variance.

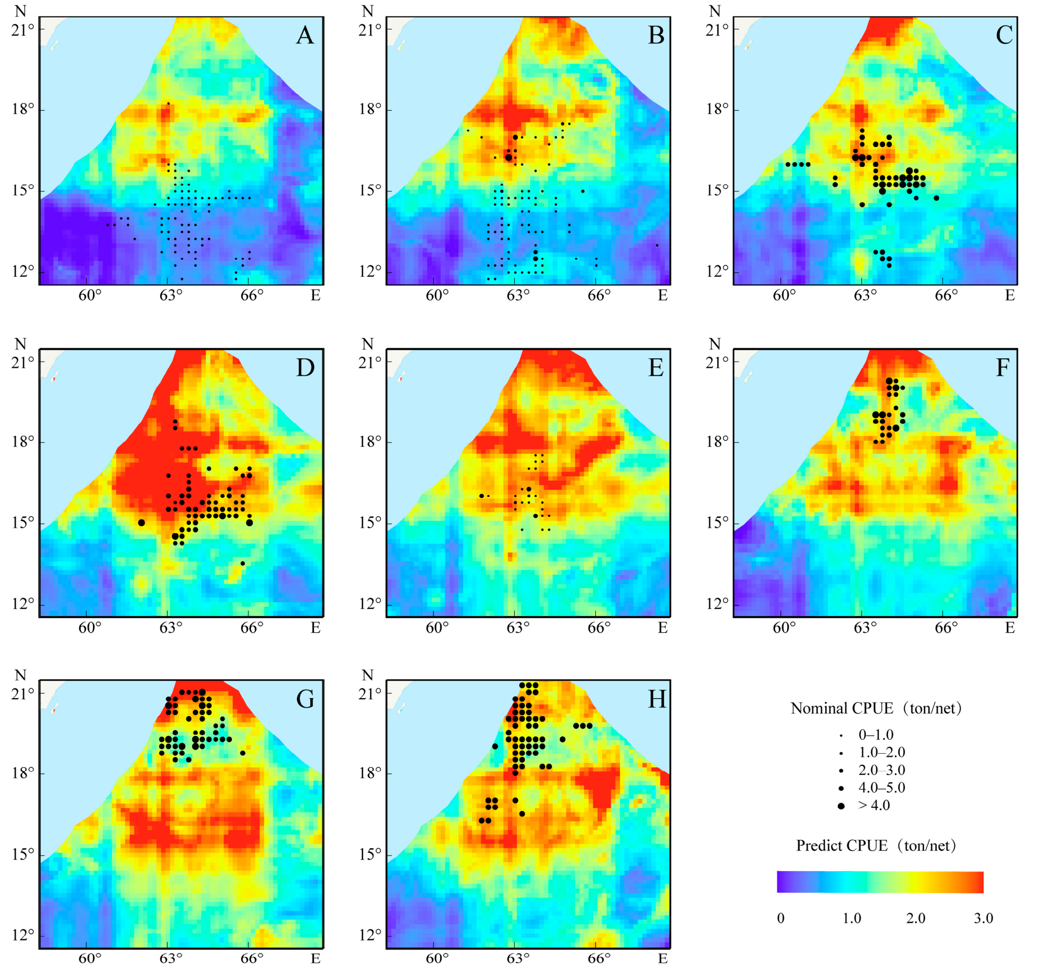

Figure 7. The monthly observed CPUE and forecast results distribution of purpleback flying squid based on GAM in the Northwest Indian Ocean in 2021: (**A**) January; (**B**) February; (**C**) March; (**D**) April; (**E**) May; (**F**) September; (**G**) October; (**H**) November.

4. Discussion

4.1. Contribution of Different Factors in Models

In both the GAM and GBT models, spatiotemporal factors consistently rank highly regarding variance explanation and relative contribution. The results shown by the GAM and GBT models in the time factors are relatively consistent. The purpleback flying squid is a species that spawns throughout the year [52] and is influenced by continuous temporal change [53]. Therefore, considering this ongoing temporal variation is necessary to comprehensively understand its impact on the biomass dynamics of the purpleback flying squid. The year and month factors showed significant performances in predicting the CPUE of squid in the models, with high contribution rates. This indicates that time changes significantly influence the biomass of purpleback flying squid resources. The analysis results of two models on the effect of spatiotemporal factors on CPUE (Figures 6 and 7) show that the purpleback flying squid biomass has been increasing by years from 2015 to 2019. And after reaching the peak biomass in 2019, it starts declining. A similar trend was also shown in studies of Chen et al. [54] and Wen et al. [55]. April and September to December are suitable for purpleback flying squid fishing, with the month effect reaching the highest value in December and the highest CPUE response value in the fourth quarter. This is consistent with the study by Wei et al. [56]. The spatial factors have a high contribution rate to the prediction performance of squid CPUE, indicating that the squid fishing ground has a strong spatial correlation. The GAM result map shows that the area between

14.5° N–18.5° N and 61.5° E–64.5° E has a significant impact on CPUE, which is consistent with the studies by Xiao et al. [18] and Zhang et al. [57], indicating that the purpleback flying squid has spatial aggregation characteristics to some extent. This phenomenon may be attributed to the seasonal activities of monsoons and ocean currents in this maritime region. The entire marine environment is influenced by the cyclical changes brought about by monsoons, subsequently impacting the distribution of purpleback flying squid [58]. Under the influence of monsoons, the movement of ocean currents gives rise to extensive upwelling zones in the area. Within these upwelling zones, abundant nutrients foster the aggregation of purpleback flying squid, creating productive fishing grounds [17].

