

Article

Predicting the Potential Habitat Distribution of *Scomber japonicus* in the High Seas of the Northwest Pacific Ocean Using MaxEnt and GARP Models

Zechen Zhu ¹ and Bilin Liu ^{1,2,3,4,*} 

¹ College of Marine Living Resource Sciences and Management, Shanghai Ocean University, Shanghai 201306, China; 13651756586@163.com

² National Engineering Research Center for Distant-Water Fisheries, Shanghai 201306, China

³ The Key Laboratory of Sustainable Exploration of Oceanic Fisheries Resources, Ministry of Education, Shanghai 201306, China

⁴ Key Laboratory of Oceanic Fisheries Exploration, Ministry of Agriculture and Rural Affairs, Shanghai 201306, China

* Correspondence: bl-liu@shou.edu.cn

Abstract

Accurate prediction of the potential habitat distribution of *Scomber japonicus*, an important target species in China's distant-water fisheries, is essential for fishing ground forecasting. Using catch data for *S. japonicus* collected from Chinese large-scale purse-seine and trawl fisheries in the Northwest Pacific Ocean from May to November during 2015–2024, this study applied the maximum entropy model (MaxEnt) and the genetic algorithm for rule-set production (GARP) model to predict the potential habitat distribution of *S. japonicus* in the Northwest Pacific Ocean. The area under the receiver operating characteristic curve (AUC) and the true skill statistic (TSS) were used to evaluate model performance. The MaxEnt model predicted a relatively concentrated highly suitable habitat, whereas the GARP model identified a broader highly suitable area. To reduce the bias and uncertainty associated with single-model predictions, the outputs of the MaxEnt and GARP models were integrated using a weighted ensemble approach, with the optimal weights for MaxEnt and GARP determined as 0.7:0.3. The ensemble model achieved higher predictive performance, with an AUC of 0.983 and a TSS of 0.840. The highly suitable habitat of *S. japonicus* was mainly concentrated within 147° E–156° E and 40° N–44° N. Chlorophyll concentration, sea surface temperature (SST), and temperatures at depths of 150 m and 200 m were the main environmental variables affecting the potential habitat distribution of *S. japonicus* in the MaxEnt model. These findings provide useful information for resource utilization, fishing ground forecasting, and sustainable management of *S. japonicus* in the high seas of the Northwest Pacific Ocean.



Academic Editors: Margarita M. Rincón Hidalgo and Heng Zhang

Received: 1 June 2026

Revised: 17 June 2026

Accepted: 24 June 2026

Published: 25 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](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: *Scomber japonicus*; habitat suitability; MaxEnt; GARP; ensemble model; Northwest Pacific Ocean; fishing ground forecasting

Key Contribution: This study integrates MaxEnt, GARP, and a MaxEnt–GARP weighted ensemble model to improve the prediction of potential suitable habitats for *Scomber japonicus* in the high seas of the Northwest Pacific Ocean. The findings identify the core suitable habitat and key environmental drivers, providing scientific support for fishing ground forecasting and sustainable resource management.

1. Introduction

S. japonicus is one of the most important commercial pelagic fish species in the Northwest Pacific Ocean. It is characterized by a wide distribution, rapid growth, large biomass, and high economic value, and it is also an important target species in China's distant-water fisheries [1]. The Northwest Pacific Ocean is strongly influenced by the Kuroshio, the Oyashio, and their extensions, which together generate complex and highly variable marine environmental conditions. This region serves as an important feeding ground and migratory corridor for many fish species [2]. As a long-distance migratory species, the fishing grounds and spatial distribution of *S. japonicus* are closely related to marine environmental variability, including SST, SSS, chlorophyll concentration and ocean circulation processes [3,4].

Ecological niche models, also referred to as species distribution models, relate species occurrence records to environmental predictors and have been widely used to predict species distributions and evaluate habitat suitability [5]. Therefore, selecting appropriate marine environmental predictors is essential for improving the accuracy of potential habitat identification for *S. japonicus* and for clarifying its responses to environmental variability. Among these models, the maximum entropy model (MaxEnt) and the genetic algorithm for rule-set production model (GARP) are two representative approaches. MaxEnt is based on the principle of maximum entropy and is particularly suitable for presence-only data. It can effectively characterize relationships between species distributions and environmental variables and generally shows high predictive performance [6]. In contrast, GARP uses genetic algorithms to generate rule sets from species occurrence records and environmental variables, thereby estimating potential species distributions and showing advantages in representing spatially continuous suitable habitats [7]. However, because of differences in algorithmic structure and spatial representation, different ecological niche models may produce divergent predictions for the potential habitat distribution of the same species, and predictions based on a single model may involve substantial uncertainty. Ensemble modelling can reduce model-dependent uncertainty by integrating predictions from algorithms with different assumptions, structures, and error-control strategies. By combining complementary model outputs, ensemble approaches often produce more robust and stable habitat predictions than single-model approaches. Previous studies have shown that ensemble models can improve predictive performance and the stability of results [8,9]. Therefore, constructing an ensemble model based on the comparison of MaxEnt and GARP can improve the reliability of potential habitat identification.

In recent years, studies on the potential habitat and fishing ground distribution of *S. japonicus* have mainly relied on single-model approaches. For example, Xue et al. used MaxEnt to predict the potential habitat distribution of *S. japonicus* and reported that the predicted suitable areas were generally consistent with actual fishing occurrence records [10]. In addition to ecological niche modeling, Han et al. developed a deep-learning-based fishing-ground prediction model and revealed pronounced seasonal shifts in the fishing grounds of *S. japonicus* in the Northwest Pacific Ocean [11]. However, systematic comparisons among different modeling approaches for *S. japonicus* remain limited, resulting in uncertainty regarding the stability and reliability of its potential habitat predictions. To address this gap, the present study constructed MaxEnt and GARP models using catch data for *S. japonicus* and marine environmental variables from the high seas of the Northwest Pacific Ocean during 2015–2024. The predictive performance and spatial differences in the two models were compared, and a MaxEnt-GARP ensemble model was further developed to determine the optimal weighting scheme. The main spatial patterns and key environmental drivers of the potential habitat of *S. japonicus* were then analyzed to pro-

vide scientific support for resource utilization, fishing ground forecasting, and sustainable fisheries management.

2. Materials and Methods

2.1. Data Sources

The fishery data used in this study were obtained from catch records of *S. japonicus* collected by large-scale purse seine and trawl fisheries provided by the National Data Centre for Distant-Water Fisheries of China. The dataset covered the period from May to November during 2015–2024 and included fishing location, fishing time, vessel information, and other relevant records. The fishing area extended from 146° E to 161° E and from 34° N to 48° N. To reduce sampling bias caused by spatial clustering and potential overfitting associated with uneven fishing effort, occurrence records were spatially thinned according to the $0.25^\circ \times 0.25^\circ$ resolution of the environmental raster layers, and only one occurrence record was retained within each grid cell. After this filtering procedure, 2414 occurrence records were used for model construction (Figure 1).