Concerning environmental factors, both the GBT and GAM underscore the significance of SST. The GAM identifies an optimal range of 24~28 °C, while the GBT model suggests an optimal range of 27~28.5 °C. Compared to other cephalopods, purpleback flying squid display a pronounced adaptability to variations in SST [41]. Zhang et al. [27] propose that the optimal SST range for Indian Ocean squid is 25~29.0 °C, while Yu et al. [19] suggest an optimal fishing ground SST range for squid of 27.0~29.0 °C. These findings align well with the results of this study. Within the GAM, the temperature difference in the 0–50 m water layer significantly contributes to the model's variance, with peak catch values observed at a temperature differential of 4–6 °C. Yan et al. [59], using grey relational analysis [60], a factor relationship analysis method suitable for small samples, found that the 5–50 m temperature gradient had the most substantial impact on purpleback flying squid CPUE in the South China Sea. This finding diverges from the results of our study. The variance in outcomes could be attributed to two principal factors. Firstly, it is worth noting that the data utilized by Yan et al. [59], encompassed survey data from 2012–2013 when the purpleback flying squid fishery in the South China Sea was still in its developmental phase, and the survey duration was relatively brief. Subsequent investigations, such as those conducted by Li et al. [61], incorporated more extensive datasets spanning 2016–2017 and 2019. Li et al. [61] emphasized that latitude and longitude emerged as the most critical factors influencing purpleback flying squid distribution, with sea surface temperature (SST) and the temperature difference between depth layers (ΔT_{0-50}) being of relative importance.

Additionally, a divergence in the findings may be attributable to regional distinctions between the South China Sea and the Indian Ocean. Monsoons and ocean currents stand as pivotal factors influencing marine organism distribution [55]. While the South China Sea is less influenced by the southwest monsoon and more affected by the northeast monsoon [62], and both of these monsoons exert substantial influence over the Northwest Indian Ocean [63]. These climatic forces impact ocean currents, consequently shaping primary productivity and subsequently influencing the distribution of the purpleback flying squid [17].

The water temperature of 100–200 m and the vertical temperature gradient of the water layers showed high importance in both models. This could be attributed to the presence of a thermocline within the 100–150 m water layer in the Northwest Indian Ocean [20], which influences the distribution of fish inhabiting the upper-middle layers [21]. Meanwhile, the Arabian Sea exhibits an oxygen minimum zone at depths ranging from 100 to 200 m [64]. In the Northwest Indian Ocean, extensive oxygen consumption occurs due to surface algae's proliferation and relatively short life cycles. This consumption happens as the algae sink to the seabed, forming a prominent anoxic layer at a depth of 100 m in the Arabian Sea [65,66]. The upwelling of deep anoxic seawater, influenced by surface winds driving the movement of surface seawater, generates a cold anoxic region characterized by high productivity. This region is juxtaposed with adjacent warm oxygenated water masses, contributing significantly to establishing productive fishing grounds [67]. Zuyev et al. [13] noted that purpleback flying squid possess the ability to utilize proteins and their decomposition products for anaerobic energy metabolism in anaerobic conditions. Purpleback flying squid is abundant in hypoxic zone with low temperature, displaying such abundance during both diurnal and nocturnal activity periods. Other environmental factors have a relatively minor contribution, exhibiting a limited ability to enhance the model fitting precision.

4.2. Analysis of the Spatiotemporal Distribution Characteristics of Predicted Fishing Grounds

Upon comparing the predicted CPUE of both the GBT model and GAM with the nominal CPUE, it becomes apparent that these two models exhibit distinct differences in their spatial performance. The GBT model demonstrates a higher degree of consistency with the spatial distribution of nominal CPUE than the GAM. The deviation of both only occurs in specific months, such as January and February. In these months, the actual fishing operations showed a lower CPUE, with only a handful of scattered fishing zones surpassing one ton/net of CPUE. Thus, these deviations in model predictions may be attributed to the scarcity of high CPUE data during the actual fishing process.