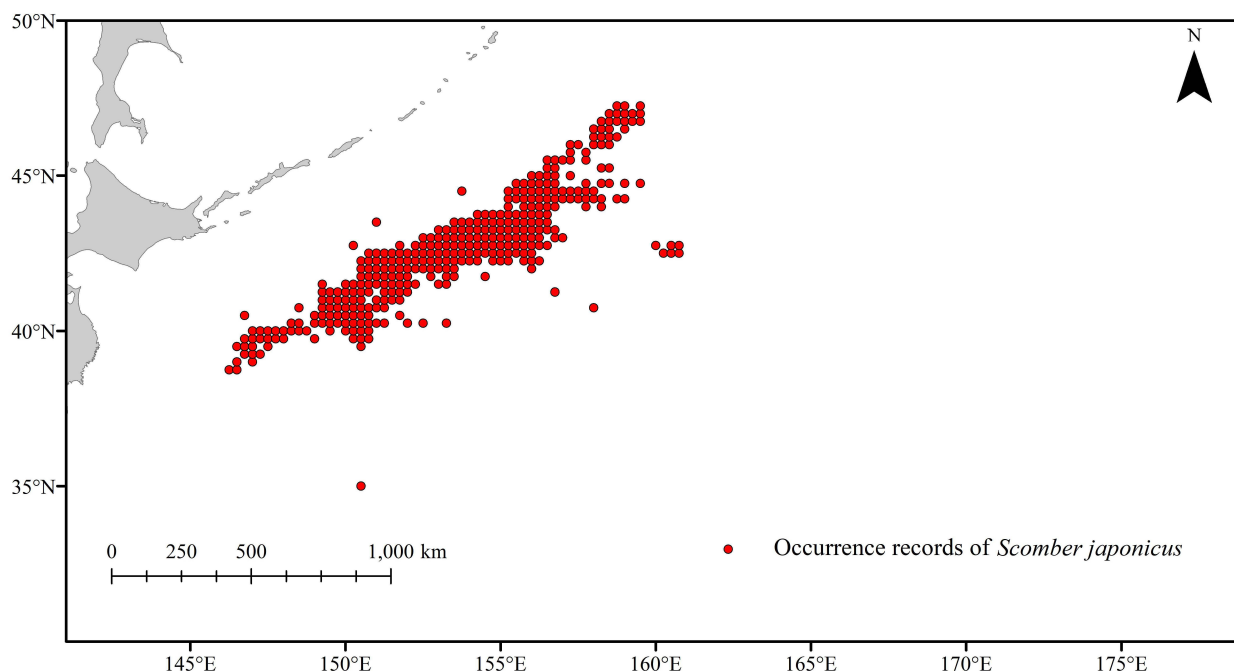


Figure 1. Fishing occurrence records of *S. japonicus* from Chinese purse-seine and trawl fleets in the high seas of the Northwest Pacific Ocean from May to November during 2015–2024.

Marine environmental variables were downloaded from the Copernicus Marine Service, operated by Mercator Ocean International (Toulouse, France), available at <https://data.marine.copernicus.eu/> (accessed on 28 February 2026). Based on the biological characteristics and habitat depth of *S. japonicus*, ten environmental variables were selected as candidate predictors for model construction, including SST, chlorophyll concentration, eastward sea water velocity (ESWV), northward sea water velocity (NSWV), sea surface height (SSH), SSS, and temperatures at depths of 50 m, 100 m, 150 m, and 200 m. These variables had a monthly temporal resolution and a spatial resolution of $0.25^\circ \times 0.25^\circ$. To match the temporal coverage of the fishery data, monthly environmental layers from May to November during 2015–2024 were processed and averaged to generate long-term seasonal mean environmental layers for the study period. Using ArcGIS 10.8, the environmental data downloaded in NetCDF format from the Copernicus Marine Environment Monitoring Service were converted into TIFF format. The TIFF files were then averaged

using the Raster Calculator and converted into ASCII raster files. Finally, mask extraction was performed to ensure consistency in the study area, grid size, and numbers of rows and columns.

2.2. Environmental Variables

To improve the robustness of variable selection, a multi-method approach was used to identify environmental predictors. For the GARP model, variable importance was evaluated using a jackknife approach based on omission rates. Specifically, changes in the average omission rate were compared between models constructed using only a given variable and models constructed after excluding that variable. This procedure was used to identify environmental variables that contributed substantially to model prediction. For the MaxEnt model, the importance of each environmental variable was comprehensively evaluated based on percent contribution, permutation importance, and jackknife test results. In addition, correlation analysis of the candidate variables was conducted using ENMTools1.3, and highly correlated variables were removed to reduce the effects of multicollinearity on model stability. Based on the combined results of these analyses, the temperature at 50 m depth and northward seawater velocity were excluded. Eight key environmental variables, namely chlorophyll concentration, SST, ESWV, SSH, SSS, T-100, T-150, and T-200, were retained for subsequent modelling (Table 1).

Table 1. Marine environmental variables.

Marine Environmental Variables	Code	Unit
Sea surface temperature	SST	°C
Temperature at 50 m depth	T-50	°C
Temperature at 100 m depth	T-100	°C
Temperature at 150 m depth	T-150	°C
Temperature at 200 m depth	T-200	°C
Chlorophyll concentration	-	mg/m ³
Eastward sea water velocity	ESWV	m/s
Northward sea water velocity	NSWV	m/s
Sea surface height	SSH	m
Sea surface salinity	SSS	PSU

2.3. Model Settings and Optimization

2.3.1. MaxEnt Model

The MaxEnt model was constructed using MaxEnt 3.4.4. The occurrence records of *S. japonicus* and the selected environmental layers were imported into the model. For model calibration, background points were randomly sampled from the masked study extent defined by the final environmental raster layers. The masked study extent was kept consistent with the prediction area in terms of spatial resolution, grid size, number of rows and columns, and coordinate system. The maximum number of background points was set to 10,000. These background points were used to characterize the available environmental conditions against which the occurrence records were contrasted. The subsampling method was used for resampling, and the model output format was set to logistic. During model construction, 75% of the occurrence records were randomly selected as the training dataset, whereas the remaining 25% were used as the testing dataset. The model was run 20 times, and the mean raster output of the 20 replicates was used as the final prediction.

To reduce overfitting associated with default parameter settings, the feature class (FC) and regularization multiplier (RM) were optimized using the kuenm package 1.1.10 in R 4.5.2. The optimal parameter combination was determined as RM = 1.3 and FC = LQH. In the initial MaxEnt model, all ten candidate environmental variables were included, and

the percent contribution and permutation importance of each variable were calculated. This MaxEnt-based screening was conducted together with the correlation analysis and the GARP jackknife test. Specifically, variables were retained or excluded based on a combined evaluation of multicollinearity, MaxEnt variable importance, and changes in GARP 1.1.6 omission rate after variable removal, rather than on a single criterion alone. Variables with very low percent contribution or permutation importance in MaxEnt were considered for removal only after confirming their limited or redundant contribution using the correlation analysis and GARP jackknife results.