The distribution of the predicted CPUE across fishing grounds reveals a state of higher values in the northern region and lower ones in the southern part. Within the area spanning 15°–19° N and 61°–65° E, the CPUE of purpleback flying squid is notably higher, corroborating previous research findings [68]. During October and November, the fishing grounds appear more dispersed, with a tendency to move southward. Regions with high-value production, located at 13° N–17° N, 60° E–64° E and 18° N–21° N, 61° E–64° E, generally align with conclusions that Shao et al. [69] and Chen et al. [54] drew. The observed southward shift in purpleback flying squid distribution can be attributed to the influence of the winter monsoon, characterized by a counterclockwise flow in the southwest direction [70]. This monsoon generates a substantial movement of ocean currents towards the south. This dynamic process prompts the upwelling of nutrient-rich deep waters to the surface, thereby elevating primary productivity levels. The influx of these nutrient-rich waters effectively enhances the availability of food resources within the ocean's upper layers [17]. As a result, purpleback flying squid are drawn towards these productive zones in search of abundant foraging opportunities.

In the predictive maps generated by the GBT model, a distinct boundary phenomenon is evident during certain months, i.e., January and February, along specific latitude and longitude coordinates. This boundary remains relatively stable, situated around 15° N and 61° E. This phenomenon can be partly attributed to the inherent characteristics of the tree model, which formulates IF-ELSE rules by creating numerous branches for each explanatory factor. Hence, the splitting strategy of tree model is fundamentally based on the sample features [71]. In this study, latitude and longitude significantly influenced the results of the model. They caused the GBT model to prioritize these features during the splitting process, resulting in the horizontal and vertical boundaries in the resultant tree model. Concurrently, Han et al. [14] found that, in the Arabian Sea, CPUE increases at higher latitude, exhibiting a trend of initial decrease followed by an increase, with the lowest CPUE value observed around 15°25′ N. Between 15°25′ N and 16°25′ N, CPUE drastically rises with an increasing latitude.

Similarly, between 61°25′ E and 61°75′ E, CPUE also rose fast with an increasing longitude. Studies by Chen et al. [20] and Lin et al. [68] identified that the primary fishing grounds for Northwest Indian Ocean purpleback flying squid are predominantly situated near 15°–16° N, 60°–62° E. The results of Chen et al. [54], Wen et al. [57], and Zhang et al. [27] also found that the CPUE of purpleback flying squid increased in this latitude and longitude range. The distinct boundaries of the latitude and longitude in the GBT model's predictive results also reflect, to some degree, the spatial distribution patterns of the actual fishing grounds.

4.3. Comparative Analysis of the Prediction Performance of Two Models

A comparison between the predicted CPUE results of the GAM and the actual ones reveals significant discrepancies in the graphical representations of the GBT model and GAM (Figures 6 and 7). The GAM's predictions delineate clear high-value and low-value production zones, with an overrepresentation of high-value zones that contradicts the actual production scenario. This discrepancy arises because the GAM is susceptible to extremes, and extrapolation beyond the operational area can yield unrealistic results [30]. Conversely, the prediction results of GBT model are more evenly distributed due to its lower sensitivity to outliers [32], contributing to its robustness. However, this trait may also cause the model to overlook potential high-value production zones.

From the results in Table 6, when the GBT model and GAM were employed to forecast CPUE of Indian Ocean purpleback flying squid, the GBT model demonstrated superior fitness and predictive accuracy. Consequently, the GBT model is more feasible than the GAM for predicting Indian Ocean purpleback flying squid fishing grounds. Furthermore, the predicted results of GBT model align more closely with the actual production outcomes.

5. Conclusions

The purpleback flying squid in the Northwest Indian Ocean have strong fluctuations, and the environment greatly affects their resource distribution. This study used the purpleback flying squid as the research object, constructed a GBT model and an optimal GAM, and predicted the fishing ground in 2021. Comparing the prediction results of the GBT model and GAM 100 times using five-fold cross-validation, all of the model evaluation indicators of the GBT model were better than those of the GAM, indicating that the GBT has good predictive performance. The spatial distributions predicted by the GBT model from January to May and September to November in 2021 were consistent with the observed values, proving the model's potential in predicting the distribution of purpleback flying squid resources in the Northwest Indian Ocean.

Currently, no regional fishery organization is responsible for overseeing this fishery, and established measures for conservation, management, monitoring, control, and surveillance are lacking. Taking these circumstances into account, our study's objective is to employ modeling techniques to identify potential fishing grounds for purpleback flying squid. Through the differentiation of areas with high and low catch rates, curtail energy expenditure during fishing activities and alleviate the threat of overfishing. This strategy has been formulated to guarantee the population's sustainability while also offering a foundation for subsequent evaluations of the purpleback flying squid's resources.

This study is only an analysis of the vertical structure of the water layer in the Northwest Indian Ocean, so the model has some shortcomings. In fact, the distribution of squid resources is influenced by temperature and factors such as chlorophyll concentration, dissolved oxygen, and sea surface height [21,72]. Dietary factors, represented by the presence and distribution of zooplankton, also play a significant role, as squids seek out optimal feeding environments [73,74]. Therefore, fluctuations in primary productivity can elucidate variations in squid fishing grounds. Large-scale environmental shifts, exemplified by events like the El Niño phenomenon can further induce changes in purpleback flying squid resources [75–77]. Consequently, future research should consider the integration of additional environmental factors.