Finally, eight key environmental variables, namely chlorophyll concentration, SST, ESWV, SSH, SSS, T-100, T-150, and T-200, were selected to predict the suitable habitat of *S. japonicus*. To facilitate comparison of the spatial prediction results among these three models, habitat suitability values generated by the model were classified into four levels: values below 0.2 were defined as unsuitable habitat, values from 0.2 to 0.4 as low-suitability habitat, values from 0.4 to 0.6 as moderately suitable habitat, and values above 0.6 as highly suitable habitat.

2.3.2. GARP Model

The GARP model was constructed using Desktop GARP 1.1.6. The preprocessed environmental variables were imported into the Dataset Manager module, and the occurrence records of *S. japonicus* were imported into the Upload Data Points module and matched with the corresponding environmental layers at the same temporal scale. Because the fishery dataset consisted of presence-only fishing occurrence records, no independently surveyed true-absence dataset was available. Therefore, no external absence dataset was introduced for GARP modelling. The GARP model was calibrated using the occurrence records and environmental layers within the same masked study extent as that used for MaxEnt, thereby ensuring that the two models were constructed within a consistent environmental domain. To reduce the influence of the stochasticity in the GARP model on prediction results, the Rule Types method was used to test different combinations of Atomic, Range, Negated Range, and Logistic Regression rules. Model performance under different rule structures was compared, and the rule combination with the highest predictive performance was selected for final model construction.

Each model group was run 100 times, with the maximum number of iterations set to 1000 and the convergence threshold set to 0.01. As in the MaxEnt model, 75% of the occurrence records were used as the training dataset, and the remaining 25% were used as the testing dataset. Among the tested rule-type combinations, Range Rules showed the best predictive performance and were therefore selected for the GARP model.

The jackknife test was then performed by sequentially removing one environmental variable at a time and rebuilding the model to calculate the average omission rate after variable removal. In total, ten variable combinations were constructed, each excluding one environmental variable. Each combination was run 20 times, and the average omission rate and model outputs were recorded. A marked decrease in the average omission rate after a variable was removed indicated that the variable contributed relatively little to the model or contained redundant information. Conversely, an increase in the omission rate after removing a variable indicated that the variable played an important role in model fitting (Table 2).

Table 2. Changes in the average omission rate of the GARP model after removing each environmental variable.

Excluded Environmental Variable	Average Omission (Ext)	Difference from Baseline
Full model	1.718909	-
Sea surface temperature	1.863636	0.144727
Temperature of the 50 m depth	1.602272	−0.116637
Temperature of the 100 m depth	1.718181	−0.000728
Temperature of the 150 m depth	1.736363	0.017454
Temperature of the 200 m depth	1.852272	0.133363
Chlorophyll concentration	1.909090	0.190181
Eastward sea water velocity	1.954545	0.235636
Northward sea water velocity	1.704545	−0.014364
Sea surface height	1.840909	0.122000
Sea surface salinity	1.727271	0.008362

Finally, eight key environmental variables, namely chlorophyll concentration, SST, ESWV, SSH, SSS, T-100, T-150, and T-200, were selected, consistent with those used in the MaxEnt model. The GARP model was run 100 times using Range Rules to predict the suitable habitat of *S. japonicus*. The final prediction was obtained by overlaying the raster outputs from the 100 model runs. Habitat suitability values were classified into four levels: values below 0.2 were defined as unsuitable habitat, values from 0.2 to 0.4 as low-suitability habitat, values from 0.4 to 0.6 as moderately suitable habitat, and values above 0.6 as highly suitable habitat.

2.3.3. MaxEnt-GARP Ensemble Model

To reduce the bias inherent in single-model predictions and improve the robustness of the results, the outputs of the MaxEnt and GARP models were integrated using a weighted ensemble approach. The logistic output of MaxEnt represents habitat suitability values ranging from 0 to 1, whereas the 100 overlaid outputs of the GARP model can be normalized to the same range. The outputs of both models were converted into raster-based habitat suitability values on a unified scale. Assuming that the weight assigned to MaxEnt is w and that assigned to GARP is $1 - w$, the ensemble model can be expressed as follows:

$$HSI = W \cdot HSI_{MaxEnt} + (1 - w) \cdot HSI_{GARP} \quad (1)$$

In this study, w was set to 0.1, 0.2, 0.3, ..., and 0.9, resulting in nine ensemble schemes. The receiver operating characteristic (ROC) curve and the area under the curve (AUC) were calculated for each scheme, and the scheme with the highest AUC value was selected as the optimal ensemble model. The MaxEnt:GARP weight ratio of 0.7:0.3 achieved the highest predictive performance and was therefore used as the optimal weighting scheme.

2.3.4. Model Evaluation

To compare the predictive performance of the GARP, MaxEnt, and ensemble models for the potential habitat of *S. japonicus*, the area under the receiver operating characteristic curve (AUC) and the true skill statistic (TSS) were used as model evaluation metrics [12,13]. An AUC value closer to 1 indicates a stronger ability to distinguish suitable habitats from unsuitable habitats. In general, AUC values greater than 0.9 indicate excellent model performance, values between 0.8 and 0.9 indicate good performance, values between 0.7 and 0.8 indicate fair performance, and values below 0.6 indicate poor performance [14].

TSS is widely used to evaluate the performance of species distribution models because it integrates both sensitivity and specificity and is not affected by species prevalence [15].

TSS values range from -1 to 1 , where -1 indicates completely incorrect prediction, 0 indicates random prediction, and 1 indicates perfect prediction. In general, TSS values greater than 0.6 indicate good predictive performance, whereas values greater than 0.8 indicate excellent predictive performance [16].

In this study, AUC and TSS were calculated based on the random training–testing split of the occurrence dataset, with 75% of the occurrence records used for model calibration and 25% used for testing. Therefore, these metrics primarily reflect the internal discriminatory performance of the models. Because the occurrence records used in this study were derived from fishery-dependent data, independent-year validation or spatial block cross-validation would provide a more rigorous assessment of model transferability.

3. Results

3.1. Comparison of Model Predictions and Predictive Performance

The MaxEnt model identified a concentrated highly suitable habitat for *S. japonicus* within 147° E– 156° E and 40° N– 44° N. This habitat formed a continuous southwest–northeast-oriented core distribution belt and accounted for 2.3% of the total study area. Moderately suitable habitat was mainly distributed around the periphery of the highly suitable habitat, accounting for 4.0% of the total study area. Low-suitability habitat was primarily located along the outer edge of the potential habitat, forming a belt-like distribution surrounding the moderately and highly suitable habitats. Several scattered low-suitability areas were also observed in the eastern and central-eastern parts of the study area, accounting for 4.9% of the total study area (Figure 2).

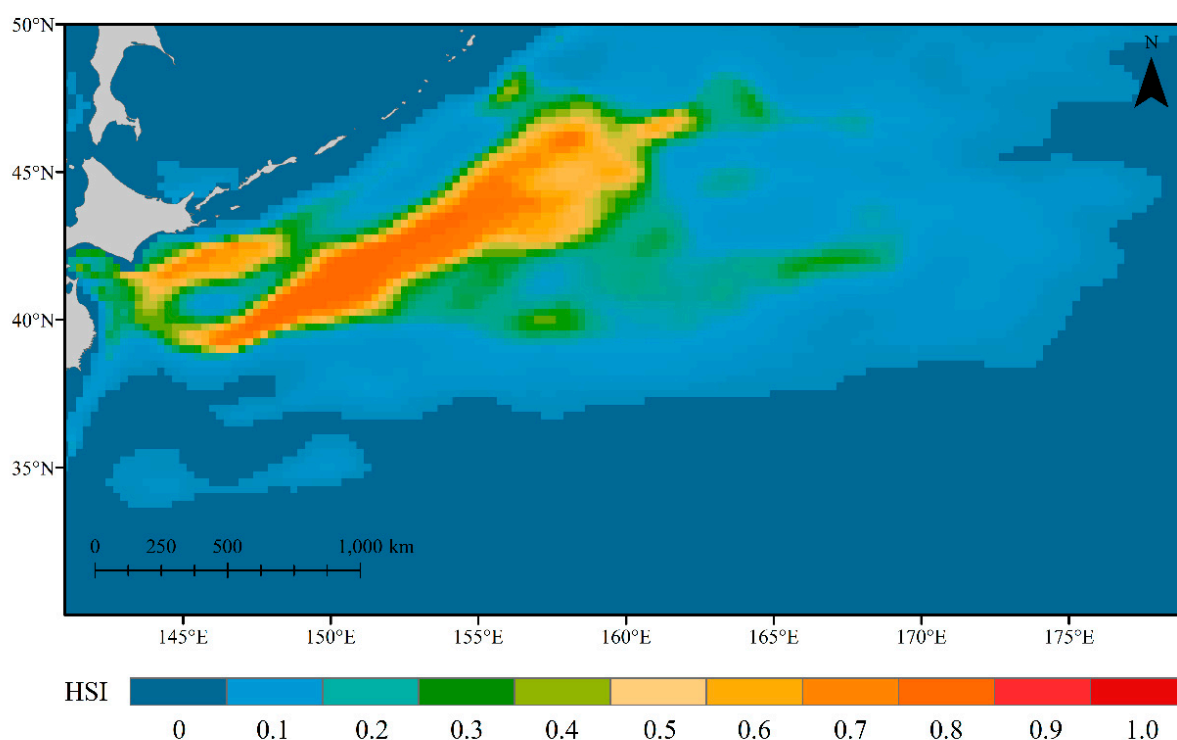


Figure 2. Suitable habitat distribution of *S. japonicus* in the Northwest Pacific Ocean from MaxEnt.

The GARP model predicted a broader highly suitable habitat for *S. japonicus*, mainly within 148° E– 170° E and 40° N– 46° N. This area showed a large, continuous distribution pattern and accounted for 32.2% of the total study area. The moderately suitable habitat was mainly distributed outside the highly suitable habitat, particularly in the northeastern and eastern waters, where it served as a transitional zone along the edge of the highly suitable

habitat and accounted for 5.6% of the total study area. The low-suitability habitat was mainly scattered along the margins of the potential habitat distribution area, accounting for 2.3% of the total study area (Figure 3).

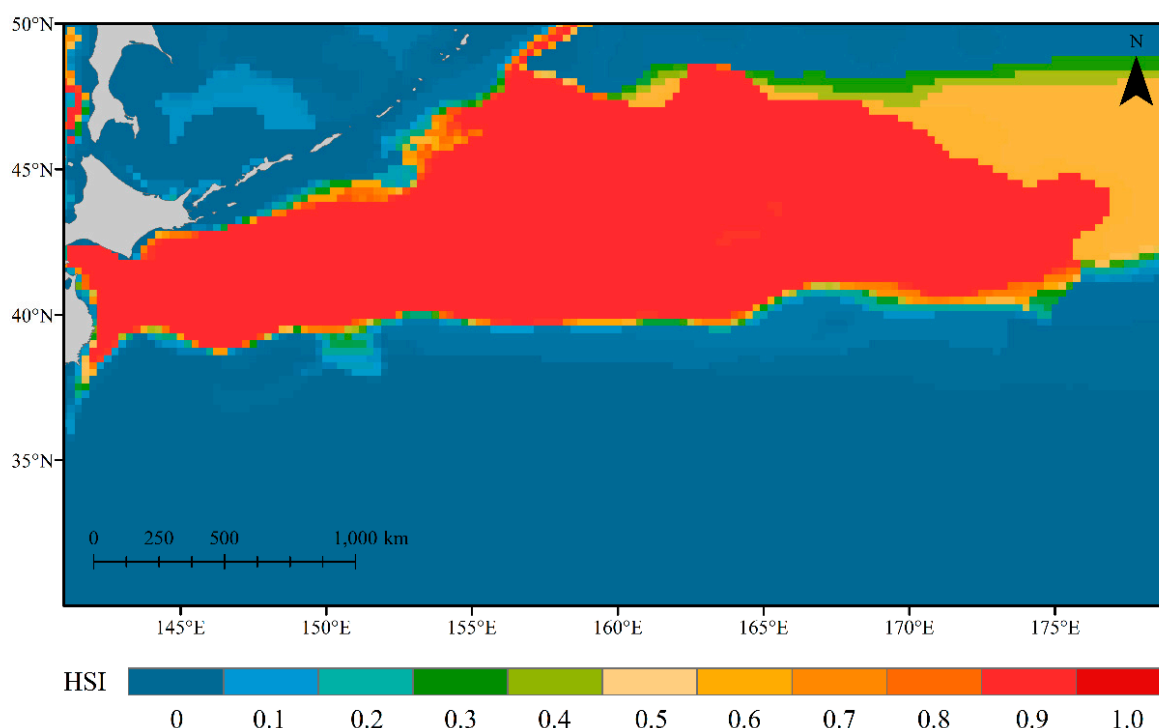


Figure 3. Suitable habitat distribution of *S. japonicus* in the Northwest Pacific Ocean from GARP.

The MaxEnt and GARP models produced markedly different spatial predictions. The suitable habitat predicted by MaxEnt was relatively concentrated, whereas the highly suitable habitat predicted by GARP covered a broader area. To integrate the strengths of both algorithms, a MaxEnt-GARP ensemble model was developed to predict the potential habitat distribution of *S. japonicus*. The ensemble model achieved the highest predictive performance when the MaxEnt and GARP outputs were weighted at 70% and 30%, with an AUC of 0.983 and a TSS of 0.840. Compared with the MaxEnt model, the ensemble model increased the AUC from 0.956 to 0.983 and the TSS from 0.811 to 0.840. Compared with the GARP model, the ensemble model increased the AUC from 0.943 to 0.983 and the TSS from 0.796 to 0.840 (Table 3). The spatial pattern of habitat suitability predicted by the ensemble model was generally similar to that predicted by MaxEnt. Compared with the MaxEnt prediction, the proportions of highly suitable, moderately suitable, and low-suitability habitats in the total study area increased by 3.4, 4.6, and 13.3 percentage points, respectively (Figure 4).