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

Funding: This research was funded by Laoshan Laboratory (No. LSKJ202201804).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The authors declare that the data supporting the findings of this study are available within the article or are available from the corresponding authors upon request.

Acknowledgments: The authors wish to express sincere gratitude to all editors and reviewers.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chembian, A.J.; Mathew, S. Growth and Mortality of the Oceanic Squid *Sthenoteuthis oualaniensis* (Lesson, 1830) off South-West Coast of India. *Indian J. Fish.* **2016**, *63*, 27–34. [CrossRef]
2. Lu, H.; Chen, Z.; Liu, K.; Ou, Y.; Zhao, M.; Sun, T. Statolith Microstructure Estimates of the Age, Growth, and Population Structure of Purpleback Flying Squid (*Sthenoteuthis oualaniensis*) in the Waters of the Xisha Islands of the South China Sea. *Fishes* **2022**, *7*, 234. [CrossRef]
3. Zhao, C. Biology and Habitat Characteristics of the Purpleback Flying Squid *Sthenoteuthis oualaniensis* in the South China Sea. Ph.D. Thesis, Jimei University, Xiamen, China, 2021.
4. Chen, Z.; Lu, H.; Liu, W.; Liu, K.; Chen, X. Beak Microstructure Estimates of the Age, Growth, and Population Structure of Purpleback Flying Squid (*Sthenoteuthis oualaniensis*) in the Xisha Islands Waters of the South China Sea. *Fishes* **2022**, *7*, 187. [CrossRef]
5. Jaquemet, S.; Potier, M.; Cherel, Y.; Kojadinovic, J.; Bustamante, P.; Richard, P.; Catry, T.; Ramos, J.A.; Le Corre, M. Comparative Foraging Ecology and Ecological Niche of a Superabundant Tropical Seabird: The Sooty Tern Sterna Fuscata in the Southwest Indian Ocean. *Mar. Biol.* **2008**, *155*, 505–520. [CrossRef]
6. Ménard, F.; Potier, M.; Romanov, E.; Jaquemet, S.; Sabatié, R.; Cherel, Y. New Information from Predator Diets on the Importance of Two Ommastrephidae: *Sthenoteuthis oualaniensis* in the Indian Ocean and Hyaloteuthis Pelagica in the Atlantic Ocean. *Global Ocean Ecosyst. Dyn.* **2007**, *24*, 49–52.
7. Chen, X.J.; Lu, H.J.; Liu, B.L.; Chen, Y.; Li, S.; Jin, M. Species Identification of *Ommastrephes bartramii*, *Dosidicus gigas*, *Sthenoteuthis oualaniensis* and *Illex argentinus* (Ommastrephidae) Using Beak Morphological Variables. *Sci. Mar.* **2012**, *76*, 473–481. [CrossRef]
8. Ragesh, N. Biology and Life History of the Purpleback Squid *Sthenoteuthis oualaniensis* from the Arabian Sea. Ph.D. Thesis, ICAR-Central Marine Fisheries Research Institute, Kochi, India, 2021.
9. Jereb, P.; Roper, C.F. *Cephalopods of the World-an Annotated and Illustrated Catalogue of Cephalopod Species Known to Date. Vol 2. Myopsid and Oegopsid Squids*; FAO: Rome, Italy, 2010; ISBN 92-5-106720-1.
10. Fan, J.; Fang, Z.; Ma, S.; Zhang, P.; Chen, Z. Age, Growth, and Population Characteristics of *Sthenoteuthis oualaniensis* in the South China Sea. *Reg. Stud. Mar. Sci.* **2022**, *55*, 102517. [CrossRef]
11. Zhang, Y. *Fisheries Acoustic Studies on the Purpleback Flying Squid Resource in the South China Sea*; National Taiwan University Publication: Taiwan, China, 2005.
12. Zhang, J.; Guobao, C.; Peng, Z.; Zuozhi, C.; Jiangtao, F. Estimation of Purpleback Flying Squid (*Sthenoteuthis oualaniensis*) Re_source in the Central and Southern South China Sea Based on Fisheries Acoustics and Light-Falling Net. *J. Fish. Sci. China* **2014**, *21*, 822–831.
13. Zuyev, G.; Nigmatullin, C.; Chesalin, M.; Nesis, K. Main Results of Long-Term Worldwide Studies on Tropical Nektonic Oceanic Squid Genus Sthenoteuthis: An Overview of the Soviet Investigations. *Bull. Mar. Sci.* **2002**, *71*, 1019–1060.
14. Han, H.; Yang, C.; Zhang, H.; Fang, Z.; Jiang, B.; Su, B.; Sui, J.; Yan, Y.; Xiang, D. Environment Variables Affect CPUE and Spatial Distribution of Fishing Grounds on the Light Falling Gear Fishery in the Northwest Indian Ocean at Different Time Scales. *Front. Mar. Sci.* **2022**, *9*, 939334. [CrossRef]
15. Vinayachandran, P.N.M.; Masumoto, Y.; Roberts, M.J.; Huggett, J.A.; Halo, I.; Chatterjee, A.; Amol, P.; Gupta, G.V.; Singh, A.; Mukherjee, A. Reviews and Syntheses: Physical and Biogeochemical Processes Associated with Upwelling in the Indian Ocean. *Biogeosciences* **2021**, *18*, 5967–6029. [CrossRef]
16. Sajikumar, K.K.; Ragesh, N.; Venkatesan, V.; Koya, K.P.; Sasikumar, G.; Kripa, V.; Mohamed, K.S. Morphological Development and Distribution of Paralarvae Juveniles of Purple Back Flying Squid *Sthenoteuthis oualaniensis* (Ommastrephidae), in the South Eastern Arabian Sea. *Vie Milieu-Life Environ.* **2018**, *68*, 75–86. [CrossRef]
17. Lu, H.; Wang, H.; He, J.; Chen, X.; Liu, K.; Chen, X. The mechanism of influence of monsoon changes on the fisheries biology and oceanography of *Sthenoteuthis oualaniensis* in northwest Indian Ocean. *J. Fish. Sci. China* **2022**, *29*, 1669–1678.
18. Xiao, G.; Xu, B.; Zhang, H.; Tang, F.; Chen, F.; Zhu, W. A study on spatial-temporal distribution and marine environmental elements of *Symplectoteuthis oualaniensis* fishing grounds in outer sea of Arabian Sea. *S. China Fish. Sci.* **2022**, *18*, 10–19.
19. Yu, W.; Chen, X. Analysis on Habitat Suitability Index of *Sthenoteuthis oualaniensis* in Northwestern Indian Ocean from September to October. *J. Guangdong Ocean. Univ.* **2012**, *32*, 74–80.
20. Chen, X.; Ye, X. Preliminary study on the relationship between fishing ground of *Symlectoteuthis oualaniensis* and environmental factors in northwestern Indian Ocean. *J. Shanghai Fish. Univ.* **2005**, *14*, 55–60. [CrossRef]
21. Sánchez-Velasco, L.; Ruvalcaba-Aroche, E.D.; Beier, E.; Godínez, V.M.; Barton, E.D.; Díaz-Viloria, N.; Pacheco, M.R. Paralarvae of the Complex *Sthenoteuthis oualaniensis-Dosidicus gigas* (Cephalopoda: Ommastrephidae) in the Northern Limit of the Shallow Oxygen Minimum Zone of the Eastern Tropical Pacific Ocean (April 2012). *J. Geophys. Res. Ocean.* **2016**, *121*, 1998–2015. [CrossRef]
22. Ribeiro, J.; Carvalho, G.M.; Gonçalves, J.M.; Erzini, K. Fish Assemblages of Shallow Intertidal Habitats of the Ria Formosa Lagoon (South Portugal): Influence of Habitat and Season. *Mar. Ecol. Prog. Ser.* **2012**, *446*, 259–273. [CrossRef]
23. França, S.; Cabral, H.N. Predicting Fish Species Richness in Estuaries: Which Modelling Technique to Use? *Environ. Model. Softw.* **2015**, *66*, 17–26. [CrossRef]
24. Zhao, J.; Cao, J.; Tian, S.; Chen, Y.; Zhang, S.; Wang, Z.; Zhou, X. A Comparison between Two GAM Models in Quantifying Relationships of Environmental Variables with Fish Richness and Diversity Indices. *Aquat. Ecol.* **2014**, *48*, 297–312. [CrossRef]