Table 3. AUC and TSS values of the GARP, MaxEnt and ensemble models.

Models	AUC	TSS
MaxEnt	0.956	0.811
GARP	0.943	0.796
MaxEnt-GARP	0.983	0.840

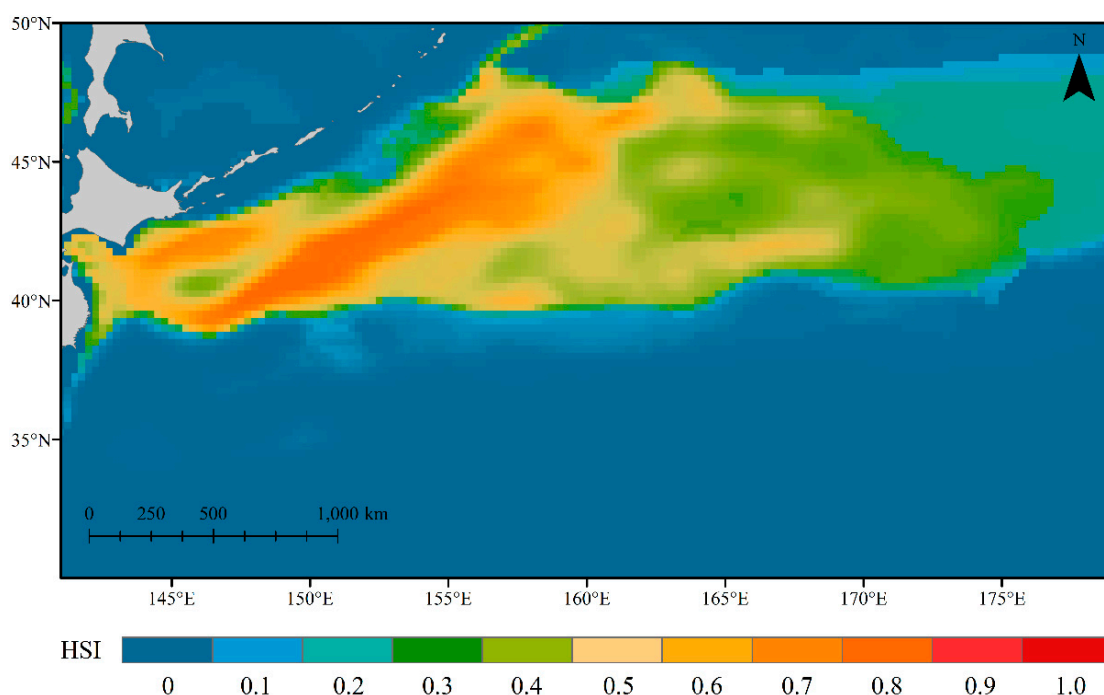


Figure 4. Suitable habitat distribution of *S. japonicus* in the Northwest Pacific Ocean from MaxEnt-GARP.

3.2. Analysis of Environmental Variables

The GARP jackknife test showed that the average omission rate decreased markedly only after the exclusion of T-50 and NSWV, suggesting that these two variables contributed relatively little to model performance or contained redundant information. Therefore, eight environmental variables, namely chlorophyll concentration, SST, ESWV, SSH, SSS, T-100, T-150, and T-200, were retained for subsequent modelling.

The percent contribution and permutation importance of each environmental variable were calculated using the MaxEnt model. Chlorophyll concentration made the largest contribution to the model, with a percent contribution of 48.8% and a permutation importance of 30.2%. SST ranked second, with a percent contribution of 16.7% and a permutation importance of 20.4%. In addition to surface environmental variables, T-150 and T-200 also showed relatively high importance. In particular, T-150 had the highest permutation importance, indicating strong independent explanatory power in the model (Table 4). ESWV showed a moderate percent contribution but low permutation importance, whereas SSH, T-100, and SSS had relatively weak overall effects.

Table 4. Percent contribution and permutation importance of environmental variables in the MaxEnt model.

Environmental Variable	Percent Contribution	Permutation Importance
Chlorophyll concentration	48.8	30.2
SST	16.7	20.4
ESWV	12.7	1.3
T-200	11	5.3
SSH	6.7	8.3
T-150	3.6	31.7
SSS	0.3	1
T-100	0.2	1.9

Based on the percent contribution and permutation importance values, chlorophyll concentration, SST, T-150, and T-200 were identified as the main environmental variables affecting the potential habitat distribution of *S. japonicus* in the MaxEnt model. Therefore, the response curves of these four variables were selected for further analysis. The response curves of chlorophyll concentration and SST were unimodal. The predicted occurrence probability reached its maximum when chlorophyll concentration ranged from 0.40 to 0.45 mg/m³ and when SST ranged from 14 to 15 °C, with maximum occurrence probabilities of approximately 0.60 and 0.64, respectively. Occurrence probability decreased markedly when chlorophyll concentration was below 0.35 mg/m³, or when SST was below 9 °C or above 20 °C. In contrast, the response curves of T-150 and T-200 were both bimodal. High occurrence probabilities for T-150 occurred at approximately 2.7–2.9 °C and 6.0–6.5 °C, whereas those for T-200 occurred at approximately 3.0–3.2 °C and 5.5–6.3 °C (Figure 5). These results indicate that *S. japonicus* exhibits a strong preference for specific mid-depth thermal structures.