25. Salazar, J.E.; Benavides, I.F.; Portilla Cabrera, C.V.; Guzmán, A.I.; Selvaraj, J.J. Generalized Additive Models with Delayed Effects and Spatial Autocorrelation Patterns to Improve the Spatiotemporal Prediction of the Skipjack (*Katsuwonus pelamis*) Distribution in the Colombian Pacific Ocean. *Reg. Stud. Mar. Sci.* **2021**, *45*, 101829. [CrossRef]
26. Stock, B.C.; Ward, E.J.; Eguchi, T.; Jannot, J.E.; Thorson, J.T.; Feist, B.E.; Semmens, B.X. Comparing Predictions of Fisheries Bycatch Using Multiple Spatiotemporal Species Distribution Model Frameworks. *Can. J. Fish. Aquat. Sci.* **2020**, *77*, 146–163. [CrossRef]
27. Zhang, B.; Lu, H.; Zhao, M.; Sun, T.; Guo, R. Standardization of Catch per Unit Effort(Cpue) in Northwest Indian Ocean *Sthenoteuthis oualaniensis* Based on Generalized Additive Model. *Oceanol. Limnol. Sin.* **2023**, *54*, 259–265.
28. Yan, L.; Li, J.; Zhang, P.; Yang, B.; Wang, T. Effects of spatiotemporal and environmental factors on the fishing ground of *Sthenoteuthis oualaniensis* in the South China Sea based on the Generalized Additive Model. *Mar. Sci. Bull.* **2021**, *40*, 217–223.
29. Georges, B.; Michez, A.; Latte, N.; Lejeune, P.; Brostaux, Y. Water Stream Heating Dynamics around Extreme Temperature Events: An Innovative Method Combining GAM and Differential Equations. *J. Hydrol.* **2021**, *601*, 126600. [CrossRef]
30. Elith, J.; Kearney, M.; Phillips, S. The Art of Modelling Range-shifting Species. *Methods Ecol. Evol.* **2010**, *1*, 330–342. [CrossRef]
31. Song, L.; Li, T.; Zhang, T.; Sui, H.; Li, B.; Zhang, M. Comparison of Machine Learning Models within Different Spatial Resolutions for Predicting the Bigeye Tuna Fishing Grounds in Tropical Waters of the Atlantic Ocean. *Fish. Oceanogr.* 2023, *early view*. [CrossRef]
32. Elith, J.; Leathwick, J.R.; Hastie, T. A Working Guide to Boosted Regression Trees. *J. Anim. Ecol.* **2008**, *77*, 802–813. [CrossRef]
33. Lyashevska, O.; Harma, C.; Minto, C.; Clarke, M.; Brophy, D. Long-Term Trends in Herring Growth Primarily Linked to Temperature by Gradient Boosting Regression Trees. *Ecol. Inform.* **2020**, *60*, 101154. [CrossRef]
34. Shirk, P.L.; Richerson, K.; Banks, M.; Tuttle, V. Predicting Bycatch of Chinook Salmon in the Pacific Hake Fishery Using Spatiotemporal Models. *ICES J. Mar. Sci.* **2023**, *80*, 133–144. [CrossRef]
35. Mohamed, K.S.; Sajikumar, K.K.; Ragesh, N.; Ambrose, T.V.; Jayasankar, J.; Said Koya, K.P.; Sasikumar, G. Relating Abundance of Purpleback Flying Squid *Sthenoteuthis oualaniensis* (Cephalopoda: Ommastrephidae) to Environmental Parameters Using GIS and GAM in South-Eastern Arabian Sea. *J. Nat. Hist.* **2018**, *52*, 1869–1882. [CrossRef]
36. Zhao, C.; Shen, C.; Bakun, A.; Yan, Y.; Kang, B. Purpleback Flying Squid *Sthenoteuthis oualaniensis* in the South China Sea: Growth, Resources and Association with the Environment. *Water* **2020**, *13*, 65. [CrossRef]
37. Tian, S.; Chen, X.; Yang, X. Study on the fishing ground distribution of *Symlectoteuthis oualaniensis* and its relationship with the environmental factors in the high sea of the northern arabian sea. *Trans. Oceanol. Limnol.* **2006**, *1*, 51–57. [CrossRef]
38. Fan, X.; Cui, X.; Tang, F.; Fan, W.; Wu, Y.; Zhang, H. Research on the prediction model of spatial distribution of *Sthenoteuthis oualaniensis* in the open sen Arabian Sea based on PCA-GAM. *J. Fish. China* **2022**, *46*, 2340–2348.
39. Yu, W.; Chen, X.; Chen, Y.; Yi, Q.; Zhang, Y. Effects of Environmental Variations on the Abundance of Western Winter-Spring Cohort of Neon Flying Squid (*Ommastrephes bartramii*) in the Northwest Pacific Ocean. *Acta Oceanol. Sin.* **2015**, *34*, 43–51. [CrossRef]
40. Ravindra, K.; Rattan, P.; Mor, S.; Aggarwal, A.N. Generalized Additive Models: Building Evidence of Air Pollution, Climate Change and Human Health. *Environ. Int.* **2019**, *132*, 104987. [CrossRef]
41. Yu, J.; Hu, Q.; Tang, D.; Zhao, H.; Chen, P. Response of *Sthenoteuthis oualaniensis* to Marine Environmental Changes in the North-Central South China Sea Based on Satellite and in Situ Observations. *PLoS ONE* **2019**, *14*, e0211474. [CrossRef]
42. Wood, S.N.; Pya, N.; Säfken, B. Smoothing Parameter and Model Selection for General Smooth Models. *J. Am. Stat. Assoc.* **2016**, *111*, 1548–1563. [CrossRef]
43. Callens, A.; Morichon, D.; Abadie, S.; Delpey, M.; Liquet, B. Using Random Forest and Gradient Boosting Trees to Improve Wave Forecast at a Specific Location. *Appl. Ocean. Res.* **2020**, *104*, 102339. [CrossRef]
44. Welchowski, T.; Maloney, K.O.; Mitchell, R.; Schmid, M. Techniques to Improve Ecological Interpretability of Black-Box Machine Learning Models: Case Study on Biological Health of Streams in the United States with Gradient Boosted Trees. *J. Agric. Biol. Environ. Stat.* **2022**, *27*, 175–197. [CrossRef] [PubMed]
45. Ahmad, G.N.; Fatima, H.; Ullah, S.; Saidi, A.S. Efficient Medical Diagnosis of Human Heart Diseases Using Machine Learning Techniques with and without GridSearchCV. *IEEE Access* **2022**, *10*, 80151–80173. [CrossRef]
46. Ranjan, G.S.K.; Verma, A.K.; Radhika, S. K-Nearest Neighbors and Grid Search Cv Based Real Time Fault Monitoring System for Industries. In Proceedings of the 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India, 29–31 March 2019; pp. 1–5.
47. Alhakeem, Z.M.; Jebur, Y.M.; Henedy, S.N.; Imran, H.; Bernardo, L.F.; Hussein, H.M. Prediction of Ecofriendly Concrete Compressive Strength Using Gradient Boosting Regression Tree Combined with GridSearchCV Hyperparameter-Optimization Techniques. *Materials* **2022**, *15*, 7432. [CrossRef] [PubMed]
48. Jung, Y.; Hu, J. A K-Fold Averaging Cross-Validation Procedure. *J. Nonparametric Stat.* **2015**, *27*, 167–179. [CrossRef] [PubMed]
49. Martínez-Rincón, R.O.; Ortega-García, S.; Vaca-Rodríguez, J.G. Comparative Performance of Generalized Additive Models and Boosted Regression Trees for Statistical Modeling of Incidental Catch of Wahoo (*Acanthocybium solandri*) in the Mexican Tuna Purse-Seine Fishery. *Ecol. Model.* **2012**, *233*, 20–25. [CrossRef]

50. Wu, J.; Dai, L.; Dai, X.; Tian, S.; Liu, J.; Chen, J.; Wang, X.; Wang, J. Comparison of generalized additive model and boosted regression tree in predicting fish community diversity in the Yangtze River Estuary, China. *Chin. J. Appl. Ecol.* **2019**, *30*, 644–652. [CrossRef]
51. Pöyry, J.; Luoto, M.; Heikkinen, R.K.; Saarinen, K. Species Traits Are Associated with the Quality of Bioclimatic Models. *Global Ecol. Biogeogr.* **2008**, *17*, 403–414. [CrossRef]
52. Lu, H.-J.; Ou, Y.-Z.; He, J.-R.; Zhao, M.-L.; Chen, Z.-Y.; Chen, X.-J. Age, Growth and Population Structure Analyses of the Purpleback Flying Squid *Sthenoteuthis oualaniensis* in the Northwest Indian Ocean by Beak Microstructure. *J. Mar. Sci. Eng.* **2022**, *10*, 1094. [CrossRef]
53. Liu, B.L.; Chen, X.J.; Wang, X.H.; Du, F.Y.; Fang, Z.; Xu, L.L. Geographic, Intraspecific and Sexual Variation in Beak Morphology of Purpleback Flying Squid (*Sthenoteuthis oualaniensis*) throughout Its Distribution Range. *Mar. Freshw. Res.* **2018**, *70*, 417–425. [CrossRef]
54. Chen, J.; Zhao, G.; Zhang, S.; Cui, X.; Tang, F.; Chen, F.; Han, H. Study on temporal and spatial distribution characteristics of *Symplectoteuthis oualaniensis* in high seas fishing ground of northwest Indian Ocean. *J. Fish. China.* 2023; *early view*.
55. Wen, L.; Zhang, H.; Fang, Z.; Chen, X. Spatial and Temporal Distribution of Fishing Ground of *Sthenoteuthis oualaniensis* in Northern Indian Ocean with Different Fishing Methods. *J. Shanghai Ocean Univ.* **2021**, *30*, 1079–1089.
56. Wei, J.; Cui, G.; Xuan, W.; Tao, Y.; Su, S.; Zhu, W. Effects of SST and Chl-a on the spatiotemporal distribution of *Sthenoteuthis oualaniensis* fishing ground in the Northwest Indian Ocean. *J. Fish. Sci. China* **2022**, *29*, 388–397.
57. Wen, L.; Zhang, H.; Fang, Z.; Chen, X. Preliminary standardization of *Sthenoteuthis oualaniensis* in northern Indian Ocean. *Trans. Oceanol. Limnol.* **2022**, *44*, 89–97. [CrossRef]
58. Hood, R.R.; Beckley, L.E.; Wiggert, J.D. Biogeochemical and Ecological Impacts of Boundary Currents in the Indian Ocean. *Prog. Oceanogr.* **2017**, *156*, 290–325. [CrossRef]
59. Yan, L.; Zhang, P.; Yang, B.; Chen, S.; Li, Y.; Li, Y.; Song, P.; Lin, L. Relationship between the catch of *Symplectoteuthis oualaniensis* and surface temperature and the vertical temperature structure in the South China Sea. *J. Fish. Sci. China* **2016**, *23*, 469–477.
60. Wang, T.; Yang, L. Combining GRA with a Fuzzy QFD Model for the New Product Design and Development of Wickerwork Lamps. *Sustainability* **2023**, *15*, 4208. [CrossRef]
61. Li, J.; Zhang, P.; Yan, L.; Wang, T.; Yang, B. Factors That Influence the Catch per Unit Effort of *Sthenoteuthis oualaniensis* in the Central-Southern South China Sea Based on a Generalized Additive Model. *J. Fish. Sci. China* **2020**, *27*, 906–915.
62. Wang, X.; Zhang, W.; Wang, P.; Yang, J.; Wang, H. Research on Mid-Depth Current of Basin Scale in the South China Sea Based on Historical Argo Observations. *Acta Oceanol. Sin.* **2018**, *40*, 1–14. [CrossRef]
63. Fu, C.; Luo, Y.; Yang, L.; Wang, D.; Zhou, F. Variations of Extreme Wave Climate of Tropical Ocean and Monsoon Region in the North Indian Ocean-South China Sea. *J. Nanjing Univ. Inf. Sci. Technol. Nat. Sci. Ed.* **2018**, *10*, 379–385.