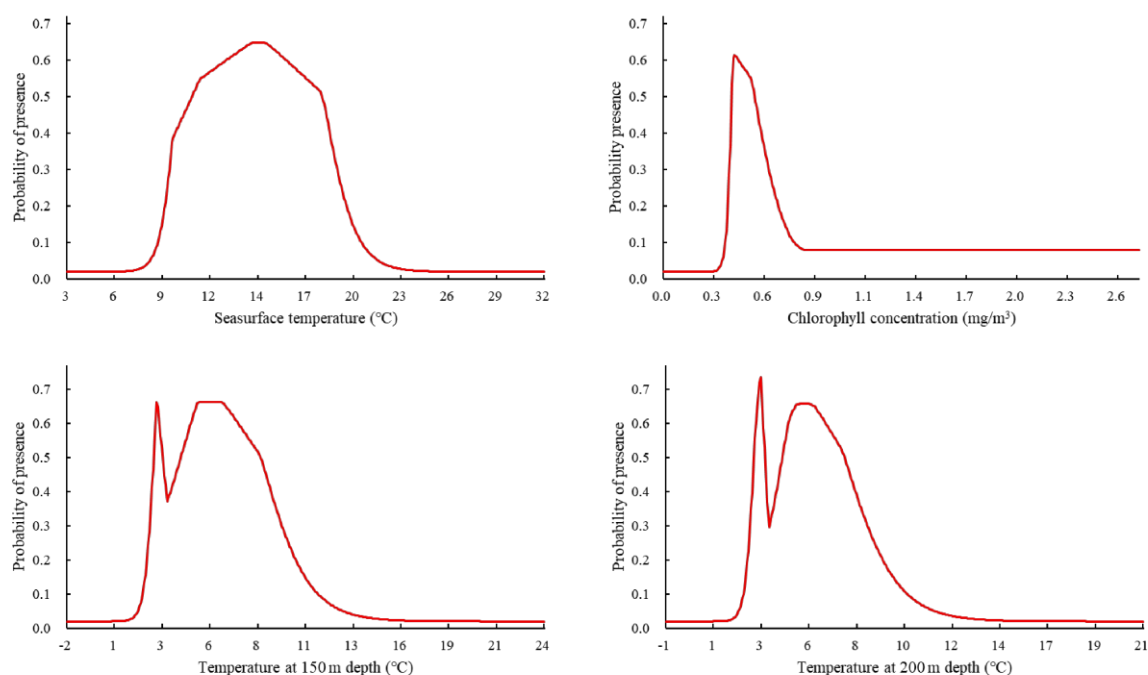


Figure 5. Response curves of predicted occurrence probability of *S. japonicus* to chlorophyll concentration, sea surface temperature, and temperatures at 150 m and 200 m depths.

4. Discussion

4.1. Comparison Among Different Models

The MaxEnt and GARP models produced generally consistent broad-scale distribution patterns for *S. japonicus*, both identifying the central-to-central-eastern part of the study area as the main potential suitable habitat, with an overall southwest–northeast orientation. However, the two models differed markedly in the spatial extent and classification of suitable habitats. The highly suitable habitat predicted by MaxEnt was relatively concentrated, with moderately suitable and low-suitability habitats gradually extending outward from the core area. In contrast, the highly suitable habitat predicted by the GARP model covered a broader area, with moderately suitable and low-suitability habitats mainly distributed along its outer margins.

The southwest–northeast distribution pattern identified in this study is broadly consistent with previous ensemble-model predictions for *S. japonicus*, which also identified the central Northwest Pacific Ocean as a major highly suitable habitat with a distinct belt-like

distribution [17]. This agreement supports the reliability of the broad-scale spatial pattern predicted in the present study. Although the overall spatial pattern was consistent with previous findings, the MaxEnt and GARP models differed substantially in their detailed predictions, mainly because of differences in their algorithmic mechanisms and error-control strategies. MaxEnt is suitable for presence-only data and places greater emphasis on reducing commission errors; therefore, it tends to identify relatively concentrated core suitable habitats [18]. In contrast, GARP generates rule sets using a genetic algorithm and provides a broader environmental envelope for suitable conditions. It also places greater emphasis on reducing omission errors, thereby highlighting the overall spatial connectivity of suitable habitats [19]. These differences explain why GARP generated a much broader highly suitable area, whereas MaxEnt produced a more spatially restricted and clearly delineated core habitat.

The divergence between the two single-model outputs indicates that predictions based on a single algorithm may involve considerable uncertainty. Ensemble models can integrate the predictive advantages of different algorithms, reduce the bias and uncertainty associated with single models, and thereby improve the stability and accuracy of habitat prediction [20]. The advantage of ensemble modelling is not limited to improved statistical performance; it can also reduce extreme spatial predictions produced by individual models, such as overly conservative or overly expansive habitat estimates, thereby generating a distribution pattern that more closely reflects the species' realized habitat [21]. In this study, the MaxEnt-GARP ensemble model outperformed the single models in terms of both AUC and TSS. Spatially, the ensemble prediction retained the core habitat identified by MaxEnt while moderately expanding the surrounding suitable areas, thereby reducing the overly expansive pattern generated by GARP.

The MaxEnt-GARP ensemble model identified the highly suitable habitat of *S. japonicus* mainly within 147° E–156° E and 40° N–44° N, with an overall southwest–northeast orientation. This spatial pattern is ecologically plausible because it is consistent with independent fishing-ground prediction studies based on actual fishery data. Independent evidence from fishing-ground prediction studies further supports this spatial pattern. Tang et al. developed a potential fishing-ground prediction model for *S. japonicus* in the Northwest Pacific using Chinese fishery catch data and marine environmental variables, and they validated the model with observed fishery data from April to November 2020 [22]. The close agreement between predicted and observed fishing locations in their study further supports the ecological plausibility of the MaxEnt-GARP ensemble prediction in the present study. Future studies should incorporate spatiotemporal machine-learning methods to reduce biases associated with unfished areas and temporal distributional shifts, thereby improving the stability of habitat identification and fishing-ground prediction for *S. japonicus*.

Although the ensemble model showed higher AUC and TSS values than the two single models, these evaluation metrics were derived from a random split of the same occurrence dataset and should therefore be interpreted mainly as measures of internal discriminatory performance. Because the available fishery data from 2015 to 2024 were used to characterize the long-term potential habitat distribution of *S. japonicus*, a strict independent-year validation was not conducted in the present study. In addition, fishery-dependent occurrence records may contain preferential sampling bias because fishing vessels tend to operate in known or expected fishing grounds rather than sampling the entire marine environment randomly. Therefore, future studies should incorporate independent fishery datasets, observer-based survey data, or spatial and temporal block cross-validation to further evaluate the transferability and generalization ability of the models.

The choice of background or pseudo-absence data is another important source of uncertainty in presence-only habitat modelling. In MaxEnt, a broader background extent may emphasize environmental contrasts between occurrence locations and the wider oceanic domain, whereas a narrower background extent may focus more strongly on habitat differentiation within the accessible area of the species. Similarly, different pseudo-absence or background treatments in GARP may affect omission and commission errors and may partly explain the broader suitable habitat predicted by GARP. In the present study, a consistent masked study extent was used for both models to improve comparability. Nevertheless, future work should test alternative background sampling strategies, such as spatially stratified background sampling, target-group background, or bias-corrected background sampling, to better account for fishery-dependent sampling bias.

As a post hoc spatial concordance check, the final MaxEnt–GARP ensemble prediction was overlaid with the observed fishing occurrence records. The results showed that 91.1% of the occurrence records were located within highly suitable habitats ($HSI > 0.6$), and 97.3% were located within moderate-to-high suitability habitats ($HSI > 0.4$). This high overlap indicates that the predicted suitable habitat belt was broadly consistent with the observed fishing distribution. However, because these occurrence records were not fully independent from the model calibration data, this result should be interpreted as supporting spatial concordance evidence rather than as independent validation.

4.2. Effects of Environmental Variables on Habitat Distribution

The Northwest Pacific Ocean is influenced by the confluence of the Kuroshio Extension and Oyashio waters, producing complex water-mass structures and pronounced spatiotemporal variability in marine environmental conditions that substantially affect the spatial distribution of pelagic migratory fishes. In this study, chlorophyll concentration, SST, T-150, and T-200 were identified as the main environmental variables affecting the potential habitat distribution of *S. japonicus* in the high seas of the Northwest Pacific Ocean, which is consistent with previous studies [23,24].