64. Chesalin, M.V.; Zuyev, G.V. Pelagic Cephalopods of the Arabian Sea with an Emphasis on *Sthenoteuthis oualaniensis*. *Bull. Mar. Sci.* **2002**, *71*, 209–221.
65. Lachkar, Z.; Lévy, M.; Smith, K.S. Strong Intensification of the Arabian Sea Oxygen Minimum Zone in Response to Arabian Gulf Warming. *Geophys. Res. Lett.* **2019**, *46*, 5420–5429. [CrossRef]
66. Lachkar, Z.; Lévy, M.; Smith, S. Intensification and Deepening of the Arabian Sea Oxygen Minimum Zone in Response to Increase in Indian Monsoon Wind Intensity. *Biogeosciences* **2018**, *15*, 159–186. [CrossRef]
67. Yang, X.; Chen, X.; Zhou, Y.; Tian, S. A marine remote sensing-based preliminary analysis on the fishing ground of purple flying squid *Sthenoteuthis oualaniensis* in the northwest Indian Ocean. *J. Fish. China* **2006**, *30*, 669–675.
68. Lin, D.; Chen, X. Fishing Ground Distribution of Symplectoteuthis ouslsniensis and Its Relations to SST in the Northwestern Indian Ocean. *Adv. Mar. Sci.* **2006**, *24*, 546–551.
69. Shao, F.; Chen, X. Relationship between fishing ground of *Symlectoteuthis oualaniensis* and sea surface height in the northwest Indian ocean. *Mar. Sci.* **2008**, *32*, 88–92.
70. Jayarathna, W.N.D.S.; Du, Y.; Zhang, H.; Sun, Q. Seasonal and Interannual Variability of Sea Surface Salinity in the Central North Arabian Sea Based on Satellite and Argo Observations. *J. Nanjing Univ. Inf. Sci. Technol.* **2018**, *10*, 311–323.
71. Jiang, S.; Li, J.; Zhang, S.; Gu, Q.; Lu, C.; Liu, H. Landslide Risk Prediction by Using GBRT Algorithm: Application of Artificial Intelligence in Disaster Prevention of Energy Mining. *Process Saf. Environ. Prot.* **2022**, *166*, 384–392. [CrossRef]
72. Fan, J.; Yu, W.; Ma, S.; Chen, Z. Spatio-temporal variability of habitat distribution of *Sthenoteuthis oualaniensis* in South China Sea and its interannual variation. *S. China Fish. Sci.* **2022**, *18*, 1–9.
73. Xu, H. Analysis on distribution of habitant and key environmental factors for the purpleback flying squid (*Sthenoteuthis oualaniensis*) in the open South China Sea. Master's Thesis, Shanghai Ocean University, Shanghai, China, 2017.
74. Zhou, W.F.; Xu, H.Y.; Li, A.Z.; Cui, X.S.; Chen, G.B. Comparison of Habitat Suitability Index Models for Purpleback Flying Squid (*Sthenoteuthis oualaniensis*) in the Open South China Sea. *Appl. Ecol. Environ. Res.* **2019**, *17*, 4903–4913. [CrossRef]
75. Fan, J.; Chen, Z.; Feng, X.; Yu, W. Climate-Related Changes in Seasonal Habitat Pattern of *Sthenoteuthis oualaniensis* in the South China Sea. *Ecosyst. Health Sustain.* **2021**, *7*, 1926338. [CrossRef]

76. Zhou, X.; Ma, S.; Cai, Y.; Yu, J.; Chen, Z.; Fan, J. The Influence of Spatial and Temporal Scales on Fisheries Modeling—An Example of *Sthenoteuthis oualaniensis* in the Nansha Islands, South China Sea. *J. Mar. Sci. Eng.* **2022**, *10*, 1840. [CrossRef]
77. Lu, H.; Tong, Y.; Liu, W.; Liu, K.; Dong, Z.; Cheng, X.; Chen, X. Fisheries biological characteristics of *Sthenoteuthis oualaniensis* in the spring season in the El Nino year of 2016 in the Zhongsha Islands waters of South China Sea. *J. Fish. China* **2018**, *42*, 912–921.