Chlorophyll concentration is widely used as an indirect indicator of primary productivity and, more broadly, lower-trophic-level productivity. The association between suitable habitats and relatively high chlorophyll concentration suggests that productivity-related food availability may play an important role in habitat selection by *S. japonicus*. This interpretation is supported by Yang et al., who reported a nonlinear response indicating that chlorophyll-related productivity conditions may influence migration pathways [25]. At broader spatial and temporal scales, chlorophyll concentration has also been recognized as an important indicator of marine productivity and fishery resource distribution [26].

Temperature reflects water-mass thermal conditions and can affect fish metabolism, activity, and migratory behavior [27]. In this study, the highly suitable habitat of *S. japonicus* corresponded to a relatively narrow SST range, indicating that surface thermal conditions strongly constrained habitat suitability. This finding is consistent with habitat suitability analyses in Korean waters, where SST variation was closely associated with shifts in the habitat center of *S. japonicus* [4]. At broader temporal scales, variability in *S. japonicus* abundance in the Northwest Pacific Ocean has been linked to SST anomalies and western boundary currents, including the Kuroshio and Oyashio [28]. These findings suggest that suitable thermal zones are important for the spatial aggregation and fishing ground formation of *S. japonicus*.

In addition to SST, the bimodal response curves of T-150 and T-200 indicate that subsurface thermal structure also contributes to habitat suitability. This vertical component is supported by evidence from the western North Pacific, where seasonal changes in the school depth and habitat temperature of *S. japonicus* have been reported based on purse-

seine fishery records [29]. Therefore, the importance of T-150 and T-200 suggests that habitat selection by *S. japonicus* is not controlled solely by surface temperature but is also associated with the thermal structure of the upper and middle water layers.

Beyond local thermal and productivity conditions, the potential habitat of *S. japonicus* in the high seas of the Northwest Pacific Ocean is also shaped by large-scale oceanographic processes associated with the Kuroshio–Oyashio transition zone. The interaction between the warm Kuroshio Extension and the cold, nutrient-rich Oyashio waters can generate strong thermal gradients, water-mass convergence, frontal systems, and enhanced lower-trophic-level productivity [30]. These physical–biological processes may promote prey aggregation and improve foraging conditions, thereby contributing to the formation of suitable habitats and fishing grounds for *S. japonicus*.

Frontal systems and mesoscale eddies may further influence the spatial aggregation and migration behavior of *S. japonicus* by modifying local temperature, nutrient supply, retention, and transport processes. Such processes can affect prey distribution patterns and may increase prey encounter rates and residence time in suitable habitats. The importance of T-150 and T-200 in this study also suggests that the habitat suitability of *S. japonicus* is related not only to surface thermal conditions but also to subsurface thermal structure, which may reflect thermocline conditions, vertical habitat accessibility, and prey depth distribution. This interpretation is consistent with evidence that larvae and juveniles of *S. japonicus* can be transported to downstream feeding and nursery areas in the Kuroshio–Oyashio transition region, where favorable prey conditions support early growth and survival [31]. Future studies should incorporate high-resolution dynamic oceanographic variables, such as sea surface height anomaly, geostrophic current velocity, eddy kinetic energy, frontal intensity, and prey-field indicators, to further clarify how frontal systems, mesoscale eddies, and prey distribution regulate habitat formation and migration behavior of *S. japonicus*.

5. Conclusions

S. japonicus is an important commercial fishery resource in the Northwest Pacific Ocean, and identifying its potential habitat distribution is essential for fishing ground forecasting and sustainable resource management. In this study, MaxEnt, GARP, and a MaxEnt-GARP ensemble model were used to predict the potential suitable habitat of *S. japonicus* in the high seas of the Northwest Pacific Ocean based on catch data and marine environmental variables from 2015 to 2024. The MaxEnt and GARP models produced markedly different spatial predictions, whereas the ensemble model improved predictive performance and provided a more balanced representation of suitable habitats. The highly suitable habitat was mainly concentrated within 147° E–156° E and 40° N–44° N, extending in a southwest–northeast direction. Chlorophyll concentration, SST, and temperatures at depths of 150 m and 200 m were identified as the main environmental variables affecting habitat suitability. These findings provide useful information for identifying core fishing grounds, forecasting potential fishing areas, optimizing fishing operations, and supporting the sustainable management of *S. japonicus* resources in the high seas of the Northwest Pacific Ocean.

Author Contributions: B.L.: Conceptualization; resources; funding acquisition; supervision; project administration; writing—review and editing. Z.Z.: Methodology; formal analysis; investigation; writing—original draft; software; validation; data curation; visualization; resources. All authors have read and agreed to the published version of the manuscript.

Funding: This work was sponsored by the Nation Key R&D Program of China (2023YFD2401302); Follow-up program for the Professor of Special Appointment (Eastern Scholar) at Shanghai Institu-

tions of Higher Learning under Contract (GZ2022011); Monitoring and Assessment of Global Fishery Resources (comprehensive scientific survey of fisheries' resources at the high seas).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The marine environmental data used in this study are publicly available from the Copernicus Marine Service. The fishery catch data are not publicly available due to restrictions from the data provider, but they may be made available from the corresponding author upon reasonable request and with permission from the National Data Centre for Distant-Water Fisheries of China. The data supporting the main findings of this study are included in the article.

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

References

- Shiraishi, T.; Okamoto, K.; Yoneda, M.; Sakai, T.; Ohshimo, S.; Onoe, S.; Yamaguchi, A.; Matsuyama, M. Age Validation, Growth and Annual Reproductive Cycle of Chub Mackerel *Scomber japonicus* off the Waters of Northern Kyushu and in the East China Sea. *Fish. Sci.* **2008**, *74*, 947–954. [\[CrossRef\]](#)
- Watanabe, C.; Nishida, H. Development of Assessment Techniques for Pelagic Fish Stocks: Applications of Daily Egg Production Method and Pelagic Trawl in the Northwestern Pacific Ocean. *Fish. Sci.* **2002**, *68*, 97–100. [\[CrossRef\]](#) [\[PubMed\]](#)
- Li, G.; Chen, X.; Lei, L.; Guan, W. Distribution of Hotspots of Chub Mackerel Based on Remote-Sensing Data in Coastal Waters of China. *Int. J. Remote Sens.* **2014**, *35*, 4399–4421. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Elith, J.; Leathwick, J.R. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annu. Rev. Ecol. Evol. Syst.* **2009**, *40*, 677–697. [\[CrossRef\]](#)
- Elith, J.; Phillips, S.J.; Hastie, T.; Dudik, M.; Chee, Y.E.; Yates, C.J. A Statistical Explanation of MaxEnt for Ecologists. *Divers. Distrib.* **2011**, *17*, 43–57. [\[CrossRef\]](#)
- Anderson, R.P.; Lew, D.; Peterson, A.T. Evaluating Predictive Models of Species' Distributions: Criteria for Selecting Optimal Models. *Ecol. Model.* **2003**, *162*, 211–232. [\[CrossRef\]](#)
- Koc, D.E.; Ustaoglu, B.; Biltekin, D. Effect of Climate Change on the Habitat Suitability of the Relict Species *Zelkova carpinifolia* Spach Using Ensembled Species Distribution Modelling. *Sci. Rep.* **2024**, *14*, 27967. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hao, T.; Elith, J.; Guillera-Aroita, G.; Lahoz-Monfort, J.J. A Review of Evidence about Use and Performance of Species Distribution Modelling Ensembles like BIOMOD. *Divers. Distrib.* **2019**, *25*, 839–852. [\[CrossRef\]](#)
- Xue, J.; Fan, W.; Tang, F.; Guo, G.; Wei, T.; 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. [\[CrossRef\]](#)
- Han, H.; Yang, C.; Jiang, B.; Shang, C.; Sun, Y.; Zhao, X.; Xiang, D.; Zhang, H.; Shi, Y. Construction of Chub Mackerel (*Scomber japonicus*) Fishing Ground Prediction Model in the Northwestern Pacific Ocean Based on Deep Learning and Marine Environmental Variables. *Mar. Pollut. Bull.* **2023**, *193*, 115158. [\[CrossRef\]](#) [\[PubMed\]](#)
- Fielding, A.H.; Bell, J.F. A Review of Methods for the Assessment of Prediction Errors in Conservation Presence/Absence Models. *Environ. Conserv.* **1997**, *24*, 38–49. [\[CrossRef\]](#)
- Liu, C.; White, M.; Newell, G. Measuring and Comparing the Accuracy of Species Distribution Models with Presence-Absence Data. *Ecography* **2011**, *34*, 232–243. [\[CrossRef\]](#)
- Manel, S.; Williams, H.C.; Ormerod, S.J. Evaluating Presence-Absence Models in Ecology: The Need to Account for Prevalence. *J. Appl. Ecol.* **2001**, *38*, 921–931. [\[CrossRef\]](#)
- Allouche, O.; Tsoar, A.; Kadmon, R. Assessing the Accuracy of Species Distribution Models: Prevalence, Kappa and the True Skill Statistic (TSS). *J. Appl. Ecol.* **2006**, *43*, 1223–1232. [\[CrossRef\]](#)
- Liu, C.; White, M.; Newell, G. Selecting Thresholds for the Prediction of Species Occurrence with Presence-Only Data. *J. Biogeogr.* **2013**, *40*, 778–789. [\[CrossRef\]](#)
- Sun, Y.; Zhang, H.; Jiang, K.; Xiang, D.; Shi, Y.; Huang, S.; Li, Y.; Han, H. Simulating the Changes of the Habitats Suitability of Chub Mackerel (*Scomber japonicus*) in the High Seas of the North Pacific Ocean Using Ensemble Models under Medium to Long-Term Future Climate Scenarios. *Mar. Pollut. Bull.* **2024**, *207*, 116873. [\[CrossRef\]](#) [\[PubMed\]](#)
- Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum Entropy Modeling of Species Geographic Distributions. *Ecol. Model.* **2006**, *190*, 231–259. [\[CrossRef\]](#)
- Stockwell, D.; Peters, D. The GARP Modelling System: Problems and Solutions to Automated Spatial Prediction. *Int. J. Geogr. Inf. Sci.* **1999**, *13*, 143–158. [\[CrossRef\]](#)

20. Araujo, M.B.; New, M. Ensemble Forecasting of Species Distributions. *Trends Ecol. Evol.* **2007**, *22*, 42–47. [[CrossRef](#)] [[PubMed](#)]
21. Ramirez-Reyes, C.; Nazeri, M.; Street, G.; Jones-Ferrand, D.T.; Vilella, F.; Evans, K.O. Embracing Ensemble Species Distribution Models to Inform At-Risk Species Status Assessments. *J. Fish. Wildl. Manag.* **2021**, *12*, 98–111. [[CrossRef](#)]
22. 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](#)]
23. Chen, X.; Li, G.; Feng, B.; Tian, S. Habitat Suitability Index of Chub Mackerel (*Scomber japonicus*) from July to September in the East China Sea. *J. Oceanogr.* **2009**, *65*, 93–102. [[CrossRef](#)]
24. Yasuda, T.; Kinoshita, J.; Niino, Y.; Okuyama, J. Vertical Migration Patterns Linked to Body and Environmental Temperatures in Chub Mackerel. *Prog. Oceanogr.* **2023**, *213*, 103017. [[CrossRef](#)]
25. Yang, S.; Wan, L.; Yu, L.; Shi, J.; Zhou, W.; Zhang, S.; Wang, F.; Wu, Z.; Dai, Y.; Jiang, K.; et al. The Environmental Niche of the Light Purse Seine Fleet in the Northwest Pacific Ocean Based on Automatic Identification System Data. *J. Mar. Sci. Eng.* **2024**, *12*, 1944. [[CrossRef](#)]
26. Han, H.; Shang, C.; Jiang, B.; Wang, Y.; Li, Y.; Xiang, D.; Zhang, H.; Shi, Y.; Jiang, K. A New Modeling Strategy for the Predictive Model of Chub Mackerel (*Scomber japonicus*) Central Fishing Grounds in the Northwest Pacific Ocean Based on Machine Learning and Operational Characteristics of the Light Fishing Vessels. *Front. Mar. Sci.* **2024**, *11*, 1451104. [[CrossRef](#)]
27. Volkoff, H.; Rønnestad, I. Effects of Temperature on Feeding and Digestive Processes in Fish. *Temperature* **2020**, *7*, 307–320. [[CrossRef](#)] [[PubMed](#)]
28. Li, J.; Zhou, W.; Dai, Y.; Tang, F.; Wu, Y.; Zhang, H.; Fan, X.; Cui, X. Effects of Western Boundary Currents and Sea Surface Temperature Anomalies on Interannual Variability of Chub Mackerel Abundance in the Northwest Pacific. *J. Sea Res.* **2025**, *203*, 102561. [[CrossRef](#)]
29. Kamimura, Y.; Taga, M.; Yukami, R.; Watanabe, C.; Furuichi, S. Intra- and Inter-Specific Density Dependence of Body Condition, Growth, and Habitat Temperature in Chub Mackerel (*Scomber japonicus*). *ICES J. Mar. Sci.* **2021**, *78*, 3254–3264. [[CrossRef](#)]
30. Yatsu, A.; Chiba, S.; Yamanaka, Y.; Ito, S.; Shimizu, Y.; Kaeriyama, M.; Watanabe, Y. Climate Forcing and the Kuroshio/Oyashio Ecosystem. *ICES J. Mar. Sci.* **2013**, *70*, 922–933. [[CrossRef](#)]
31. Guo, C.; Ito, S.; Kamimura, Y.; Xiu, P. Evaluating the Influence of Environmental Factors on the Early Life History Growth of Chub Mackerel (*Scomber japonicus*) Using a Growth and Migration Model. *Prog. Oceanogr.* **2022**, *206*, 102821. [[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.